New Results in the Simultaneous Message Passing Model via Information Theoretic Techniques

Rahul Jain
Centre for Quantum Technologies
and Department of Computer Science
National University of Singapore
Email: rahul@comp.nus.edu.sg

Hartmut Klauck
Centre for Quantum Technologies
National University of Singapore
Email: hklauck@gmail.com

***Abstract*—Consider the following *Simultaneous Message Passing* (SMP) model for computing a relation $f \subseteq \mathcal{X} \times \mathcal{Y} \times \mathcal{Z}$. In this model Alice, on input $x \in \mathcal{X}$ and Bob, on input $y \in \mathcal{Y}$, send one message each to a third party Referee who then outputs a $z \in \mathcal{Z}$ such that $(x, y, z) \in f$. We first show optimal *Direct sum* results for all relations f in this model, both in the quantum and classical settings, in the situation where we allow shared resources (shared entanglement in quantum protocols and public coins in classical protocols) between Alice and Referee and Bob and Referee and no shared resource between Alice and Bob. This implies that, in this model, the communication required to compute k simultaneous instances of f, with constant success overall, is at least k-times the communication required to compute one instance with constant success.**

This in particular implies an earlier Direct sum result, shown by Chakrabarti, Shi, Wirth and Yao [CSWY01] for the Equality function (and a class of other so-called robust functions), in the classical SMP model with no shared resources between any parties.

Furthermore we investigate the gap between the SMP model and the one-way model in communication complexity and exhibit a partial function that is exponentially more expensive in the former if quantum communication with entanglement is allowed, compared to the latter even in the deterministic case.

***Keywords*-Direct Sum, Simultaneous Message Passing, Quantum, Communication Complexity, Information Theory.**

I. Introduction

A. The Direct sum problem

The Direct sum question asks if computing k instances of a given function or relation together, with constant success overall, requires k-times the resources required for computing one instance, with constant success. It is a widely studied question and its resolution in some settings can lead to important consequences. Karchmer, Raz, and Wigderson [KRW95] show that a Direct sum result for deterministic communication complexity of certain relations would probably imply $\mathsf{NC}^1 \neq \mathsf{NC}^2$. Bar-Yossef, Jayram, Kumar, and Sivakumar [BYJKS04] use Direct sum results to prove space lower bounds in the datastream model [BYJKS04]. Pătraşcu and Thorup [PT06] use Direct sum type results to prove stronger lower bounds for approximate near-neighbor (ANN) search in the cell probe model. Work on the Direct sum property has also inspired earlier lower bounds for ANN due to Chakrabarti and Regev [CR04].

In this paper we concentrate on the Direct sum question in communication complexity. Although they seem highly plausible, it is well-known that Direct sum results fail to hold for some modes of communication. For example, testing the equality of $k = \log n$ pairs of n-bit strings with a constant-error private-coin communication protocol has complexity $O(k \log k + \log n) = O(\log n \log \log n)$ (see, e.g., [KN97, Example 4.3, page 43]), where we might expect a complexity of $\Omega(k \log n) = \Omega(\log^2 n)$.

The Direct sum property is known to hold for the nondeterministic communication complexity of total Boolean functions, and consequently a Direct sum theorem with a quadratic loss holds also for the (many round) deterministic communication complexity of such functions [KN97]. No general Direct sum theorem is known for the randomized or quantum (many round) communication complexity, although for specific functions such results are sometimes known (e.g. [KSdW07] implies a Direct sum theorem for the quantum communication complexity of Disjointness, [LSS08] for all functions for which the discrepancy bound is tight).

In the one-way model of communication complexity (in which only one message is sent from player Alice to Bob) the deterministic complexity can easily be seen to satisfy the Direct sum property, and Jain et al. [JRS05] have established Direct sum theorems for the randomized and quantum case (with public coins and entanglement respectively).

A stronger form of the Direct sum property is the so-called Strong direct product property, in which one requires that even with about k times the resources to solve one instance the success probability of solving k instances goes down exponentially in k. Such results are even more elusive than Direct sum results, and in communication complexity are known only under restrictions or for specific classes of functions/relations, see [JKN08] for a discussion.

Our results: In this paper we consider Direct sum question in certain Simultaneous Message Passing models of communication complexity and answer in the affirmative. To be more precise, let $f \subseteq \mathcal{X} \times \mathcal{Y} \times \mathcal{Z}$ be a relation, where $\mathcal{X}, \mathcal{Y}, \mathcal{Z}$ are finite sets. For a positive integer k, let us define the *k-fold product* of f, $f^{\otimes k} \subseteq \mathcal{X}^k \times \mathcal{Y}^k \times \mathcal{Z}^k$ as $f^{\otimes k} \overset{\mathsf{def}}{=} \{(x_1, \ldots x_k, y_1, \ldots, y_k, z_1, \ldots, z_k) : \forall i \quad (x_i, y_i, z_i) \in f\}$. This relation captures the problem of solving k independent instances of the relation f. Details of the SMP models we consider and the definitions of corresponding communication complexities appear in Sec. II-B. We show the following result.

Theorem 1 (Direct sum): Let $f \subseteq \mathcal{X} \times \mathcal{Y} \times \mathcal{Z}$ be a relation. Let k be a positive integer. Let $\epsilon, \delta \in (0, 1/4)$. Then,

$$1) \quad \mathsf{Q}_\epsilon^{\widetilde{\|,\mathsf{priv}}}(f^{\otimes k}) \quad \geq \quad \Omega(k \cdot \delta^3 \cdot \mathsf{Q}_{\epsilon+\delta}^{\widetilde{\|,\mathsf{priv}}}(f)) \ ,$$

$$2) \quad \mathsf{R}_\epsilon^{\widetilde{\|,\mathsf{priv}}}(f^{\otimes k}) \quad \geq \quad \Omega(k \cdot \delta^3 \cdot \mathsf{R}_{\epsilon+\delta}^{\widetilde{\|,\mathsf{priv}}}(f)) \ .$$

Here $\mathsf{Q}_\epsilon^{\widetilde{\|,\mathsf{priv}}}(f)$ denotes the communication complexity of a relation f in the quantum simultaneous message passing model with no shared resources between Alice and Bob, but shared entanglement between Alice and Referee resp. Bob and Referee. Similarly, for $\mathsf{R}_\epsilon^{\widetilde{\|,\mathsf{priv}}}(f)$ Alice and Bob share no resources, but Alice and Referee have shared access to a source of random bits (not seen by Bob), Bob and Referee access to a different source (not seen by Alice).

Using standard arguments due to Newman [New91] one can show that for any relation $f \subseteq \mathcal{X} \times \mathcal{Y} \times \mathcal{Z}$,

$$\mathsf{R}^{\widetilde{\|,\mathsf{priv}}}(f) \quad \geq \quad \Omega(\mathsf{R}^{\|,\mathsf{priv}}(f) - O(\log|\mathcal{X}| + \log|\mathcal{Y}|)) \ .$$

The super-script priv represents the model in which there is no shared resource between any pair among Alice/Bob/Referee. Hence we obtain the following corollary of Thm. 1:

Corollary 1: Let $f \subseteq \mathcal{X} \times \mathcal{Y} \times \mathcal{Z}$ be a relation and let k be a positive integer. Then,

$$\begin{aligned} \mathsf{R}^{\|,\mathsf{priv}}(f^{\otimes k}) \quad &\geq \quad \mathsf{R}^{\widetilde{\|,\mathsf{priv}}}(f^{\otimes k}) \quad \geq \quad \Omega(k \cdot \mathsf{R}^{\widetilde{\|,\mathsf{priv}}}(f)) \\ &\geq \quad \Omega(k \cdot (\mathsf{R}^{\|,\mathsf{priv}}(f) - O(\log|\mathcal{X}| + \log|\mathcal{Y}|))). \end{aligned}$$

Note that a similar result to Newman's is unknown for the quantum model (and probably does not hold), so we do not get a corresponding tight Direct sum result in the quantum case for the SMP model where no entanglement is shared between any pair among Alice/Bob/Referee.

Previous work on Direct sum in the SMP *model:* Babai and Kimmel [BK97], using arguments similar to those in Newman [New91], show the following.

Fact 1 ([BK97]): For a relation $f \subseteq \mathcal{X} \times \mathcal{Y} \times \mathcal{Z}$, let $\mathsf{D}^{\|}(f)$ represent the deterministic communication complexity for computing f in the SMP model. Then, $\mathsf{R}^{\|,\mathsf{priv}}(f) = \Omega(\sqrt{\mathsf{D}^{\|}(f)})$.

The Direct sum result for $\mathsf{D}^{\|}(f)$ is easy to show and hence one can derive the following Direct sum result for $\mathsf{R}^{\|,\mathsf{priv}}(f)$[1].

Fact 2 (Implicit from [BK97]): Let $f \subseteq \mathcal{X} \times \mathcal{Y} \times \mathcal{Z}$ be a relation and let k be a positive integer. Then,

$$\begin{aligned} \mathsf{R}^{\|,\mathsf{priv}}(f^{\otimes k}) \quad &= \quad \Omega(\sqrt{\mathsf{D}^{\|}(f^{\otimes k})}) \\ &= \quad \Omega(\sqrt{k \cdot \mathsf{D}^{\|}(f)}) \\ &= \quad \Omega(\sqrt{k \cdot \mathsf{R}^{\|,\mathsf{priv}}(f)}) \ . \end{aligned}$$

Chakrabarti, Shi, Wirth and Yao [CSWY01] consider the Direct sum problem in the private coins SMP model and show the following result. For a Boolean function $f : \{0,1\}^n \times \{0,1\}^n \to \{0,1\}$, define

$$\tilde{\mathsf{R}}^{\|,\mathsf{priv}}(f) \overset{\mathsf{def}}{=} \min_S \mathsf{R}^{\|,\mathsf{priv}}(f|_{S \times S}),$$

where S ranges over all subsets of $\{0,1\}^n$ of size at least $(\frac{2}{3})2^n$ and $f|_{S \times S}$ denotes the function f restricted to inputs x, y both from the set S. It is easily seen that $\tilde{\mathsf{R}}^{\|,\mathsf{priv}}(f) \leq \mathsf{R}^{\|,\mathsf{priv}}(f)$.

Fact 3 ([CSWY01]): Let k be a positive integer. Then,

$$\mathsf{R}^{\|,\mathsf{priv}}(f^{\otimes k}) = \Omega(k \cdot (\tilde{\mathsf{R}}^{\|,\mathsf{priv}}(f) - O(\log n))) \ .$$

For the Equality function $\mathrm{EQ}_n : \{0,1\}^n \times \{0,1\}^n \to \{0,1\}$, in which the Referee outputs 1 iff the inputs of Alice and Bob are equal, it can easily be seen that $\tilde{\mathsf{R}}^{\|,\mathsf{priv}}(\mathrm{EQ}_n) = \Theta(\mathsf{R}^{\|,\mathsf{priv}}(\mathrm{EQ}_n))$. Hence the above result, Fact 3, provides an optimal Direct sum result for EQ_n. Note that our Direct Sum result Cor. 1 in comparison holds for arbitrary relations and hence is a generalization of the above result due to [CSWY01].

In the SMP models in which Alice and Bob share public coins, optimal Direct sum results have been shown earlier by Jain, Radhakrishnan and Sen [JRS05].

Fact 4 (Direct sum [JRS05]): Let $f \subseteq \mathcal{X} \times \mathcal{Y} \times \mathcal{Z}$ be a relation. Let k be a positive integer. Let $\epsilon, \delta \in (0, 1/4)$, then

$$1) \quad \mathsf{Q}_\epsilon^{\|,\mathsf{pub}}(f^{\otimes k}) \quad \geq \quad \Omega\left(k \cdot \delta^3 \cdot \mathsf{Q}_{\epsilon+\delta}^{\|,\mathsf{pub}}(f)\right).$$

$$2) \quad \mathsf{R}_\epsilon^{\|,\mathsf{pub}}(f^{\otimes k}) \quad \geq \quad \Omega\left(k \cdot \delta^3 \cdot \mathsf{R}_{\epsilon+\delta}^{\|,\mathsf{pub}}(f)\right).$$

Note that for the Equality function there is an exponential gap between the classical randomized public coin SMP-complexity $\mathsf{R}^{\|,\mathsf{pub}}(\mathrm{EQ}_n)$, which is $\Theta(\log n)$, and the private-public coin SMP-complexity $\mathsf{R}^{\widetilde{\|,\mathsf{priv}}}(\mathrm{EQ}_n)$, which is $\Theta(\sqrt{n})$. (see e.g. [BK97]). A similar exponential gap is known for a relation in the quantum model [GKRdW06], i.e. there is a relation g with $\mathsf{R}^{\|,\mathsf{pub}}(g) = O(\log n)$ and $\mathsf{Q}^{\widetilde{\|,\mathsf{priv}}}(g) =$

[1] Note that this result is weaker than our result Cor. 1, whenever $\mathsf{R}^{\|,\mathsf{priv}}(f) = \Omega(\log|\mathcal{X}| + \log|\mathcal{Y}|)$.

$\Omega(n^{1/3})$.[2] Hence the previous results in the public model do not imply ours, and in particular any approach using arguments about distributional communication complexity cannot establish Thm. 1, due to the inherent connection between deterministic distributional complexity and public coin randomized complexity.

Our techniques: A rough overview of our proof of Direct sum for quantum protocols is as follows (the direct sum for classical protocols also follows using analogous arguments). We view Alice's map from her k inputs to her message quantum state, as a *classical-quantum channel* on k coordinates. We then use an information theoretic fact, the *superadditivity* of such channels due to [Jai05], Fact 6, to find a suitable coordinate and a *derived channel* on that coordinate which has only $1/k$ times the original *capacity*. We do the same for Bob. Now running the original protocol, with suitably fixed inputs for other coordinates, gives us a protocol for one input, which is put in the chosen coordinate, in which the capacity of derived channel has decreased by a factor of k. Hence the *mutual information* between messages and inputs is small for any distribution on the inputs (due to the definition of capacity). In this situation, we can employ an information theoretic result due to [Jai06], Fact 7, to find a single quantum state that has "good overlap" with every message state. This allows for a message compression, due to Jain, Radhakrishnan and Sen [JRS05], that works for all messages, decreasing the worst case communication to the channel capacity, which is at most $1/k$ times the original communication, while keeping the worst case error bounded. This now immediately implies the Direct sum result.

Although both we and [JRS05] derive the Direct sum result via message compression arguments, there is an important difference in our techniques. Since [JRS05] are concerned with the model in which Alice and Bob share public coins, they are able to translate to the distributional setting via the well known relationship between public coin communication complexity and the distributional communication complexity due to Yao. Hence for them message compression with average error being bounded suffices. However since we cannot translate to the distibutional setting, we have to impliment a message compression with worst case error being bounded and hence we use different arguments and techniques, involving channel capacity, as mentioned in the previous paragraph.

B. One-way vs. simultaneous messages

The other main result in this paper concerns the question of how different the simultaneous message passing model and the one-way communication complexity model really are. It is clear that one-way protocols, in which either Alice or Bob sends one message to the other player, who then outputs the result, can easily simulate simultaneous message passing protocols by "merging" either Alice or Bob with the Referee. It is well known that the one-way complexities of a function can vary exponentially depending on which player sends the message even in the quantum case, e.g. for the Index function [Nay99]. So the natural measure to compare to the SMP-complexity is the maximum of the one-way complexities over the choice of the player sending the message. We denote this complexity by $\mathsf{R}^{1,\mathsf{max}}(f)$ (we will use similar notations for the other modes of communication).

Clearly $\mathsf{R}^{1,\mathsf{max}}(f) \leq \mathsf{R}^{\|}(f)$. But how much smaller can the left hand side be compared to the SMP-complexity? Besides being interesting in its own right this question is also relevant to Direct sum results in the simultaneous message passing model, because in a situation where the two complexities coincide, a Direct sum result in the one-way model might imply a Direct sum result in the SMP-model. We point out that optimal Direct sum results in the one-way classical public and private coins models and the quantum model with entanglement, are already known [JRS05].

For deterministic complexity it is easy to see that $\mathsf{D}^{1,\mathsf{max}}(f) = \Theta(\mathsf{D}^{\|}(f))$ for all total functions f (in this case we could immediately conclude a Direct sum result for the SMP-model from the Direct Sum result in the one-way model). On the other hand Bar-Yossef et al. [BYJKS02] exhibit a total function g for which $\mathsf{R}^{1,\mathsf{max}}(g) = O(\log n)$, while $\mathsf{R}^{\|}(g) = \Omega(\sqrt{n})$.

We first generalize this result to the quantum case, showing that $\mathsf{Q}^{\|,\mathsf{pub}}(g) = \Omega(\sqrt{n})$ as well. Just like in [BYJKS02] the lower bound is based on giving a lower bound for the Generalized Addressing Function of [BGKL03]. In fact all known lower bounds for this function are based on a certain subfunction, for which $\Omega(\sqrt{n})$ is tight, whereas the exact complexity of the Generalized Addressing Function is open, see [AL00] for the best known upper bound. However, the proof of the above lower bound fails, when we allow entanglement between Alice and Bob. So we consider a different partial function f which has the desired behavior even if we allow arbitrary tripartite entanglement.

Theorem 2: There is a partial Boolean function f on n inputs such that $\mathsf{D}^{1,\mathsf{max}}(f) \leq \log n$, while $\mathsf{Q}^{\|,\mathsf{ent}}(f) \geq \Omega(\sqrt{n})$.

Note that a similar result cannot be true for a total function (the function g above only has a randomized upper bound for one-way protocols).

So at least for a partial function we get the strongest possible separation, and indeed the simultaneous message passing model turns out to be fundamentally different from the one-way model.

[2] The paper states only a $\mathsf{Q}^{\|,\mathsf{priv}}$ bound, but the proof can be extended easily. Note that the random access code arguments in the proof are still valid if Alice/Bob and Referee share entanglement by using our Fact 5. The rest of the argument goes through by considering the complete states that Referee owns (Alice/Bob's message together with the corresponding entanglement), which are still independent of each other.

C. Organization

In the next section we present the necessary definitions and facts that are subsequently used in our proofs. In Sec. III we present the proofs of our Direct sum results. Sec. IV contains the results comparing one-way- to SMP-complexity. We conclude in Sec. V with some open problems. For completeness, in Sec. A, we present the proofs of the earlier known facts that we use in this work.

II. Preliminaries

A. Information theory

For an operator A, its *trace norm* is defined to be $\|A\|_{\mathrm{tr}} \overset{\mathsf{def}}{=} \mathsf{Tr}\sqrt{A^{\dagger}A}$. We use the *bra-ket* notation in which a vector is represented as $|\phi\rangle$ and its adjoint is represented as $\langle\phi|$. A *quantum state* is a positive semi definite trace one operator. A *pure state* is a quantum state of rank one and is often represented by its sole eigenvector with non-zero eigenvalue. For a quantum state ρ in Hilbert space $\mathcal{H}$, a pure state $|\phi\rangle \in \mathcal{H}\otimes\mathcal{K}$ is called its *purification* if $\mathsf{Tr}_{\mathcal{K}}|\phi\rangle\langle\phi| = \rho$. For a quantum state ρ, its *von-Neumann entropy* is defined as $S(\rho) \overset{\mathsf{def}}{=} \sum_i -\lambda_i \log \lambda_i$, where λ_i's represent the various eigenvalues of ρ. It is easily seen that for an l qubit quantum system A with state ρ_A, $S(A) \overset{\mathsf{def}}{=} S(\rho_A) \leq l$. For systems A, B their *mutual information* is defined as $I(A:B) \overset{\mathsf{def}}{=} S(A) + S(B) - S(AB)$. Given quantum states ρ, σ, their *relative entropy* is defined as $S(\rho\|\sigma) \overset{\mathsf{def}}{=} \mathsf{Tr}\rho(\log\rho - \log\sigma)$. For a joint classical-quantum system XM, where X is a classical random variable, let state of $M|(X=x)$ be ρ_x. Let $\rho \overset{\mathsf{def}}{=} \mathbb{E}_{x\leftarrow X}[\rho_x]$. Then we have an alternate characterization of $I(X:M)$ as follows:

$$I(X:M) = \mathbb{E}_{x\leftarrow X}[S(\rho_x\|\rho)] \ . \tag{1}$$

For classical random variables the analogous definitions and facts hold *mutatis mutandis*.

B. Communication complexity

Quantum communication complexity: In a Simultaneous Message Passing (SMP) quantum communication protocol $\mathcal{P}$ for computing a relation $f \subseteq \mathcal{X}\times\mathcal{Y}\times\mathcal{Z}$, Alice and Bob get inputs $x \in \mathcal{X}$ and $y \in \mathcal{Y}$ respectively. They each send a message to a third party called Referee. The Referee then outputs a $z \in \mathcal{Z}$ such that $(x,y,z) \in f$. The internal computations and messages send by the parties can be quantum. On any input pair (x,y), the protocol can err with a small probability. 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 $(x,y,z) \in f$. Given $\epsilon \in (0,1/2)$, the communication complexity in any given model is defined to be the communication of the best SMP protocol in that model, with error at most ϵ on all inputs. There are four models of quantum SMP protocols that we consider. In the first model there is no shared resource between any of the parties and the communication complexity in this model is denoted by $\mathsf{Q}_{\epsilon}^{\|,\mathsf{priv}}(f)$. In the second model we allow prior entanglement to be shared between Alice and Referee, Bob and Referee, but no shared resource between Alice and Bob. The entangled state for Alice and Referee is independent of the entangled state for Bob and Referee. The communication complexity in this model is denoted by $\mathsf{Q}_{\epsilon}^{\|,\widetilde{\mathsf{priv}}}(f)$. In the third model, we allow prior entanglement to be shared between Alice and Referee, Bob and Referee, and public coins to be shared between Alice and Bob. The communication complexity in this model is denoted by $\mathsf{Q}_{\epsilon}^{\|,\mathsf{pub}}(f)$. Finally, in Sec. IV we will also consider the model, in which Alice, Bob, and Referee share an arbitrary entangled tripartite state and the communication complexity in this model is denoted by $\mathsf{Q}_{\epsilon}^{\|,\mathsf{ent}}(f)$. Whenever the error parameter ϵ is not specified it is assumed to be $1/3$.

Classical communication complexity: In the classical models, the internal computations by the parties and the messages sent are classical. Similar to the quantum case, we consider three models of classical SMP protocols. In the first model, there is no shared resource between any of the parties and the communication complexity is denoted by $\mathsf{R}_{\epsilon}^{\|,\mathsf{priv}}(f)$. In the second model, we let the public coins to be shared between Alice and Referee, Bob and Referee and no shared resource between Alice and Bob. The communication complexity in this model is denoted by $\mathsf{R}_{\epsilon}^{\|,\widetilde{\mathsf{priv}}}(f)$. In the third model, we let public coins to be shared between Alice and Referee, Bob and Referee and between Alice and Bob. The communication complexity in this model is denoted by $\mathsf{R}_{\epsilon}^{\|,\mathsf{pub}}(f)$. As before whenever error parameter ϵ is not specified it is assumed to be $1/3$.

C. Useful facts

Here we present some known facts that will subsequently be useful in our proofs. We provide proofs for some of them in Sec. A for completeness. We state them here in the quantum case. In the classical case, these hold *mutatis mutandis* by replacing quantum states by probability distributions and we avoid making explicit statements and proofs.

The following fact is probably folklore and appears among other places for example in [JRS05].

Fact 5: Let XMN be a tri-partite system. If $I(X:M) = 0$ then $I(X:MN) \leq 2S(N)$.

Let $\mathcal{X}$ be a finite set and let $\mathcal{S}$ be the set of all quantum states. A classical-quantum (c-q) channel E is a map from $\mathcal{X}$ to $\mathcal{S}$. A classical-classical (c-c) channel is a map from $\mathcal{X}$ to the set of probability distributions. Unless explicitly stated otherwise, all the channels we consider will be c-q channels and we will avoid mentioning c-q explicitly from now on. For a probability distribution μ over $\mathcal{X}$, let E_μ be the bipartite state $\mathbb{E}_{x\leftarrow\mu}[|x\rangle\langle x| \otimes E(x)]$. Let $I(E_\mu)$ be the mutual information between the two systems in E_μ. The channel capacity of such a channel is defined as follows.

Definition 1 (Channel capacity): Channel capacity of the channel $E : \mathcal{X} \mapsto \mathcal{S}$ is defined as $C(E) \overset{\mathsf{def}}{=} \max_\mu I(E_\mu)$.

A *derived channel* is defined as follows.

Definition 2 (Derived channel): Let $\mathcal{X}$ and $\mathcal{Y}$ be finite sets. Let $E : \mathcal{X} \times \mathcal{Y} \to \mathcal{S}$ be a channel. For a collection $\{\mu_x : x \in \mathcal{X}\}$, where each μ_x is a probability distribution on $\mathcal{Y}$, let $F : \mathcal{X} \to \mathcal{S}$ be a channel given by $F(x) \overset{\mathsf{def}}{=} \mathbb{E}_{y \leftarrow \mu_x}[E(x, y)]$. Such a channel F is referred to as an E-derived channel on $\mathcal{X}$. Similarly we can define E-derived channels on $\mathcal{Y}$ using collections of probability distributions on $\mathcal{X}$.

We will need the following result from Jain [Jai05].

Fact 6 (Super-additivity [Jai05]): Let k be a positive integer and let $\mathcal{X}_1, \mathcal{X}_2, \ldots, \mathcal{X}_k$ be finite sets. Let $E : \mathcal{X}_1 \times \cdots \times \mathcal{X}_k \to \mathcal{S}$ be a channel. For $i \in [k]$, let $\mathcal{C}_i$ be the set of all E-derived channels on $\mathcal{X}_i$. Then,

$$C(E) \quad \geq \quad \sum_{i=1}^{k} \min_{F_i \in \mathcal{C}_i} C(F_i) \ .$$

This fact might seem somewhat unusual, because the motivation here is to fix inputs to a channel on a product set, in all but one coordinate, in such a way that the capacity of the channel in the remaining coordinate becomes as small as possible. This may not appear to be a standard task in information theory where usually the motivation is to maximize channel capacity as much as possible. However for our purpose reducing the capacity of the messages on a k-fold input is exactly what we need to enable appropriate message compression and consequently establish the direct sum property.

We will also use the following result from Jain [Jai06]. An alternate proof of this fact for the special case of classical-classical (c-c) channels, can be found in [HJMR07].

Fact 7 ([Jai06]): Let $\mathcal{X}$ be a finite set and let $E : \mathcal{X} \to \mathcal{S}$ be a channel. There exists a quantum state τ such that

$$\forall x \in \mathcal{X}, \quad S(E(x) \| \tau) \quad \leq \quad C(E) \ .$$

If E is a classical-classical (c-c) channel then τ above is a classical distribution.

The above fact allows worst case message compression when the channel capacity is small: given τ we can reconstruct **any** $E(x)$ using the following compression result of [JRS05].

Fact 8 (Compression [JRS05]): Let Alice and Referee share sufficiently many copies of a bi-partite pure state $|\phi\rangle$ between them, such that the marginal of $|\phi\rangle$ on Referee's part is τ. For any state ρ and for any $\delta > 0$, Alice can measure her part of the states and send $O(\frac{1}{\delta^3} \cdot S(\rho \| \tau))$ bits to Referee, enabling Referee to pick state ρ' with him such that $\|\rho - \rho'\|_{\mathrm{tr}} \leq \delta$.

We explicitly state the classical version of the above result for clarity.

Fact 9 (Compression [JRS05]): Let Alice and Referee, using public coins, be able to sample from distribution Q, sufficiently many times. For any distribution P and for any $\delta > 0$, Alice can send $O(\frac{1}{\delta^2} \cdot S(P \| Q))$ bits to Referee, at the end of which Referee can sample from a distribution P' such that $\|P - P'\| \leq \delta$.

We will use the following relation between relative entropy and trace distance from [KNTSZ07].

Fact 10 ([KNTSZ07]): For density matrices ρ, σ :

$$\|\rho - \sigma\|_{\mathrm{tr}} \leq \sqrt{2} S(\rho \| \sigma)^{1/2}.$$

Finally, we need the *quantum random access code* bound due to Nayak [Nay99] (here also stated for the case where entanglement is allowed).

Fact 11 ([Nay99]): Assume Alice receives a uniformly random string $x \in \{0, 1\}^n$ and Bob a uniformly random index $i \in \{1, \ldots, n\}$. Alice and Bob may share entanglement, and Alice sends one message to Bob, which allows him to decode x_i with probability $1 - \epsilon$ (averaged over the inputs). Then Alice's message needs to have $(1 - H(\epsilon))n/2$ qubits, where H denotes the binary entropy function. Without entanglement the bound is $(1 - H(\epsilon))n$.

The above result is essentially a lower bound in the quantum one-way communication complexity model for a function known as the Index function. Alternatively we will refer to Alice's message as the random access code of the strings x.

III. Direct sum

In this section, we restate and subsequently prove our main result about Direct sum.

Theorem 3 (Direct sum): Let $f \subseteq \mathcal{X} \times \mathcal{Y} \times \mathcal{Z}$ be a relation. Let k be a positive integer. Let $\epsilon, \delta \in (0, 1/4)$. Then,

1) $\mathsf{Q}_{\epsilon}^{\|,\widetilde{\mathsf{priv}}}(f^{\otimes k}) \quad \geq \quad \Omega(k \cdot \delta^3 \cdot \mathsf{Q}_{\epsilon+\delta}^{\|,\widetilde{\mathsf{priv}}}(f))$.
2) $\mathsf{R}_{\epsilon}^{\|,\widetilde{\mathsf{priv}}}(f^{\otimes k}) \quad \geq \quad \Omega(k \cdot \delta^2 \cdot \mathsf{R}_{\epsilon+\delta}^{\|,\widetilde{\mathsf{priv}}}(f))$.

Proof: We state the proof of part 1 above. The proof of part 2 follows very similarly by using the classical versions of the facts used.

Let $c \overset{\mathsf{def}}{=} \mathsf{Q}_{\epsilon}^{\|,\widetilde{\mathsf{priv}}}(f^{\otimes k})$. Let $\mathcal{P}$ be an SMP protocol for $f^{\otimes k}$ with communication c and its error on all inputs being at most ϵ. Let ρ_x be the combined state of the qubits received by Referee from Alice, when Alice's input is x, and Referee's part of the shared entangled state with Alice. Similarly let σ_y be the combined state of the qubits received by Referee from Bob, when Bob's input is y, and Referee's part of the shared entangled state with Bob. Let $\mathcal{S}$ be the set of all quantum states. Let $A : \mathcal{X} \to \mathcal{S}$ be a channel given by $A(x) \overset{\mathsf{def}}{=} \rho_x$ and let $B : \mathcal{Y} \to \mathcal{S}$ be a channel given by $B(y) \overset{\mathsf{def}}{=} \sigma_y$. Using Fact 5 and the fact that for an l qubit quantum system M, $S(M) \leq l$, it can be seen that $C(A) \leq 2c$ and $C(B) \leq 2c$. From Fact 6 and using Markov's inequality, we have that there exists a coordinate $i \in [k]$ and an A-derived channel A_i on the input on the i-th coordinate and a B-derived channel B_i on the input on the i-th coordinate, such that $C(A_i) \leq \frac{4c}{k}$ and $C(B_i) \leq \frac{4c}{k}$.

We will now present a protocol $\mathcal{P}'$ for f. In $\mathcal{P}'$, Alice on input x, sends state $A_i(x)$ to Referee. Similarly Bob on input y, sends state $B_i(y)$ to Referee. Referee performs the same actions as in $\mathcal{P}$ and outputs the result corresponding to the i-th coordinate. It can be seen that the error in $\mathcal{P}'$, on any input pair (x, y) is bounded by ϵ.

Now we present the final protocol $\mathcal{P}''$. Let τ_a be the state obtained from Fact 7 such that $\forall x \in \mathcal{X}, \quad S(A_i(x)||\tau_a) \leq C(A_i)$. Similarly let τ_b be the state obtained from Fact 7 such that $\forall y \in \mathcal{Y}, \quad S(B_i(y)||\tau_b) \leq C(B_i)$. Let $|\phi_a\rangle$ be a purification of τ_a and let $|\phi_b\rangle$ be a purification of τ_b. Alice and Referee share several copies of $|\phi_a\rangle$ as shared entanglement in $\mathcal{P}''$. Bob and Referee share several copies of $|\phi_b\rangle$ as shared entanglement in $\mathcal{P}''$. Alice, on receiving input x, using Fact 8 sends $O(\frac{1}{\delta^3} \cdot S(A_i(x)\|\tau_a))$ bits to Referee at the end of which Referee has a state ρ'_x such that $\|A_i(x) - \rho'_x\|_{\mathrm{tr}} \leq \delta$. Similarly Bob, on receiving input y, using Fact 8 sends $O(\frac{1}{\delta^3} \cdot S(B_i(y)\|\tau_b))$ bits to Referee at the end of which Referee has a state σ'_x such that $\|A_i(x) - \sigma'_x\|_{\mathrm{tr}} \leq \delta$. It can be seen that the error of protocol $\mathcal{P}''$ on any input pair (x, y) is bounded by $\epsilon + \delta$.[3] Also the communication for any input pair is bounded by $O(\frac{c}{k\delta^3})$. Hence we can conclude part 1 from the definitions of $\mathsf{Q}_{\epsilon}^{\widetilde{\|,\mathsf{priv}}}(f^{\otimes k})$ and $\mathsf{Q}_{\epsilon+\delta}^{\widetilde{\|,\mathsf{priv}}}(f)$. ■

IV. Comparing simultaneous messages and one-way communication

Recall that $\mathsf{D}^{1,\mathsf{max}}(f)$ denotes the maximum of the deterministic one-way communication complexities over Alice and Bob sending the message. It is easy to see that $\mathsf{D}^{1,\mathsf{max}}(f) = \Theta(\mathsf{D}^{\|}(f))$ for all total functions f. Bar-Yossef et al. [BYJKS02] describe a total function g for which $\mathsf{R}^{1,\mathsf{max}}(g) = O(\log n)$, while $\mathsf{R}^{\|}(g) = \Omega(\sqrt{n})$. This function is a variant of the Generalized Addressing Function investigated in [BGKL03].

For g Alice receives inputs $x \in \{0,1\}^n$ and $i \in \{1, \ldots, n\}$, Bob $y \in \{0,1\}^n$ and $j \in \{1, \ldots, n\}$, and $g(x, i, y, j) = 1 \iff x = y$ and $x_{i \oplus j} = 1$, where $\oplus$ denotes the bit-wise addition modulo 2 (here assume n to be a power of two and view i, j as $\log n$ bit strings). The upper bound on $\mathsf{R}^{1,\mathsf{max}}(g)$ is straightforward and based on fingerprinting. For the lower bound one can restrict the inputs to $x = y$, and arrive at an equivalent of the 3-party number on the forehead Generalized Addressing Function from [BGKL03] over Z_2^n, for which the corresponding lower bound is $\Omega(\sqrt{n})$. In fact this lower bound can be shown for the easier problem h defined like g, except that i and j are strings of length $\log(n)/2$, and we are interested in the bit x_k where k is the concatenation of i and j. While for h the resulting lower bound is obviously tight, the exact complexity of the Generalized Addressing Function remains open [AL00].

We will describe a partial function for which $\mathsf{D}^{1,\mathsf{max}}(f) \leq \log n$, while the quantum SMP-complexity with entanglement is still $\Omega(\sqrt{n})$. But first let us generalize the result of [BYJKS02] to the quantum case. The lower bound builds on and simplifies the information theoretic part of the proof in [BGKL03]. In fact we simply reduce the problem to random access coding.

Theorem 4: $\mathsf{Q}^{\|,\mathsf{pub}}(h) = \Omega(\sqrt{n})$, while $\mathsf{R}^{1,\mathsf{max}}(h) = O(\log n)$.

Proof: We restrict the inputs to the set where $x = y$. For clarity let us first present a lower bound on $\mathsf{Q}^{\|,\mathsf{priv}}(h)$. The plan is to construct a short quantum random access code from the messages in the protocol. For fixed x Alice is left with $\sqrt{n}$ different inputs i, similarly Bob has only $\sqrt{n}$ different inputs j. Let the messages of Alice be denoted by σ_i and the messages of Bob by ρ_j. We claim that the collection of all these messages forms a random access code for x. By the correctness of the protocol Referee has a measurement that, applied to $\sigma_i \otimes \rho_j$, produces x_{ij} with high probability for all i, j, which is exactly what we require. Hence, using Fact 11 we can conclude that all the $2\sqrt{n}$ messages together must have $(1 - H(\epsilon))n$ qubits to achieve success probability $1 - \epsilon$. Consequently at least one message of the SMP-protocol must have length $\Omega(\sqrt{n})$. Note that the above lower bound holds even if the average error (over inputs distributed uniformly) is upper bounded by ϵ (instead of the worst case error being bounded by ϵ), since Fact 11 also holds for inputs distributed uniformly.

To establish the same bound in the case Alice and Referee as well as Bob and Referee share entanglement, and Alice and Bob a classical public coin, first note that since we are concerned with success probability when inputs are uniformly distributed, we can fix the public coins between Alice and Bob suitably and hence we are back to the situation where messages of Alice and Bob are independent. We can then produce a one-way protocol with entanglement for the Index function in the same way as above by composing the different messages of Alice and Bob (with Referee holding the additional entanglement). ■

The same lower bound obviously extends to the Generalized Addressing function over Z_2^n. It is easy to see that the proof can be generalized to the Generalized Addressing function over other groups and to the multiparty setting along the lines of the arguments in [BGKL03].

Now note that the above proof fails if we allow entanglement between Alice and Bob, since the messages σ_i and ρ_j will in general be entangled and so we cannot simply collect all of them while preserving the pairwise entanglement. We still conjecture the lower bound to hold for the quantum case with entanglement, but have not yet been able to show this. Instead we will construct a partial function (on n^2 inputs) for which $\mathsf{D}^{1,\mathsf{max}}(f) \leq \log n$ while every quantum SMP

[3] This follows due to the standard fact of monotonicity of trace distance under measurements.

protocol needs communication $\Omega(n)$, even if Alice, Bob, and Referee share arbitrary tripartite entanglement. Note that such a result cannot not hold for total functions.

In fact the separation we seek is easily established for the following relation s: Let Alice be given $x \in \{0,1\}^n$ and $i \in \{1,\ldots,n\}$, while Bob gets $y \in \{0,1\}^n$ and $j \in \{1,\ldots,n\}$. Solving the relation requires Referee to output either x_j or y_i (and to indicate which). Clearly, $\mathsf{D}^{1,\max}(s) \leq \log n$. On the other hand a lower bound for the quantum SMP-model can be argued along the following lines: For each input one of the two allowed outputs must be made with probability at least $(1-\epsilon)/2$ (assuming error ϵ). Hence under the uniform distribution on all inputs Referee is able to compute either x_j or y_i with probability $1/2 - \epsilon/2$. If Referee, say, can compute x_j under the uniform distribution then Referee may toss a coin in case the protocol produces the other output and this leads to a simultaneous message protocol that computes the Index function with success probability at least $3/4 - \epsilon/2$. Hence the communication must be $\Omega(n)$, even with quantum messages and arbitrary entanglement, see Fact 11.

We now describe a partial Boolean function with the same separation.

Definition 3: Let Alice receive inputs $x \in \{0,1\}^n$ and $i \in \{1,\ldots,n\}$, while Bob receives n inputs $y_1,\ldots,y_n \in \{0,1\}^n$, and $j \in \{1,\ldots,n\}$. The promise is that $y_i = x$ and the desired function value is $f(x,i,y,j) = x_j$.

Note that this function is essentially the Index function, but with enough side-information to allow it being computable by one-way protocols in both directions. Furthermore, this side-information is obfuscated in such a way as to make it useless in the SMP-model.

Theorem 5: $\mathsf{D}^{1,\max}(f) \leq \log n$, while $\mathsf{Q}^{\|,\mathsf{ent}}(f) \geq \Omega(n)$.

Proof: For the upper bound note that there are deterministic SMP-protocols, in which either Alice or Bob sends only $\log n$ bits, and the other player n bits. These protocols can be easily simulated in the one-way model.

For the lower bound we show that if Bob sends δn qubits only and the error is ϵ, then Alice must send $(1 - H(\epsilon + \sqrt{\delta}))n/2$ qubits. Hence for constant ϵ, one of the messages has length $\Omega(n)$.

Assuming that Bob sends δn qubits only, we show that an SMP protocol $\mathcal{P}$ for f (with worst case error ϵ on inputs satisfying the promise $y_i = x$) can be turned into an SMP protocol $\mathcal{P}'$ for the Index function. In protocol $\mathcal{P}'$ Alice gets input x, Bob gets input j (there are no inputs i, y) and they compute x_j with slightly larger error than ϵ (averaged over the uniform distribution on (x,j)). To achieve this we choose i in a suitable way and fix it in $\mathcal{P}$. We then show that choosing y uniformly and independent of x (instead with the correlation $y_i = x$) can cause only small extra error in computing x_j in $\mathcal{P}$. Hence we get an SMP protocol $\mathcal{P}'$ for the Index function (with y acting as private randomness of Bob). This implies the bound on Alice's message length via Fact 11, since it is easy to convert an SMP protocol to a one-way protocol. Details follow.

Let the registers X, I hold Alice's inputs, and the registers Y, J hold Bob's inputs. Denote by E_A, E_B, E_R the registers which contain the initial entangled state for Alice, Bob, and the Referee. These registers may hold an arbitrary state independent of the input. Let register M_A contain Alice's message and register M_B contain Bob's message.

Let the distribution μ be such that y, i and j are chosen uniformly and independently from their respective domains, and $x = y_i$. Let us put distribution μ on (X,I,Y,J). Now consider the situation when Bob has created his message, but neither Alice nor Referee have done anything yet (this can be assumed since Alice and Bob's operations act on different qubits). In this situation by Fact 5 we have $I(JE_AE_RM_B : Y) \leq 2|M_B|$ and hence

$$\mathbb{E}_{i\leftarrow I}[I(JE_AE_RM_B : Y_i)] \leq 2|M_B|/n = 2\delta \ . \qquad (2)$$

This can be shown using easy generalization of Fact 12 to multiple random variables and then noting that the collection $\{Y_i \ : \ i \in [n]\}$ is independent.

Denote by $\sigma_{i,x}$ the joint state of J, E_A, E_R, M_B when $I = i$ and $X = x$ (and hence $Y_i = x$). Setting $\sigma_i \overset{\mathsf{def}}{=} \mathbb{E}_{x\leftarrow X}[\sigma_{i,x}]$ we get from Eq. 2 and Eq. 1: $\mathbb{E}_{i\leftarrow I}\mathbb{E}_{x\leftarrow X}[S(\sigma_{i,x}\|\sigma_i)] \leq 2\delta$. Let $\tilde{\imath} \in [n]$ be such that $\mathbb{E}_{x\leftarrow X}[S(\sigma_{\tilde{\imath},x}\|\sigma_{\tilde{\imath}})] \leq 2\delta$. Fact 10 and concavity of the square root function now implies:

$$\mathbb{E}_{x\leftarrow X}\left\|\sigma_{\tilde{\imath},x} - \sigma_{\tilde{\imath}}\right\|_{\mathrm{tr}} \leq 2\sqrt{\delta} \ . \qquad (3)$$

Let the distribution $\mu_{\tilde{\imath}}$ be obtained from μ by fixing $i = \tilde{\imath}$. Let $\rho_{\tilde{\imath}}$ be the joint state of J, E_A, E_R, M_B, X, just after Bob has created his message in the protocol $\mathcal{P}$, when we start with distribution $\mu_{\tilde{\imath}}$ on (X,I,Y,J). Let the distribution $\mu'_{\tilde{\imath}}$ be such that all of x,y,j are chosen uniformly and independently (without any correlation between $y_{\tilde{\imath}}$ and x) and i fixed to $\tilde{\imath}$. Let $\theta_{\tilde{\imath}}$ be the joint state of J, E_A, E_R, M_B, X, just after Bob has created his message in $\mathcal{P}$, when we start with distribution $\mu'_{\tilde{\imath}}$ on (X,I,Y,J). Note that, using Eq. 3 we have,

$$\left\|\rho_{\tilde{\imath}} - \theta_{\tilde{\imath}}\right\|_{\mathrm{tr}} = \mathbb{E}_{x\leftarrow X}\left\|\sigma_{\tilde{\imath},x} - \sigma_{\tilde{\imath}}\right\|_{\mathrm{tr}} \leq 2\sqrt{\delta} \ . \qquad (4)$$

Note that the "relevant" registers for correctness of the protocol $\mathcal{P}$ (after Bob's message is generated) are only X, J, E_A, E_R, M_B (since the output needs to be X_J). When we start with distribution $\mu_{\tilde{\imath}}$ on (X,I,Y,J), the protocol $\mathcal{P}$ would be correct with probability $1-\epsilon$ (since all inputs with positive probability under $\mu_{\tilde{\imath}}$ satisfy the promise $y_i = x$), and changing the state of all "relevant" registers from $\rho_{\tilde{\imath}}$ to $\theta_{\tilde{\imath}}$ can introduce an average extra error of at most $\sqrt{\delta}$ in computing X_J (due to Eq. 4)[4].

Now consider the protocol $\mathcal{P}'$ for the Index function in which on inputs (x,j) to Alice and Bob respectively (with x,j drawn uniformly and independently), Alice fixes input i

[4] This is a standard fact that follows due to monotonicity of trace distance under admissible quantum operations.

in $\mathcal{P}$ to $\tilde{\imath}$, Bob generates a y uniformly and independent of (x, j) using private coins, and then Alice, Bob and Referee proceed with the rest of the protocol $\mathcal{P}$. Note that in this case registers (X, I, Y, J) have distribution $\mu'_{\tilde{\imath}}$ on them. Due to our earlier observation, distributional error of $\mathcal{P}'$, under $\mu'_{\tilde{\imath}}$, is at most $\epsilon + \sqrt{\delta}$. Now $\mathcal{P}'$ can trivially be turned into a one-way quantum protocol $\mathcal{P}''$ with entanglement between Alice and Bob (by letting Bob do also the role of Referee), and Alice sending the message of same length as in $\mathcal{P}'$. By Fact 11, $\mathcal{P}''$ needs communication $(1 - H(\epsilon + \sqrt{\delta}))n/2$, hence Alice's message in $\mathcal{P}'$ must be that long. ■

V. Conclusions and open problems

We have shown a tight (up to an additive log factor) Direct sum result for the randomized SMP-complexity with private coins, and a tight Direct sum result for the $\mathsf{Q}^{\|,\widetilde{\mathsf{priv}}}$ model. While for some relations like one investigated in [GKRdW06] lower bounds known for $\mathsf{Q}^{\|,\mathsf{priv}}$ can be extended to the $\mathsf{Q}^{\|,\widetilde{\mathsf{priv}}}$ model, the general relation between those models remains unknown, and is related to the general open question of how useful entanglement is in quantum communication. The main open problems here are, however, to show a Direct Sum result for the $\mathsf{Q}^{\|,\mathsf{ent}}$ model, and for the $\mathsf{Q}^{\|,\mathsf{priv}}$ model, or disprove such statements.

Furthermore we have investigated the gap between the SMP model and the one-way model. We have described an exponential gap between the fully entangled quantum SMP model and the deterministic one-way model for a partial function, which is optimal in the sense that such a gap does not hold for total functions. However, most likely there is an exponential gap between the $\mathsf{Q}^{\|,\mathsf{ent}}$ model and randomized one-way complexity for the (total function variant) Generalized Addressing function, but we have only been able to lower bound the $\mathsf{Q}^{\|,\mathsf{pub}}$ complexity of this problem. Finally, lower bounds for this function in any mode that exceed the $\sqrt{n}$ barrier, or improved upper bounds would be very interesting.

References

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

[AL00] A. Ambainis and S. V. Lokam. Improved upper bounds on the simultaneous messages complexity of the generalized addressing function. In *Proceedings of LATIN'2000*, pages 135–147, 2000.

[BGKL03] L. Babai, A. Gal, P. G. Kimmel, and S. V. Lokam. Simultaneous messages vs. communication. *SIAM Journal on Computing*, 33 No.1:137–166, 2003.

[BK97] L. Babai and P.G. Kimmel. Randomized simultaneous messages. In *Proceedings of the 12th Annual IEEE Symposium on Computational Complexity*, pages 239–246, 1997.

[BYJKS02] Ziv Bar-Yossef, T. S. Jayram, R. Kumar, and S. Sivakumar. Information theory methods in communication complexity. In *Proceedings of the 17th Annual IEEE Conference on Computational Complexity*, pages 93–102, 2002.

[BYJKS04] Ziv Bar-Yossef, T. S. Jayram, Ravi 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] Amit Chakrabarti and Oded Regev. An optimal randomised cell probe lower bound for approximate nearest neighbour searching. In *Proceedings of the 45th 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.

[GKRdW06] D. Gavinsky, J. Kempe, O. Regev, and R. de Wolf. Bounded-error quantum state identification and exponential separations in communication complexity. In *Proceedings of the 38th Annual ACM Symposium on Theory of Computing*, pages 594–605, 2006.

[HJMR07] P. Harsha, R. Jain, D. McAllester, and J. Radhakrishnan. The communication complexity of correlation. In *Proceedings of the Twenty-Second Annual IEEE Conference on Computational Complexity (CCC)*, 2007.

[Jai05] R. Jain. A super-additivity inequality for channel capacity of classical-quantum channels. arXiv:quant-ph/0507088, 2005.

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

[JKN08] R. Jain, H. Klauck, and A. Nayak. Direct product theorems for classical communication complexity via subdistribution bounds. In *Proceedings of the 40th Annual ACM Symposium on Theory of Computing*, pages 599–608, 2008.

[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.

[JRS08] R. Jain, J. Radhakrishnan, and P. Sen. A theorem about relative entropy of quantum states with an application to privacy in quantum communication. *Journal of ACM*, 2008. To appear. Extended abstract of the paper appeared previously in Proceedings of the 43rd Annual IEEE Symposium on Foundations of Computer Science, 2002.

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

[KNTSZ07] H. Klauck, A. Nayak, A. Ta-Shma, and D. Zuckerman. Interaction in quantum communication. *IEEE Transactions on Information Theory*, 53 No.6:1970–1982, 2007.

[KRW95] Mauricio Karchmer, Ran Raz, and Avi Wigderson. Super-logarithmic depth lower bounds via direct sum in communication complexity. *Computational Complexity*, 5:191–204, 1995.

[KSdW07] H. Klauck, R. Spalek, and R. de Wolf. Quantum and classical strong direct product theorems and optimal time-space tradeoffs. *SIAM Journal on Computing*, 36 No.5:1472–1493, 2007.

[LSS08] T. Lee, A. Shraibman, and R. Spalek. A direct product theorem for discrepancy. In *Proceedings of the 23rd Annual IEEE Conference on Computational Complexity*, pages 71–80, 2008.

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

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

[PT06] Mihai Pătraşcu and Mikkel 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, 2006.

APPENDIX

Proof of Fact 5: 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:

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

Now,

$$\begin{aligned} I(X : MN) &= I(X : M) + I(XM : N) - I(M : N) \\ &\leq I(XM : N) \quad \leq \quad 2S(N) \ . \end{aligned}$$

The equality above follows easily due to definition of mutual information and the first inequality follows since $I(X : M) = 0$ and $I(M : N) \geq 0$. ■

Proof of Fact 6: We show the fact for $k = 2$, which easily implies the same for larger k. Let $\mathcal{X} \overset{\mathsf{def}}{=} \mathcal{X}_1$ and $\mathcal{Y} \overset{\mathsf{def}}{=} \mathcal{X}_2$. For each $x \in \mathcal{X}$, let $E^x : \mathcal{Y} \to \mathcal{S}$ be an E-derived channel on $\mathcal{Y}$ given by $E^x(y) \overset{\mathsf{def}}{=} E(x, y)$. For each $x \in \mathcal{X}$, let μ_x be a probability distribution on $\mathcal{Y}$ such that $I(E^x_{\mu_x}) = C(E^x)$. Now let $E^{\mathcal{X}} : \mathcal{X} \to \mathcal{S}$ be an E-derived channel on $\mathcal{X}$ given by $E^{\mathcal{X}}(x) \overset{\mathsf{def}}{=} \mathbb{E}_{y \leftarrow \mu_x}[E(x, y)]$. Let $\mu_{\mathcal{X}}$ be a distribution on $\mathcal{X}$ such that $I(E^{\mathcal{X}}_{\mu_{\mathcal{X}}}) = C(E^{\mathcal{X}})$. Let μ be the distribution on $\mathcal{X} \times \mathcal{Y}$ arising by sampling from $\mathcal{X}$ according to $\mu_{\mathcal{X}}$, and conditioned on sampling x, sampling from $\mathcal{Y}$ according to μ_x. Now the following *chain rule property* holds for mutual information.

Fact 12: Let X, Y, Z be a tripartite system where X is a classical system. Let P be the distribution of X. Then,

$$I(XY : Z) = I(X : Z) + \mathbb{E}_{x \leftarrow P}[I((Y : Z) \mid X = x)] \ .$$

Now we have,

$$\begin{aligned} C(E) &\geq I(E_\mu) \quad \text{(from definition of capacity)} \\ &= I(E^{\mathcal{X}}_{\mu_{\mathcal{X}}}) + \mathbb{E}_{x \leftarrow \mu_{\mathcal{X}}}[I(E^x_{\mu_x})] \quad \text{(from Fact 12)} \\ &= C(E^{\mathcal{X}}) + \mathbb{E}_{x \leftarrow \mu_{\mathcal{X}}}[C(E^x)] \\ &\geq \min_{F_1 \in \mathcal{C}_1} C(F_1) + \min_{F_2 \in \mathcal{C}_2} C(F_2) \ . \end{aligned}$$

This finishes the proof. ■

Proof of Fact 7: We will need the following *joint convexity* property of relative entropy. For quantum states $\rho_1, \rho_2, \sigma_1, \sigma_2$ and $p \in [0, 1]$ we have:

$$\begin{aligned} &S(p\rho_1 + (1-p)\rho_2 \| p\sigma_1 + (1-p)\sigma_2) \\ &\leq p \cdot S(\rho_1 \| \sigma_1) + (1-p) \cdot S(\rho_2 \| \sigma_2) \ . \end{aligned}$$

We will require the following minimax theorem from game theory, which is a consequence of the Kakutani fixed point theorem in real analysis.

Fact 13: Let A_1, A_2 be non-empty, convex and compact subsets of $\mathbb{R}^n$ ($\mathbb{R}$ stands for the set of real numbers) for some positive integer n. Let $u : A_1 \times A_2 \to \mathbb{R}$ be a continuous function, such that

1) $\forall a_2 \in A_2$, the set $\{a_1 \in A_1 : u(a_1, a_2) = \max_{a'_1 \in A_1} u(a'_1, a_2)\}$ is convex; and
2) $\forall a_1 \in A_1$, the set $\{a_2 \in A_2 : u(a_1, a_2) = \min_{a'_2 \in A_2} u(a_1, a'_2)\}$ is convex.

Then, there is an $(a_1^*, a_2^*) \in A_1 \times A_2$ such that

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

Let $A_1 = A_2$ be the set of all distributions on the set $\mathcal{X}$. Since $\mathcal{X}$ is finite, A_1, A_2 are convex and compact subsets of $\mathbb{R}^n$ for some n. Let $\rho_x \overset{\mathsf{def}}{=} E(x)$. For distribution μ on $\mathcal{X}$, let $\rho_\mu \overset{\mathsf{def}}{=} \mathbb{E}_{x \leftarrow \mu}[\rho_x]$. Let the function $u : A_1 \times A_2 \mapsto \mathbb{R}$ be such that $u(\lambda, \mu) \overset{\mathsf{def}}{=} \mathbb{E}_{x \leftarrow \lambda}[S(\rho_x \| \rho_\mu)]$. The condition 1 of Fact 13 can be easily seen to be satisfied since $u(\cdot, \cdot)$ is linear in the first argument. For condition 2 consider the following. Fix $\lambda \in A_1$. Let $\mu_1, \mu_2 \in A_2$ be such that $u(\lambda, \mu_1) = u(\lambda, \mu_2) = \min_{\mu'} u(\lambda, \mu')$. Let $p \in [0, 1]$; we need to show that $\mu_p \overset{\mathsf{def}}{=} p\mu_1 + (1-p)\mu_2$ satisfies $u(\lambda, \mu_p) = \min_{\mu'} u(\lambda, \mu')$. We have from joint convexity of relative entropy:

$$\begin{aligned} & \mathbb{E}_{x\leftarrow\lambda}[S(\rho_x\|\rho_{\mu_p})] \\ \leq\ & \mathbb{E}_{x\leftarrow\lambda}[p\cdot S(\rho_x\|\rho_{\mu_1}) + (1-p)\cdot S(\rho_x\|\rho_{\mu_2})] \\ =\ & p\cdot \mathbb{E}_{x\leftarrow\lambda}[S(\rho_x\|\rho_{\mu_1})] + (1-p)\cdot \mathbb{E}_{x\leftarrow\lambda}[S(\rho_x\|\rho_{\mu_2})] \\ =\ & p\cdot u(\lambda,\mu_1) + (1-p)\cdot u(\lambda,\mu_2) = \min_{\mu'} u(\lambda,\mu') \ . \end{aligned}$$

Therefore we have:

$$\begin{aligned} \min_{\mu}\max_{x} S(\rho_x||\rho_\mu) &= \min_{\mu}\max_{\lambda} \mathbb{E}_{x\leftarrow\lambda}[S(\rho_x||\rho_\mu)] \\ &= \max_{\lambda}\min_{\mu} \mathbb{E}_{x\leftarrow\lambda}[S(\rho_x||\rho_\mu)] \quad \text{(Fact 13)} \\ &\leq \max_{\lambda} \mathbb{E}_{x\leftarrow\lambda}[S(\rho_x||\rho_\lambda)] \\ &= \max_{\lambda} I(E_\lambda) = C(E) \end{aligned}$$

Therefore there exists $\tilde{\mu} \in A_2$ such that $\max_{x\in\mathcal{X}} S(\rho_x||\rho_{\tilde{\mu}}) \leq C(E)$. We let $\tau \overset{\mathsf{def}}{=} \rho_{\tilde{\mu}}$ and conclude our proof. Also it is easily seen that if E is a classical-classical (c-c) channel, then τ will be a probability distribution. ■

Proof of Fact 8: We use the following information-theoretic result called the *substate theorem* due to Jain, Radhakrishnan, and Sen [JRS08].

Fact 14 (Substate theorem [JRS08]): 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 ρ,τ be density matrices in $\mathcal{H}$ such that $S(\rho\|\tau) < \infty$. Let $|\overline{\rho}\rangle$ be a purification of ρ in $\mathcal{H}\otimes\mathcal{K}$. Then, for $r > 1$, there exist pure states $|\psi\rangle, |\theta\rangle \in \mathcal{H}\otimes\mathcal{K}$ and $|\overline{\tau}\rangle \in \mathcal{H}\otimes\mathcal{K}\otimes\mathbb{C}^2$, depending on r, such that $|\overline{\tau}\rangle$ is a purification of τ and $\||\overline{\rho}\rangle\langle\overline{\rho}| - |\psi\rangle\langle\psi|\|_{\mathrm{tr}} \leq \frac{2}{\sqrt{r}}$, where

$$|\overline{\tau}\rangle \overset{\mathsf{def}}{=} \sqrt{\frac{r-1}{r2^{rk}}}\,|\psi\rangle|1\rangle + \sqrt{1-\frac{r-1}{r2^{rk}}}\,|\theta\rangle|0\rangle$$

and $k \overset{\mathsf{def}}{=} 8S(\rho\|\tau) + 14$.

We will also require the following fact that is easily shown using *Schmidt decompositions* of pure states.

Fact 15 (Local-transition): Let ρ be a quantum state in $\mathcal{K}$. Let $|\phi_1\rangle$ and $|\phi_2\rangle$ be two purification of ρ in $\mathcal{H}\otimes\mathcal{K}$. There is a local unitary transformation U acting on $\mathcal{H}$ such that $(U\otimes I)|\phi_1\rangle = |\phi_2\rangle$.

Let $c \overset{\mathsf{def}}{=} S(\rho\|\tau)$. Let us invoke Fact 14 with $|\overline{\rho}\rangle$ being any purification of ρ and $r \overset{\mathsf{def}}{=} 16/\delta^2$. Let $|\overline{\tau}\rangle$ be the purification of τ as given by Fact 14. Let Alice and Referee start with $2^{\frac{2rk}{\delta}}$ ($k \overset{\mathsf{def}}{=} 8c+14$) copies of the pure state $|\phi\rangle$, such that marginal of $|\phi\rangle$ on Referee's side is τ. Since the reduced quantum state on Referee's part in both $|\psi\rangle$ and $|\overline{\tau}\rangle$ is the same, from Fact 15 there exists a transformation acting only in Alice's side which takes $|\phi\rangle$ to $|\overline{\tau}\rangle$. Alice transforms each $|\psi\rangle$ to $|\overline{\tau}\rangle$ and measures the first bit. If she obtains 1 in any copy of $|\overline{\tau}\rangle$ she communicates the number of that copy to Referee. In case she fails to obtain 1 in $2^{\frac{2rk}{\delta}}$ trials, she communicates this to Referee and Referee assumes the state $|0\rangle\langle 0|$. It is easily seen that the communication from Alice is at most $O(\frac{c}{\delta^3})$. Also since $\Pr(\text{Alice observes } 1) = \frac{r-1}{r2^{rk}}$, and Alice makes $2^{\frac{2rk}{\delta}}$ tries she succeeds with probability at least $1-\delta/2$. In case she succeeds, let the state with Referee in which Alice succeeds be $\tilde{\rho}$. From Fact 14 and monotonicity of trace-norm, $\|\tilde{\rho}-\rho\|_{\mathrm{tr}} \leq \delta/2$. So for the final state ρ' produced with Referee, it follows that $\|\rho'-\rho\|_{\mathrm{tr}} \leq \delta$ (using triangle inequality for trace norm). ■

Proof of Fact 9: This proof follows on very similar lines as that of Fact 8. We use the following classical substate theorem [JRS08].

Fact 16 (Classical substate theorem): Let P,Q be probability distributions on the same set such that $S(P\|Q) < \infty$. For every $r > 1$, there exist distributions $\tilde{P},R$ such that $\|P-\tilde{P}\| \leq 2/r$ and $Q = \frac{r-1}{r2^{rk}}\tilde{P} + (1-\frac{r-1}{r2^{rk}})R$, where $k \overset{\mathsf{def}}{=} S(\rho\|\tau) + 1$.

We will also need the following easily verifiable fact.

Fact 17: Let X be a random variable distributed according to Q. Let $p \in [0,1]$ and Q_1, Q_2 be distributions such that $Q = pQ_1 + (1-p)Q_2$. There exists a binary random variable $Z \in \{0,1\}$, correlated with X, with $\Pr[Z=1] = p$, such that the distribution of X conditioned on $Z=1$ is Q_1 and the distribution of X conditioned on $Z=0$ is Q_2.

Let $c \overset{\mathsf{def}}{=} S(P\|Q)$. Let us invoke Fact 16 with $r \overset{\mathsf{def}}{=} 4/\delta$ and let $\tilde{P},R$ be as obtained by Fact 16. Let X be a random variable distributed according to Q. Let Alice and Referee share $2^{\frac{2rk}{\delta}}$ ($k \overset{\mathsf{def}}{=} c+1$) copies of X as public randomness. Let Z be a random variable, correlated with X, obtained from Fact 17 by letting $Q_1 \overset{\mathsf{def}}{=} \tilde{P}$ and $Q_2 \overset{\mathsf{def}}{=} R$. Alice generates the random variable Z for each copy of X, measures Z and sends the number of the first copy in which she succeeds to obtain 1 to Referee. In case she fails to obtain a 1 in $2^{\frac{2rk}{\delta}}$ trials, she communicates this to Referee and Referee assumes single point distribution concentrated on 0. It is easily seen that the communication from Alice is at most $O(\frac{c}{\delta^2})$. Also since $\Pr(\text{Alice observes } 1) = \frac{r-1}{r2^{rk}}$, and Alice makes $2^{\frac{2rk}{\delta}}$ tries she succeeds with probability at least $1-\delta/2$. In case she succeeds, the copy which she communicates to Referee will be distributed according to $\tilde{P}$. From Fact 14, $\|\tilde{P}-P\| \leq \delta/2$. So for the final distribution P' produced with Referee, it follows that $\|P'-P\| \leq \delta$. ■