Distributed compression and multiparty squashed entanglement

2008 J. Phys. A: Math. Theor. 41 115301

(http://iopscience.iop.org/1751-8121/41/11/115301)

IOP PUBLISHING JOURNAL OF PHYSICS A: MATHEMATICAL AND THEORETICAL

J. Phys. A: Math. Theor. **41** (2008) 115301 (25pp) doi:10.1088/1751-8113/41/11/115301

Distributed compression and multiparty squashed entanglement

David Avis[1], **Patrick Hayden**[1] **and Ivan Savov**[2]

[1] School of Computer Science, McGill University, Montreal, Quebec, H3A 2A7, Canada
[2] Physics Department, McGill University, Montreal, Quebec, H3A 2A7, Canada

E-mail: avis@cs.mcgill.ca, patrick@cs.mcgill.ca and ivan.savov@mail.mcgill.ca

Received 4 October 2007, in final form 4 February 2008
Published 4 March 2008
Online at stacks.iop.org/JPhysA/41/115301

Abstract

We study a protocol in which many parties use quantum communication to transfer a shared state to a receiver without communicating with each other. This protocol is a multiparty version of the fully quantum Slepian–Wolf protocol for two senders and arises through the repeated application of the two-sender protocol. We describe bounds on the achievable rate region for the distributed compression problem. The inner bound arises by expressing the achievable rate region for our protocol in terms of its vertices and extreme rays and, equivalently, in terms of facet inequalities. We also prove an outer bound on all possible rates for distributed compression based on the multiparty squashed entanglement, a measure of multiparty entanglement.

PACS numbers: 03.67.Hk, 03.67.Ac, 03.65.Ud

(Some figures in this article are in colour only in the electronic version)

1. Introduction

Quantum information theory studies the interconversion of information resources like quantum channels, states and entanglement for the purpose of accomplishing communication tasks [1–4]. This approach is rendered possible by the substantial body of results characterizing quantum channels [5–8] and quantum communication resources like entanglement [9–11].

In classical information theory, distributed compression is the search for the optimal rates at which two parties Alice and Bob can compress and transmit information faithfully to a third-party Charlie. If the senders are allowed to communicate among themselves then they can obviously use the correlations between their sources to achieve better rates. The more interesting problem is to ask what rates can be achieved if no communication is allowed between the senders. The classical version of this problem was solved by Slepian and Wolf [12]. The quantum version of this problem was first approached in [13, 14] and more recently

in [4], which describes the fully quantum Slepian–Wolf (FQSW) protocol and partially solves the distributed compression problem for two senders.

In this paper, we generalize the results of the FQSW protocol to a multiparty scenario where m senders, Alice 1 through Alice m, send quantum information to a single receiver, Charlie. We exhibit a set of achievable rates as well as an outer bound on the possible rates based on a new measure of multiparty entanglement that generalizes squashed entanglement [15]. Our protocol is optimal for input states that have zero squashed entanglement, notably separable states.

The multiparty squashed entanglement is interesting in its own right, and we develop a number of its properties in the paper. (It was also found independently by Yang *et al* and described in a recent paper [16].) While there exist several measures for bipartite entanglement with useful properties and applications [1, 17–19], the theory of multiparty entanglement, despite considerable effort [20–23], remains comparatively undeveloped. Multiparty entanglement is fundamentally more complicated because it cannot be described by a single number even for pure states. We can, however, define *useful* entanglement measures for particular applications, and the multiparty squashed entanglement seems well suited to application in the distributed compression problem.

The structure of the paper is as follows. In section 2, we describe the quantum distributed compression problem and present our protocol. Our results are twofold. In theorem 2.1, we give the formula for the achievable rate region using this protocol and in theorem 2.2 we provide a bound on the best possible rates for any protocol. The proof of theorem 2.1 is in section 3. The proof of theorem 2.2 is given in section 6 but before we get to it we need to introduce and describe the properties of the multiparty information quantity in section 4 and multiparty squashed entanglement in section 5.

Notation. We will denote quantum systems as A, B, R and the corresponding Hilbert spaces $\mathcal{H}^A, \mathcal{H}^B, \mathcal{H}^R$ with respective dimensions d_A, d_B, d_R. We denote pure states of the system A by a ket $|\varphi\rangle^A$ and the corresponding density matrices as $\varphi^A = |\varphi\rangle\langle\varphi|^A$. We denote by $H(A)_\rho = -\mathrm{Tr}(\rho^A \log \rho^A)$ the von Neumann entropy of the state ρ^A. For a bipartite state σ^{AB} we define the conditional entropy $H(A|B)_\sigma = H(AB)_\sigma - H(B)_\sigma$ and the mutual information $I(A;B)_\sigma = H(A)_\sigma + H(B)_\sigma - H(AB)_\sigma$. The trace distance between states σ and ρ is $\|\sigma - \rho\|_1 = \mathrm{Tr}|\sigma - \rho|$ where $|X| = \sqrt{X^\dagger X}$. The fidelity is defined to be $F(\sigma, \rho) = \mathrm{Tr}(\sqrt{\sqrt{\rho}\sigma\sqrt{\rho}})^2$. Two states that are very similar have fidelity close to 1 whereas states with little similarity will have low fidelity. Throughout this paper, logarithms and exponents are taken base 2 unless otherwise specified.

2. Multiparty distributed compression

Distributed compression of classical information involves many parties collaboratively encoding their sources $X_1, X_2, \ldots, X_m$ and sending the information to a common receiver [24]. In the quantum setting, the parties are given a quantum state $\varphi^{A_1 A_2 \cdots A_m} \in \mathcal{H}^{A_1 A_2 \cdots A_m}$ and are asked to individually compress their shares of the state and transfer them to the receiver while sending as few qubits as possible [13]. The objective is to successfully transmit the quantum information stored in the A systems, meaning any entanglement with an external reference system, to the receiver. No communication between the senders is allowed and, unlike [14], in this paper there is no classical communication between the senders and the receiver.

In our analysis, we work in the case where we have many copies of the input state, so that the goal is to send shares of the purification $|\psi\rangle^{A_1 A_2 \cdots A_m R} = (|\varphi\rangle^{A_1 A_2 \cdots A_m R})^{\otimes n}$, where A_i's

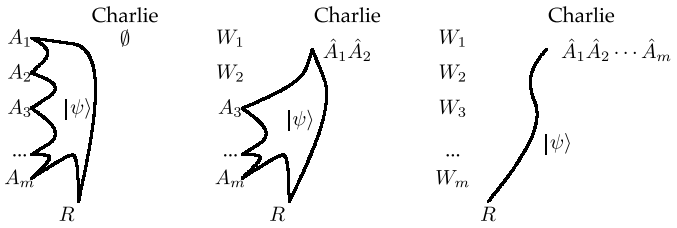


Figure 1. Pictorial representation of the quantum correlations between the systems at three stages of the protocol. Originally the state $|\psi\rangle$ is shared between $A_1A_2\cdots A_m$ and R. The middle picture shows the protocol in progress. Finally, all systems are received by Charlie and $|\psi\rangle$ is now shared between Charlie's systems $\widehat{A}_1\widehat{A}_2\cdots\widehat{A}_m$ and R.

denote the m different systems and R denotes the reference system, which does not participate in the protocol. Note that we use A_i to denote both the individual system associated with state φ as well the n-copy version associated with ψ; the meaning should be clear from the context. We also use the shorthand notation $A = A_1A_2\cdots A_m$ to denote all the senders.

The objective, as we have mentioned, is for the participants to transfer their R-entanglement to a third-party Charlie as illustrated in figure 1. Note that any other type of correlation the A systems could have with an external subsystem is automatically preserved in this case, which implies for example that if φ were written as a convex combination $\varphi = \sum_i p_i\varphi_i$ then a successful protocol would automatically send φ_i with high fidelity on average [25].

An equivalent way of thinking about quantum distributed compression is to say that the participants are attempting to decouple their systems from the reference R solely by sending quantum information to Charlie. Indeed, if we assume that originally R is the purification of $A_1A_2\cdots A_m$, and at the end of the protocol there are no correlations between the remnant W systems (see figure 1) and R, then the purification of R must have been transferred to Charlie's laboratory since none of the original information was discarded.

To perform the distributed compression task, each of the senders independently encodes her share before sending part of it to Charlie. The encoding operations are modeled by quantum operations, that is, completely positive trace-preserving (CPTP) maps E_i with outputs C_i of dimension 2^{nQ_i}. Once Charlie receives the systems that were sent to him, he will apply a decoding CPTP map D with output system $\widehat{A} = \widehat{A}_1\widehat{A}_2\ldots\widehat{A}_m$ isomorphic to the original $A = A_1A_2\cdots A_m$.

Definition 2.1 (the rate region)**.** *We say that a rate tuple* $\vec{Q} = (Q_1, Q_2, \ldots, Q_m)$ *is achievable if for all* $\epsilon > 0$ *there exists* $N(\epsilon)$ *such that for all* $n \geqslant N(\epsilon)$ *there exist* n*-dependent maps* $(E_1, E_2, \ldots, E_m, D)$ *with domains and ranges as in the previous paragraph for which the fidelity between the original state* $|\psi\rangle^{A^nR^n} = (|\varphi\rangle^{A_1A_2\cdots A_mR})^{\otimes n}$ *and the final state* $\sigma^{\widehat{A}_1\widehat{A}_2\ldots\widehat{A}_mR} = \sigma^{\widehat{A}^n\widehat{R}^n}$ *satisfies*

$$F\left(|\psi\rangle^{A^nR^n}, \sigma^{\widehat{A}^n\widehat{R}^n}\right) = {}^{\widehat{A}^n\widehat{R}^n}\langle\psi|\,(D\circ(E_1\otimes\cdots\otimes E_m))\left(\psi^{A^nR^n}\right)|\psi\rangle^{\widehat{A}^n\widehat{R}^n} \geqslant 1-\epsilon. \tag{1}$$

We call the closure of the set of achievable rate tuples the rate region.

At this point it is illustrative to review the results of the two-party state transfer protocol [4], which form a key building block for the multiparty distributed compression protocol presented in section 2.2.

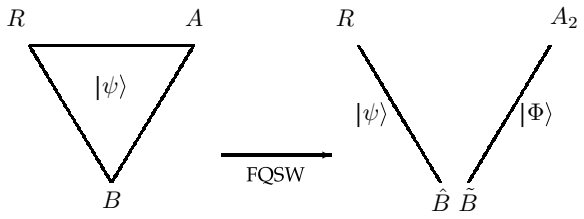


Figure 2. Diagram representing the ABR correlations before and after the FQSW protocol. Alice manages to decouple completely from the reference R. The $\widehat{B}$ system is isomorphic to AB.

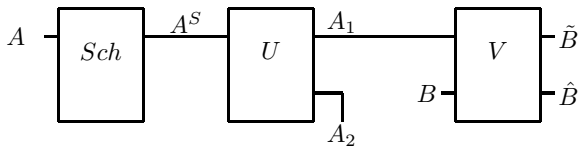


Figure 3. A circuit diagram that shows the Schumacher compression and unitary encoding done by Alice and the decoding done by Bob.

2.1. The FQSW protocol

The fully quantum Slepian–Wolf protocol [4] describes a procedure for simultaneous quantum state transfer and entanglement distillation. This communication task can be used as a building block for nearly all the other protocols of quantum information theory [2], yet despite its powerful applications it is fairly simple to implement.

Consider a setup where the state $|\psi\rangle^{ABR} = (|\varphi\rangle^{ABR})^{\otimes n}$ is shared between Alice, Bob and a reference system R. The FQSW protocol describes a procedure for Alice to transfer her R-entanglement to Bob while at the same time generating ebits with him. Alice can accomplish this by encoding and sending part of her system, denoted A_1, to Bob. The state after the protocol can approximately be written as $|\Phi\rangle^{A_2\widetilde{B}}(|\varphi\rangle^{R\widehat{B}})^{\otimes n}$, where the systems $\widetilde{B}$ and $\widehat{B}$ are held in Bob's lab while A_2 remains with Alice. The state $|\Phi\rangle^{A_2\widetilde{B}}$ is a maximally entangled state shared between Alice and Bob, a handy side-product which can be used to build more advanced protocols [26, 27]. Figure 2 illustrates the entanglement structure before and after the protocol.

The protocol, represented graphically in figure 3, consists of the following steps:

(1) Alice performs Schumacher compression on her system A to obtain the output system A^S.
(2) Alice then applies a random unitary U_A to A^S.
(3) Next, she splits her system into two parts: $A_1A_2 = A^S$ with $d_{A_1} = 2^{nQ_A}$ and

$$Q_A > \tfrac{1}{2}I(A;R)_\varphi. \tag{2}$$

She sends the system A_1 to Bob.
(4) Bob, in turn, performs a decoding operation $V_B^{A_1B\to\widehat{B}\widetilde{B}}$ which splits his system into a $\widehat{B}$ part purifying R and a $\widetilde{B}$ part which is fully entangled with Alice.

The best way to understand the mechanism behind this protocol is by thinking about destroying correlations. If, at the end of the protocol, Alice's system A_2 is nearly decoupled from the reference in the sense that $\sigma^{A_2R} \approx \sigma^{A_2} \otimes \sigma^R$, then Alice must have succeeded in

sending her R entanglement to Bob because it is Bob alone who then holds the R purification. We can therefore guess the lower bound on how many qubits Alice will have to send before she can decouple from the reference. Originally, Alice and R share $I(A;R)_\varphi$ bits of information per copy of $|\varphi\rangle^{ABR}$. Since one qubit can carry away at most two bits of quantum mutual information, this means that the minimum rate at which Alice must send qubits to Bob is

$$Q_A > \tfrac{1}{2} I(A;R)_\varphi. \tag{3}$$

It is shown in [4] that this rate is achievable in the limit of many copies of the state. Therefore the FQSW protocol is optimal for the state transfer task.

2.2. The multiparty FQSW protocol

Like the original FQSW protocol, the multiparty version relies on Schumacher compression and the mixing effect of random unitary operations for the encoding. The only additional ingredient is an agreed upon permutation of the participants. The temporal order in which the participants will perform their encoding is of no importance. However, the permutation determines how much information each participant is to send to Charlie.

For each permutation π of the participants, the protocol consists of the following steps:

(1) Each Alice-i performs Schumacher compression on her system A_i reducing its effective size to the entropy bound of roughly $H(A_i)$ qubits per copy of the state.
(2) Each participant applies a known, pre-selected random unitary to the compressed system.
(3) Participant i sends to Charlie a system C_i of dimension 2^{nQ_i} where

$$Q_i > \tfrac{1}{2} I(A_i; A_{\mathcal{K}_i} R)_\varphi, \tag{4}$$

where $\mathcal{K}_i = \{\pi(j) : j > \pi^{-1}(i)\}$ is the set of participants who come after participant i according to the permutation.
(4) Charlie applies a decoding operation D consisting of the composition of the decoding maps $D_{\pi(m)} \circ \cdots \circ D_{\pi(2)} \circ D_{\pi(1)}$ defined by the individual FQSW steps in order to recover $\sigma^{\widehat{A}_1\widehat{A}_2\ldots\widehat{A}_m}$ nearly identical to the original $\psi^{A_1A_2\cdots A_m}$ and purifying R.

2.3. Statement of results

This subsection contains our two main theorems about multiparty distributed compression. In theorem 2.1, we give the formula for the set of achievable rates using the multiparty FQSW protocol (sufficient conditions). Then, in theorem 2.2 we specify another set of inequalities for the rates Q_i which must be true for any distributed compression protocol (necessary conditions). In what follows, we consistently use $\mathcal{K} \subseteq \{1, 2, \ldots m\}$ to denote any subset of the senders in the protocol.

Theorem 2.1. *Let $|\varphi\rangle^{A_1A_2\cdots A_mR}$ be a pure state. If the inequality*

$$\sum_{k\in\mathcal{K}} Q_k \geqslant \frac{1}{2}\left[\sum_{k\in\mathcal{K}} [H(A_k)_\varphi] + H(R)_\varphi - H(RA_{\mathcal{K}})_\varphi\right] \tag{5}$$

holds for all $\mathcal{K} \subseteq \{1, 2, \ldots, m\}$, then the rate tuple $(Q_1, Q_2, \ldots, Q_m)$ is achievable for distributed compression of the A_i systems.

Because theorem 2.1 expresses a set of sufficient conditions for the protocol to succeed, we say that these rates are contained in the rate region. The proof is given in the following section.

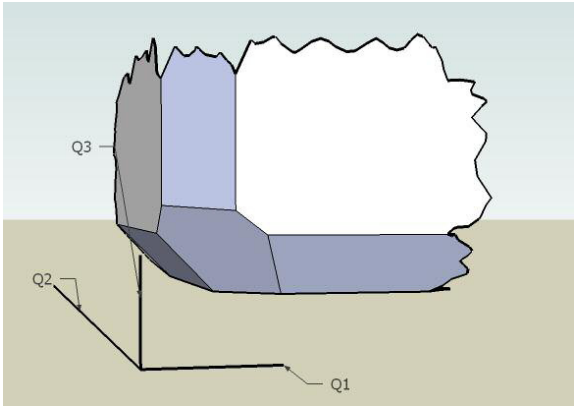


Figure 4. Sketch of the rate region for the multiparty FQSW protocol for three senders.

In the m-dimensional space of rate tuples $(Q_1, Q_2, \ldots, Q_m) \in \mathbb{R}^m$, the inequalities (5) define a convex polyhedron [28] whose facets are given by the corresponding hyperplanes, as illustrated in figure 4.

In order to characterize the rate region further we also derive an outer bound which all rate tuples must satisfy.

Theorem 2.2. *Let $|\varphi\rangle^{A_1 A_2 \cdots A_m R}$ be a pure-state input to a distributed compression protocol which achieves the rate tuple $(Q_1, Q_2, \ldots, Q_m)$, then it must be true that*

$$\sum_{k\in\mathcal{K}} Q_k \geqslant \frac{1}{2}\left[\sum_{k\in\mathcal{K}} [H(A_k)_\varphi] + H(R)_\varphi - H(RA_\mathcal{K})_\varphi\right] - E_{\mathrm{sq}}\left(A_{k_1}; A_{k_2}; \ldots; A_{k_{|\mathcal{K}|}}\right)_\varphi, \tag{6}$$

for all $\mathcal{K} \subseteq \{1, 2, \ldots, m\}$, where E_{sq} is the multiparty squashed entanglement.

The multiparty squashed entanglement, independently discovered in [16], is a measure of multipartite entanglement which generalizes the bipartite squashed entanglement of [15]. Sections 4 and 5 define the quantity and investigate some of its properties. The proof of theorem 2.2 is given in section 6.

Note that theorems 2.1 and 2.2 both provide bounds of the same form and only differ by the presence of the E_{sq} term. The rate region is squeezed somewhere between these two bounds as illustrated in figure 5. For states which have zero squashed entanglement, the inner and outer bounds on the region coincide so that in those cases our protocol is an optimal solution to the multiparty distributed compression problem.

One can verify that when only two parties are involved ($m = 2$), the inequalities in (5) reduce to the 2-party bounds in the original FQSW paper:

$$\begin{aligned} Q_1 &\geqslant \tfrac{1}{2} I(A_1; R), \\ Q_2 &\geqslant \tfrac{1}{2} I(A_2; R), \\ Q_1 + Q_2 &\geqslant \tfrac{1}{2}[H(A_1) + H(A_2) + H(A_1A_2)]. \end{aligned}$$

The family of inequalities (6) similarly reduce to the corresponding expressions in [4] with the multiparty squashed entanglement being replaced by the original two-party squashed entanglement of [15].

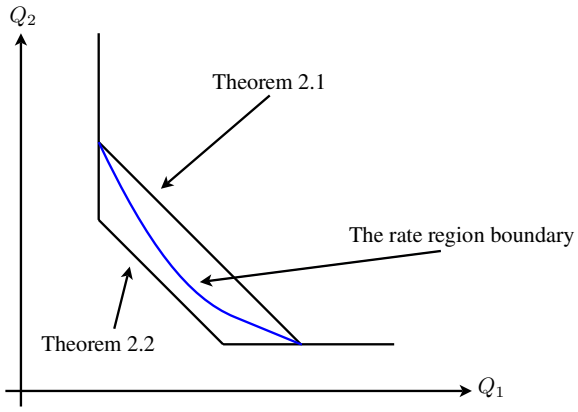


Figure 5. A two-dimensional diagram showing the inner bound from theorem 2.1 and the outer bound from theorem 2.2. The boundary of the real rate region must lie somewhere in between.

3. Proof of the achievable rates

The multiparty fully quantum Slepian–Wolf protocol can be constructed through the repeated application of the two-party FQSW protocol [4]. In the multiparty case, however, the geometry of the rate region is more involved and some concepts from the theory of polyhedra [28] prove helpful in giving it a precise characterization. Multiparty rate regions in information theory have previously appeared in [24, 29].

For every permutation $\pi \in S_m$ of the m senders, there is a different rate tuple $\vec{q}_\pi = (Q_1, Q_2, \ldots, Q_m)_\pi \in \mathbb{R}^m$ which is achievable in the limit of many copies of the state. By time-sharing we can achieve any rate that lies in the *convex hull* of these points. We will show that the rate region for an input state $|\varphi\rangle^{A_1 \cdots A_m R}$ can equivalently be described by the set of inequalities from theorem 2.1, that is

$$\sum_{k \in \mathcal{K}} Q_k \geqslant \frac{1}{2}\left[\sum_{k \in \mathcal{K}} H(A_k)_\varphi + H(R)_\varphi - H(RA_{\mathcal{K}})_\varphi\right] =: C_{\mathcal{K}}, \tag{7}$$

where $\mathcal{K} \subseteq \{1, 2, \ldots, m\}$ ranges over all subsets of participants and $C_{\mathcal{K}}$ is the name we give to the constant on the right-hand side of the inequality. The proof of theorem 2.1 proceeds in two steps. First we show the set of rate tuples $\{\vec{q}_\pi\}$ is contained in the rate region and then we prove that the set of inequalities (7) is an equivalent description of the rates obtained by time sharing and resource wasting of the rates $\{\vec{q}_\pi\}$.

Consider the m-dimensional space of rate tuples $(Q_1, \ldots, Q_m) \in \mathbb{R}^m$. We begin by a formal definition of a corner point $\vec{q}_\pi$.

Definition 2 (corner point)**.** *Let $\pi \in S_m$ be a permutations of the senders in the protocol. The corresponding rate tuple $\vec{q}_\pi = (Q_1, Q_2, \ldots, Q_m)$ is a corner point if*

$$Q_{\pi(k)} = \tfrac{1}{2} I(A_{\pi(k)}; A_{\pi(k+1)} \cdots A_{\pi(m)} R), \tag{8}$$

where the set $A_{\pi(k+1)} \cdots A_{\pi(m)}$ denotes all the systems which come after k in the permutation π.

We define $\mathcal{Q} := \{\vec{q}_\pi : \pi \in S_m\}$, the set of all corner points. Clearly, $|\mathcal{Q}| \leqslant m!$ but since some permutations might lead to the same rate tuple, the inequality may be strict.

Lemma 3.2. *The set of corner points, $\mathcal{Q} = \{\vec{q}_\pi : \pi \in S_m\}$, is contained in the rate region.*

Proof (sketch for lemma 3.2)**.** We will now exhibit a protocol that achieves one such point. In order to simplify the notation, but without loss of generality, we choose the reversed-order permutation $\pi = (m, \ldots, 2, 1)$. This choice of permutation corresponds to Alice-m sending her information first and Alice-1 sending last.

We will repeatedly use the FQSW protocol in order to send the m systems to Charlie:

(1) The first-party Schumacher compresses her system A_m and sends it to Charlie. She succeeds provided
$$Q_m \geqslant \tfrac{1}{2} I(A_m; A_1 A_2 \cdots A_{m-1} R) + \delta = H(A_m) + \delta$$
for any $\delta > 0$. The above rate is dictated by the FQSW inequality (3) because we are facing the same type of problem except that the 'reference' consists of R as well as the remaining participants $A_1 A_2 \cdots A_{m-1}$. The fact that the formula reduces to $Q_m > H(A_m)$ should also be expected since there are no correlations that the first participant can take advantage of; she is just performing Schumacher compression.

(2) The second party also faces an instance of an FQSW problem. The task is to transmit the system A_{m-1} to Charlie, who is now assumed to hold A_m. The purifying system consists of $A_1 A_2 \cdots A_{m-2} R$. According to inequality (3) the rate must be
$$Q_{m-1} \geqslant \tfrac{1}{2} I(A_{m-1}; A_1 A_2 \cdots A_{m-2} R) + \delta$$
for any $\delta > 0$.

(3) The last person to be merging with Charlie will have a purifying system consisting of only R. Her transfer will be successful if
$$Q_1 \geqslant \tfrac{1}{2} I(A_1; R) + \delta$$
for any $\delta > 0$.

On the receiving end of the protocol, Charlie will apply the decoding map D consisting of the composition of the decoding maps $D_1 \circ D_2 \circ \cdots \circ D_m$ defined by the individual FQSW steps to recover the state $\sigma^{\widehat{A}_1 \widehat{A}_2 \cdots \widehat{A}_m}$, which will be such that the fidelity between $|\psi\rangle^{A^n R^n}$ and $\sigma^{\hat{A}^n R^n}$ is high, essentially by the triangle inequality. Finally, because we can make δ arbitrarily small, the rate tuple $(Q_1, \ldots, Q_m)$, with
$$Q_k = \tfrac{1}{2} I(A_k; A_1 \cdots A_{k-1} R), \tag{9}$$
must be contained in the rate region. The same argument applies for each permutation $\pi \in S_m$, leading to the conclusion that the full set $\mathcal{Q}$ is contained in the rate region. □

Each one of the corner points $\vec{q}_\pi$ can also be described by an equivalent set of equations involving sums of the rates.

Lemma 3.3. *The rate tuple $(Q_1, Q_2, \ldots, Q_m)$ is a corner point if and only if for some $\pi \in S_m$ and for all l such that $1 \leqslant l \leqslant m$,*
$$\sum_{m-l+1 \leqslant k \leqslant m} Q_{\pi(k)} = \frac{1}{2}\left[\sum_{m-l+1 \leqslant k \leqslant m} H(A_{\pi(k)}) + H(R) - H(A_{\pi[m-l+1,m]} R)\right] = C_{\pi[m-l+1,m]}, \tag{10}$$
where $A_{\pi[m-l+1,m]} := A_{\pi(m-l+1)} A_{\pi(m-l+2)} \cdots A_{\pi(m)}$ denotes the last l participants according to the permutation π.

Proof (of lemma 3.3)**.** The proof follows trivially from lemma 3.2 by considering sums of the rates. If we again choose the permutation $\pi = (m, \ldots, 2, 1)$ for simplicity, we see that the sum of the rates of the last l participants is

$$\begin{aligned} Q_1 + \cdots + Q_l &= \frac{1}{2}\Big[I(A_1; R) + I(A_2; A_1 R) + \cdots + I(A_l; A_1 \cdots A_{l-1} R) \Big] \\ &= \frac{1}{2}\Big[\sum_{1 \leqslant k \leqslant l} H(A_k) + H(R) - H(A_1 \cdots A_l R) \Big] = C_{12\ldots l}. \end{aligned} \quad (11)$$

A telescoping effect occurs and most of the inner terms cancel so we are left with a system of equations identical to (10). Moreover, this system is clearly solvable for the individual rates Q_k. The analogous simplification occurs for all other permutations. □

So far, we have shown that the set of corner points $\mathcal{Q}$ is contained in the rate region of the multiparty fully quantum Slepian–Wolf protocol. The convex hull of a set of points $\mathcal{Q}$ is defined to be

$$\mathrm{conv}(\mathcal{Q}) := \left\{ \vec{x} \in \mathbb{R}^m : \vec{x} = \sum \lambda_i \vec{q}_i, \vec{q}_i \in \mathcal{Q}, \lambda_i \geqslant 0, \sum \lambda_i = 1 \right\}. \quad (12)$$

Because of the possibility of time-sharing between the different corner points, the entire convex hull $\mathrm{conv}(\mathcal{Q})$ must be achievable. Furthermore, by simply allowing any one of the senders to waste resources, we know that if a rate tuple $\vec{q}$ is achievable, then so is $\vec{q} + \vec{w}$ for any vector $\vec{w}$ with non-negative coefficients. More formally, we say that any $\vec{q} + \mathrm{cone}(\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_m)$ is also inside the rate region, where $\{\vec{e}_i\}$ is the standard basis for $\mathbb{R}^m$: $\vec{e}_i = (\underbrace{0, 0, \ldots, 0, 1}_{i}, 0, 0)$

and

$$\mathrm{cone}(\vec{e}_1, \ldots, \vec{e}_m) := \left\{ \vec{x} \in \mathbb{R}^m : \vec{x} = \sum \lambda_i \vec{e}_i, \lambda_i \geqslant 0 \right\}. \quad (13)$$

Thus, we have demonstrated that the set of rates

$$P_{\mathcal{V}} := \mathrm{conv}(\mathcal{Q}) + \mathrm{cone}(\vec{e}_1, \ldots, \vec{e}_m) \quad (14)$$

is achievable. To complete the proof of theorem 2.1, we will need to show that $P_{\mathcal{V}}$ has an equivalent description as

$$P_{\mathcal{H}} := \left\{ (Q_1, \ldots, Q_m) \in \mathbb{R}^m : \sum_{k \in \mathcal{K}} Q_k \geqslant C_{\mathcal{K}}, \forall \mathcal{K} \subseteq \{1, 2, \ldots, m\} \right\}, \quad (15)$$

where the constants $C_{\mathcal{K}}$ are as defined in equation (7). This equivalence is an explicit special case of the Minkowski–Weyl theorem on convex polyhedra.

Theorem 3.1 (Minkowski–Weyl theorem) [28, p 30]**.** *For a subset $P \subseteq \mathbb{R}^m$, the following two statements are equivalent:*

- *P is a $\mathcal{V}$-polyhedron: the sum of a convex hull of a finite set of points $\mathcal{P} = \{\vec{p}_i\}$ plus a conical combination of vectors $\mathcal{W} = \{\vec{w}_i\}$*

  $$P = \mathrm{conv}(\mathcal{P}) + \mathrm{cone}(\mathcal{W}), \quad (16)$$

  where $\mathrm{conv}(\mathcal{P})$ *and* $\mathrm{cone}(\mathcal{W})$ *are defined in (12) and (13) respectively.*
- *P is a $\mathcal{H}$-polyhedron: an intersection of n closed halfspaces*

  $$P = \{\vec{x} \in \mathbb{R}^m : A\vec{x} \geqslant \vec{a}\} \quad (17)$$

  for some matrix $A \in \mathbb{R}^{n \times m}$ and some vector $\vec{a} \in \mathbb{R}^n$. Each of the n rows in equation (17) defines one halfspace.

Preliminaries. Before we begin the equivalence proof in earnest, we make two useful observations which will be instrumental to our subsequent argument. First, we prove a very important property of the constants $C_{\mathcal{K}}$ which will dictate the geometry of the rate region.

Lemma 3.4 (superadditivity)**.** *Let $\mathcal{K}, \mathcal{L} \subseteq \{1, 2, \ldots, m\}$ be any two subsets of the senders. Then*

$$C_{\mathcal{K}\cup\mathcal{L}} + C_{\mathcal{K}\cap\mathcal{L}} \geqslant C_{\mathcal{K}} + C_{\mathcal{L}}. \tag{18}$$

Proof (of lemma 3.4)**.** We expand the C terms and cancel the $\frac{1}{2}$-factors to obtain

$$\begin{array}{l} \sum_{k\in\mathcal{K}\cup\mathcal{L}} H(A_k) + H(R) - H(RA_{\mathcal{K}\cup\mathcal{L}}) \\ + \sum_{k\in\mathcal{K}\cap\mathcal{L}} H(A_k) + H(R) - H(RA_{\mathcal{K}\cap\mathcal{L}}) \end{array} \geqslant \begin{array}{l} \sum_{k\in\mathcal{K}} H(A_k) + H(R) - H(RA_{\mathcal{K}}) \\ + \sum_{k\in\mathcal{L}} H(A_k) + H(R) - H(RA_{\mathcal{L}}). \end{array}$$

After canceling all common terms we find that the above inequality is equivalent to

$$H(RA_{\mathcal{K}}) + H(RA_{\mathcal{L}}) \geqslant H(RA_{\mathcal{K}\cup\mathcal{L}}) + H(RA_{\mathcal{K}\cap\mathcal{L}}), \tag{19}$$

which is true by the strong subadditivity (SSA) inequality of quantum entropy [30]. □

As a consequence of this lemma, we can derive an equivalence property for the saturated inequalities.

Corollary 3.5. *Suppose that the following two equations hold for a given point of $P_{\mathcal{H}}$:*

$$\sum_{k\in\mathcal{K}} Q_k = C_{\mathcal{K}} \qquad \text{and} \qquad \sum_{k\in\mathcal{L}} Q_k = C_{\mathcal{L}}. \tag{20}$$

Then the following equations must also be true:

$$\sum_{k\in\mathcal{K}\cup\mathcal{L}} Q_k = C_{\mathcal{K}\cup\mathcal{L}} \qquad \text{and} \qquad \sum_{k\in\mathcal{K}\cap\mathcal{L}} Q_k = C_{\mathcal{K}\cap\mathcal{L}}. \tag{21}$$

Proof (of corollary 3.5)**.** The proof follows from the equation

$$\sum_{k\in\mathcal{K}} Q_k + \sum_{k\in\mathcal{L}} Q_k = C_{\mathcal{K}} + C_{\mathcal{L}} \leqslant C_{\mathcal{K}\cup\mathcal{L}} + C_{\mathcal{K}\cap\mathcal{L}} \leqslant \sum_{k\in\mathcal{K}\cup\mathcal{L}} Q_k + \sum_{k\in\mathcal{K}\cap\mathcal{L}} Q_k, \tag{22}$$

where the first inequality comes from lemma 3.4. The second inequality is true by the definition of $P_{\mathcal{H}}$ since $\mathcal{K}\cup\mathcal{L}$ and $\mathcal{K}\cap\mathcal{L}$ are subsets of $\{1, 2, \ldots, m\}$. Because the leftmost terms and rightmost terms are identical, we must have equality throughout equation (22), which in turn implies the union and the intersection equations are saturated. □

An important consequence of lemma 3.4 is that it implies that the polyhedron $P_{\mathcal{H}}$ has a very special structure. It is known as a supermodular polyhedron or contra-polymatroid. The fact that $\text{conv}(Q) = P_{\mathcal{H}}$ was proved by Edmonds [31], whose ingenious proof makes use of linear programming duality. Below we give an elementary proof that does not use duality.

A *vertex* is a zero-dimensional face of a polyhedron. A point $\bar{Q} = (\bar{Q}_1, \bar{Q}_2, \ldots, \bar{Q}_m) \in P_{\mathcal{H}} \subset \mathbb{R}^m$ is a vertex of $P_{\mathcal{H}}$ if and only if it is the unique solution of a set of linearly independent equations

$$\sum_{k\in\mathcal{L}_i} Q_k = C_{\mathcal{L}_i}, \qquad 1 \leqslant i \leqslant m \tag{23}$$

for some subsets $\mathcal{L}_i \subseteq \{1, 2, \ldots, m\}$. In the remainder of the proof we require only a specific consequence of linear independence, which we state in the following lemma.

Lemma 3.6 (no co-occurrence)**.** *Let $\mathcal{L}_i \subseteq \{1, 2, \ldots, m\}$ be a collection of m sets such that the system (23) has a unique solution. Then there is no pair of elements j, k such that $j \in \mathcal{L}_i$ if and only if $k \in \mathcal{L}_i$ for all i.*

Proof. If there was such a pair j and k, then the corresponding columns of the left-hand side of (23) would be linearly dependent. □

Armed with the above tools, we will now show that there is a one-to-one correspondence between the corner points $\mathcal{Q}$ and the vertices of the $\mathcal{H}$-polyhedron $P_{\mathcal{H}}$. We will then show that the vectors that generate the cone part of the $\mathcal{H}$-polyhedron correspond to the resource wasting vectors $\{\vec{e}_i\}$.

Step 1. $\mathcal{Q} \subseteq vertices(P_{\mathcal{H}})$. We know from lemma 3.3 that every point $\vec{q}_\pi \in \mathcal{Q}$ satisfies the m equations

$$\sum_{m-i+1 \leqslant k \leqslant m} Q_{\pi(k)} = C_{\pi[m-i+1,m]}, \qquad 1 \leqslant i \leqslant m. \tag{24}$$

Equations (24) are linearly independent since the left-hand side is triangular, and have the form of the inequalities in (15) that are used to define $P_{\mathcal{H}}$. They have the unique solution:

$$Q_{\pi(m)} = C_{\pi(m)}, \qquad Q_{\pi(i)} = C_{\pi[i,m]} - C_{\pi[i+1,m]}, \qquad 1 \leqslant i \leqslant m-1. \tag{25}$$

We need to show that this solution satisfies all the inequalities used to define $P_{\mathcal{H}}$ in (15). We proceed by induction on $|\mathcal{K}|$. The case $|\mathcal{K}| = 1$ follows from (25) and the superadditivity property (18). For $|\mathcal{K}| \geqslant 2$ we can write $\mathcal{K} = \{\pi(i)\} \cup \mathcal{K}'$ for some $\mathcal{K}' \subseteq \{\pi(i+1), \pi(i+2), \ldots, \pi(m)\}$. Then

$$\begin{aligned}
\sum_{k \in \mathcal{K}} Q_k &= Q_{\pi(i)} + \sum_{k \in \mathcal{K}'} Q_k \\
&\geqslant C_{\pi[i,m]} - C_{\pi[i+1,m]} + \sum_{k \in \mathcal{K}'} Q_k \\
&\geqslant C_{\pi[i,m]} - C_{\pi[i+1,m]} + C_{\mathcal{K}'} \qquad \text{(induction)} \\
&\geqslant C_{\mathcal{K}},
\end{aligned}$$

where we again used superadditivity to get the last inequality.

Step 2. *Vertices* $(P_{\mathcal{H}}) \subseteq \mathcal{Q}$. In order to prove the opposite inclusion, we will show that every vertex of $P_{\mathcal{H}}$ is of the form of lemma 3.3. More specifically, we want to prove the following proposition.

Proposition 3.7 (existence of a maximal chain)**.** *Every vertex of $P_{\mathcal{H}}$, that is, the intersection of m linearly independent hyperplanes*

$$\sum_{k \in \mathcal{L}_i} Q_k = C_{\mathcal{L}_i}, \qquad 1 \leqslant i \leqslant m, \tag{26}$$

defined by the family of sets $\{\mathcal{L}_i; 1 \leqslant i \leqslant m\}$ can be described by an equivalent set of equations

$$\sum_{k \in \mathcal{K}_i} Q_k = C_{\mathcal{K}_i}, \qquad 1 \leqslant i \leqslant m, \tag{27}$$

for some family of sets distinct $\mathcal{K}_i \subseteq \{1, 2, \ldots, m\}$ that form a maximal chain in the sense of

$$\emptyset = \mathcal{K}_0 \subset \mathcal{K}_1 \subset \mathcal{K}_2 \subset \cdots \subset \mathcal{K}_{m-1} \subset \mathcal{K}_m = \{1, 2, \ldots, m\}. \tag{28}$$

Since there exists a permutation π such that $\forall i, \pi[m-i+1, m] = \mathcal{K}_i$ this implies that all the vertices of $P_{\mathcal{H}}$ are in $\mathcal{Q}$. The main tool we have at our disposal in order to prove this proposition is corollary 3.5, which we will use extensively.

Proof of proposition 3.7. Let $\{\mathcal{L}_i\}_{i=1}^m$ be the subsets of $\{1, 2, \ldots, m\}$ for which the inequalities are saturated and define $\mathcal{L}_i^{\mathcal{S}} := \mathcal{L}_i \cap \mathcal{S}$, the intersection of $\mathcal{L}_i$ with some set $\mathcal{S} \subseteq \{1, 2, \ldots, m\}$.
Construct the directed graph $G = (V, E)$, where

- $V = \{1, 2, \ldots, m\}$, i.e. the vertices are the numbers from 1 to m;
- $E = \{(j, k) : (\forall i)\, j \in \mathcal{L}_i \longrightarrow k \in \mathcal{L}_i\}$, i.e. there is an edge from vertex j to vertex k if whenever vertex j occurs in the given subsets, then so does vertex k.

Now G has to be acyclic by lemma 3.6, so it has a topological sorted order. Let us call this order ν. Let $\mathcal{K}_0 = \emptyset$ and

$$\mathcal{K}_l = \{\nu_{m-l+1}, \ldots, \nu_m\} \tag{29}$$

for $l \in \{1, \ldots, m\}$. The sets $\mathcal{K}_l$, which consist of the last l vertices according to the ordering ν, form a maximal chain $\mathcal{K}_0 \subset \mathcal{K}_1 \subset \cdots \subset \mathcal{K}_{m-1} \subset \mathcal{K}_m$ by construction.

We claim that all the sets $\mathcal{K}_l$ can be constructed from the sets $\{\mathcal{L}_i\}$ by using unions and intersections as dictated by corollary 3.5. The statement is true for $\mathcal{K}_m = \{1, 2, \ldots, m\}$ because every variable must appear in some constraint equation, giving $\mathcal{K}_m = \cup_i \mathcal{L}_i$. The statement is also true for $\mathcal{K}_{m-1} = \{\nu_2, \ldots, \nu_m\}$ since the vertex ν_1 has no in-edges in G by the definition of a topological sort, which means that

$$\mathcal{K}_{m-1} = \bigcup_{\nu_1 \notin \mathcal{L}_i^{\mathcal{K}_m}} \mathcal{L}_i^{\mathcal{K}_m}. \tag{30}$$

For the induction statement, let $l \in \{m-1, \ldots, 2, 1\}$ and assume that $\mathcal{K}_l = \bigcup_i \mathcal{L}_i^{\mathcal{K}_l}$. Since the vertex ν_{m-l} has no in-edges in the induced subgraph generated by the vertices $\mathcal{K}_l$ by the definition of the topological sort, $\mathcal{K}_{l-1}$ can be obtained from the union of all the sets not containing ν_{m-l}:

$$\mathcal{K}_{l-1} = \bigcup_{\nu_{m-l} \notin \mathcal{L}_i^{\mathcal{K}_l}} \mathcal{L}_i^{\mathcal{K}_l}. \tag{31}$$

In more detail, we claim that for all $\omega \neq \nu_{m-l} \in \mathcal{K}_{l-1}$ there exists i such that $\nu_{m-l} \notin \mathcal{L}_i^{\mathcal{K}_l}$ and $\omega \in \mathcal{L}_i^{\mathcal{K}_l}$. If it were not true, that would imply the existence of $\omega \neq \nu_{m-l} \in \mathcal{K}_{l-1}$ such that for all i, $\nu_{m-l} \in \mathcal{L}_i^{\mathcal{K}_l}$ or $\omega \notin \mathcal{L}_i^{\mathcal{K}_l}$. This last condition implies that whenever $\omega \in \mathcal{L}_i^{\mathcal{K}_l}$ it is also true that $\nu_{m-l} \in \mathcal{L}_i^{K_l}$, which corresponds to an edge (ω, ν_{n-l}) in the induced subgraph. □

We have shown that every vertex can be written in precisely the same form as lemma 3.3 and is therefore a point in $\mathcal{Q}$. This proves vertices$(P_{\mathcal{H}}) \subseteq \mathcal{Q}$, which together with the result of step 1, implies vertices$(P_{\mathcal{H}}) = \mathcal{Q}$.

Step 3. Cone part. The final step is to find the set of direction vectors that correspond to the cone part of $P_{\mathcal{H}}$. The generating vectors of the cone are all vectors that satisfy the homogeneous versions of the halfspace inequalities (17), which in our case gives

$$\sum_{k \in \mathcal{K}} Q_k \geqslant 0 \tag{32}$$

for all $\mathcal{K} \subset \{1, 2, \ldots, m\}$. These inequalities are satisfied if and only if $Q_k \geqslant 0$ for all k. We can therefore conclude that the cone part of $P_{\mathcal{H}}$ is cone$(\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_m)$.

This completes our demonstration that $P_{\mathcal{V}}$ is the $\mathcal{V}$-polyhedron description of the $\mathcal{H}$-polyhedron $P_{\mathcal{H}}$. Thus we arrive at the statement we were trying to prove; if the inequalities

$$\sum_{k\in\mathcal{K}} Q_k \geqslant C_{\mathcal{K}} = \frac{1}{2}\left[\sum_{k\in\mathcal{K}} H(A_k)_\varphi + H(R)_\varphi - H(RA_{\mathcal{K}})_\varphi\right] \tag{33}$$

are satisfied for any $\mathcal{K} \subseteq \{1, 2, \ldots, m\}$, then the rate tuple $(Q_1, Q_2, \ldots, Q_m)$ is inside the rate region. This completes the proof of theorem 2.1.

An important discovery by Edmonds [31] is that optimizing a linear function over a supermodular polyhedron can be done in an almost trivial manner by the greedy algorithm. Indeed, let $c_1, c_2, \ldots, c_m$ be any given scalars, and suppose we wish to solve the linear program,

$$\min \sum_{i=1}^{m} c_i Q_i \qquad \text{for} \qquad (Q_1, Q_2, \ldots, Q_m) \in P_{\mathcal{H}}.$$

Let π be the permutation such that

$$c_{\pi(1)} \geqslant c_{\pi(2)} \geqslant \cdots \geqslant c_{\pi(m)}.$$

Edmonds showed that (25) gives an optimum solution to the above linear program. We note in passing that we have no idea how hard it is to optimize over the region described by theorem 2.2.

4. Multiparty information

In this section and the following, we present some tools that we will need in order to prove the outer bound on the rate region stated in theorem 2.2. The following quantity is one possible generalization of the mutual information $I(A; B)$ for multiple parties.

Definition 4.1 (multiparty information)**.** *Given the state $\rho^{X_1X_2\cdots X_m}$ shared between m systems, we define the multiparty information as the following quantity:*

$$\begin{aligned} I(X_1; X_2; \cdots; X_m)_\rho &:= H(X_1) + H(X_2) + \cdots + H(X_m) - H(X_1X_2\cdots X_m) \\ &= \sum_{i=1}^{m} H(X_i) - H(X_1X_2\cdots X_m). \end{aligned} \tag{34}$$

The subadditivity inequality for quantum entropy ensures that the multiparty information is zero if and only if ρ has the tensor product form $\rho^{X_1} \otimes \rho^{X_2} \otimes \cdots \otimes \rho^{X_m}$. The conditional version of the multiparty mutual information is obtained by replacing all the entropies by conditional entropies

$$\begin{aligned} I(X_1; X_2; \cdots; X_m|E)_\rho &:= \sum_{i=1}^{m} H(X_i|E) - H(X_1X_2\cdots X_m|E) \\ &= \sum_{i=1}^{m} H(X_iE) - H(X_1X_2\cdots X_mE) - (m-1)H(E) \\ &= I(X_1; X_2; \cdots; X_m; E) - \sum_{i=1}^{m} I(X_i; E). \end{aligned} \tag{35}$$

This definition of multiparty information has appeared previously in [32–34] and more recently in [16], where many of its properties were investigated.

Next we investigate some formal properties of the multiparty information which will be useful in our later analysis.

Lemma 4.2 (merging of multiparty information terms)**.** *Arguments of the multiparty information can be combined by subtracting their mutual information*

$$I(A; B; X_1; X_2; \cdots; X_m) - I(A; B) = I(AB; X_1; X_2; \cdots; X_m). \tag{36}$$

Proof. This identity is a simple calculation. It is sufficient to expand the definitions and cancel terms. □

Discarding a subsystem inside the conditional multiparty information cannot lead it to increase. This property, more than any other, justifies its use as a measure of correlation.

Lemma 4.3 (monotonicity of conditional multiparty information)**.**

$$I(AB; X_1; \ldots X_m|E) \geqslant I(A; X_1; \ldots X_m|E). \tag{37}$$

Proof. This follows easily from strong subadditivity of quantum entropy (SSA).

$$\begin{aligned}
I(AB; X_1; X_2; \ldots; X_m|E) &= H(ABE) + \sum_i H(X_iE) - H(ABX_1X_2\ldots X_mE) - mH(E) \\
&= H(ABE) + \sum_i H(X_iE) - H(ABX_1X_2\ldots X_mE) - mH(E) \\
&\quad + \underbrace{H(AE) - H(AE)}_{=0} + \underbrace{H(AX_1X_2\ldots X_mE) - H(AX_1X_2\ldots X_mE)}_{=0} \\
&= H(AE) + \sum_i H(X_iE) - H(AX_1X_2\ldots X_m) - mH(E) \\
&\quad + \underbrace{[H(ABE) + H(AX_1X_2\ldots X_mE) - H(AE) - H(ABX_1X_2\ldots X_mE)]}_{\geqslant 0 \text{ by SSA}} \\
&\geqslant H(AE) + \sum_i H(X_iE) - H(AX_1X_2\ldots X_mE) - mH(E) \\
&= I(A; X_1; X_2; \ldots; X_m|E).
\end{aligned}$$

□

We will now prove a multiparty information property that follows from a more general chain rule, but is all that we will need for applications.

Lemma 4.4 (Chain-type Rule)**.**

$$I(AA'; X_1; \ldots; X_m|E) \geqslant I(A; X_1; \ldots; X_m|A'E) \tag{38}$$

Proof.

$$\begin{aligned}
I(AA'; X_1; \ldots; X_m|E) &= H(AA'E) + \sum_{i=1}^{m} H(X_iE) - H(AA'X_1, \ldots, X_m) - mH(E) \\
&= I(A; X_1; \ldots; X_m|A'E) + \sum_{i=1}^{m} [H(A'E) + H(X_iE) - H(E) - H(A'X_iE)] \\
&\geqslant I(A; X_1; \ldots; X_m|A'E).
\end{aligned}$$

The inequality is true by strong subadditivity. □

Remark. It is interesting to note that we have two very similar reduction-of-systems formulae derived from different perspectives. From lemma 4.3 (monotonicity of the multiparty information) we have that

$$I(AB; X_1; \ldots; X_m|E) \geqslant I(A; X_1; \ldots; X_m|E), \tag{39}$$

but we also know from lemma 4.4 (chain-type rule) that

$$I(AB; X_1; \ldots; X_m|E) \geqslant I(A; X_1; \ldots; X_m|BE). \tag{40}$$

The two expressions are inequivalent; one is not strictly stronger than the other. We use both of them depending on whether we want to keep the deleted system around for conditioning.

5. Squashed entanglement

Using the definition of the conditional multiparty information from the previous section, we can define a multiparty squashed entanglement analogous to the bipartite version [35, 36, 15]. The multiparty squashed entanglement has been investigated independently by Yang *et al* [16]. For the convenience of the readers and authors alike, we will provide full proofs of all the E_{sq} properties relevant to the distributed compression problem.

Definition 5.1 (multiparty squashed entanglement)**.** *Consider the state $\rho^{X_1X_2\cdots X_m}$ shared by m parties. We define the multiparty squashed entanglement in the following manner:*

$$\begin{aligned} E_{\mathrm{sq}}(X_1; X_2; \ldots; X_m)_\rho &:= \frac{1}{2} \inf_E \left[\sum_{i=1}^{m} H(X_i|E)_{\tilde{\rho}} - H(X_1X_2\cdots X_m|E)_{\tilde{\rho}} \right] \\ &= \frac{1}{2} \inf_E I(X_1; X_2; \ldots; X_m|E)_{\tilde{\rho}}, \end{aligned} \tag{41}$$

where the infimum is taken over all states $\tilde{\rho}^{X_1X_2\cdots X_mE}$ such that $\mathrm{Tr}_E(\tilde{\rho}^{X_1X_2\cdots X_mE}) = \rho^{X_1X_2\cdots X_m}$. *(We say $\tilde{\rho}$ is an extension of ρ.)*

The dimension of the extension system E can be arbitrarily large, which is in part what makes calculations of the squashed entanglement very difficult except for simple systems. The motivation behind this definition is that we can include a copy of all classical correlations inside the extension E and thereby eliminate them from the multiparty information by conditioning. Since it is impossible to copy quantum information, we know that taking the infimum over all possible extensions E we will be left with a measure of the purely quantum correlations.

Example. It is illustrative to calculate the squashed entanglement for separable states, which are probabilistic mixtures of tensor products of local pure states. Consider the state

$$\rho^{X_1X_2\cdots X_m} = \sum_j p_j |\alpha_j\rangle\langle\alpha_j|^{X_1} \otimes |\beta_j\rangle\langle\beta_j|^{X_2} \otimes \cdots |\zeta_j\rangle\langle\zeta_j|^{X_m},$$

which we choose to extend by adding a system E containing a record of the index j as follows:

$$\tilde{\rho}^{X_1X_2\cdots X_mE} = \sum_j p_j |\alpha_j\rangle\langle\alpha_j|^{X_1} \otimes |\beta_j\rangle\langle\beta_j|^{X_2} \otimes \cdots |\zeta_j\rangle\langle\zeta_j|^{X_m} \otimes |j\rangle\langle j|^{E}.$$

When we calculate conditional entropies we note that for any subset $\mathcal{K} \subseteq \{1, 2, \ldots m\}$,

$$H(X_{\mathcal{K}}|E)_{\tilde{\rho}} = 0. \tag{42}$$

Knowledge of the classical index leaves us with a pure product state for which all the relevant entropies are zero. Therefore, separable states have zero squashed entanglement,

$$E_{\mathrm{sq}}(X_1; X_2; \ldots; X_m)_\rho = \frac{1}{2}\left[\sum_i^m H(X_i|E)_{\tilde{\rho}} - H(X_1X_2\ldots X_m|E)_{\tilde{\rho}}\right] = 0. \tag{43}$$

We now turn our attention to the properties of E_{sq}. Earlier we showed that the squashed entanglement measures purely quantum contributions to the mutual information between systems, in the sense that it is zero for all separable states. In this section, we will show that the multiparty squashed entanglement cannot increase under the action of local operations and classical communication, that is, that E_{sq} is an LOCC-monotone. We will also show that E_{sq} has other desirable properties; it is convex, subadditive and continuous.

Proposition 5.2. *The quantity E_{sq} is an entanglement monotone, i.e. it does not increase on average under local quantum operations and classical communication (LOCC).*

Proof. In order to show this we will follow the argument of [15], which in turn follows the approach described in [37]. We will show that E_{sq} has the following two properties:

(1) Given any unilocal quantum instrument $\mathcal{E}_k$ (a collection of completely positive maps such that $\sum_k \mathcal{E}_k$ is trace preserving [38]) and any quantum state $\rho^{X_1\ldots X_m}$, then

$$E_{\mathrm{sq}}(X_1; X_2; \ldots X_m)_\rho \geqslant \sum_k p_k E_{\mathrm{sq}}(X_1; X_2; \ldots X_m)_{\tilde{\rho}_k}, \tag{44}$$

where

$$p_k = \mathrm{Tr}\,\mathcal{E}_k(\rho^{X_1\ldots X_m}) \qquad \text{and} \qquad \tilde{\rho}_k^{X_1\ldots X_m} = \frac{1}{p_k}\mathcal{E}_k(\rho^{X_1\ldots X_m}). \tag{45}$$

(2) E_{sq} is convex.

Without loss of generality, we assume that $\mathcal{E}_k$ acts on the first system. We will implement the quantum instrument by appending to X_1 environment systems X_1' and X_1'' prepared in standard pure states, applying a unitary U on $X_1X_1'X_1''$, and then tracing out over X_1''. We store k, the classical record of which $\mathcal{E}_k$ occurred, in the X_1' system. More precisely, for any extension of $\rho^{X_1X_2\cdots X_m}$ to $X_1X_2\cdots X_mE$,

$$\rho^{X_1X_2\cdots X_mE} \mapsto \tilde{\rho}^{X_1X_1'X_2\cdots X_mE} := \sum_k \mathcal{E}_k \otimes I_E(\rho^{X_1X_2\cdots X_mE}) \otimes |k\rangle\langle k|^{X_1'}. \tag{46}$$

The argument is then as follows:

$$\begin{aligned}
\frac{1}{2}I(X_1; X_2; \ldots X_m|E)_\rho &= \frac{1}{2}I(X_1X_1'X_1''; X_2; \ldots; X_m|E)_\rho && (47)\\
&= \frac{1}{2}I(X_1X_1'X_1''; X_2; \ldots; X_m|E)_{\tilde{\rho}} && (48)\\
&\geqslant \frac{1}{2}I(X_1X_1'; X_2; \ldots; X_m|E)_{\tilde{\rho}} && (49)\\
&\geqslant \frac{1}{2}I(X_1; X_2; \ldots; X_m|EX_1')_{\tilde{\rho}} && (50)\\
&= \frac{1}{2}\sum_k p_k I(X_1; X_2; \ldots; X_m|E)_{\tilde{\rho}_k} && (51)\\
&\geqslant \sum_k p_k E_{\mathrm{sq}}(X_1; X_2; \ldots; X_m)_{\tilde{\rho}_k}. && (52)
\end{aligned}$$

The equality (47) is true because adding an uncorrelated ancilla does not change the entropy of the system. The transition $\rho \to \tilde{\rho}$ is unitary and does not change entropic quantities so (48) is true. For (49) we use the monotonicity of conditional multiparty information, lemma 4.3. In (50) we use the chain-type rule from lemma 4.4. In (51) we use the index information k contained in X_1'. Finally, since E_{sq} is the infimum over all extensions, it must be no more than the particular extension E, so (52) must be true. Now since the extension E in (47) was arbitrary, it follows that $E_{\mathrm{sq}}(X_1; X_2; \ldots; X_m)_\rho \geqslant \sum_k p_k E_{\mathrm{sq}}(X_1; X_2; \ldots; X_m)_{\tilde{\rho}_k}$ which completes the proof of property 1.

To show the convexity of E_{sq}, we again follow the same route as in [15]. Consider the states $\rho^{X_1X_2\cdots X_m}$ and $\sigma^{X_1X_2\cdots X_m}$ and their extensions $\tilde{\rho}^{X_1X_2\cdots X_mE}$ and $\tilde{\sigma}^{X_1X_2\cdots X_mE}$ defined over the same system E. We can also define the weighted sum of the two states $\tau^{X_1X_2\cdots X_m} = \lambda\rho^{X_1X_2\cdots X_m} + (1-\lambda)\sigma^{X_1X_2\cdots X_m}$ and the following valid extension:

$$\tilde{\tau}^{X_1X_2\cdots X_mEE'} = \lambda\rho^{X_1X_2\cdots X_mE} \otimes |0\rangle\langle 0|^{E'} + (1-\lambda)\sigma^{X_1X_2\cdots X_mE} \otimes |1\rangle\langle 1|^{E'}. \quad (53)$$

Using the definition of squashed entanglement we know that

$$\begin{aligned} E_{\mathrm{sq}}(X_1; X_2; \ldots; X_m)_\tau &\leqslant \tfrac{1}{2} I(X_1; X_2; \ldots; X_m|EE')_{\tilde{\tau}} \\ &= \tfrac{1}{2}[\lambda I(X_1; X_2; \ldots; X_m|E)_{\tilde{\rho}} + (1-\lambda) I(X_1; X_2; \ldots; X_m|E)_{\tilde{\sigma}}]. \end{aligned}$$

Since the extension system E is completely arbitrary we have

$$E_{\mathrm{sq}}(X_1; X_2; \ldots; X_m)_\tau \leqslant \lambda E_{\mathrm{sq}}(X_1; X_2; \ldots; X_m)_\rho + (1-\lambda) E_{\mathrm{sq}}(X_1; X_2; \ldots; X_m)_\sigma, \quad (54)$$

so E_{sq} is convex.

We have shown that E_{sq} satisfies both properties 1 and 2. Therefore, it must be an entanglement monotone. □

Subadditivity on product states. Another desirable property for measures of entanglement is that they should be additive or at least subadditive on tensor products of the same state. Subadditivity of E_{sq} is easily shown from the properties of multiparty information.

Proposition 3. *E_{sq} is subadditive on tensor product states, i.e.,*

$$E_{\mathrm{sq}}(\rho^{X_1Y_1;X_2Y_2;\ldots;X_mY_m}) \leqslant E_{\mathrm{sq}}(\rho^{X_1;X_2;\ldots;X_m}) + E_{\mathrm{sq}}(\rho^{Y_1;Y_2;\ldots;Y_m}), \quad (55)$$

where $\rho^{X_1Y_1X_2Y_2\cdots X_mY_m} = \rho^{X_1X_2\cdots X_m} \otimes \rho^{Y_1Y_2\cdots Y_m}$.

Proof. Assume that $\rho^{X_1X_2\cdots X_mE}$ and $\rho^{Y_1Y_2\cdots Y_mE'}$ are extensions. Together they form an extension $\rho^{X_1Y_1X_2Y_2\cdots X_mY_mEE'}$ for the product state,

$$2E_{\mathrm{sq}}(X_1Y_1; X_2Y_2; \ldots; X_mY_m)_\rho \leqslant I(X_1Y_1; X_2Y_2; \ldots; X_mY_m|EE') \quad (56)$$

$$= \sum_i H(X_iY_iEE') - H(X_1Y_1X_2Y_2\cdots X_mY_mEE') - (m-1)H(EE') \quad (57)$$

$$= I(X_1; X_2; \ldots; X_m|E) + I(Y_1; Y_2; \ldots; Y_m|E'). \quad (58)$$

The first line holds because the extension for the XY system that can be built by combining the X and Y extensions is not the most general extension. The proposition then follows because the inequality holds for all extensions of ρ and σ. □

The question of whether E_{sq} is additive, meaning superadditive in addition to subadditive, remains an open problem. Indeed, if it were possible to show that correlation between the X and Y extensions is unnecessary in the evaluation of the squashed entanglement of $\rho \otimes \sigma$, then

E_{sq} would be additive. This is provably true in the bipartite case [15] but the same method does not seem to work with three or more parties.

Continuity. The continuity of bipartite E_{sq} was conjectured in [15] and proved by Alicki and Fannes in [39]. We will follow the same argument here to prove the continuity of the multiparty squashed entanglement. The key to the continuity proof is the following lemma which makes use of an ingenious geometric construction.

Lemma 5.4 (Continuity of conditional entropy [39]). *Given density matrices ρ^{AB} and σ^{AB} on the space $\mathcal{H}^A \otimes \mathcal{H}^B$ such that*

$$\|\rho - \sigma\|_1 = \tfrac{1}{2}\,\mathrm{Tr}|\rho - \sigma| \leqslant \epsilon, \tag{59}$$

it is true that

$$\left|H(A|B)_\rho - H(A|B)_\sigma\right| \leqslant 4\epsilon \log d_A + 2h(\epsilon), \tag{60}$$

where $d_A = \dim \mathcal{H}^A$ and $h(\epsilon) = -\epsilon \log \epsilon - (1-\epsilon)\log(1-\epsilon)$ is the binary entropy.

This seemingly innocuous technical lemma makes it possible to prove the continuity of E_{sq} in spite of the unbounded dimension of the extension system.

Proposition 5.5 (E_{sq} is continuous). *For all $\rho^{X_1X_2\cdots X_m}, \sigma^{X_1X_2\cdots X_m}$ with $\|\rho - \sigma\|_1 \leqslant \epsilon$, $\|E_{\mathrm{sq}}(\rho) - E_{\mathrm{sq}}(\sigma)\| \leqslant \epsilon'$ where ϵ' depends on ϵ and vanishes as $\epsilon \to 0$.*

The precise form of ϵ' can be found in equation (67).

Proof. Proximity in trace distance implies proximity in fidelity distance [40], in the sense that

$$F(\rho^{X_1X_2\cdots X_m}, \sigma^{X_1X_2\cdots X_m}) \geqslant 1 - \epsilon, \tag{61}$$

but by Uhlmann's theorem [41] this means that we can find purifications $|\rho\rangle^{X_1X_2\cdots X_mR}$ and $|\sigma\rangle^{X_1X_2\cdots X_mR}$ such that

$$F(|\rho\rangle^{X_1X_2\cdots X_mR}, |\sigma\rangle^{X_1X_2\cdots X_mR}) \geqslant 1 - \epsilon. \tag{62}$$

Now if we imagine some general operation Λ that acts only on the purifying system R

$$\rho^{X_1X_2\cdots X_mE} = (I^{X_1X_2\cdots X_m} \otimes \Lambda^{R\to E})\,|\rho\rangle\langle\rho|^{X_1X_2\cdots X_mR} \tag{63}$$

$$\sigma^{X_1X_2\cdots X_mE} = (I^{X_1X_2\cdots X_m} \otimes \Lambda^{R\to E})\,|\sigma\rangle\langle\sigma|^{X_1X_2\cdots X_mR}, \tag{64}$$

we have from the monotonicity of fidelity for quantum channels that

$$F(\rho^{X_1X_2\cdots X_mE}, \sigma^{X_1X_2\cdots X_mE}) \geqslant F(|\rho\rangle^{X_1X_2\cdots X_mR}, |\sigma\rangle^{X_1X_2\cdots X_mR}) \geqslant 1 - \epsilon, \tag{65}$$

which in turn implies [40] that

$$\|\rho^{X_1X_2\cdots X_mE} - \sigma^{X_1X_2\cdots X_mE}\|_1 \leqslant 2\sqrt{\epsilon}. \tag{66}$$

Now we can apply lemma 5.4 to each term in the multiparty information to obtain

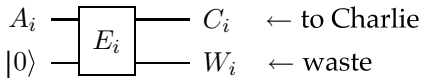
$$\begin{aligned}
&|I(X_1; X_2; \ldots X_m|E)_\rho - I(X_1; X_2; \ldots X_m|E)_\sigma| \\
&\quad\leqslant \sum_{i=1}^{m} |H(X_i|E)_\rho - H(X_i|E)_\sigma| + |H(X_1X_2\cdots X_m|E)_\rho - H(X_1X_2\cdots X_m|E)_\sigma| \\
&\quad\leqslant \sum_{i=1}^{m} [8\sqrt{\epsilon}\log d_i + 2h(2\sqrt{\epsilon})] + 8\sqrt{\epsilon}\log\left(\prod_{i=1}^{m} d_i\right) + 2h(2\sqrt{\epsilon}) \\
&\quad= 16\sqrt{\epsilon}\log\left(\prod_{i=1}^{m} d_i\right) + (m+1)2h(2\sqrt{\epsilon}) =: \epsilon',
\end{aligned} \tag{67}$$

where $d_i = \dim \mathcal{H}^{X_i}$ and $h(.)$ is as defined in lemma 5.4. Since we have shown the above inequalities for *any* extension E and the quantity ϵ' vanishes as $\epsilon \to 0$, we have proved that E_{sq} is continuous. □

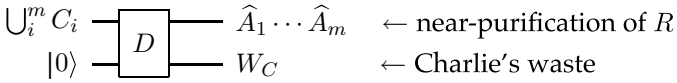
6. Proof of outer bound on the rate region

Armed with the new tools of multiparty information and squashed entanglement, we are now ready to give the proof of theorem 2.2. We want to show that *any* distributed compression protocol which works must satisfy all of the inequalities of type (6) from theorem 2.2. We break the proof into three steps.

Step 1. Decoupling formula. We know that the input system $|\psi\rangle^{A^n R^n}$ is a pure state. If we account for the Stinespring dilations of each encoding and decoding operation, then we can view any protocol as implemented by unitary transformations with ancilla and waste. Therefore, the output state (including the waste systems) should also be pure. More specifically, the encoding operations are modeled by CPTP maps E_i with outputs C_i of dimension 2^{nQ_i}. In our analysis, we will keep the Stinespring dilations of the CPTP maps W_i so the evolution as a whole will be unitary.



Once Charlie receives the systems that were sent to him, he will apply a decoding CPTP map D with output system $\widehat{A} = \widehat{A}_1\widehat{A}_2\ldots\widehat{A}_m$ isomorphic to the original $A = A_1A_2\cdots A_m$.



In what follows we will use figure 6 extensively in order to keep track of the evolution and purity of the states at various points in the protocol.

The starting point of our argument is the fidelity condition (1) for successful distributed compression, which we restate below for convenience

$$F\left(|\psi\rangle^{A^nR^n}, \sigma^{\widehat{A}^nR^n}\right) \geqslant 1-\epsilon, \tag{68}$$

where $|\psi\rangle^{A^nR^n} = (|\varphi\rangle^{A_1A_2\cdots A_mR})^{\otimes n}$ is the input state to the protocol and $\sigma^{\widehat{A}^nR^n}$ is the output state of the protocol. Since $\sigma^{\widehat{A}^nR^n}$ has high fidelity with a rank-1 state, it must have one large eigenvalue

$$\lambda_{\max}\left(\sigma^{\widehat{A}^nR^n}\right) \geqslant 1-\epsilon. \tag{69}$$

Therefore, the full output state $|\sigma\rangle^{\widehat{A}R^nW_1\cdots W_mW_C}$ has Schmidt decomposition of the form

$$|\sigma\rangle^{\widehat{A}^nR^nW_1\cdots W_mW_C} = \sum_i \sqrt{\lambda_i}\,|e_i\rangle^{\widehat{A}^nR^n} \otimes |f_i\rangle^{W_1\cdots W_mW_C}, \tag{70}$$

where $|e_i\rangle, |f_i\rangle$ are orthonormal bases and $\lambda_1 = \lambda_{\max} \geqslant 1-\epsilon$.

Next we show that the output state $|\sigma\rangle^{\widehat{A}^nR^nW_1\cdots W_mW_C}$ is very close in fidelity to a totally decoupled state $\sigma^{\widehat{A}^nR^n} \otimes \sigma^{W_1\cdots W_mW_C}$, which is a tensor product of the marginals of $|\sigma\rangle$ on the subsystems $\widehat{A}^nR^n$ and $W_1\cdots W_mW_C$:

$$\begin{aligned} &F\left(|\sigma\rangle^{\widehat{A}^nR^nW_1\cdots W_mW_C}, \sigma^{\widehat{A}^nR^n} \otimes \sigma^{W_1\cdots W_mW_C}\right) \\ &\quad = \mathrm{Tr}\left[|\sigma\rangle\langle\sigma|^{\widehat{A}^nR^nW_1\cdots W_mW_C}\left(\sigma^{\widehat{A}^nR^n} \otimes \sigma^{W_1\cdots W_mW_C}\right)\right] \\ &\quad = \sum_i \lambda_i^3 \geqslant (1-\epsilon)^3 \geqslant 1-3\epsilon. \end{aligned} \tag{71}$$

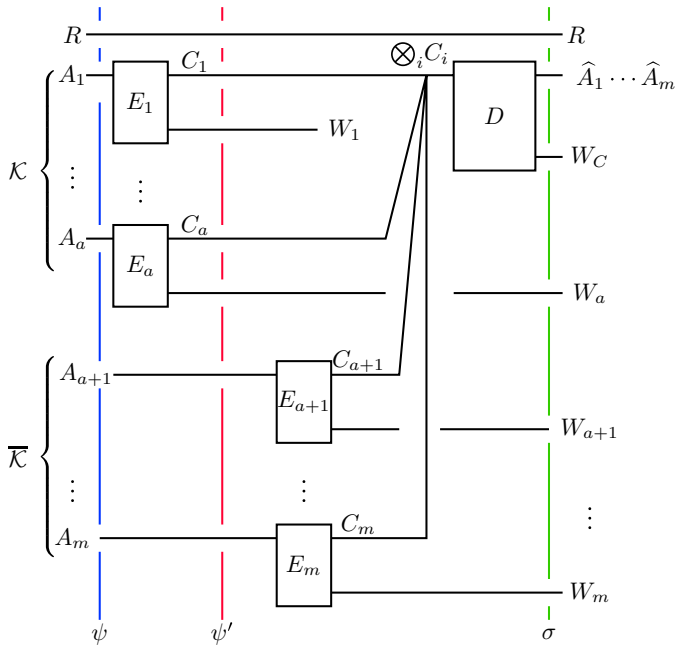


Figure 6. A general distributed compression circuit diagram showing the encoding operations E_i with output systems C_i (compressed data) and W_i (waste). The decoding operation takes all the compressed data $\bigotimes_i C_i$ and applies the decoding operation D to output a state $\sigma^{\widehat{A}^n R^n}$ which has high fidelity with the original $|\psi\rangle^{A^n R^n}$.

Using the relationship between fidelity and trace distance [40], we can transform (71) into the trace distance bound

$$\| |\sigma\rangle \langle\sigma|^{\widehat{A}^n R^n W_1 \cdots W_m W_C} - \sigma^{\widehat{A}^n R^n} \otimes \sigma^{W_1 \cdots W_m W_C} \|_1 \leqslant 2\sqrt{3\epsilon}. \tag{72}$$

By the contractivity of trace distance, the same equation must be true for any subset of the systems. This bound combined with the Fannes inequality implies that the entropies taken with respect to the output state are nearly additive:

$$\begin{aligned} |H(R^n W_{\mathcal{K}})_\sigma - H(R^n)_\sigma + H(W_{\mathcal{K}})_\sigma| &\leqslant 2\sqrt{3\epsilon} \log(d_{R^n} d_{W_{\mathcal{K}}}) + \eta(2\sqrt{3\epsilon}) \\ &\leqslant 2\sqrt{3\epsilon} \log\left(d_{A^n} d_{A_{\mathcal{K}}^{2n}}\right) + \eta(2\sqrt{3\epsilon}) \\ &\leqslant 2\sqrt{3\epsilon} n \log\left(d_A^3\right) + \eta(2\sqrt{3\epsilon}) = f_1(\epsilon, n). \end{aligned} \tag{73}$$

for any subset $\mathcal{K} \subseteq \{1, 2, \ldots, m\}$ with $\epsilon \leqslant \frac{1}{12e^2}$ and $\eta(x) = -x \log x$. In the second line we have used the fact that $d_A = d_R$ and exploited the fact that $d_{W_{\mathcal{K}}}$ can be taken less than or equal to $d_{A_{\mathcal{K}}^{2n}}$, the maximum size of an environment required for a quantum operation with inputs and outputs of dimension no larger than $d_{A_{\mathcal{K}}^n}$.

Step 2. Dimension counting. The entropy of any system is bounded above by the logarithm of its dimension. In the case of the systems that participants send to Charlie, this implies that

$$n \sum_{k \in \mathcal{K}} Q_k \geqslant H(C_{\mathcal{K}})_{\psi'}. \tag{74}$$

We can use this fact and the diagram of figure 6 to bound the rates Q_i. First we add $H(A_{\bar{\mathcal{K}}})_\psi = H(A_{\bar{\mathcal{K}}})_{\psi'}$ to both sides of equation (74) and obtain the inequality

$$H(A_{\bar{\mathcal{K}}})_\psi + n\sum_{k\in\mathcal{K}} Q_k \geqslant H(C_{\mathcal{K}})_{\psi'} + H(A_{\bar{\mathcal{K}}})_{\psi'} \geqslant H(C_{\mathcal{K}}A_{\bar{\mathcal{K}}})_{\psi'}. \tag{75}$$

For each encoding operation, the input system A_i is unitarily related to the outputs $C_i W_i$ so we can write

$$H(A_i)_\psi = H(W_i C_i)_{\psi'} \leqslant H(W_i)_{\psi'} + H(C_i)_{\psi'} \leqslant H(W_i)_{\psi'} + nQ_i, \tag{76}$$

where in the last inequality we have used the dimension bound $H(C_i) \leqslant nQ_i$. If we collect all the Q_i terms from equations (75) and (76), we obtain the inequalities

$$n\sum_{i\in\mathcal{K}} Q_i \geqslant H(C_{\mathcal{K}}A_{\bar{\mathcal{K}}})_{\psi'} - H(A_{\bar{\mathcal{K}}})_\psi \tag{77}$$

$$n\sum_{i\in\mathcal{K}} Q_i \geqslant \sum_{i\in\mathcal{K}} H(A_i)_\psi - \sum_{i\in\mathcal{K}} H(W_i)_{\psi'}. \tag{78}$$

Now add equations (77) and (78) to get

$$\begin{aligned}
2n\sum_{i\in\mathcal{K}} Q_i &\geqslant \sum_{i\in\mathcal{K}} H(A_i)_\psi - \sum_{i\in\mathcal{K}} H(W_i)_{\psi'} + H(C_{\mathcal{K}}A_{\bar{\mathcal{K}}})_{\psi'} - H(A_{\bar{\mathcal{K}}})_\psi \\
&=^{(1)} \sum_{i\in\mathcal{K}} H(A_i)_\psi - \sum_{i\in\mathcal{K}} H(W_i)_{\psi'} + H(W_{\mathcal{K}}R^n)_{\psi'} - H(R^n A_{\mathcal{K}})_\psi \\
&\geqslant^{(2)} \sum_{i\in\mathcal{K}} H(A_i)_\psi - \sum_{i\in\mathcal{K}} H(W_i)_{\psi'} + H(W_{\mathcal{K}})_{\psi'} \\
&\quad + H(R^n)_{\psi'} - H(R^n A_{\mathcal{K}})_\psi - f_1(\epsilon, n) \\
&= \left[\sum_{i\in\mathcal{K}} H(A_i) + H(R^n) - H(R^n A_{\mathcal{K}})\right]_\psi \\
&\quad + H(W_{\mathcal{K}})_{\psi'} - \sum_{i\in\mathcal{K}} H(W_i)_{\psi'} - f_1(\epsilon, n),
\end{aligned} \tag{79}$$

where the equality $^{(1)}$ comes about because the systems $|\psi\rangle^{A_{\mathcal{K}}A_{\bar{\mathcal{K}}}R^n}$ and $|\psi'\rangle^{C_{\mathcal{K}}W_{\mathcal{K}}A_{\bar{\mathcal{K}}}R^n}$ are pure. The inequality (73) from step 1 was used in $^{(2)}$.

Step 3. Squashed entanglement. We would like to have a bound on the extra terms in equation (79) that does not depend on the encoding and decoding maps. We can accomplish this if we bound the waste terms $\sum_{i\in\mathcal{K}} H(W_i)_\sigma - H(W_{\mathcal{K}})_\sigma$ by the squashed entanglement $2E_{\mathrm{sq}}(A_{k_1}; \ldots; A_{k_l})_\psi$ of the input state for each $\mathcal{K} = \{k_1, k_2, \ldots, k_l\} \subseteq \{1, \ldots, m\}$ plus some small corrections. The proof requires a continuity statement analogous to (73), namely that

$$|H(W_i) - H(W_i|R)| \leqslant f_2(\epsilon, n), \tag{80}$$

where f_2 is some function such that $f_2(\epsilon, n)/n \to 0$ as $\epsilon \to 0$. The proof is very similar to that of (73) so we omit it.

Furthermore, if we allow an arbitrary transformation $\mathcal{E}^{R\to E}$ to be applied to the R system, we will obtain some general extension but the analog of equation (80) will

remain true by the contractivity of the trace distance under CPTP maps. We can therefore write:

$$\begin{aligned}
\sum_{i\in\mathcal{K}} H(W_i)_\psi - H(W_\mathcal{K})_\psi &\leqslant \sum_{i\in\mathcal{K}} H(W_i|E) - H(W_\mathcal{K}|E) + [|\mathcal{K}|+1]f_2(\epsilon,n)\\
&= I\big(W_{k_1}; W_{k_2}; \ldots; W_{k_l}; E\big) - I\big(W_{k_1}; E\big) - \sum_{i\in\{\mathcal{K}\setminus k_1\}} I(W_i; E) + f_2'(\epsilon,n)\\
&=^{(1)} I\big(W_{k_1}E; W_{k_2}; \ldots; W_{k_l}\big) - \sum_{i\in\{\mathcal{K}\setminus k_1\}} I(W_i; E) + f_2'(\epsilon,n)\\
&\leqslant^{(2)} I\big(A_{k_1}E; W_{k_2}; \ldots; W_{k_l}\big) - \sum_{i\in\{\mathcal{K}\setminus k_1\}} I(W_i; E) + f_2'(\epsilon,n)\\
&=^{(1)} I\big(A_{k_1}; W_{k_2}; \ldots; W_{k_l}, E\big) - I(A_{k_1}; E) - \sum_{i\in\{\mathcal{K}\setminus k_1\}} I(W_i; E) + f_2'(\epsilon,n)\\
&\leqslant^{(3)} I\big(A_{k_1}; A_{k_2}; \ldots; A_{k_l}; E\big) - \sum_{i\in\mathcal{K}} I(A_i; E) + f_2'(\epsilon,n)\\
&\leqslant I\big(A_{k_1}; A_{k_2}; \ldots; A_{k_l}|E\big) + f_2'(\epsilon,n),
\end{aligned}$$

where we have used the shorthand $f_2'(\epsilon,n) = [|\mathcal{K}|+1]f_2(\epsilon,n)$ for brevity. Equations marked$^{(1)}$ use lemma 4.2 and inequality $^{(2)}$ comes about from lemma 4.3, the monotonicity of the multiparty information. Inequality $^{(3)}$ is obtained when we repeat the steps for $k_2,\ldots,k_l$. The above result is true for any extension E but we want to find the tightest possible lower bound for the rate region so we take the infimum over all possible extensions E thus arriving at the definition of squashed entanglement.

Putting together equation (79) from step 2 and the bound from step 3 we have

$$\begin{aligned}
2n\sum_{i\in\mathcal{K}} Q_i &\geqslant \left[\sum_{i\in\mathcal{K}} H(A_i) + H(R^n) - H(R^nA_\mathcal{K})\right]_\psi - \left(\sum_{i\in\mathcal{K}} H(W_i)_{\psi'} - H(W_\mathcal{K})_{\psi'}\right) - f_1(\epsilon,n)\\
&\geqslant \left[\sum_{i\in\mathcal{K}} H(A_i) + H(R^n) - H(R^nA_\mathcal{K})\right]_\psi\\
&\quad - 2E_{\mathrm{sq}}\big(A_{k_1}; \ldots; A_{k_l}\big)_\psi - f_1(\epsilon,n) - f_2'(\epsilon,n).
\end{aligned}$$

We can simplify the expression further by using the fact that $|\psi\rangle = |\varphi\rangle^{\otimes n}$ to obtain

$$\begin{aligned}
\sum_{k\in\mathcal{K}} Q_k &\geqslant \frac{1}{2}\left[\sum_{k\in\mathcal{K}} H(A_k) + H(R) - H(RA_\mathcal{K})\right]_\varphi\\
&\quad - E_{\mathrm{sq}}\big(A_{k_1}; A_{k_2}; \ldots A_{k_l}\big)_\varphi - \frac{f_1(\epsilon,n)}{2n} - \frac{f_2'(\epsilon,n)}{2n},
\end{aligned}$$

where we used explicitly the additivity of the entropy for tensor product states and the subadditivity of squashed entanglement demonstrated in proposition 3.

Theorem 2.2 follows from the above since $\epsilon > 0$ was arbitrary and $(f_1(\epsilon,n) + f_2'(\epsilon,n))/n \to 0$ as $\epsilon \to 0$.

7. Discussion

We have shown how to build protocols for multiparty distributed compression out of the two-party fully quantum Slepian–Wolf protocol. The resulting achievable rates generalize those

found in [4] for the two-party case and, for the most part, the arguments required are direct generalizations of those required for two parties. The most interesting divergence is to be found in section 3, where we characterize the multiparty rates that can be achieved starting from sequential applications of the two-party protocol. These rates are most easily expressed in terms of the vertices of the associated polyhedron and we use a graph-theoretic argument to describe the polyhedron instead in terms of facet inequalities. We note that it is possible to give a direct proof [42] that this multiparty rate region is achievable by mimicking the proof techniques of [4], but in the spirit of that paper, we wanted to demonstrate that the more complicated multiparty compression protocols can themselves be built out of the simpler near-universal building block of two-party FQSW. Multiparty compression thus joins entanglement distillation, entanglement-assisted communication, channel simulation, communication over quantum broadcast channels, state redistribution [27, 43] and many other protocols in the FQSW matriarchy.

Multiparty FQSW can then itself be used as a building block for other multiparty protocols. For example, when classical communication between the senders and the receiver is free, combining multiparty FQSW with teleportation reproduces the multiparty state merging protocol of [14]. Running the protocol backwards in time yields an optimal reverse Shannon theorem for broadcast channels [44].

The achievable rates we describe here, however, are only known to be optimal in the case when the source density operator is separable. Otherwise, we proved an outer bound on the rate region of the same form as the achievable rate region but with a correction term equal to the multiparty squashed entanglement of the source. In order to perform our analysis, we developed a number of basic properties of this quantity, notably that it is a convex, subadditive, continuous entanglement measure, facts that were established independently in [16]. We are thus left with some compelling open problems. The most obvious is, of course, to close the gap between our inner and outer bounds on distributed compression. While that may prove to be difficult, some interesting related questions may be easier. For example, can the gap between the rate region we have presented here and the true distributed compression region be characterized by an entanglement measure? That is, while we have used the multiparty squashed entanglement as a correction term, could it be that the true correction term is an entanglement monotone? Also, focusing on the squashed entanglement, the two-party version is known to be not just subadditive but additive. Is the same true of the multiparty version?

Acknowledgments

We would like to thank Leonid Chindelevitch, Frédéric Dupuis, Michal Horodecki, Jonathan Oppenheim and Andreas Winter for helpful comments on the subjects of distributed compression and squashed entanglement. The authors gratefully acknowledge funding from the Alfred P Sloan Foundation, the Canada Research Chairs program, CIFAR, FQRNT, MITACS and NSERC.

Note added in proof. The additivity of the multiparty squashed entanglement was recently proved in an updated version of [16] which now includes W Song in the author list.

References

[1] Bennett C H, Bernstein H J, Popescu S and Schumacher B 1996 Concentrating partial entanglement by local operations *Phys. Rev.* A **53** 2046–52 (*Preprint* quant-ph/9511030)

[2] Devetak I, Harrow A W and Winter A 2004 A family of quantum protocols *Phys. Rev. Lett.* **93** 230504 (*Preprint* quant-ph/0308044)

[3] Devetak I, Harrow A W and Winter A 2005 *A Resource Framework for Quantum Shannon Theory* (*Preprint* quant-ph/0512015)
[4] Abeyesinghe A, Devetak I, Hayden P and Winter A 2006 The mother of all protocols: restructuring quantum information's family tree *Preprint* quant-ph/0606225
[5] Holevo A S 1998 The capacity of the quantum channel with general signal states *IEEE Trans. Inf. Theory* **44** 269–73 (*Preprint* quant-ph/9611023)
[6] Schumacher B and Westmoreland M D 1997 Sending classical information via noisy quantum channels *Phys. Rev.* A **56** 131–8
[7] Bennett C H, Shor P W, Smolin J A and Thapliyal A V 1999 Entanglement-assisted classical capacity of noisy quantum channels *Phys. Rev. Lett.* **83** 3081 (*Preprint* quant-ph/9904023)
[8] Devetak I 2005 The private classical capacity and quantum capacity of a quantum channel *IEEE Trans. Inf. Theory* **51** 44 (*Preprint* quant-ph/0304127)
[9] Bennett C H, DiVincenzo D P, Smolin J A and Wootters W K 1996 Mixed-state entanglement and quantum error correction *Phys. Rev.* A **54** 3824 (*Preprint* quant-ph/9604024)
[10] Devetak I and Winter A 2005 Distillation of secret key and entanglement from quantum states *Proc. R. Soc. Lond. Ser.* A **461** 207–35 (*Preprint* quant-ph/0306078)
[11] Plenio M B and Virmani S 2007 An introduction to entanglement measures *Quantum Inf. Comp.* **7** 1 (*Preprint* quant-ph/0504163)
[12] Slepian D and Wolf J 1973 Noiseless coding of correlated information sources *IEEE Trans. Inf. Theory* **19** 471–80
[13] Ahn C, Doherty A, Hayden P and Winter A 2006 On the distributed compression of quantum information *IEEE Trans. Inf. Theory* **52** 4349 (*Preprint* quant-ph/0403042)
[14] Horodecki M, Oppenheim J and Winter A 2005 Quantum information can be negative *Nature* **436** 673 (doi:10.1038/nature03909)
[15] Christandl M and Winter A 2004 Squashed entanglement—an additive entanglement measure *J. Math. Phys.* **45** 829 (*Preprint* quant-ph/0308088)
[16] Yang D, Horodecki K, Horodecki M, Horodecki P and Oppenheim J 2007 Squashed entanglement for multipartite states and entanglement measures based on the mixed convex roof *Preprint* quant-ph/0704.2236
[17] Hayden P M, Horodecki M and Terhal B M 2001 The asymptotic entanglement cost of preparing a quantum state *J. Phys. A: Math. Gen.* **34** 6891–8
[18] Rains E M 1999 A rigorous treatment of distillable entanglement *Phys. Rev.* A **60** 173 (*Preprint* quant-ph/9809078)
[19] Vedral V and Plenio M B 1998 Entanglement measures and purification procedures *Phys. Rev.* A **57** 1619 (*Preprint* quant-ph/9707035)
[20] Linden N, Popescu S, Schumacher B and Westmoreland M 2005 Reversibility of local transformations of multiparticle entanglement *Quantum Inf. Proc.* **4** 241–50 (*Preprint* quant-ph/9912039)
[21] Dur W, Cirac J I and Tarrach R 1999 Separability and distillability of multiparticle quantum systems *Phys. Rev. Lett.* **83** 3562 (*Preprint* quant-ph/9903018)
[22] Coffman V, Kundu J and Wootters W K 2000 Distributed entanglement *Phys. Rev.* A **61** 052306 (*Preprint* quant-ph/9907047)
[23] Bennett C H, Popescu S, Rohrlich D, Smolin J A and Thapliyal A V 2000 Exact and asymptotic measures of multipartite pure-state entanglement *Phys. Rev.* A **63** 012307 (*Preprint* quant-ph/9908073)
[24] Cover T 1975 A proof of the data compression theorem of Slepian and Wolf for ergodic sources *IEEE Trans. Inf. Theory* **21** 226–8
[25] Schumacher B 1996 Sending entanglement through noisy quantum channels *Phys. Rev.* A **54** 2614–28 (*Preprint* quant-ph/9604023)
[26] Dupuis F and Hayden P 2006 A father protocol for quantum broadcast channels *Preprint* quant-ph/0612155
[27] Devetak I and Yard J 2006 The operational meaning of quantum conditional information *Preprint* quant-ph/0612050
[28] Ziegler G M 1995 *Lectures on Polytopes* (New York: Springer)
[29] Tse D and Hanley S 1998 Multiaccess fading channels: polymatroid structure, optimal resource allocation and throughput capacities *IEEE Trans. Inf. Theory* **44** 2796–815
[30] Lieb E H and Ruskai M B 1973 Proof of the strong subaddivity of quantum-mechanical entropy *J. Math. Phys.* **14** 1938–41
[31] Edmonds J 1969 Submodular functions, matroids and certain polyhedra *Proc. Calgary Int. Conf. Combinatorial Structures and Algorithms (June 1969)* pp 69–87
Edmonds J 2003 *LNCS* **2570** 11–26 (reprinted)
[32] Lindblad G 1973 Entropy, information and quantum measurements *Commun. Math. Phys.* **33** 305–22

[33] Horodecki R 1994 Informationally coherent quantum systems *Phys. Lett.* A **187** 145–50
[34] Groisman B, Popescu S and Winter A 2005 Quantum, classical and total amount of correlations in a quantum state *Phys. Rev.* A **72** 032317
[35] Tucci R R 1999 Quantum entanglement and conditional information transmission *Preprint* quant-ph/9909041
[36] Tucci R R 2002 Entanglement of distillation and conditional mutual information *Preprint* quant-ph/0202144
[37] Vidal G 2000 Entanglement monotones *J. Mod. Opt.* **47** 355 (*Preprint* quant-ph/9807077)
[38] Davies E and Lewis J 1970 An operational approach to quantum probability *Commun. Math. Phys.* **17** 239–60
[39] Alicki R and Fannes M 2003 Continuity of quantum mutual information *Preprint* quant-ph/0312081
[40] Fuchs C A and Graaf J van de 1999 Cryptographic distinguishability measures for quantum-mechanical states *IEEE Trans. Inf. Theory* **45** 1216 (doi:10.1109/18.761271)
[41] Uhlmann A 1976 The 'transition probability' in the state space of a *-algebra *Rep. Math. Phys.* **9** 273
[42] Hayden P and Winter A 2006 Achievable rates for multiparty distributed compression (unpublished)
[43] Oppenheim J 2007 Redistributing quantum information from fully quantum Slepian–Wolf (private communication)
[44] Hayden P and Dupuis F 2007 An optimal reverse Shannon theorem for quantum broadcast channels (in preparation)