We present a protocol for performing state merging when multiple parties share a single copy of a mixed state, and analyze the entanglement cost in terms of min- and max-entropies. Our protocol allows for interpolation between corner points of the rate region without the need for time-sharing, a primitive which is not available in the one-shot setting. We also compare our protocol to the more naive strategy of repeatedly applying a single-party merging protocol one party at a time, by performing a detailed analysis of the rates required to merge variants of the embezzling states. Finally, we analyze a variation of multiparty merging, which we call split-transfer, by considering two receivers and many additional helpers sharing a mixed state. We give a protocol for performing a split-transfer and apply it to the problem of assisted entanglement distillation.

I. INTRODUCTION

An important part of quantum information theory is concerned with the design and analysis of quantum communication protocols. The subject has flourished over the past two decades, with early discoveries like teleportation \cite{1} and superdense coding \cite{2} laying the groundwork for a series of major advances over the last five years. (See, for example, \cite{3-10}.) Another early result, Schumacher compression \cite{11}, studies the amount of quantum communication required to transmit to another location a sequence of quantum states $|\psi_1\rangle|\psi_2\rangle|\psi_3\rangle\ldots$ emitted by a statistical source. If we assume the states coming from the source are independent and identically distributed (i.e., an i.i.d source), we get the quantum analogue of Shannon compression, and the optimal rate of compression is given by the von Neumann entropy $S(\rho)$ of the density matrix $\rho = \sum_j p_j \psi_j$ associated with the source \cite{11}. This gives an informational meaning to the von Neumann entropy, whose original definition was motivated by the desire to extend the Gibbs entropy, a thermodynamical concept, to the quantum setting. Schumacher compression is often used in more complex protocols as a preliminary preprocessing step.

Other information theoretic quantities, such as the conditional von Neumann entropy $S(A|B)_\psi$ \cite{12} and the conditional mutual information $I(A : B|R)_\psi$ \cite{13}, were only more recently given meaning \cite{4, 6, 13}. If we consider an i.i.d. source $S$ emitting an unknown sequence of states $|\psi_1^{AB}\rangle|\psi_2^{AB}\rangle\ldots|\psi_n^{AB}\rangle$, distributed to two spatially separated parties $A$ (Alice) and $B$ (Bob), an interpretation of $S(A|B)_\psi$ can be obtained \cite{4, 13} as the optimal rate at which pure entanglement needs to be consumed in order to transfer the entire sequence to Bob’s location. Whenever $S(A|B)_\psi$ is negative, it is understood that entanglement is gained instead of consumed and that the transfer can be accomplished using only local operations and classical communication (LOCC). The first protocol \cite{4, 13} for achieving this task, also known as state merging, was based on a random measurement strategy, a popular approach when designing quantum communication protocols. Examples of other tasks that can be achieved using this approach are distributed

\footnote{Electronic address: ndutil@cs.mcgill.ca}

\footnote{Electronic address: patrick@cs.mcgill.ca}
compression \cite{14} and assisted distillation \cite{15}. In assisted distillation, \( m \) helpers \( C_1, C_2, \ldots, C_m \) and two recipients \( A \) and \( B \) share a multipartite pure state \( \psi_{C_1C_2\ldots C_mB} \), and the objective is to extract an optimal amount of pure entanglement between \( A \) and \( B \) by using LOCC operations and classical information broadcasted by the \( m \) helpers \( C_1, C_2, \ldots, C_m \).

Both distributed compression and assisted distillation involve multiple parties (i.e. more than two) sharing a multipartite state \( \psi \). In the case of distributed compression, we can use the state merging primitive to perform compression at an optimal rate by transferring each sender’s share one at a time \cite{4}. This strategy will work for any rate which is a corner point of the boundary of the rate region associated with distributed compression. To achieve compression at rates which are not corner points, however, we need to use a time-sharing strategy as the decoding operation performed by the receiver can only recover the shares one at a time. One contribution of this paper is to present a protocol for the more general task of multiparty state merging which will eliminate the need for time-sharing. That is, we consider \( m \) senders \( C_1, C_2, \ldots, C_m \) and a decoder \( B \) sharing a multipartite mixed state \( \psi_{C_1C_2\ldots C_mB} \), potentially with additional entanglement in the form of EPR pairs (ebits) distributed between the decoder and each of the \( m \) senders. Given many copies of the input state \( \psi_{C_1C_2\ldots C_mB} \), the task is to transfer the shares \( C_1, C_2, \ldots, C_m \) to the receiver \( B \) with high fidelity using only LOCC operations.

If only a single copy of the state \( \psi_{C_1C_2\ldots C_mB} \) is available to the parties, we can use our protocol to achieve merging within an error tolerance \( \epsilon \) if we distribute enough initial entanglement between each of the senders and the receiver. In this regime, a more naive strategy consisting of repeatedly applying a one-shot state merging protocol \cite{16} on one sender at a time will generally require more initial entanglement to perform the state transfer than does our protocol. In addition, this strategy only yields a handful of achievable combinations of entanglement costs and does not permit interpolating between them. A full characterization of the entanglement cost in the one-shot regime when \( m = 1 \) was performed by \cite{14} using smooth min- and max-entropies. By applying the random measurement strategy of \cite{4} and by using the min- and max-entropy formalism of \cite{17}, we generalize some of the results of \cite{16} to the multipartite case \( m \geq 2 \). This work complements other recent attempts to study quantum information theory in the one-shot setting \cite{18–22}.

To perform assisted distillation in the context of multiple parties, we introduce a second decoder, whom we label \( A \) (Alice), and consider a variation of multiparty merging. Given a partition of the helpers \( C_1, C_2, \ldots, C_m \) into a set \( \mathcal{T} \) and its complement \( \overline{\mathcal{T}} := \{C_1C_2\ldots C_m\} \backslash \mathcal{T} \), we want to transfer the shares \( \mathcal{T} \) and \( \overline{\mathcal{T}} \) to the locations of the decoders \( A \) and \( B \) respectively. We call this task a split-transfer of the state \( \psi_{C_1C_2\ldots C_mB} \). A protocol for performing a split-transfer can be obtained by using the random measurement strategy on \( C_1, C_2, \ldots, C_m \), followed by appropriate decodings \( U_A \) and \( V_B \) by the decoders \( A \) and \( B \). The optimal achievable rate for assisted distillation was found in \cite{4} by using a recursive argument. By using a split-transfer protocol, we give a simpler demonstration which does not rely on a recursive argument.

**Structure of the paper:** In Section II, we introduce the definition for multiparty merging of a state \( \psi_{C_1C_2\ldots C_mB} \), and review the known results for the i.i.d setting. In Section III, we formulate a condition that a set of instruments performed by the senders \( C_1, C_2, \ldots, C_m \) must satisfy in order to accomplish merging within a fixed error tolerance. In Section IV, we consider random measurements performed by the senders \( C_1, C_2, \ldots, C_m \) and prove an upper bound to the quantum error when a single copy of the input state is available. We analyze the asymptotic setting in Section V, recovering the main theorem of Section II without the need for time-sharing. In Section VI, we reformulate the bounds obtained in Section III in terms of min-entropies and give necessary and sufficient conditions for merging in the one-shot regime. Section VII is devoted to analyzing the rates achievable for variants of the embezzling states, comparing our protocol to a
strategy of merging the shares one at a time. We introduce a split-transfer of the state \( \psi^{C_1C_2...C_mB} \) in Section VII and show the existence of a protocol for performing this task. We use this protocol to recover the optimal distillation rate for the problem of assisted distillation. Appendices, containing relevant folklore material, appear at the end.

**Notation:** In this paper, we restrict our attention to finite dimensional Hilbert spaces. Quantum systems under consideration will be denoted \( A, B, \ldots \), and are freely associated with their Hilbert spaces, whose (finite) dimensions are denoted \( d_A, d_B, \ldots \). If \( A \) and \( B \) are two Hilbert spaces, we write \( AB \equiv A \otimes B \) for their tensor product and write \( A^n \) for the tensor product \( \bigotimes_{i=1}^n A \). If we have \( m \) Hilbert spaces \( A_1, A_2, \ldots, A_m \), we write \( A_M \) for the tensor product \( \bigotimes_{i=1}^m A_i \). An ancilla augmenting the system \( A_i \) is denoted as \( A_i^0 \). We write \( A_M^0 \) for the tensor product \( \bigotimes_{i=1}^m A_i^0 \) of \( m \) ancillas \( A_1^0, A_2^0, \ldots, A_m^0 \). The maximally entangled state \( \frac{1}{\sqrt{K_1}} \sum_{k=1}^{K_1} |k\rangle \langle k| \) of Schmidt rank \( K_i \) is denoted as \( |\Phi_i\rangle \). We write \( \Phi \) for the density operator \( |\psi_{k1}\rangle \langle \psi_{k1}| \otimes |\psi_{k2}\rangle \langle \psi_{k2}| \otimes \ldots \otimes |\psi_{km}\rangle \langle \psi_{km}| \). The space of linear operators acting on the Hilbert space \( A \) is denoted by \( \mathcal{L}(A) \). The identity operator acting on \( A \) is denoted by \( I_A \). The symbol \( id_A \) denotes the identity map acting on \( \mathcal{L}(A) \). Unless otherwise stated, a “state” can be either pure or mixed. The symbol for such a state (such as \( \psi \) and \( \rho \)) also denotes its density operator. The density operator \( |\psi\rangle \langle \psi| \) of a pure state will frequently be written as \( \psi \). We denote by \( \tau_{A1} \) the maximally mixed state of dimension \( d_A \). We write \( S(A) \psi = -\text{Tr}(\rho_A \log \rho_A) \) to denote the von Neumann entropy of a density matrix \( \psi_A \) for the system \( A \). The function \( F(\rho, \sigma) := \text{Tr} \sqrt{\sqrt{\rho} / 2 \sigma \rho / 2} \) is the fidelity [23] between the two states \( \rho \) and \( \sigma \). The trace norm of an operator, \( ||X||_1 \) is defined to be \( \text{Tr} |X| = \text{Tr} \sqrt{X^†X} \). We will use the terms “receiver” and “decoder” interchangeably throughout the following sections.

**II. DEFINITIONS AND MAIN RESULT**

For a bipartite state \( \rho^{AB} \), the operation known as quantum state merging can be viewed in two different ways. The original formulation of the problem [4] was in terms of a statistical source emitting (unknown) states \( |\psi_{AB}\rangle, |\psi_{AB}'\rangle, \ldots \), with average density operator \( \rho^{AB} \) assumed to be known by the parties. The objective was then to transfer the entire sequence to the location of the decoder (Bob) using as little quantum communication as possible. An equivalent view of the problem is to consider a purification \( \psi^{ABR} \) of the density matrix \( \rho^{AB} \) and regard the process of merging as that of transferring all the correlations between Alice’s share and the purification system \( R \) to the location of the decoder \( B \). This means decoupling Alice’s system from the reference \( R \) while leaving the state \( \psi^R \) intact (up to some arbitrarily small perturbation) in the process. The receiver will hold a purification \( \phi^{BR} \) of the system \( R \), and since all purifications are equivalent up to an isometry on the purification system, he can recover the original state \( \psi^{ABR} \) by applying an appropriate isometry to \( \phi^{BR} \). Additional entanglement between Alice and Bob might be distilled in the process.

To analyze the multipartite scenario, where \( m \) senders and a decoder/receiver share a state \( \psi^{C_1C_2...C_mBR} \), with purifying system \( R \), we adopt the second view and look at the transformations that can be performed by the senders on their shares to allow the receiver to recover the purified state \( \psi^{C_1C_2...C_mBR} \) with high fidelity. The resources at their disposal will be pure entanglement, in the form of maximally entangled states shared between each of the senders \( C_1, C_2, \ldots, C_m \) and the receiver Bob, and noiseless classical channels, which will be used to transmit measurement outcomes to the receiver. Any transformation applied by the senders will need to decouple the reference \( R \) from the senders’ shares \( C_1, C_2, \ldots, C_m \) and leave the reference unchanged. Otherwise, the receiver might hold a purification that cannot be taken, by means of an isometry, to the original state. If each of the senders \( C_1, C_2, \ldots, C_m \) perform an incomplete measurement, de-
The receiver also has ancillas in mind, we can give a definition of a multiparty state merging for a state with subsystem entangled states \( \Phi \). Let \( \Lambda_m : C_M C_M^0 \otimes B_B^0 B_M^0 \rightarrow C_M^1 B_B^1 B_M^1 \) be an LOCC quantum operation performed by the senders \( C_1, C_2, \ldots, C_m \) and the decoder \( B \). Initially, each sender \( C_i \) is given an ancilla \( C_i^0 \). The receiver also has ancillas \( B_M^0 := B_1^0 B_2^0 \ldots B_m^0 \) with \( d_{B_i} = d_{C_i} \), and \( B_M^1 := B_1^1 B_2^1 \ldots B_m^1 \) with \( d_{B_i} = d_{C_i} \). Before the map \( \Lambda_m \) is applied, the systems \( C_M^0 \) and \( B_M^0 \) will hold maximally entangled states \( \Phi K_1 \otimes \Phi K_2 \otimes \ldots \otimes \Phi K_m \), where \( \Phi K_i \) has Schmidt rank \( K_i = d_{C_i} \) and is shared between the sender \( C_i \) and the receiver \( B \). After the map \( \Lambda_m \) is applied, the senders share a subsystem \( C_M^1 := C_1^1 C_2^1 \ldots C_m^1 \) of \( C_M \), and the receiver holds three systems: \( B_1, B_M^1 \) and \( B_M^1 \), with \( B_M \) being an ancillary system of dimension of the same size as the system \( C_M \).

This operation will implement merging, as illustrated in Figure 1, if the output state of the map \( \text{id}_R \otimes \Lambda_m \) is approximately a tensor product of the initial state \( \psi_{C_1 C_2 \ldots C_m B_B} \) and maximally entangled states \( \Phi L := \Phi L_1 \otimes \Phi L_2 \otimes \ldots \otimes \Phi L_m \) shared between the senders and the decoder. Each \( \Phi L_i \) is a maximally entangled state of Schmidt rank \( L_i \) on the tensor space \( C_i^1 B_i^1 \). More formally, we have

**Definition 1 (m-Party State Merging)** Let \( \Lambda_m \) be defined as in the previous paragraphs. We say that \( \Lambda_m \) is an \( m \)-party state merging protocol for the state \( \psi_{C_1 C_2 \ldots C_m B_B} \) with error \( \epsilon \) and entanglement cost
\[ \overline{E} := (\log K_1 - \log L_1, \log K_2 - \log L_2, \ldots, \log K_m - \log L_m) \]

\[ \left\| (\text{id}_R \otimes \Lambda^m_{\psi})(\psi^{C_1C_2\ldots C_mBR} \otimes \Phi^K) - \psi^{B_MBR} \otimes \Phi^L \right\|_1 \leq \epsilon, \]

where the state \( \psi^{B_MBR} \) corresponds to the initial state \( \psi^{C_1C_2\ldots C_mBR} \) with the system \( B_M \) substituted for \( C_M \). If we are given \( n \) copies of the same state, \( \psi = (\sigma)^{\otimes n} \), the entanglement rate \( \overline{R}(\sigma) \) is defined as \( \frac{1}{n} \overline{E}(\psi) \).

Before stating the main theorem, we need to define what it means for a rate-tuple \( \overline{R} \) to be achievable for multiparty merging using LOCC operations.

**Definition 2 (The Rate Region)** We say that the rate-tuple \( \overline{R} := (R_1, R_2, \ldots, R_m) \) is achievable for multiparty merging of the state \( \psi^{C_1C_2\ldots C_mB} \) if, for all \( \epsilon > 0 \), we can find an \( N(\epsilon) \) such that for every \( n \geq N(\epsilon) \) there exists an \( m \)-party state merging protocol \( \Lambda^m_{\psi} \) acting on \( \psi^{\otimes n} \otimes \Phi^K \), with error \( \epsilon \) and entanglement rate \( \overline{R}_n := \frac{1}{n}(\log K_1 - \log L_1, \log K_2 - \log L_2, \ldots, \log K_m - \log L_m) \) approaching \( \overline{R} \). We call the closure of the set of achievable rate-tuples the rate region.

Suppose \( m \) systems \( C_1, C_2, \ldots, C_m \) in the state \( \psi^{C_1C_2\ldots C_mR} \), where \( R \) is a purifying system, are distributed to \( m \) senders, spatially separated from each other. To recover the purified state \( \psi^{C_1C_2\ldots C_mR} \) at the receiver’s end, a task called distributed compression, the senders need an enough supply of initial entanglement. It was shown in [4] that the rate region associated with distributed compression is characterized by the inequalities

\[ \sum_{i \in \mathcal{T}} R_i \geq S(\mathcal{T} | \mathcal{T})_\psi \quad \text{for all nonempty sets} \ \mathcal{T} \subseteq \{1, 2, \ldots, m\}. \]

Here, the symbol \( \mathcal{T} \) also denotes the tensor product space \( \mathcal{T} := \bigotimes_{i \in \mathcal{T}} C_i \) associated with the set \( \mathcal{T} \). The set \( \mathcal{T} \) is defined as \( \{1, 2, \ldots, m\} \setminus \mathcal{T} \) and the tensor product space \( \mathcal{T} \) as \( \bigotimes_{i \in \mathcal{T}^c} C_i \). If a rate-tuple \( (R_1, R_2, \ldots, R_m) \) is achievable for distributed compression and some of the rates \( R_i \) are negative, then the senders \( C_i \) will be able to transfer their shares to the receiver using only LOCC operations, and furthermore, they will gain a potential for future communication in the form of maximally entangled states. Allowing the receiver to have side information \( B \) as well, leads to a similar set of equations describing the rate region associated with the task of multiparty state merging.

**Theorem 3 (\( m \)-Party Quantum State Merging [4])** Let \( \psi^{C_1C_2\ldots C_mBR} \) be a pure state shared between \( m \) senders \( C_1, C_2, \ldots, C_m \) and a receiver Bob, with purifying system \( R \). Then, the rate \( \overline{R} := (R_1, R_2, \ldots, R_m) \) is achievable for multiparty merging iff the inequality

\[ \sum_{i \in \mathcal{T}} R_i \geq S(\mathcal{T} | \mathcal{T} B)_\psi \]

holds for all nonempty subsets \( \mathcal{T} \subseteq \{1, 2, \ldots, m\} \).

The theorem was proved in [4] by showing that the corner points of the region are achievable and then using time-sharing to interpolate between them. In addition to recovering the result without time-sharing, we will extend it to the one-shot setting. Time-sharing, which consists of partitioning a large supply of states and applying different protocols to each subset, is impossible if only a single copy of a state is available. Instead, we will construct a single protocol which merges all shares at once.
III. CONDITIONS FOR MERGING MANY PARTIES

Let’s try to construct an LOCC operation \( \Lambda^m : C_M C_M^0 \otimes B B_M^0 \rightarrow C_M^1 \otimes B_M^1 B_M \) that when applied to the state \( \psi^{C_1 C_2 \ldots C_m B} \otimes \Phi^K \) will achieve merging, and destroy all existing correlations between the senders’ shares \( C_1, C_2, \ldots, C_m \) at the same time. It can be seen as a three step process:

- First, each sender \( C_i \) applies a quantum instrument \( I_i := \{ \mathcal{E}_j \}_{j=1}^X \) to his share of the state \( \psi^{C_1 C_2 \ldots C_m B} \otimes \Phi^K \). This will yield both quantum and classical outputs. Each operator \( \mathcal{E}_j \) for the instrument \( I_i \) is completely positive, and maps the space \( C_i C_i^0 \) to the subspace \( C_i^1 \).
- Secondly, the senders \( C_1, C_2, \ldots, C_m \) send their classical outputs \( J := (j_1, j_2, \ldots, j_m) \) to the decoder \( B \).
- Finally, the decoder will use his side information \( \psi^B \), his share of the maximally entangled states \( \{ \Phi^K \} \), and the classical information \( J \) to perform a decoding operation \( D_J : B B_M^0 \rightarrow B_M^1 B_M \) (i.e. a trace-preserving completely positive map (TP-CPM)) and recover the state \( \psi^{C_1 C_2 \ldots C_m B} \otimes \Phi^L \).

The state of the systems \( C_M^1 B_M^0 B X \) after steps 1 and 2 are performed can be written as:

\[
\psi^{C_M^1 B_M^0 B X} := \sum_{J := (j_1, j_2, \ldots, j_m)} [ (\text{id} B_M^0 B \otimes \mathcal{E}_J) (\psi^{C_1 C_2 \ldots C_m B} \otimes \Phi^K) ] C_M^1 B_M^0 B \otimes |J\rangle \langle J|^X,
\]

where \( \mathcal{E}_J := \mathcal{E}_1 \otimes \mathcal{E}_2 \otimes \ldots \otimes \mathcal{E}_m \) and \( \psi^{C_M^1 B_M^0 B X} \) is the normalized state given by \( (\text{id} B_M^0 B \otimes \mathcal{E}_J) (\psi^{C_1 C_2 \ldots C_m B} \otimes \Phi^K) \). The system \( X \) is an ancillary system held by the receiver which contains the classical outputs of the instruments \( I_1, I_2, \ldots, I_m \). If we restrict the operators \( \mathcal{E}_j \) to consist of only one Kraus operator (i.e \( \mathcal{E}_j (\rho) = A_j \rho A_j^\dagger \) for all \( i, j \)) and to satisfy \( \sum_j (A_j^\dagger A_j) = I \), the outcome states \( \{ \psi^{C_M^1 B_M^0 B X}_J \} \) are pure and are the result of performing \( m \) incomplete measurements, one for each sender \( C_i \).

After the senders have finished performing their instruments, we would ideally like for the state \( \psi^{C_M^1 B_M^0 B X}_j \) to be in the product form

\[
\psi^{C_M^1 B_M^0 B X}_j = \tau^{C_1} \otimes \tau^{C_2} \otimes \ldots \otimes \tau^{C_m} \otimes \psi^R,
\]

where \( \tau^{C_i} \) is the maximally mixed state of dimension \( L_i = d_{C_i} \) on the space \( C_i^1 \). Suppose, for the moment, that this property is satisfied for all \( \{ \psi^{C_M^1 B_M^0 B X}_j \} \). Then, the state \( \psi^{C_M^1 B_M^0 B X}_j \) purifies \( \tau^{C_M} \otimes \psi^R \), with purification systems \( B_M^0 B \).

Another purification of \( \tau^{C_M} \otimes \psi^R \) is also given by \( \Phi^L \otimes \psi^{B_M B X} \), where the state \( \psi^{B_M B X} \) corresponds to the original state \( \psi^{C_1 C_2 \ldots C_m B} \) with the system \( B_M \) substituted for \( C_M \). It follows from the Schmidt decomposition that these two purifications are related by an isometry \( U_J : B_M^0 B \rightarrow B_M^1 B_M \) on Bob’s side such that

\[
(I^{C_M^1 B_M^0 B X})^\dagger U_J \psi^{C_M^1 B_M^0 B X} (I^{C_M^1 B_M^0 B X \otimes U_J})^\dagger = \Phi^L \otimes \Phi^L \otimes \ldots \otimes \Phi^L \otimes \psi^{B_M B X},
\]

where \( \Phi^L \) is the maximally mixed state of dimension \( L \).
Hence, if the senders can perfectly decouple their systems $C_M C_M^0$ from the reference, their "R-entanglement" will be transferred to Bob’s location. Furthermore, by applying $U_J$, the receiver will recover the original state and distill some pure bipartite entanglement.

The previous scenario was ideal, and in general, will not be feasible for most states $\psi_{C_1 C_2 \ldots C_m BR}$. Hence, we relax our decoupling requirement and accept that the measurements performed by the senders will perturb the reference $\psi^R$ up to some tolerable amount, and that a small dose of correlations between the senders’ shares might still be present. In more formal terms, we have

**Proposition 4 (Compare to Proposition 4 of [4])** Let $\psi^{C_M^1 B_M^0 BR}_J$ be defined as in eq. (10), with reduced density matrix $\psi^{C_M^1 R}_J$. Define the following quantity:

$$Q_I(\psi^{C_1 C_2 \ldots C_m BR}_J \otimes \Phi^K) := \sum_{p_{J}} p_{J} \left\| \psi^{C_M^1 R}_J - \tau^{C_M^1} \otimes \psi^R \right\|_1,$$

(9)

where $p_{J}$ is the probability of obtaining the state $\psi^{C_M^1 B_M^0 BR}_J$ after all the senders have performed their instruments. If $Q_I(\psi^{C_1 C_2 \ldots C_m BR}_J \otimes \Phi^K) \leq \epsilon$, then there exists an LOCC operation $\Lambda^m$ which is an $m$-party state merging protocol for the state $\psi^{C_1 C_2 \ldots C_m BR}_J$ with error $2\sqrt{\epsilon}$ and entanglement cost $E = (\log K_1 - \log L_1, \log K_2 - \log L_2, \ldots, \log K_m - \log L_m)$.

**Proof** The proof of the above statement is very similar to the proof of Proposition 4 in [4]. We give the full proof here for completeness. Using Lemma 21 (see Appendix A), we have

$$\sum_{p_{J}} p_{J} F(\psi^{C_M^1 R}_J, \tau^{C_M^1} \otimes \psi^R) \geq 1 - \frac{\epsilon}{2}.$$

(10)

By Ulhmann’s theorem, we know there exist an isometry (i.e. a decoding) $U_J : B_M^0 B \rightarrow B_M^1 B M$ implementable by Bob such that

$$F(\psi^{C_M^1 R}_J, \tau^{C_M^1} \otimes \psi^R) = F\left((I^{C_M^1 R} \otimes U_J) \psi^{C_M^1 B_M^0 BR}_J (I^{C_M^1 R} \otimes U_J)^\dagger, \Phi^L \otimes \psi^{B_M BR}_J\right).$$

(11)

Thus, using the concavity of $F$ (see [24] for a proof) in its first argument, we have

$$F(\psi^{C_M^1 B_M^1 B_M^0 BR}_J, \Phi^L \otimes \psi^{B_M BR}_J) \geq \sum_{p_{J}} p_{J} F\left((I^{C_M^1 R} \otimes U_J) \psi^{C_M^1 B_M^0 BR}_J (I^{C_M^1 R} \otimes U_J)^\dagger, \Phi^L \otimes \psi^{B_M BR}_J\right) \geq 1 - \frac{\epsilon}{2},$$

(12)

where

$$\psi^{C_M^1 B_M^1 B_M^0 BR}_J = (\text{id}_R \otimes \Lambda^m_J) (\psi^{C_1 C_2 \ldots C_m BR}_J \otimes \Phi^K) := \sum_{p_{J}} p_{J} (I^{C_M^1 R} \otimes U_J) |\psi_J\rangle \langle \psi_J|^{C_M^1 B_M^0 BR}_J (I^{C_M^1 R} \otimes U_J)^\dagger$$

is the output state of the protocol. Using the relation between fidelity and trace distance once more, we arrive at

$$\left\| \psi^{C_M^1 B_M^1 B_M^0 BR}_J - \Phi^L \otimes \psi^{B_M BR}_J \right\|_1 \leq 2\sqrt{\epsilon - \epsilon^2/4} \leq 2\sqrt{\epsilon}.$$

(14)
IV. ONE-SHOT MERGING BY RANDOM MEASUREMENTS

One possible strategy for decoupling the system $C_i$ from the reference $R$ and the other systems $\{C_j : j \neq i\}$ is to perform a random von Neumann measurement on $C_i$ with $N_i = \lfloor \frac{d_{C_i}}{L_i} \rfloor$ projectors of rank $L_i$, and a little remainder, followed by a unitary $U_i$ mapping the outcome state to a subspace $C_i^1$ of $C_iC_i^0$. For such measurements, we can bound the quantum error $Q_T(\psi^{C_1C_2...C_mBR} \otimes \Phi^K)$ as follows:

Proposition 5 (One-Shot Multiparty Merging) Let $\psi^{C_1C_2...C_mBR}$ be a multipartite state, with local dimensions $d_B, d_R$ and $d_{C_i}, 1 \leq i \leq m$, and let $\Phi^K$ be some additional pure entanglement, as defined in the previous sections, shared between the receiver and the senders. For each sender $C_i$, there exists an instrument $I_i$ consisting of $N_i := \lfloor \frac{d_{C_i}K_i}{L_i} \rfloor$ CP maps

$$\mathcal{E}_j^i(\rho) := P_j^i \rho (P_j^i)^\dagger \quad 1 \leq j \leq N_i,$$

(15)

where $P_j^i : C_i \to C_i^1$ is a partial isometry of rank $L_i$ (i.e. $(P_j^i)^\dagger P_j^i$ is a projector onto an $L_i$ dimensional subspace of $C_i$), and one map $\mathcal{E}_j^i(\rho) := P_0^i \rho (P_0^i)^\dagger$, where $P_0^i$ is of rank $L_i' = d_{C_i}K_i - N_iL_i < L_i$, such that the overall quantum error $Q_T(\psi^{C_1C_2...C_mBR} \otimes \Phi^K)$ is bounded by

$$Q_T(\psi^{C_1C_2...C_mBR} \otimes \Phi^K) \leq 2 \sum_{\substack{T \subseteq \{1,2,...,m\} \atop T \neq \emptyset}} \prod_{i \in T} \frac{L_i}{d_{C_i}K_i} + 2 \left[ \frac{d_R}{L} \sum_{\substack{T \subseteq \{1,2,...,m\} \atop T \neq \emptyset}} \prod_{i \in T} \frac{L_i}{K_i} \text{Tr}[(\psi^{RT})^2] \right] := \Delta_T,$$

(16)

and there is a merging protocol with error at most $2\sqrt{\Delta_T}$. In fact, for each sender $C_i$, if we perform a random von Neumann measurement on $C_i$ followed by a unitary $U$ mapping the outcome state to a subspace $C_i^1$, the left hand side of eq. (16) is bounded from above on average by the right hand side.

To prove this proposition, we will need the following technical lemma, which generalizes Lemma 6 in [4] to the case of $m$ senders. The proof will follow a similar line of reasoning.

Lemma 6 (Compare to Lemma 6 in [4]) For each sender $C_i$, let $P_i : C_i \to C_i^1$ be a random partial isometry of rank $L_i$. One way to construct such an isometry is to fix some rank $L_i$-projector $Q_i$ onto a subspace $C_i^1$ of $C_i$ and precede it with a Haar distributed unitary $U_i$ on $C_i$ (i.e $P_i := Q_iU_i$). Define the subnormalized density matrix

$$\omega^{C_iR}_M(U) := (Q_1U_1 \otimes Q_2U_2 \otimes \ldots \otimes Q_mU_m \otimes I_R)\psi^{C_iR}_M(Q_1U_1 \otimes Q_2U_2 \otimes \ldots \otimes Q_mU_m \otimes I_R)^\dagger,$$

(17)

where $U := U_1 \otimes U_2 \otimes \ldots \otimes U_m$. Then, we have

$$\int_{U(C_1)} \int_{U(C_2)} \ldots \int_{U(C_m)} \left\| \omega^{C_iR}_M(U) - \frac{L}{d_{C_M}} \tau^{C_iM} \otimes \psi^R \right\|_1 dU \leq \frac{L}{d_{C_M}} \sqrt{d_R \sum_{\substack{T \subseteq \{1,2,...,m\} \atop T \neq \emptyset}} \prod_{i \in T} L_i \text{Tr}[(\psi^{RT})^2]},$$

(18)

where $dU := dU_1dU_2\ldots dU_m$, $\int dU_i = 1$ and $L := \prod_i L_i$. 
Proof For the remainder of this proof, we will write $\int f(U)dU$ as $\mathbb{E}[f(U)]$, indicating expectation, and abbreviate $\omega^{C^1_M R}(U)$ by $\omega^{C^1_M R}$. We have

$$
\mathbb{E} \left[ \left\| \omega^{C^1_M R} - \frac{L}{d_{C_M}} \sigma_{C^1_M} \otimes \psi^R \right\|_2^2 \right] = \mathbb{E} \left[ \left\| \omega^{C^1_M R} - \mathbb{E} [\omega^{C^1_M R}] \right\|_2^2 \right]
= \mathbb{E} \left[ \text{Tr}[(\omega^{C^1_M R})^2] - \text{Tr} \left[ \mathbb{E} [\omega^{C^1_M R}]^2 \right] \right].
$$  

(19)

To evaluate the average of $\text{Tr}[(\omega^{C^1_M R})^2]$, we use the following property:

$$
\text{Tr}[(\omega^{C^1_M R})^2] = \text{Tr} \left( (\omega^{C^1_M R} \otimes \omega^{C^1_M R'})(F^{C^1_M C^1_M} \otimes F^{RR'}) \right),
$$  

(20)

where $F^{C^1_M C^1_M} := \bigotimes_{i=1}^m F^{C^1_i C^1_i}$ is a tensor product of swap operators $F^{C^1_i C^1_i} := Q_i \otimes Q_i, F^{C^1_i C^1_i} Q_i \otimes Q_i$, exchanging the system $C^1_i$ and a copied version $C^1_i$. The expectation of $\text{Tr}[(\omega^{C^1_M R})^2]$ then becomes equal to

$$
\mathbb{E} \left[ \text{Tr}[(\omega^{C^1_M R})^2] \right]
= \mathbb{E} \left[ \text{Tr} \left( (\omega^{C^1_M R} \otimes \omega^{C^1_M R'})(F^{C^1_M C^1_M} \otimes F^{RR'}) \right) \right]
= \mathbb{E} \left[ \text{Tr} \left( (U_{CM} \otimes U_{CM'} \otimes I_{RR'}) (\psi^{CM R} \otimes \psi^{CM'R'}) (U_{CM} \otimes U_{CM'} \otimes I_{RR'}) \right) \right]
= \text{Tr} \left( (\psi^{CM R} \otimes \psi^{CM'R'}) \mathbb{E} \left[ (U_{CM} \otimes U_{CM'})(F^{C^1_M C^1_M} \otimes U_{CM'} \otimes U_{CM}) \right] \otimes F^{RR'} \right),
$$  

(21)

where we have used the shorthand $U_{CM} := U_1 \otimes U_2 \otimes \ldots U_m$, with $U_i$ being a Haar distributed unitary on $C_i$. The unitary $U_{CM'}$ is identical to $U_{CM}$ but acts on $C'_M$. Observe that the projections $\{Q_i\}$ from the state $\omega^{C^1_M R}$ are absorbed by the swap operators $\{F^{C^1_i C^1_i}\}$ (i.e. $F^{C^1_i C^1_i} = Q_i \otimes Q_i, F^{C^1_i C^1_i} Q_i \otimes Q_i$). The expectation $\mathbb{E} \left[ (U_{CM} \otimes U_{CM'})^\dagger F^{C^1_M C^1_M} (U_{CM} \otimes U_{CM'}) \right]$ can then be expanded as

$$
\mathbb{E} \left[ (U_{CM} \otimes U_{CM'})^\dagger (F^{C^1_M C^1_M}) (U_{CM} \otimes U_{CM'}) \right] = \bigotimes_{i=1}^m \mathbb{E} \left[ (U_i^{\otimes 2})^\dagger F^{C^1_i C^1_i} U_i^{\otimes 2} \right],
$$  

(22)

where we have used the shorthand $U_i^{\otimes 2} := U_i \otimes U_i$. Each of the expected values $\mathbb{E} \left[ (U_i^{\otimes 2})^\dagger F^{C^1_i C^1_i} U_i^{\otimes 2} \right]$ can be re-expressed, using an argumentation similar to the one found in Appendix B of [1], as

$$
\mathbb{E} \left[ (U_i^{\otimes 2})^\dagger F^{C^1_i C^1_i} U_i^{\otimes 2} \right] = r_i I_{C_i C'_i} + s_i F^{C^1_i C^1_i},
$$  

(23)

where the coefficients $r_i$ and $s_i$ are defined as

$$
r_i := \frac{L_i}{d_{C_i}} - \frac{L_i}{d_{C_i}^2} \leq \frac{L_i}{d_{C_i}^2},
$$  

$$
s_i := \frac{L_i^2}{d_{C_i}} - \frac{L_i}{d_{C_i}^2} \leq \frac{(L_i)^2}{d_{C_i}^2}.
$$  

(24)
Substituting eqs. (22), (23) and (24) into eq. (21), we get
\[
\mathbb{E} \left[ \text{Tr}(\psi^{C_M} R)^2 \right] = \text{Tr} \left[ (\psi^{CM} R \otimes \psi^{CM} R') \bigotimes_{i=1}^{m} \left( r_i I_{C_i C_i'} + s_i E_{C_i C_i'} \right) \otimes E_{RR'} \right]
\]
\[
= \sum_{\mathcal{T} \subseteq \{1, 2, \ldots, m\}} \prod_{i \notin \mathcal{T}} r_i \prod_{i \in \mathcal{T}} s_i \text{Tr} \left[ (\psi^{RT})^2 \right],
\]
(25)
where \( \mathcal{T} \) appearing in \( \psi^{RT} \) denotes the system \( \otimes_{i \in \mathcal{T}} C_i \). When \( \mathcal{T} \) is the empty set, the last expression in eq. (25) reduces to \( \prod_{i=1}^{m} r_i \text{Tr}[(\psi^R)^2] \). From eq. (24), we can bound the quantity \( \prod_{i=1}^{m} r_i \text{Tr}[(\psi^R)^2] \) from above by:
\[
\prod_{i=1}^{m} r_i \text{Tr}[(\psi^R)^2] \leq \frac{L}{d^{2}_{CM}} \text{Tr}[(\psi^R)^2]
\]
\[
= \text{Tr} \left[ \frac{L^2}{d^{2}_{CM}} (T^{C_M})^2 \otimes (\psi^R)^2 \right]
\]
\[
= \text{Tr} \left[ \left( \frac{L}{d_{CM}} T^{C_M} \otimes \psi^R \right)^2 \right]
\]
\[
= \text{Tr} \left[ \mathbb{E}[(\psi^{C_M} R)^2] \right].
\]
(26)
Hence, using eqs. (16), (20), (21) and the previous bound, we have
\[
\mathbb{E} \left[ \left\| \frac{L}{d_{CM}} T^{C_M} \otimes \psi^R \right\|_2^2 \right] \leq \sum_{\mathcal{T} \subseteq \{1, 2, \ldots, m\}} \prod_{i \notin \mathcal{T}} L_i \prod_{i \in \mathcal{T}} \left( \frac{L_i}{d_{C_i}} \right)^2 \text{Tr} \left[ (\psi^{RT})^2 \right].
\]
(27)
To obtain a bound on \( \mathbb{E} \left[ \left\| \frac{L}{d_{CM}} T^{C_M} \otimes \psi^R \right\|_2 \right] \), we use the Cauchy-Schwarz inequality:
\[
\mathbb{E} \left[ \left\| \frac{L}{d_{CM}} T^{C_M} \otimes \psi^R \right\|_2 \right] \leq L d_R \mathbb{E} \left[ \left\| \frac{L}{d_{CM}} T^{C_M} \otimes \psi^R \right\|_2 \right]
\]
\[
\leq L d_R \sum_{\mathcal{T} \subseteq \{1, 2, \ldots, m\}} \prod_{i \notin \mathcal{T}} \frac{L_i}{d_{C_i}} \prod_{i \in \mathcal{T}} \left( \frac{L_i}{d_{C_i}} \right)^2 \text{Tr} \left[ (\psi^{RT})^2 \right]
\]
\[
\leq L^2 \frac{d_R}{d_{CM}} \sum_{\mathcal{T} \subseteq \{1, 2, \ldots, m\}} \prod_{\mathcal{T} \neq \emptyset} L_i \text{Tr} \left[ (\psi^{RT})^2 \right].
\]
(28)
And thus,
\[
\mathbb{E} \left[ \left\| \frac{L}{d_{CM}} T^{C_M} \otimes \psi^R \right\|_2 \right] \leq \frac{L}{d_{CM}} \left( d_R \sum_{\mathcal{T} \subseteq \{1, 2, \ldots, m\}} \prod_{\mathcal{T} \neq \emptyset} L_i \text{Tr} \left[ (\psi^{RT})^2 \right] \right).
\]
(29)

Proof of Proposition

Fix a random measurement by choosing, for each sender \( C_i, N_i := \left\lfloor \frac{d_{C_i} K_i}{L_i} \right\rfloor \) fixed orthogonal subspaces of \( C_i C_i^0 \) of dimension \( L_i \) and one of dimension \( L_i' = d_{C_i} K_i - N_i L_i < L_i \).
The projectors onto these subspaces followed by a fixed unitary mapping it to $C_i$, we denote by $Q_i$, $j = 0,...,N_i$. Note that $Q_i^0$ projects onto a subspace of dimension $L_i < L_i$. Set $P_i := Q_i^0 U_i$ with a Haar distributed random unitary $U_i$ on $C_i C_i^0$. Applying Lemma 6 for a measurement outcome $J = (j_1,j_2,...,j_m)$, with $C_M^R = (Q_i^1 U_1 \otimes Q_i^2 U_2 \otimes \ldots \otimes Q_i^m U_m \otimes I_R) \psi^C_M R \otimes \tau^{C_M} (Q_i^1 U_1 \otimes Q_i^2 U_2 \otimes \ldots \otimes Q_i^m U_m \otimes I_R)^\dagger$, we have

$$
\mathbb{E} \left[ \sum_{j_1=1}^{N_1} \sum_{j_2=1}^{N_2} \ldots \sum_{j_m=1}^{N_m} \left| \left| \omega_{j}^{C_M^R} - \frac{L}{d_{C_M} K} \tau^{C_M^R} \otimes \psi^R \right| \right|_1 \right] \\
\leq \left( \prod_{i=1}^{m} N_i \right) \frac{L}{d_{C_M} K} \sqrt{d_R \sum_{T \subseteq \{1,2,...,m\}} \prod_{i \in T} \frac{L_i}{K_i} \text{Tr} \left[ (\psi^{RT})^2 \right]} \leq \frac{d_R}{d_{C_M}} \sum_{T \subseteq \{1,2,...,m\}} \prod_{i \in T} \frac{L_i}{K_i} \text{Tr} \left[ (\psi^{RT})^2 \right].
$$

Taking the normalisation into account, with $p_j = \text{Tr} (\omega_{j}^{C_M^R})$ and $\psi_{j}^{C_M^R} = \frac{1}{p_j} \omega_{j}^{C_M^R}$, we need to show that on average, the $p_j$ are close to $\frac{1}{d_{C_M} K}$. Looking at eq. (30) and tracing out, we get

$$
\mathbb{E} \left[ \sum_{j_1=1}^{N_1} \sum_{j_2=1}^{N_2} \ldots \sum_{j_m=1}^{N_m} p_j \left| \left| \psi_{j}^{C_M^R} - \tau^{C_M^R} \otimes \psi^R \right| \right|_1 \right] \leq \frac{d_R}{d_{C_M}} \sum_{T \subseteq \{1,2,...,m\}} \prod_{i \in T} \frac{L_i}{K_i} \text{Tr} \left[ (\psi^{RT})^2 \right].
$$

Hence we obtain, using the triangle inequality,

$$
\mathbb{E} \left[ \sum_{j_1=1}^{N_1} \sum_{j_2=1}^{N_2} \ldots \sum_{j_m=1}^{N_m} p_j \left| \left| \psi_{j}^{C_M^R} - \tau^{C_M^R} \otimes \psi^R \right| \right|_1 \right] \leq 2 \frac{d_R}{d_{C_M}} \sum_{T \subseteq \{1,2,...,m\}} \prod_{i \in T} \frac{L_i}{K_i} \text{Tr} \left[ (\psi^{RT})^2 \right] =: \Gamma_{\psi^{R}} \Phi_{\tau^{R}}.
$$

Lastly, we need to consider what happens when at least one sender $i$ obtains a measurement outcome $j_i$ equal to 0. For an outcome $J = (j_1,j_2,...,j_m)$, define the subset $T(J) \subseteq \{1,2,...,m\}$ such that $i \in T(J)$ iff $j_i = 0$. Also, define the set $Z = \{J : |T(J)| > 0\}$. Then, it is easy to show that the cardinality of the set $Z$ is

$$
|Z| = \sum_{T \subseteq \{1,2,...,m\}} \prod_{i \notin T} N_i.
$$

For an outcome $J \in Z$, the expected probability of the state $\omega_{j}^{C_M^R}$ is given by

$$
\mathbb{E}_{U_1 U_2...U_m} \left[ \text{Tr} (\omega_{j}^{C_M^R}) \right] = \text{Tr} \left[ \mathbb{E}_{U_1 U_2...U_m} (\omega_{j}^{C_M^R}) \right] = \text{Tr} \left[ \bigotimes_{i \in T(J)} Q_i^0 \tau^{C_M^0} (Q_i^0)^\dagger \bigotimes_{i \notin T(J)} Q_i^j \tau^{C_M^0} (Q_i^j)^\dagger \right] = \prod_{i \in T(J)} L_i \prod_{i \notin T(J)} L_i \frac{L_i}{d_{C_M} K}.
$$
With this formula in hand and the fact that the trace norm between two states is at most 2, we can bound the expected value of the quantum error $Q_I(ψ^C_1ψ_{c_2...c_mBR} ⊗ Φ^K)$ as follows:

\[
\mathbb{E}
\left[
\sum_{j_1=0}^{N_1} \sum_{j_2=0}^{N_2} \ldots \sum_{j_m=0}^{N_m} p_j \left\| ψ^C_{j_1} ⊗ ψ_{c_2...c_m} - τ^C_{j_1} ⊗ ψ_R \right\|_1
\right]
\leq 2 \sum_{\mathcal{T} \subseteq \{1,2,...,m\}} \prod_{\mathcal{T} \neq \emptyset} \frac{\prod_{i \in \mathcal{T}} L_i}{d_{C_i} K_i} + \Gamma_{ψ ⊗ Φ^K}
\leq 2 \sum_{\mathcal{T} \subseteq \{1,2,...,m\}} \prod_{\mathcal{T} \neq \emptyset} \frac{L_i}{d_{C_i} K_i} + \Gamma_{ψ ⊗ Φ^K}.
\] (35)

V. MULTIPARTY STATE MERGING: I.I.D VERSION

In this section, we analyze the case where the parties have at their disposal arbitrarily many copies of the state $ψ^{C_1C_2...C_mB,R}$. We give a proof of Theorem 3 and then look at the case of distributed compression as an application. As mentioned earlier, the rates characterized by eq. (5) will be achievable without the need for a time-sharing strategy. Indeed, we will show the existence of multiparty merging protocols where each sender performs a single measurement on his share of the input state and communicates the outcome to the receiver. If the parties were to employ a time-sharing strategy, on the other hand, the many initial copies of the input state $ψ^{C_1C_2...C_mB,R}$ would need to be divided into blocks, and for each of these blocks, the senders would have to perform a different measurement.

A. Proof of Theorem 3

Proof To prove the direct statement of the theorem, we use Proposition 5 in combination with Shumacher compression [11]. For $n$ copies of the state $ψ^{C_1C_2...C_mB,R}$, consider the Schumacher compressed state

\[
|Ω⟩ := (Π_B ⊗ Π_{C_1} ⊗ Π_{C_2} ⊗ \ldots \otimes Π_{c_m} ⊗ Π_R)|ψ⟩^n,
\] (36)

and its normalized version $|Ψ⟩ := \frac{1}{\sqrt{|Ω|}} |Ω⟩$. Here, the systems $B, C_1, ..., C_m, R$ are the typical subspaces (see [5] for detailed definitions) of $B^n, C_1^n, ..., C_m^n, R^n$ and $Π_B, Π_{C_1}, ..., Π_{C_m}, Π_R$ are the projection operators onto these typical subspaces. In particular, we have that

\[
⟨Ω|Ω⟩ = ⟨ψ⟩^n Π_B ⊗ Π_{C_1} ⊗ \ldots \otimes Π_{c_m} ⊗ Π_R|ψ⟩^n \geq 1 - \epsilon
\] (37)

for any $ε > 0$ and large enough $n$. Furthermore, we can set $ε$ to be equal to $(m + 2)exp(-cδ^2n)$ for some constant $c$, where $δ > 0$ is a typicality parameter. This follows from the fact (see Appendix B) that

\[
Π_B ⊗ Π_{C_1} ⊗ \ldots ⊗ Π_{c_m} ⊗ Π_R \geq Π_B + Π_{C_1} + \ldots + Π_{c_m} + Π_R - (m + 1)I_{B\hat{C}_1...\hat{C}_m\hat{R}}
\] (38)
and, that by typicality, we have \( \text{Tr}(\psi^{\otimes n} \Pi_B) \geq 1 - \exp(-c_0^2n) \) for all \( 1 \leq i \leq m \) (see [25] for the exponential bounds). Note that we have omitted some identity operator factors on the right hand side for the sake of clarity. The operator \( \Pi_B \) on the right hand side of eq. (38) is in fact \( (I^m \otimes \Pi_B) \), and the same applies for all the other projectors on that side of the inequality.

The properties for the typical projectors \( \Pi_B, \Pi_{\mathcal{C}_1}, \ldots, \Pi_{\mathcal{C}_m} \) allow us to tightly bound the various dimensions and purities appearing in Proposition 5 by appropriate “entropic” formulas. In particular, we have [5] for any system \( F = C_i, B, R \):

\[
(1 - \epsilon)2^{n(S(F)_\Psi - \delta)} \leq \text{Tr}[\Pi_F] \leq 2^{n(S(F)_\Psi + \delta)} \leq 2^{-n(S(F)_\Psi - \delta)}.
\]

(39)

Hence, all parties follow a merging protocol as if they had \( \Psi \), with additional entanglement \( \Phi^K := \Phi^{K_1} \otimes \Phi^{K_2} \otimes \ldots \otimes \Phi^{K_m} \). If each sender \( C_i \) performs a random measurement on his system, as in Proposition 5, with projectors of rank \( L_i \) (and one of rank \( L'_i \leq L_i \)) such that

\[
\prod_{i \in \mathcal{T}} \frac{L_i}{K_i} \leq 2^{n(S(RT)_\Psi - S(R)_\Psi - 3\delta |\mathcal{T}|)}
\]

(40)

holds for all nonempty subsets \( \mathcal{T} \subseteq \{1, 2, \ldots, m\} \), then the expected value of the quantum error \( Q_{\mathcal{T}}(\Psi \otimes \Phi^K) \) is bounded from above by

\[
\mathbb{E}_{U_1U_2\ldots U_m}[Q_{\mathcal{T}}(\Psi \otimes \Phi^K)] \leq 2 \sum_{\mathcal{T} \subseteq \{1, 2, \ldots, m\}} \prod_{\mathcal{T} \neq \emptyset} \frac{L_i}{d_{\mathcal{C}_i}K_i} + 2 \sqrt{\sum_{\mathcal{T} \subseteq \{1, 2, \ldots, m\}} \prod_{\mathcal{T} \neq \emptyset} \frac{L_i}{K_i} \text{Tr}[(\Psi_{\mathcal{T}}^R)^2]}
\]

\[
\leq 2 \sum_{\mathcal{T} \subseteq \{1, 2, \ldots, m\}} \frac{2^{n(S(RT)_\Psi - S(R)_\Psi - \sum_{i \in \mathcal{T}} S(C_i)_\Psi - 2\delta |\mathcal{T}|)}}{(1 - \epsilon)^{|\mathcal{T}|}} + 2 \sqrt{\sum_{\mathcal{T} \subseteq \{1, 2, \ldots, m\}} \prod_{\mathcal{T} \neq \emptyset} \frac{2^{-n\delta}}{2^{-n\delta} - 2^{-n\delta}}} = O(2^{-n\delta/2}).
\]

(41)

To bound the first term on the right hand side of eq. (41), we have used subadditivity twice: \( S(RT)_\Psi \leq S(R)_\Psi + S(T)_\Psi \) and \( S(T)_\Psi \leq \sum_{i \in \mathcal{T}} S(C_i)_\Psi \). Hence, by Proposition 4, we can conclude that there exists a merging protocol with error \( O(2^{-n\delta/4}) \) and entanglement cost \( \frac{\overline{E}}{n} := (\log K_1 - \log L_1, \log K_2 - \log L_2, \ldots, \log K_m - \log L_m) \). From eq. (40), the entanglement rate \( \frac{1}{n} \overline{E} \) must satisfy

\[
\sum_{i \in \mathcal{T}} \frac{1}{n} (\log K_i - \log L_i) \geq (S(R)_\Psi - S(RT)_\Psi + 3\delta |\mathcal{T}|)
\]

\[
= S(T | \mathcal{T} \text{B})_\Psi + 3\delta |\mathcal{T}|
\]

(42)

for all non empty subsets \( \mathcal{T} \subseteq \{1, 2, \ldots, m\} \). Since we have a vanishing error for this protocol as \( n \) goes to infinity, all rate-tuples \( \overline{R} \) satisfying the preceding set of inequalities are achievable for merging for the state \( \Psi \). However, by the gentle measurement lemma and the triangle inequality,

\[
\left\| \left( \psi_{\mathcal{C}_1\mathcal{C}_2\ldots\mathcal{C}_mBR} \right)^{\otimes n} - \Psi \right\|_1 \leq 4 \sqrt{\epsilon},
\]

(43)
and so, if we apply the same merging protocol on the state \( (\psi^{C_1C_2...C_mBR}) \otimes \Phi^K \), we get an error of \( O(2^{-n\delta/4}) + O(2^{-cn\delta^2/2}) \). This error also vanishes as \( n \) goes to infinity, and since \( \delta \) was arbitrarily chosen, we can conclude that any rate-tuple \( \overrightarrow{R} = (R_1, R_2, \ldots, R_m) \) satisfying

\[
\sum_{i \in T} R_i \geq S(T|TB)_{\psi}
\]

for all non-empty \( T \subseteq \{1, 2, \ldots, m\} \) must be contained in the rate region. This proves the direct part of Theorem 3.

The converse was established in \cite{4} so we are done.

\[\Box\]

### B. Distributed compression for two senders

To illustrate some of the properties of our protocol, let’s consider the problem of distributed compression for two senders sharing a state \( \psi^{C_1C_2R} \), with purifying system \( R \). The objective is the same as in state merging, except that the decoder has no prior information about the state. Quantum communication will be achieved using pre-shared EPR pairs and classical communication. The decoder will recover the original state by applying an appropriate decoding operation. Pure entanglement, shared between the decoder and the involved senders, might also be distilled in the process. If we let \( R_i \) denote the net amount of entanglement consumed (or generated if \( R_i \) is negative) in a distributed compression scheme, it was found in \cite{4} that the rates must obey

\[
\begin{align*}
R_1 &\geq S(C_1|C_2)_{\psi} \\
R_2 &\geq S(C_2|C_1)_{\psi} \\
R_1 + R_2 &\geq S(C_1C_2)_{\psi}
\end{align*}
\]
Observe that this is just a special case of eq. \[(44)\] with \(m = 2\) and Bob having no side information. Figure 2 shows the achievable rate region when the conditional entropies have positive values.

One way to perform distributed compression is to apply the original state merging protocol as many times as needed, adjusting the amount of pre-shared entanglement required depending on the information the decoder has after each application of the protocol. For instance, if we wish to first transfer \(C_1\)'s share to Bob, we can apply the state merging protocol using an entanglement rate of \(S(C_1)\), which amounts to Schumacher compressing the state \(\psi^{C_1}\) since the receiver has no prior information about the state \(\psi^{C_1C_2R}\). Then, to transfer \(C_2\)'s share of the state, we perform another state merging, this time, with an entanglement rate of \(S(C_2|C_1)\). This will correspond to one specific corner of the boundary in Figure 2. Transferring \(C_2\) first, and then \(C_1\) will give us the other corner. To attain all other points on the boundary using this approach, time-sharing will be required.

The techniques used to prove Theorem 3, however, demonstrated that time sharing is not essential to the task of multiparty state merging. Let \((R_1, R_2)\) be any point in the rate region. Then, \(R_1\) and \(R_2\) must satisfy eq. \[(45)\], and so by Theorem 3 the rate-tuple \((R_1, R_2)\) is achievable for multiparty merging of the state \(\psi^{C_1C_2R}\). That is, given a large number of copies of \(\psi^{C_1C_2R}\), there exist multiparty state merging protocols \(\Lambda^{n}_1\) of vanishing error and entanglement rate \(\frac{1}{n}(\log K_1^n - \log L_1^n - \log K_2^n - \log L_2^n)\) approaching \(R_1\) and \(R_2\) respectively. In the proof of Theorem 3 we have shown the existence of merging protocols of a specific kind. For these protocols, each sender performs a single measurement with projectors of rank \(L_i\), and one of rank \(L'_i \leq L_i\) on his share \((C_iC''_i)\otimes n\). The amount of pre-shared entanglement required and the rank of the projectors will need to satisfy eq. \[(40)\]. The receiver will then apply a decoding \(U_J\) once he receives the outcome of the measurements. These protocols do not partition the input state \((\psi^{C_1C_2R})\otimes n\) to achieve the desired rates \((R_1, R_2)\). Hence, time-sharing is not required and the parties can perform merging at any rate \((R_1, R_2)\) lying in the rate region if they were supplied with enough initial entanglement.

VI. MIN-ENTROPIES AND ONE-SHOT MERGING

A. Review of Min- and Max-Entropies

Quantum min- and max-entropies are adaptations of the classical Rényi entropies of order \(\alpha\) when \(\alpha \to \infty\) and \(\alpha = 1/2\) respectively. The Rényi entropies were introduced by Rényi [26] in 1961 as alternatives to the Shannon entropy as measures of information. Although introduced in an operational way, the Shannon entropy can also be regarded as the unique function which satisfies a set of prescribed postulates. Rényi showed that by generalizing some of the postulates, other information-theoretic quantities could be obtained, and this gave rise to the family of Rényi entropies, parameterized by a positive number \(\alpha\). Rényi entropies and their quantum generalizations have found applications in areas such as cryptography [27, 28] and statistics [29, 30]. For our purposes, only the definitions and some basic properties of the min- and max- entropies will actually be needed.

Let \(S_\leq(AR)\) be the set of sub-normalized density operators (i.e. \(\text{Tr}(\bar{\rho}^{AR}) \leq 1\)) on the space \(AR\). The quantum min-entropy [17] of an operator \(\rho^{AR} \in S_\leq(AR)\) relative to a density operator \(\sigma^R\) is given by

\[
H_{\min}(\rho^{AR}|\sigma^R) := -\log \lambda, \tag{46}
\]

where \(\lambda\) is the minimum real number such that \(\lambda(I^A \otimes \sigma^R) - \rho^{AR}\) is positive semidefinite. The
conditional min-entropy \( H_{\text{min}}(\rho^{AR}|R) \) is obtained by maximizing the previous quantity over all density operators \( \sigma^R \):

\[
H_{\text{min}}(\rho^{AR}|R) := \sup_{\sigma^R} H_{\text{min}}(\rho^{AR}|\sigma^R). \tag{47}
\]

For two sub-normalized states \( \rho \) and \( \bar{\rho} \), we define the purified distance between \( \rho \) and \( \bar{\rho} \) as

\[
P(\rho, \bar{\rho}) := \sqrt{1 - \mathcal{F}(\rho, \bar{\rho})^2}, \tag{48}
\]

where \( \mathcal{F}(\rho, \bar{\rho}) \) is the generalized fidelity between \( \rho \) and \( \bar{\rho} \) (see [18] for the definition). It is related to the trace distance \( D(\rho, \bar{\rho}) := \frac{1}{2} \| \rho - \bar{\rho} \|_1 \) as follows

\[
D(\rho, \bar{\rho}) \leq P(\rho, \bar{\rho}) \leq 2\sqrt{D(\rho, \bar{\rho})}. \tag{49}
\]

A proof of this fact can be found in Lemma 6 of [18]. (Lemma 6 actually relates the purified distance to the generalized distance \( \bar{D}(\rho, \bar{\rho}) \). However, \( \bar{D}(\rho, \bar{\rho}) \) is always greater than or equal to the trace distance and bounded above by \( \|\rho - \bar{\rho}\|_1 \).)

Using the purified distance as our measure of closeness, we obtain a family of smooth min-entropies \( \{H_{\text{min}}^\epsilon\} \) by optimizing over all sub-normalized density operators close to \( \rho^{AB} \) with respect to \( P(\bar{\rho}, \rho) \):

\[
H_{\text{min}}^\epsilon(\rho^{AR}|R) := \sup_{\bar{\rho}^{AR}} H_{\text{min}}(\bar{\rho}^{AR}|R), \tag{50}
\]

where the supremum is taken over all \( \bar{\rho}^{AR} \) such that \( P(\bar{\rho}^{AR}, \rho^{AR}) \leq \epsilon \). If we use the trace distance instead as our measure of closeness, we obtain the family \( \{\bar{H}_{\text{min}}^\epsilon\} \):

\[
\bar{H}_{\text{min}}^\epsilon(\rho^{AR}|R) := \sup_{\bar{\rho}^{AR}} H_{\text{min}}(\bar{\rho}^{AR}|R), \tag{51}
\]

where the supremum is taken over all sub-normalized \( \bar{\rho}^{AR} \) such that \( D(\bar{\rho}^{AR}, \rho^{AR}) \leq \epsilon \). From eq. (49), the smooth min-entropy \( H_{\text{min}}^\epsilon \) can be bounded by \( \bar{H}_{\text{min}}^\epsilon \) in the following way:

\[
H_{\text{min}}^\epsilon(\rho^{AR}|R) \leq \bar{H}_{\text{min}}^\epsilon(\rho^{AR}|R) \leq H_{\text{min}}^\epsilon(\rho^{AR}|R). \tag{52}
\]

Given a purification \( \rho^{ABR} \) of \( \rho^{AR} \), with purifying system \( B \), the family of smooth max entropies \( \{H_{\text{max}}^\epsilon\} \) is defined as

\[
H_{\text{max}}^\epsilon(\rho^{AB}|B) := -H_{\text{min}}^\epsilon(\bar{\rho}^{AR}|R), \tag{53}
\]

for any \( \epsilon \geq 0 \). The smooth max-entropies can also be expressed as

\[
H_{\text{max}}^\epsilon(\rho^{AB}|B) = \inf_{\bar{\rho}^{AB}} H_{\text{max}}(\bar{\rho}^{AB}|B), \tag{54}
\]

where the infimum is taken over all \( \bar{\rho}^{AB} \) such that \( P(\bar{\rho}, \rho) \leq \epsilon \). We refer to [18] for a proof of this fact. When \( \epsilon = 0 \), an alternative expression for the max-entropy \( H_{\text{max}}(\psi^{AB}|B) \) was obtained in [31]:

\[
H_{\text{max}}(\rho^{AB}|B) = \sup_{\sigma^B} \log F^2(\rho^{AB}, I^A \otimes \sigma^B), \tag{55}
\]
where the supremum is taken over all density operators $\sigma^B$ on the space $B$. From this last expression, the smooth max-entropy $H^e_{\max}(\rho^A)$ of a sub-normalized operator $\rho^A \in S_{\leq}(A)$ reduces to
\begin{equation}
H^e_{\max}(\rho^A) = 2 \log \sum_x \sqrt{\bar{r}_x},
\end{equation}
where $\bar{r}_x$ are the eigenvalues of the sub-normalized density operator $\bar{\rho}^A$ which optimizes the right hand side of eq. (54).

When defining the smooth min- and max-entropies using the purified distance, other useful properties such as quantum data processing inequalities and concavity of the max-entropy are also known to hold. A detailed analysis of these properties can be found in [18].

We will also need, for technical reasons, another entropic quantity called the conditional collision entropy $H_2(\rho^{AB}\mid\sigma^B)[18]$. It is defined as
\begin{equation}
H_2(\rho^{AB}\mid\sigma^B) := -\log \text{Tr}
\left[
\left(I_A \otimes \sigma^{-1/4}_B\right)\rho^{AB}\left(I_A \otimes \sigma^{-1/4}_B\right)
\right]^2,
\end{equation}
It is a quantum adaptation of the classical conditional collision entropy. We have the following lemma relating the min-entropy to the collision entropy:

**Lemma 7** [18] For density operators $\rho^{AB}$ and $\sigma^B$ with $\text{supp}\{\text{Tr}_A(\rho^{AB})\} \subseteq \text{supp}\{\sigma^B\}$, we have
\begin{equation}
H_{\min}(\rho^{AB}\mid\sigma^B) \leq H_2(\rho^{AB}\mid\sigma^B).
\end{equation}

The last two results we will need are the additivity of the min-entropy and the following lemma which relates the trace norm of an hermitian operator $S$ to its Hilbert-Schmidt norm, with respect to a positive semidefinite operator $\sigma$:

**Lemma 8** Let $S$ be an hermitian operator acting on some space $X$ and $\sigma$ be a positive semidefinite operator on $X$. Then, we have
\begin{equation}
\|S\|_1 \leq \sqrt{\text{Tr}(\sigma)}\|\sigma^{-1/4}S\sigma^{-1/4}\|_2.
\end{equation}

**Lemma 9 (Additivity)** Let $\rho^{AB}$ and $\rho^{A'B'}$ be sub-normalized operators on the spaces $AB$ and $A'B'$ respectively. For density operators $\sigma^B$ and $\sigma^{B'}$, we have
\begin{equation}
H_{\min}(\rho^{AB} \otimes \rho^{A'B'}\mid\sigma^B \otimes \sigma^{B'}) = H_{\min}(\rho^{AB}\mid\sigma^B) + H_{\min}(\rho^{A'B'}\mid\sigma^{B'}).
\end{equation}

For proofs of the preceding two lemmas, see [16].

**B. Characterizing the entanglement cost of merging using min-entropies**

In [16], Berta showed that the smooth min-entropy is the information theoretic measure which quantifies the minimal amount of entanglement necessary for performing state merging when a single copy of $\psi^{ABR}$ is available. More specifically, he proved that the minimal entanglement cost $\log K - \log L$ necessary for merging the $A$ part of a state $\psi^{ABR}$ to the location of the $B$ system is bounded from below by
\begin{equation}
\log K - \log L \geq -\bar{H}^e_{\min}(\psi^{AR}\mid R).
\end{equation}
Furthermore, he demonstrated the existence of a state merging protocol using an entanglement cost\(^1\) of

\[
\log K - \log L = -\tilde{H}_{\min}^2(\psi^{AR}|R) + 4 \log \left(\frac{1}{\epsilon}\right) + 12. \tag{61}
\]

This last result was derived by re-expressing the upper bound to the quantum error (Lemma 6 of \cite{4}) as a function of the smooth min-entropy.

In this section, we would like to generalize the main results of \cite{16} to the case of multiple parties sharing a state \(\psi_{C_1C_2\ldots C_mBR}\). Our first result is a reformulation of Lemma 6 in terms of min-entropies:

**Lemma 10 (Compare to Lemma 4.5 of \cite{16})** For each sender \(C_i\), let \(P_i : C_i \rightarrow C_i^1\) be a projector onto a subspace \(C_i^1\) of \(C_i\) and \(U_i\) be a unitary on the space \(C_i\). Define the state

\[
\omega_{C_M^R} := (P_1U_1 \otimes P_2U_2 \otimes \ldots \otimes P_mU_m \otimes I_R)\psi_{C_1C_2\ldots C_mR}(P_1U_1 \otimes P_2U_2 \otimes \ldots \otimes P_mU_m \otimes I_R)^\dagger. \tag{62}
\]

If the unitaries \(U_1, U_2, \ldots, U_m\) are distributed according to the Haar measure, then for any state \(\sigma^R\) of the system \(R\), we have

\[
\mathbb{E} \left[ \left\| \frac{\omega_{C_M^R}}{d_{C_M}} \tau_{C_M^1} \otimes \psi^R \right\|_1 \right] \leq \frac{L}{d_{C_M}} \sqrt{\sum_{\mathcal{T}\subseteq\{1,2,\ldots,m\}} 2^{-\left(\tilde{H}_{\min}(\psi^R|\sigma^R) - \log L_\mathcal{T}\right)}}, \tag{63}
\]

where \(L_\mathcal{T} = \prod_{i \in \mathcal{T}} L_i\).

**Proof** Using Lemma 8 we have, for any state \(\sigma^R\) of \(R\),

\[
\left\| \frac{\omega_{C_M^R}}{d_{C_M}} \tau_{C_M^1} \otimes \psi^R \right\|_1 \leq \sqrt{L} \left( I_{C_M^1} \otimes (\sigma^R)^{-\frac{1}{2}} \right) \left( \frac{\omega_{C_M^R}}{d_{C_M}} \tau_{C_M^1} \otimes \psi^R \right) \left( I_{C_M^1} \otimes (\sigma^R)^{-\frac{1}{2}} \right) \tag{64}
\]

Define

\[
\tilde{\omega}_{C_M^R} := (I_{C_M^1} \otimes (\sigma^R)^{-\frac{1}{2}}) (\omega_{C_M^R} \otimes (\sigma^R)^{-\frac{1}{2}})
\]

\[
\tilde{\psi}_{C_M^R} := (P_1U_1 \otimes P_2U_2 \otimes \ldots \otimes P_mU_m \otimes I_R) \tilde{\omega}_{C_M^R} \left( P_1U_1 \otimes P_2U_2 \otimes \ldots \otimes P_mU_m \otimes I_R \right)^\dagger. \tag{65}
\]

Then, the right hand side of eq. (64) can be rewritten as \(\sqrt{L} \left\| \frac{\tilde{\omega}_{C_M^R}}{d_{C_M}} \tau_{C_M^1} \otimes \tilde{\psi}^R \right\|_2\). Using eq. (27) in the proof of Lemma 6 we have

\[
\mathbb{E} \left[ \left\| \frac{\tilde{\omega}_{C_M^R}}{d_{C_M}} \tau_{C_M^1} \otimes \tilde{\psi}^R \right\|_2^2 \right] \leq L \prod_{\mathcal{T}\subseteq\{1,2,\ldots,m\}} \prod_{i \in \mathcal{T}} \frac{L_i}{d_{C_i}^2} \left[ (\tilde{\psi}^R)^2 \right] \tag{66}
\]

\[ \leq \frac{L}{d_{C_M}^2} \sum_{\mathcal{T}\subseteq\{1,2,\ldots,m\}} L_\mathcal{T} \left[ (\tilde{\psi}^R)^2 \right]. \]

\(^1\) The numbers \(K, L\) are natural numbers, and so we must choose values for \(K\) and \(L\) such that \(\log K - \log L\) is minimal, but greater or equal than the right hand side of eq. (61).
The quantity $\text{Tr}[(\tilde{\psi}^R)^2]$ can be rewritten as:

$$
\text{Tr}[(\tilde{\psi}^R)^2] = \text{Tr}\left[\\left(\text{Tr}_C(\tilde{\psi}_C^M)^2\right)\right] \\
= \text{Tr}\left[\left((I^T \otimes (\sigma^R)^{-\frac{1}{2}})(\tilde{\psi}^R)(I^T \otimes (\sigma^R)^{-\frac{1}{2}})\right)^2\right] \\
= 2^{-H_2(\psi^R|\sigma^R)},
$$

(67)

where $H_2(\psi^R|\sigma^R)$ is the conditional collision entropy of $\psi^R$ relative to $\sigma^R$. Combining eqs. (64), (66) and (67) together, and using the fact that $H_{\min}(\psi^R|\sigma^R) \leq H_2(\psi^R|\sigma^R)$, we get

$$
\mathbb{E}\left[\left\|\omega^{C^M_R} - \frac{L}{d_C} \psi^C \right\|_1\right] \leq \frac{L}{d_C} \sum_{T \subseteq \{1,2,\ldots,m\}} \frac{L_T 2^{-H_2(\psi^R|\sigma^R)}}{2^{-\left(H_{\min}(\psi^R|\sigma^R) - \log L_T\right)}}.
$$

(68)

With this result in hand, we are now ready to give an adaptation of Lemma 4.6 in [16] for our general multiparty setting.

**Theorem 11 (Compare to Proposition 4.7 of [16]):** Let $\psi^{C_1C_2\ldots C_mBR}$ be any $(m+2)$-partite pure state and fix $\epsilon > 0$. Then, for any entanglement cost $E = (\log K_1 - \log L_1, \log K_2 - \log L_2, \ldots, \log K_m - \log L_m)$ satisfying

$$
\log K_T - \log L_T := \sum_{i \in T} (\log K_i - \log L_i) \geq -H_{\min}(\psi^R|\sigma^R) + 4\log\left(\frac{1}{\epsilon}\right) + 2m + 8
$$

(69)

for all non-empty subsets $T \subseteq \{1,2,\ldots,m\}$, there exists a state merging protocol acting on $\psi^{C_1C_2\ldots C_mBR}$ with error $\epsilon$. The set $\overline{T}$ is defined as the complement of $T$.

**Proof** We proceed by fixing a random measurement for each sender $C_i$ as in Proposition 5. We can describe $C_i$’s random measurement using $N_i := \left\lfloor \frac{d_{C_i}}{L_i} \right\rfloor$ partial isometries $P_i^j = Q_i^j U_i$, where $U_i$ is a Haar distributed random unitary acting on the system $C_i C_i^0$ and $Q_i^j$ is as defined in Proposition 5. If $d_{C_i} K_i > N_i L_i$, there is an additional partial isometry $P_i^0$ of rank $L_i^0 := d_{C_i} K_i - N_i L_i < L_i$. Applying the previous lemma to the state $\psi^{C^M_R \otimes T^{K_1} \otimes T^{K_2} \otimes \ldots \otimes T^{K_m}}$, with $\sigma^R = \psi^R$, we get

$$
\mathbb{E}\left[\left\|\sum_{j=1}^{N_1} \sum_{j_2=1}^{N_2} \ldots \sum_{j_m=1}^{N_m} \omega^{C^M_R} - \frac{L}{d_C} \psi^C \right\|_1\right] \leq \frac{\prod_{i=1}^{m} N_i L_i}{d_C K} \sum_{T \subseteq \{1,2,\ldots,m\}} \frac{2^{-\left(H_{\min}(\psi^R|\sigma^R) + \log K_T - \log L_T\right)}}{2^{-H_{\min}(\psi^R|\sigma^R)}}.
$$

(70)
where $K_T = \prod_{i \in T} K_i$. Note that to get the first inequality, we have used the fact that $H_{\min} (\psi^T R | \psi^R) = H_{\min} (\psi^T R | \psi^R) + \log K_T$.

Using eq. (69), we can simplify the previous inequality and obtain

$$E \left[ \sum_{j_1=1}^{N_1} \sum_{j_2=1}^{N_2} \cdots \sum_{j_m=1}^{N_m} \left( \omega_{j}^{C_1} - \frac{L}{d_{C_M}} \tau_{C_M} \otimes \psi^R \right) \right] \leq \frac{\sum_{\substack{T \subseteq \{1,2,\ldots,m\} \neq \emptyset}} \tau_{\subseteq T} \epsilon^2}{2^{m+7}} \leq \frac{\epsilon^2}{16}.$$  

(71)

Taking normalisation into account, with $p_j = \text{Tr} (\omega_{j}^{C_1} R)$ and $\psi_{j}^{C_1} = \frac{1}{p_j} \omega_{j}^{C_1} R$, we can trace out the left hand side of the previous set of inequalities, and get

$$E \left[ \sum_{j_1=1}^{N_1} \sum_{j_2=1}^{N_2} \cdots \sum_{j_m=1}^{N_m} p_j - \frac{L}{d_{C_M}} \right] \leq \frac{\epsilon^2}{16}.$$  

(72)

By applying the triangle inequality, we obtain

$$E \left[ \sum_{j_1=1}^{N_1} \sum_{j_2=1}^{N_2} \cdots \sum_{j_m=1}^{N_m} p_j \left| \tau_{j}^{C_1} - \tau_{C_M} \otimes \psi^R \right|_1 \right] \leq \frac{\epsilon^2}{16}.$$  

(73)

Using this, and eq. (35) in the proof of Proposition 5 we can get an upper bound to the quantum error $Q_T (\psi_{C_1 C_2 \cdots C_m BR} \otimes \Phi^K)$:

$$E \left[ \sum_{j_1=1}^{N_1} \sum_{j_2=1}^{N_2} \cdots \sum_{j_m=1}^{N_m} p_j \left| \tau_{j}^{C_1} - \tau_{C_M} \otimes \psi^R \right|_1 \right] \leq 2 \sum_{\substack{T \subseteq \{1,2,\ldots,m\} \neq \emptyset}} \prod_{i \in T} \frac{L_i}{d_{C_i} K_i} + E \left[ \sum_{j_1=1}^{N_1} \sum_{j_2=1}^{N_2} \cdots \sum_{j_m=1}^{N_m} p_j \left| \tau_{j}^{C_1} - \tau_{C_M} \otimes \psi^R \right|_1 \right] \leq \frac{\epsilon^2}{16} + \frac{\epsilon^2}{8}.$$  

(74)

To get the third inequality, we have used the strong subadditivity of the min-entropy [18]:

$$H_{\min} (\psi^T R | \psi^R) \leq H_{\min} (\psi^T).$$

The last line follows from the fact that $H_{\min} (\psi^T) = -\log \lambda_{\text{max}} (\psi^T) \leq \log d_{C_T}$. From Proposition 4 we can conclude there exists a state merging protocol acting on $\psi_{C_1 C_2 \cdots C_m BR}$ with error $\epsilon$. □

When $m = 1$, the previous result yields a merging protocol of error $\epsilon$ and entanglement cost $\log K - \log L = -H_{\min} (\psi^C R | \psi^R) + 4 \log (\frac{1}{\epsilon}) + 10$. Berta [16] showed however that the min-entropy
where \( \log \psi \) is replaced by \( H \). 

Conjecture 12

Let \( \psi^{C_1 R} \) be an \((m+2)\)-partite state. For any \( \epsilon > 0 \), there exists a multiparty state merging of error \( \epsilon \) whenever the entanglement cost \( E := (\log K_1 - \log L_1, \log K_2 - \log L_2, \ldots, \log K_m - \log L_m) \) satisfies

\[
\log K_T - \log L_T := \sum_{i \in T} (\log K_i - \log L_i) \geq -H^e_{\min}(\psi^{TR | R}) + O(\log 1/\epsilon) + O(m)
\]

for all non-empty subsets \( T \subseteq \{1, 2, \ldots, m\} \).

The main difficulty in proving the conjecture is that it allows independent smoothing of each of the min-entropies. It is straightforward to modify our proof to allow smoothing using a common state for all the min-entropies, but the monolithic nature of the protocol does not naturally permit tailoring the smoothing state term-by-term. We can, however, give a partial characterization of the entanglement cost in terms of smooth min-entropies if we apply the single-shot state merging protocol of [15] on one sender at a time.

Proposition 13

For a \((m+2)\)-partite pure state \( \psi^{C_1 C_2 \ldots C_m BR} \), fix an \( \epsilon > 0 \) and let \( \pi : \{1, 2, \ldots, m\} \to \{1, 2, \ldots, m\} \) be any ordering of the \( m \)-senders \( C_1, C_2, \ldots, C_m \). Then, for any entanglement cost \( E := (\log K_1 - \log L_1, \log K_2 - \log L_2, \ldots, \log K_m - \log L_m) \) satisfying

\[
\log K_i - \log L_i \geq -H^{m^2}_{\min}(\psi^{C_i R_{\pi^{-1}(i)} | R_{\pi^{-1}(i)}}) + 4 \log \left( \frac{m}{\epsilon} \right) + 12 \quad \text{for all } 1 \leq i \leq m,
\]

where \( R_i := R \otimes_{j=1, j\neq i} C_{\pi(j)} \), there exists a multiparty state merging protocol acting on the state \( \psi^{C_1 C_2 \ldots C_m BR} \) with error \( \epsilon \).
Proof Our multiparty state merging protocol for the state $\psi^{C_{1}C_{2}...C_{m}BR}$ will consist of sending each sender’s share to the receiver one at a time according to the ordering $\pi$: The sender $C_{\pi(1)}$ will merge his part of the state first, followed by $C_{\pi(2)}$, $C_{\pi(3)}$, etc. We can view the input state $\psi^{C_{1}C_{2}...C_{m}BR}$ as a tripartite pure state $\psi^{C_{\pi(1)}R_{1}B_{1}}$, with the reference system $R_{1}$ being composed of the systems $C_{\pi(2)}C_{\pi(3)}...C_{\pi(m)}R$. According to Berta [16], there exists a state merging protocol of error $\epsilon/m$ and entanglement cost

$$
\log K_{1}' - \log L_{1} := -\frac{H_{\text{min}}^{c\otimes B}}{2} (\psi^{C_{\pi(1)}R_{1}} | R_{1} ) + 4 \log \left( \frac{m}{\epsilon} \right) + 12
$$

$$
\leq -\frac{H_{\text{min}}^{c\otimes B}}{2} (\psi^{C_{\pi(1)}R_{1}} | R_{1} ) + 4 \log \left( \frac{m}{\epsilon} \right) + 12
$$

$$
\leq \log K_{1} - \log L_{1},
$$

which will produce an output state $\rho^{C_{\pi(1)}B_{\pi(1)}BR_{1}}$ satisfying

$$
\left\| \rho_{1}^{C_{\pi(1)}B_{\pi(1)}BR_{1}} - \psi^{B_{\pi(1)}BR_{1}} \otimes \Phi^{L_{1}} \right\|_{1} \leq \frac{\epsilon}{m},
$$

where the system $B_{\pi(1)}$ is substituted for the system $C_{\pi(1)}$. After $C_{\pi(1)}$ has merged his share, the next sender $C_{\pi(2)}$ will perform a random measurement on his share of the state and send the measurement outcome to the receiver. Suppose, for the moment, that the parties share the state $\psi^{B_{\pi(1)}BR_{1}} \otimes \Phi^{L_{1}}$ instead of the output state $\rho_{1}$. The state $\psi^{B_{\pi(1)}BR_{1}}$ can be viewed as a tripartite state $\psi^{C_{\pi(1)}B_{2}R_{2}}$, with $B_{2} := B_{\pi(1)}B_{2}$ and $R_{2} := C_{\pi(3)}C_{\pi(4)}...C_{\pi(m)}R$. Using Berta’s result once more, we know there exists a state merging protocol of error $\epsilon/m$ and entanglement cost

$$
\log K_{2}' - \log L_{2} := -\frac{H_{\text{min}}^{c\otimes B}}{2} (\psi^{C_{\pi(2)}R_{2}} | R_{2} ) + 4 \log \left( \frac{m}{\epsilon} \right) + 12
$$

$$
\leq -\frac{H_{\text{min}}^{c\otimes B}}{2} (\psi^{C_{\pi(2)}R_{2}} | R_{2} ) + 4 \log \left( \frac{m}{\epsilon} \right) + 12
$$

$$
\leq \log K_{2} - \log L_{2},
$$

which produces an output state $\rho_{2}^{C_{\pi(1)}C_{\pi(2)}B_{\pi(2)}B_{\pi(2)}BR_{2}R_{2}}$ satisfying

$$
\left\| \rho_{2}^{C_{\pi(1)}C_{\pi(2)}B_{\pi(2)}B_{\pi(2)}BR_{2}R_{2}} - \psi^{B_{\pi(2)}BR_{2}R_{2}} \otimes \Phi^{L_{1}} \otimes \Phi^{L_{2}} \right\|_{1} \leq \frac{\epsilon}{m},
$$

where the system $B_{\pi(2)}$ is substituted for the system $C_{\pi(2)}$. If we apply the same protocol on the state $\rho_{1}^{C_{\pi(1)}B_{\pi(1)}BR_{1}}$ instead, we get an output state $\rho_{3}$ which satisfies

$$
\left\| \rho_{3}^{C_{\pi(1)}C_{\pi(2)}B_{\pi(2)}B_{\pi(2)}BR_{2}R_{2}} - \psi^{B_{\pi(1)}BR_{2}R_{2}} \otimes \Phi^{L_{1}} \otimes \Phi^{L_{2}} \right\|_{1} \leq \frac{2\epsilon}{m},
$$

where we have used the triangle inequality and monotonicity of the trace distance under quantum operations. The analysis for the other senders $C_{\pi(3)}$, $C_{\pi(4)}$, ..., $C_{\pi(m)}$ can be performed in a similar way, which leads to a multiparty state merging protocol of error $\epsilon$ and entanglement cost $\vec{E} := \text{entanglement cost of all senders combined}$.
\[ E_1 \leq -H_{\max}^{\max}(|\psi_{C_1 R}|)R + 4\log(2) \]

\[ E_2 \leq -H_{\max}^{\max}(|\psi_{C_2 R}|)R + 4\log(2) \]

\[ \| \rho_m - \psi^{B_{z(1)}B_{z(2)}...B_{z(m)}}R \otimes \Phi^{L_1} \otimes \Phi^{L_2} \otimes ... \otimes \Phi^{L_m} \|_1 \leq \epsilon. \] (84)

Figure 3 depicts the boundaries of the regions described by Theorem 11, Conjecture 12, and Proposition 13. Note that the hatched area is not part of the cost region described by Proposition 13.

VII. A WORKED EXAMPLE

The proof of Theorem 11 is significantly more complicated than that of Proposition 13. To illustrate the benefits accruing from the additional effort, we will compare the two results’ estimates of the costs achievable for merging \( C_1 \) and \( C_2 \) to \( R \) for states of the form

\[ |\psi\rangle^{C_1 C_2 R} = \frac{1}{\sqrt{H_d}} \sum_{j=1}^{d} \frac{1}{\sqrt{j}} |j\rangle^{C_1} |\psi_j\rangle^{C_2} |j\rangle^R, \] (85)

where \( H_d = \sum_{j=1}^{d} 1/j \) is the \( d \)th harmonic number. These are close relatives of the embezzling states introduced in [32], which are useful resources for channel simulation and other tasks [19, 33, 34]. They make interesting examples because they have sufficient variation in their Schmidt coefficients that the i.i.d. state merging rates of Theorem 3 are not achievable in the one-shot regime. Nonetheless, our results yield nontrivial one-shot rates that are significantly better than simple teleportation. We will assume that \( |\langle \psi_i | \psi_j \rangle| \leq \alpha \) for \( i \neq j \) and try to and express the rates in terms of \( \alpha \). We will assume for convenience that \( \alpha > 0 \) since, when \( \alpha = 0 \), the costs are essentially the same as when \( \alpha = \Omega(1/d) \).
Protocols from Theorem 11

Let \((E_1, E_2)\) be a pair of entanglement costs achievable according to Theorem 11. The only constraints on the costs (aside from needing to be the logs of integers) are

\[
E_1 \geq -H_{\min}(\psi^{C_1R}|\psi^R) + 4\log(1/\epsilon) + 12 \tag{86}
\]
\[
E_2 \geq -H_{\min}(\psi^{C_2R}|\psi^R) + 4\log(1/\epsilon) + 12 \tag{87}
\]
\[
E_1 + E_2 \geq -H_{\min}(\psi^{C_1C_2R}|\psi^R) + 4\log(1/\epsilon) + 12. \tag{88}
\]

To begin, we will find a sufficient condition for the \(E_1\) constraint to be satisfied, so we need to evaluate \(H_{\min}(\psi^{C_1R}|\psi^R)\). Let \(\lambda\) be the smallest real number such that \(\lambda(I^{C_1} \otimes \psi^R) - \psi^{C_1R} \geq 0\). Expanding the operators, that condition is the same as

\[
\sum_{ij} \frac{\lambda - \delta_{ij}}{j} |i\rangle\langle j| C_1^R = \sum_{i} \sum_{j \neq i} \frac{1}{\sqrt{ij}} |ii\rangle\langle jj| C_1^R \psi \geq 0,
\]

where \(\delta_{ij}\) is the Kronecker delta function. By the Gershgorin Circle Theorem \[35, 36\], the operator \[
\sum_{ij} \frac{\lambda - \delta_{ij}}{j} |i\rangle\langle j| C_1^R - \sum_{i} \sum_{j \neq i} \frac{1}{\sqrt{ij}} |ii\rangle\langle jj| C_1^R \psi \geq 0,
\]

where \(\delta_{ij}\) is the Kronecker delta function. By the Gershgorin Circle Theorem \[35, 36\], the operator will be positive semidefinite if each diagonal entry dominates the sum of the absolute values of the off-diagonal entries in the corresponding row. That condition reduces to

\[
\frac{\lambda - 1}{i} \geq \sum_{j \neq i} \frac{1}{\sqrt{ij}} |\psi_j\psi_i|
\]

holding for all \(i\), which is true provided \(\lambda - 1 \geq \alpha \sum_{j=1}^{d} \sqrt{d/j}\). But

\[
\sum_{j=1}^{d} \frac{1}{\sqrt{j}} \leq \int_{0}^{d} \frac{1}{\sqrt{x}} dx = 2\sqrt{d}.
\]

Therefore, the operator of eq. (89) will be positive semidefinite if \(\lambda \geq 2\alpha d + 1\). This in turn implies that

\[
-H_{\min}(\psi^{C_1R}|\psi^R) \leq \log(2\alpha d + 1) \leq \log(\alpha d) + 1.
\]

The lower bound of eq. (86) will therefore be satisfied provided \(E_1 \geq \log(\alpha d) + 4\log(1/\epsilon) + 14\). The interpretation is that if the states \(\{\psi_j\}\) are indistinguishable, then \(C_1\) holds the whole purification of \(R\) and must therefore be responsible for the full cost of merging. As the states \(\{\psi_j\}\) become more distinguishable, the purification of \(R\) becomes shared between \(C_1\) and \(C_2\), so the merging cost can be shared. Indeed, if \(\alpha = O(1/d)\), then the lower bound on \(E_1\) becomes a constant, independent of the size of the input state \(|\psi\rangle^{C_1C_2R}\).

Moving on to the \(E_2\) constraint, eq. (87), a similar but easier calculation shows that \(H_{\min}(\psi^{C_2R}|\psi^R) = 0\). For the sum rate \(E_1 + E_2\), it is necessary to evaluate \(H_{\min}(\psi^{C_1C_2R}|\psi^R)\). Since the rank of \(\psi^{C_1C_2}\) is \(d\), this entropy is at least \(-\log d\) \[18\].

So any pair of costs \((E_1, E_2)\) satisfying

\[
E_1 \geq \log(\alpha d) + 4\log(1/\epsilon) + 14 \tag{93}
\]
\[
E_2 \geq 4\log(1/\epsilon) + 12 \tag{94}
\]
\[
E_1 + E_2 \geq \log d + 4\log(1/\epsilon) + 12 \tag{95}
\]

will be achievable by Theorem 11. The total cost \(E_1 + E_2\) must be at least \(\log d\) plus terms independent of the size of \(|\psi\rangle^{C_1C_2R}\) and that cost can be shared between \(E_1\) and \(E_2\). The lower bound
on $E_2$ alone is independent of $d$ and should be regarded as a small “overhead” for the protocol. There is a minimal $d$-dependent cost for $E_1$, however, which encodes the fact that if $C_2$ does not carry enough of the purification of $R$ by virtue of the nonorthogonality of the $\{|\psi_j\rangle\}$, then more of the burden will fall to $C_1$.

### A. Protocols from Proposition 13

Now let us consider the costs achievable according to Proposition 13. For fixed $\epsilon$, the proposition provides two cost pairs, plus others that are simply degraded versions of those two arising from the wasteful consumption of unnecessary entanglement. Proposition 13 does not permit interpolation between the two points, as compared to Theorem 11. It might be the case, however, that Proposition 13’s freedom to smooth the entropy and vary the operator being conditioned upon could result in those two cost pairs being much better than any of those provided by Theorem 11. On the contrary, for the states of the example, the improvement achieved with the extra freedom is minimal.

Let $(E'_1, E'_2)$ be a cost pair achievable by Proposition 13. For the purposes of illustration, consider the point with the smallest possible value of $E'_2$. Letting $\delta = \epsilon^2 / 256$, that point will satisfy

\begin{align}
E'_1 &\geq -H_{\min}(\psi^{C_1C_2R}|C_2R) + 4 \log (2/\epsilon) + 12 \\
E'_2 &\geq -H_{\min}(\psi^{C_2R}|R) + 4 \log (2/\epsilon) + 12.
\end{align}

(96)  (97)

Since the state $\psi^{C_2R}$ is separable, the cost $E'_2$ cannot be negative, at least for sufficiently small $\epsilon$, so the key number is the $E'_1$ cost. Before introducing the extra complication of smoothing, consider first $H_{\min}(\psi^{C_1C_2R}|C_1R)$. By [31], this is related to the largest overlap that can be achieved with a maximally entangled state on $C_2/C_1R$ by acting with a quantum channel on the $C_1R$ part of $\psi^{C_1C_2R}$. This maximum singlet fraction is at least what is achieved by just aligning the Schmidt bases, which is

\begin{align}
\left| \sum_{j=1}^{d} \frac{1}{\sqrt{j} \cdot H_d} \frac{1}{\sqrt{d}} \right|^2 &\leq \frac{1}{d \cdot H_d} \left| \sum_{j=1}^{d} \frac{1}{\sqrt{j}} \right|^2 \\
&\geq \frac{1}{d \cdot H_d} \left| \int_1^d \frac{1}{\sqrt{x}} dx \right|^2 \\
&= \frac{4}{H_d} \left( 1 - O \left( \frac{1}{\sqrt{d}} \right) \right) \\
&\geq \frac{5}{\log d},
\end{align}

(98)  (99)  (100)

where the last line holds for sufficiently large $d$. Above and in what follows, we use the inequality $\ln(d+1) \leq H_d \leq (\ln d) + 1$, which was supplemented above by the fact that $4/(\ln d + 1) \geq 5.7/\log d$ for sufficiently large $d$. According to Theorem 2 of [31], the resulting bound on $H_{\min}$ is

\begin{equation}
-H_{\min}(\psi^{C_1C_2R}|RC_2) \geq \log d - \log \log d + 2.
\end{equation}

(102)

Therefore, ignoring smoothing, the sum cost for Proposition 13 will always satisfy

\begin{equation}
E'_1 + E'_2 \geq \log d - \log \log d + 26 + 8 \log \left( \frac{2}{\epsilon} \right),
\end{equation}

(103)
for sufficiently large \( d \), which has worse constants and even asymptotically only differs from the sum cost \( (95) \) for Theorem 11 by \( O(\log \log d) \).

Now let us introduce some smoothing. By duality of the min- and max- entropies,

\[
-H^\delta_{\min}(\psi^{C_1C_2|R} | RC_2) = H^\delta_{\max}(\psi^{C_1}).
\]

Lemma 25 of Appendix C gives that

\[
H^\delta_{\max}(\psi^{C_1}) \geq 2 \log \min \left\{ \sum_{j=1}^{k-1} \frac{1}{\sqrt{j \cdot H_d}} : k \text{ such that } \sum_{j=k+1}^{d} \frac{1}{j \cdot H_d} \leq \frac{\delta^2}{2} \right\}.
\]

Getting a lower bound on this expression requires finding large \( k \) that nonetheless fail to satisfy the tail condition. That restriction on \( k \) is equivalent to \( 1 - H_k / H_d \leq \delta^2 / 2 \), which will not be met by any \( k \) small enough to obey

\[
k \leq (d + 1)^{1-\delta^2 / 2} / e
\]

for sufficiently large \( d \). Using a similar estimate as for the maximum singlet fraction calculation, we get

\[
2 \log \sum_{j=1}^{k-1} \frac{1}{\sqrt{j \cdot H_d}} \geq \log \left( \frac{1}{\sqrt{H_d}} \int_1^{\sqrt{k}} \frac{1}{\sqrt{x}} dx \right)^2
\]

\[
\geq \log \frac{4k}{H_d} \left( 1 - O \left( \frac{1}{\sqrt{k}} \right) \right)
\]

\[
\geq \log k - \log \log d + \log 5
\]

for sufficiently large \( k \). Substituting in the largest possible \( k \) consistent with eq. (106) and \( \delta = \epsilon^2 / 256 \) gives

\[
E'_1 + E'_2 \geq \left( 1 - \frac{\epsilon^4}{512} \right) \log(d + 1) - \log \log d + 24 + 8 \log \left( \frac{2}{\epsilon} \right),
\]

for sufficiently large \( d \). The sum costs achievable using Theorem 11 compare favorably with this bound. The additional savings from smoothing are only about \( \epsilon^4 \log(d + 1) / 512 \) ebits, which is insignificant for small \( \epsilon \). These tiny savings also come at the expense of being able to interpolate between achievable costs. To be fair, these states were chosen specifically because they are known to maintain their essential character even after smoothing, as was observed in [37]. The freedom to smooth is certainly more beneficial for some other classes of states, most notably i.i.d. states. Indeed, since \( S(\psi^{C_1C_2}) = (\log d) / 2 + O(\log \log d) \), merging many copies of \( |\psi\rangle^{C_1C_2R} \) can be done at a rate roughly half the cost required for one-shot merging.

### VIII. A VARIANT OF MERGING: SPLIT TRANSFER

In the previous sections, we’ve analyzed and characterized the entanglement cost for merging the state \( \psi^{C_1C_2...C_mBR} \) to a single receiver Bob in the asymptotic setting and in the one-shot regime. Here, we modify our initial setup by introducing a second decoder \( A \) (Alice), who is spatially separated from Bob and also has side information about the input state. That is, the helpers
C_1, C_2, \ldots, C_m$ and the two receivers Alice and Bob share a global state $\psi^{C_1C_2\ldots C_mABR}$ and the objective is then to redistribute the state $\psi^{C_1C_2\ldots C_mABR}$ to Alice, Bob, and the reference $R$. The motivation for this problem comes from the multipartite entanglement of assistance problem \cite{4,5}, where the task is to distill entanglement in the form of EPR pairs from a $(m+2)$-partite pure state $\psi^{C_1C_2\ldots C_mAB}$ shared between two recipients (Alice and Bob) and $m$ other helpers $C_1, C_2, \ldots, C_m$. If many copies of the input state are available, the optimal EPR rate was shown in [4] to be equal to

$$E_\infty(\psi^{C_1C_2\ldots C_mAB}) := \min_T S(AT)_\psi,$$  \hspace{1cm} (111)

where $T \subseteq \{C_1, C_2, \ldots, C_m\}$ is a subset (i.e. a bipartite cut) of the helpers. We denote the complement by $\overline{T} := \{C_1C_2\ldots C_m\} \setminus T$. We call $\min_T \{S(AT)_\psi\}$ the minimum cut (min-cut) entanglement of the state $\psi^{C_1C_2\ldots C_mAB}$. The procedure described in the proof of [4] makes use of a multiple-blocking strategy. That is, given $n$ copies of the input state $\psi^{C_1C_2\ldots C_mAB}$, the first helper will perform $d = n/m$ random measurements, each acting on $m$ copies of $\psi^{C_1C_2\ldots C_mAB}$ and generating $J$ possible outcomes. Then, if each measurement can yield outcomes $j_1, j_2, \ldots, j_d$, we need to group together the residual states corresponding to outcome 1, then group the ones corresponding to outcome 2, etc... When this is done, the next helper will perform random measurements for each of these groups in the same way the first sender proceeded. That is, for each group, you need to divide into blocks, and so on. Needless to say, this approach fails in the one-shot setting.

It was conjectured in [4] that these layers of blocking could be removed by letting all the helpers perform simultaneous measurements on their respective typical subspaces. Such a strategy would still produce states which preserve the min-cut entanglement, thereby providing a way to prove eq. (111) without the need for a recursive argument. We will show in the remainder of this section that for a cut $T_{min}$ which minimizes $S(AT)_\psi$, there exists an LOCC protocol acting on the state $\psi^{C_1C_2\ldots C_mBR}$ which will send $T_{min}$ to Alice and its complement to Bob. The protocol will consist of two parts. First, all the helpers will perform measurements on their typical subspaces and broadcast their outcomes to both decoders. Then, Alice will use the classical information coming from the helpers which are part of the cut $T_{min}$ and apply an isometry $U$, while Bob will apply an isometry $V$ depending on the outcomes of the helpers belonging to $\overline{T}_{min}$. This will redistribute the initial state to Alice, Bob, and the reference $R$. Standard distillation protocols \cite{38,39} on the recovered state will yield EPR pairs at a rate given by eq. (111).

**Definition 14** Let $\psi^{T\overline{T}ABR}$ be an $(m+2)$-partite state, where $T$ and $\overline{T}$ are a partition of the helpers $C_1, C_2, \ldots, C_m$. Furthermore, assume that the helpers and the decoders share maximally entangled states $\Phi^K := \bigotimes_{i \in T} \Phi^K_i$ and $\Gamma^M := \bigotimes_{i \in \overline{T}} \Gamma^M_i$ on the tensor product spaces $T^0A^0_i$ and $\overline{T}^0B^0_i$.

We call the LOCC operation $\mathcal{M} : TT^0\overline{T}\overline{T}^0 \otimes AA^0_T \otimes BB^0_R \rightarrow T^1A^1_T A_T \otimes \overline{T}^1B^1_{\overline{T}} B_{\overline{T}}$ a split transfer of the state $\psi^{T\overline{T}ABR}$ with error $\epsilon$ and associated entanglement costs $\overline{E}^0_T(\psi) := \bigoplus_{i \in T}(\log K_i - \log L_i)$ and $\overline{E}^0_{\overline{T}}(\psi) := \bigoplus_{i \in \overline{T}}(\log M_i - \log N_i)$ if

$$\left\| (\text{id}_R \otimes \mathcal{M})(\psi^{T\overline{T}ABR} \otimes \Phi^K \otimes \Gamma^M) - \psi_{A_T B_{\overline{T}} R} \otimes \Phi^L \otimes \Gamma^N \right\|_1 \leq \epsilon,$$  \hspace{1cm} (112)

where $\Phi^L := \bigotimes_{i \in T} \Phi^L_i$, $\Gamma^N := \bigotimes_{j \in \overline{T}} \Gamma^N_j$, with the states $\Phi^L_i$ and $\Gamma^N_j$ being maximally entangled states of Schmidt ranks $L_i$ and $N_j$ on $C^1_i A^1_i$ and $C^1_j B^1_j$ respectively. Also, the systems $A_T$ and $B_{\overline{T}}$ are...
ancillary systems of the same size as $T$ and $\overline{T}$ and are held by Alice and Bob respectively. For the state $\Psi := (\psi^{TT^{ABR}})^\otimes n$, the entanglement rates $\overline{R}_{T}^T(\psi)$ and $\overline{R}_{T}^{\overline{T}}(\psi)$ are defined as $\frac{1}{n}E_T^T(\psi)$ and $\frac{1}{n}E_{\overline{T}}^{\overline{T}}(\psi)$.

In the above definition, we have denoted by $\bigoplus_{i \in T} (\log K_i - \log L_i)$ a vector of length $|T|$ whose components are given by $\log K_i - \log L_i$ for $i \in T$ in the lexicographical order.

The rate region where a split-transfer can be accomplished by LOCC can be defined in a manner analogous to definition\(\text{[2]}\). We omit the details here, but whenever we will say that a rate is achievable for a split-transfer of the state $\psi^{TT^{ABR}}$, it will mean that it is contained in the rate region.

Now, we’d like to specify conditions, as in Proposition\(\text{[4]}\) that the initial state should satisfy in order to allow the group $T$ (resp. $\overline{T}$) to transfer their share of the state to Alice (resp. Bob). For a pure state $\psi^{TT^{ABR}}$, suppose all the helpers perform incomplete measurements (as in Section\(\text{[3]}\)) on their respective shares of the state. For measurement outcomes $J := (j_1, j_2, \ldots, j_m)$, define the state

$$|\psi^J_T^{TT^{ABR}}\rangle := \frac{1}{\sqrt{p_J}}(P^j_1 \otimes P^j_2 \otimes \ldots \otimes P^j_m \otimes I^{ABR})|\psi^{TT^{ABR}}\rangle = \frac{1}{\sqrt{p_J}}(P^T_J \otimes P^{\overline{T}}_J \otimes I^{ABR})|\psi^{TT^{ABR}}\rangle,$$

(113)

where $p_J$ is the probability of getting outcome $J$. In the above definition, $j_T$ is a vector of length $|T|$ whose components correspond to outcomes of measurements performed by the helpers belonging to the cut $T$. The vector $j_T$ is defined similarly. Finally, the Kraus operators $P^T_{j_T} = \bigotimes_{i \in T} P^i_{j_i}$ and $P^{\overline{T}}_{j_T} = \bigotimes_{i \in T} P^i_{j_i}$ map the spaces $T$ and $\overline{T}$ to the subspaces $T^1$ and $\overline{T}^1$ respectively.

Define another state $|\varphi^J_T^{TT^{ABR}}\rangle := \frac{1}{\sqrt{p_J}}(P^T_{j_T} \otimes I^{ABR})|\psi^{TT^{ABR}}\rangle$, where $p_{J_T}$ is the probability of getting the outcome $j_T$, and suppose that we have

$$\varphi^J_T^{TT^{AR}} = \tau^T_L \otimes \psi^{TT^{BR}},$$

(114)

where $\tau^T_L = \bigotimes_{i \in T} \tau^{C_i}$ is the maximally mixed state of dimension $L$ on the system $T^1$. From the Schmidt decomposition, we know there exists an isometry $U^{A}_{j_T} : A \rightarrow A_T^1 A_T A$ which Alice can perform such that

$$(I^{T^1T^{BR}} \otimes U^{A}_{j_T})|\varphi^J_T^{TT^{ABR}}\rangle = |\Phi^L_T\rangle \otimes |\psi^{TT^{ABR}}\rangle,$$

(115)

where the state $|\psi^{TT^{ABR}}\rangle$ is the same as the original state $|\psi^{TT^{ABR}}\rangle$ with the ancillary system $A_T$ substituted for $T$. The state $|\Phi^L_T\rangle$ is a maximally entangled state on $T^1 A_T$.

Finally, define the state $|\nu^{A_T\overline{T}^{ABR}}_{j_T}\rangle := \frac{1}{\sqrt{p_T}}(P^T_{j_T} \otimes I^{ABR})|\psi^{TT^{ABR}}\rangle$ and suppose again that we have

$$\nu^{A_T\overline{T}^{1AR}}_{j_T} = \tau^{T\overline{T}}_N \otimes \psi^{TT^{1AR}},$$

(116)

where $\tau^{T\overline{T}}_N = \bigotimes_{i \in T} \tau^{C_i}$ is the maximally mixed state of dimension $N$ on the system $T^{\overline{T}}$. Applying the Schmidt decomposition once more, Bob can perform an isometry $V^B_{j_T} : B \rightarrow B^{1\overline{T}} B$ such that

$$(I^{A_T\overline{T}^{1AR}} \otimes V^B_{j_T})|\nu^{A_T\overline{T}^{1AR}}_{j_T}\rangle = |\Gamma^N\rangle \otimes |\psi^{TT^{1ABR}}\rangle,$$

(117)
where the state $|\psi^{ATB_TABR}\rangle$ is the same as the original state $\psi^{T\overline{T}ABR}$ with the ancillary systems $A_T$ and $B_T$ substituted for $T$ and $\overline{T}$.

If we apply the isometries $U_{j_T}^A$ and $V_{j_T}^B$ to the outcome state $|\psi_{j_T}^{T\overline{T}ABR}\rangle$, the resulting state is given by

$$
(I^{T\overline{T}R} \otimes U_{j_T}^A \otimes V_{j_T}^B) |\psi_{j_T}^{T\overline{T}ABR}\rangle
$$

$$
= \frac{1}{\sqrt{p_J}}(I^{T\overline{T}R} \otimes U_{j_T}^A \otimes V_{j_T}^B)(P_{j_T}^T \otimes \psi^{T\overline{T}ABR})(P_{j_T}^T \otimes I^{T\overline{T}ABR}) |\psi^{T\overline{T}ABR}\rangle
$$

$$
= \frac{1}{\sqrt{p_J}}(I^{T\overline{T}R} \otimes U_{j_T}^A \otimes V_{j_T}^B)(P_{j_T}^T \otimes I^{T\overline{T}ABR})(P_{j_T}^T \otimes U_{j_T}^A)(I^{T\overline{T}R} \otimes |\psi^{T\overline{T}ABR}\rangle)
$$

$$
= \frac{1}{\sqrt{p_J}}(I^{T\overline{T}R} \otimes U_{j_T}^A \otimes V_{j_T}^B)(P_{j_T}^T \otimes I^{T\overline{T}ABR})(P_{j_T}^T \otimes |\psi^{T\overline{T}ABR}\rangle)
$$

$$
= \sqrt{p_J}(P_{j_T}^T \otimes V_{j_T}^B)(P_{j_T}^T \otimes I^{T\overline{T}ABR}) |\psi_{j_T}^{T\overline{T}ABR}\rangle.
$$

(118)

Since the states $(I^{T\overline{T}R} \otimes U_{j_T}^A \otimes V_{j_T}^B) |\psi_{j_T}^{T\overline{T}ABR}\rangle$ and $|\Phi^L\rangle \otimes |\Gamma^N\rangle \otimes |\psi^{AATB_TBR}\rangle$ are both normalized, we must have $p_J = p_{j_T}^2$. Hence, in this ideal case, we can achieve a split transfer of the state $\psi^{C_1C_2...C_mB_{BR}}$ by letting all the helpers measure their share simultaneously. The decoding by Alice and Bob will follow once they receive the measurement outcomes.

**Proposition 15 (Conditions for a Split-Transfer)** Denote the state shared between $m$ helpers and two receivers (Alice and Bob) by $\psi^{T\overline{T}ABR}$, with purifying system $R$. Suppose all the helpers perform incomplete measurements on their share of the state $\psi^{T\overline{T}ABR}$ as in the previous paragraphs, yielding a state $|\psi_{j_T}^{T\overline{T}ABR}\rangle := \frac{1}{\sqrt{p_J}}(P_{j_T}^T \otimes P_{j_T}^T \otimes I^{ABR})(\psi^{T\overline{T}ABR})$ for an outcome $J := (j_1, j_2, \ldots, j_m)$ with probability $p_J$, where the Kraus operators $P_{j_T}^T$ and $P_{j_T}^T$ map the spaces $T$ and $\overline{T}$ to the subspaces $T^1$ and $\overline{T}^1$.

If, for the quantum errors $Q_1^T(\psi^{T\overline{T}ABR})$ and $Q_2^T(\psi^{T\overline{T}ABR})$, we have

$$
Q_1^T(\psi^{T\overline{T}ABR}) := \sum_{j_T} p_{j_T} \| \varphi_{j_T}^{T\overline{T}ABR} - \tau_{T}^{1} \otimes \psi^{T\overline{T}ABR} \|_1 \leq \epsilon
$$

$$
Q_2^T(\psi^{T\overline{T}ABR}) := \sum_{j_T} p_{j_T} \| \psi_{j_T}^{AAT\overline{T}AR} - \tau_{N}^{1} \otimes \psi^{AAT\overline{T}AR} \|_1 \leq \epsilon
$$

(119)

then there exists a split-transfer of the state $\psi^{T\overline{T}ABR}$ with error $2\sqrt{\epsilon} + 2\sqrt{\epsilon'}$ and entanglement costs $E_T^2 = \bigoplus_{i \in T} (- \log L_i)$ and $E_{\overline{T}}^2 = \bigoplus_{i \in \overline{T}} (- \log N_i)$. The states $\varphi_{j_T}^{T\overline{T}ABR}$ and $\psi_{j_T}^{AAT\overline{T}AR}$ are reduced density operators for the states $|\varphi_{j_T}^{T\overline{T}ABR}\rangle := \frac{1}{\sqrt{p_{j_T}}}(P_{j_T}^T \otimes I^{ABR}) |\psi^{T\overline{T}ABR}\rangle$ and $|\psi_{j_T}^{AAT\overline{T}AR}\rangle := \frac{1}{\sqrt{p_{j_T}}}(P_{j_T}^T \otimes A^{ABR}) |\psi^{AAT\overline{T}ABR}\rangle$.

**Proof** Since the quantum errors $Q_1^T$ and $Q_2^T$ are bounded from above by $\epsilon$ and $\epsilon'$ respectively,
Proposition 4 can be applied, which tells us of the existence of isometries $U^A_{jT}$ and $V^B_{jT}$ such that

$$
\left\| \sum_{jT} p_{jT} \left( (I^{TBR} \otimes U^A_{jT}) | \varphi^T_{jT}^{TABR} \rangle \langle \varphi^T_{jT}^{TABR} | (I^{TBR} \otimes U^A_{jT})^\dagger - \psi^{A_T ABR} \otimes \Phi^L \right) \right\|_1 \leq 2\sqrt{\epsilon}
$$

$$
\left\| \sum_{jT} p_{jT} (I^{A_T AR} \otimes V^B_{jT}) | \psi^T_{jT}^{TABR} \rangle \langle \psi^T_{jT}^{TABR} | (I^{A_T AR} \otimes V^B_{jT})^\dagger - \psi^{A_T B ABR} \otimes \Gamma^N \right\|_1 \leq 2\sqrt{\epsilon}.
$$

(120)

If we apply the isometries $U^A_{jT}$ and $V^B_{jT}$ to the state $| \psi^T_{jT}^{TABR} \rangle$ after obtaining outcome $J$, the output state $\rho$ of the protocol will be of the form

$$
\rho := \sum_{J} p_{J} \left( (I^{TBR} \otimes U^A_{jT} \otimes V^B_{jT}) \psi^T_{jT}^{TABR} (I^{TBR} \otimes U^A_{jT} \otimes V^B_{jT})^\dagger \right)
$$

$$
= \sum_{J} p_{J} \left( \frac{p_{jT}}{p_{jT}} (I \otimes V^B_{jT}) (P^T_{jT} \otimes I) (I \otimes U^A_{jT}) | \varphi^T_{jT}^{TABR} \rangle \langle \varphi^T_{jT}^{TABR} | (I \otimes U^A_{jT})^\dagger (P^T_{jT} \otimes I)^\dagger (I \otimes V^B_{jT})^\dagger \right)
$$

$$
= \sum_{jT} (I \otimes V^B_{jT}) (P^T_{jT} \otimes I) \zeta (P^T_{jT} \otimes I)^\dagger (I \otimes V^B_{jT})^\dagger
$$

$$
=: M(\zeta)
$$

(121)

where $\zeta := \sum_{jT} p_{jT} (I \otimes U^A_{jT}) | \varphi^T_{jT}^{TABR} \rangle \langle \varphi^T_{jT}^{TABR} | (I \otimes U^A_{jT})^\dagger$. It can be seen as the output state we would get if only the helpers in $\mathcal{T}$ wanted to transfer their share of the state to the decoder $A$. The map $\mathcal{M}$, as defined above, corresponds to an LOCC quantum operation acting on $\zeta$ which consists of measurements by the helpers in $\mathcal{T}$ followed by an isometry on $B$. Note that we have removed some of the superscript notation for the sake of clarity.

We would like to bound the trace distance between the output state $\rho$ and the state $\psi^{A_T B ABR} \otimes \Phi^L \otimes \Gamma^N$. To achieve this, we introduce the following intermediate state

$$
\sigma := \sum_{jT} (I \otimes V^B_{jT}) (P^T_{jT} \otimes I) (\psi^{A_T TABR} \otimes \Phi^L) (P^T_{jT} \otimes I)^\dagger (I \otimes V^B_{jT})^\dagger
$$

(122)

and apply the triangle inequality

$$
\left\| \rho - \psi^{A_T TABR} \otimes \Phi^L \otimes \Gamma^N \right\|_1 \leq \left\| \rho - \sigma \right\|_1 + \left\| \sigma - \psi^{A_T TABR} \otimes \Phi^L \otimes \Gamma^N \right\|_1.
$$

(123)

The trace norm $\left\| \sigma - \psi^{A_T TABR} \otimes \Phi^L \otimes \Gamma^N \right\|_1$ is equal to the trace norm appearing in the second line of eq. (120), and so is bounded from above by $2\sqrt{\epsilon}$. To bound $\left\| \rho - \sigma \right\|_1$, we have

$$
\left\| \rho - \sigma \right\|_1 = \left\| M(\zeta) - M(\psi^{A_T TABR} \otimes \Phi^L) \right\|_1
$$

$$
\leq \left\| \zeta - \psi^{A_T TABR} \otimes \Phi^L \right\|_1 \leq 2\sqrt{\epsilon}.
$$

(124)
The first inequality holds since the trace distance is non-increasing under quantum operations, and the second inequality is just the first part of eq. (120). Thus, we have a split-transfer of the state $\psi^{C_1 C_2 \ldots C_m ABR}$ with error $2\sqrt{\epsilon} + 2\sqrt{\epsilon'}$. □

With this result in hand, a one-shot split-transfer protocol of the state $\psi^{TTABR}$ where all the helpers perform simultaneous random measurements on their share can be obtained by two independent applications of Proposition 5 followed by an application of Proposition 15. We state the result here.

**Proposition 16 (One-Shot Split-Transfer)** Let $\psi^{TTABR}$ be an $(m + 2)$-partite pure state, with purifying system $R$ and local dimensions $d_A, d_B, d_R$. Furthermore, let $d_{\mathcal{T}} := \prod_{i \in \mathcal{T}} d_{C_i}$ and $d_{\overline{T}} := \prod_{i \notin \overline{T}} d_{C_i}$ be the dimensions of the systems $\mathcal{T}$ and $\overline{T}$. Finally, allow the helpers to share additional maximally entangled states $\Phi^K$ and $\Gamma^M$ with the decoders.

For each party $C_i$ in the cut $\mathcal{T}$, there exists an instrument $I_i = \{\mathcal{E}_i^1\}_{j=0}^{d_{C_i} K_i}$ consisting of $F_i = \frac{d_{C_i} K_i}{L_i}$ partial isometries of rank $L_i$ and one of rank $L'_i = d_{C_i} K_i - F_i L_i < L_i$ such that the overall quantum error $Q^1_{\mathcal{T}}(\psi^{TTABR} \otimes \Phi^K)$ is bounded by

$$Q^1_{\mathcal{T}}(\psi^{TTABR} \otimes \Phi^K) \leq 2 \sum_{S \subseteq \mathcal{T}} \prod_{i \in S} \frac{L_i}{d_{C_i} K_i} + 2 \sqrt{d_{AB} d_R \sum_{S \subseteq \mathcal{T}} \prod_{i \in S} \frac{L_i}{K_i} \text{Tr}\left((\psi^{TTBR})^2\right)} =: \Delta^1_{\mathcal{T}}. \quad (125)$$

Similarly, for each helper $C_i$ in the cut $\overline{T}$, there exists an instrument $I_i = \{\mathcal{E}_i^2\}_{j=0}^{G_i}$ consisting of $G_i = \frac{d_{C_i} M_i}{N_i}$ partial isometries of rank $N_i$ and one of rank $N'_i = d_{C_i} M_i - G_i N_i < N_i$ such that the overall quantum error $Q^2_{\mathcal{T}}(\psi^{TTABR} \otimes \Gamma^M)$ is bounded by

$$Q^2_{\mathcal{T}}(\psi^{TTABR} \otimes \Gamma^M) \leq 2 \sum_{S \subseteq \mathcal{T}} \prod_{i \in S} \frac{N_i}{d_{C_i} M_i} + 2 \sqrt{d_{AB} d_R \sum_{S \subseteq \mathcal{T}} \prod_{i \in S} \frac{N_i}{M_i} \text{Tr}\left((\psi^{TTBR})^2\right)} =: \Delta^2_{\mathcal{T}}. \quad (126)$$

Then, there exists a split-transfer of the state $\psi^{TTABR}$ with error $2\sqrt{\Delta^1_{\mathcal{T}}} + 2\sqrt{\Delta^2_{\mathcal{T}}}$. The left hand sides of eqs. (125) and (126) are bounded from above on average by their right hand sides if we perform random measurements on all the helpers according to the Haar measure.

**Proof** The bound on the quantum errors $Q^1_{\mathcal{T}}$ and $Q^2_{\mathcal{T}}$ given by eqs. (125) and (126) can be obtained by two independent applications of Proposition 5 to our setting. We leave the details to the reader. Note here that since the helpers have additional entanglement at their disposal, the partial isometries $P_{\mathcal{T}}^{TT}$ and $P_{\overline{T}}^{TT}$ in Proposition 15 are replaced by $P_{\mathcal{T}}^{TT0}$ and $P_{\overline{T}}^{TT0}$. These will act on the spaces $\mathcal{T}T^0$ and $\overline{T}T^0$ respectively, with output spaces corresponding to $\mathcal{T}^1$ and $\overline{T}^1$. □

Similarly, for the i. i. d. version, we can treat each quantum error independently and follow a line of reasoning similar to that in Section VI. We arrive at a variation on Theorem 3.

**Theorem 17 (m-Party Split-Transfer)** Let $\psi^{TTABR}$ be a purified state which is shared between $m$ helpers and two receivers (Alice and Bob), with purifying system $R$. For all non-empty subsets $X \subseteq \mathcal{T}$ and $Y \subseteq \overline{T}$, define $X$ and $Y$ as the tensor products $\bigotimes_{i \in X} C_i$ and $\bigotimes_{i \in Y} C_i$. Then, the rates $R_T(\psi) :=$
\( \bigoplus_{i \in T}(R_i) \) and \( R_T^>(\psi) := \bigoplus_{i \in T}(R_i) \) are achievable for a split-transfer of \( \psi^T ABR \) iff the following inequalities

\[
\begin{align*}
\sum_{i \in X} R_i & \geq S(X|\overline{X}A)_{\psi} \\
\sum_{i \in Y} R_i & \geq S(Y|\overline{Y}B)_{\psi}
\end{align*}
\]  

(127)  
(128)

hold for all non-empty subsets \( X \subseteq T \) and \( Y \subseteq \overline{T} \). The systems \( \overline{X} \) and \( \overline{Y} \) are defined as the complements of \( X \) and \( Y \) with respect to the systems \( T \) and \( \overline{T} \) respectively.

**Proof**  
To prove achievability, we can proceed exactly as in the proof of Theorem 3. That is, we Schumacher compress the state \( (\psi^T ABR)^{\otimes n} \), and then perform random measurements on the helpers with the following bounds on the ranks of the projectors and of the pre-shared entanglement:

\[
0 \leq \prod_{i \in X} \frac{L_i}{K_i} \leq 2^{n(S(X|\overline{X}A)_{\psi} - S(\overline{X}B|\overline{A}^n)_{\psi} - 3\delta|X|)}
\]  

(129)

\[
0 \leq \prod_{i \in Y} \frac{M_i}{N_i} \leq 2^{n(S(Y|\overline{Y}B)_{\psi} - S(\overline{Y}A|\overline{B}^n)_{\psi} - 3\delta|Y|)}
\]  

(130)

for all non-empty subsets \( X \subseteq T \) and \( Y \subseteq \overline{T} \). The bounds on the quantum errors \( Q^1 \) and \( Q^2 \) given in Proposition 16 can then be made arbitrarily small. That is, we will have \( Q^1 \) and \( Q^2 \) bounded from above by \( O(2^{-n\delta^2/2}) \) for some typicality parameter \( \delta \). By applying Proposition 15, we get a split-transfer of the state \( \psi^T ABR \) with error \( O(2^{-n\delta/4}) \) and entanglement costs \( E_T^>(\psi) = \bigoplus_{i \in T}(\log K_i - \log L_i) \) and \( E_T^>(\psi) = \bigoplus_{i \in T}(\log M_i - \log N_i) \). These will satisfy

\[
\begin{align*}
\sum_{i \in X} \frac{1}{n}(\log K_i - \log L_i) & \geq S(X|\overline{X}A) + 3\delta|X| \\
\sum_{i \in Y} \frac{1}{n}(\log M_i - \log N_i) & \geq S(Y|\overline{Y}B) + 3\delta|Y|
\end{align*}
\]  

(131)

for all non-empty subsets \( X \subseteq T \) and \( Y \subseteq \overline{T} \). An application of the gentle measurement lemma and the triangle inequality then tell us that we can apply the same protocol on the state \( (\psi^T ABR)^{\otimes n} \) and obtain a split-transfer with error \( O(2^{-n\delta^2/4}) + O(2^{-2n\delta^2/2}) \). Since this error goes to zero as \( n \) tends to infinity and \( \delta \) was arbitrarily chosen, we get back the direct part of the statement of the theorem.

To get the converse, we can consider any cut \( X \) of the helpers in \( T \) and look at the preservation of the entanglement across the cut \( AX \) vs \( XB \). We assume, for technical reasons, that \( L_i \leq 2^{O(n)} \) for all \( i \in T \). The initial entropy of entanglement across the cut \( AX \) vs \( XB \) is

\[
E_{in} := nS(AX)_{\psi} + \sum_{i \in X} \log K_i.
\]  

(132)

At the end of any LOCC operation on the state \( (\psi^T ABR)^{\otimes n} \), the output state can be seen as an ensemble \( \{q_k, \psi^k_T^1A^1_T A^n_T^1 A^n_T, \psi^k_T B^n_T R^n_T \} \) of pure states. Using monotonicity of the entropy of entanglement under LOCC, we have

\[
nS(AX)_{\psi} + \sum_{i \in X} \log K_i \geq \sum q_k S(AX^1_T A^n_T)_{\psi^k},
\]  

(133)
where $\mathcal{X}^i := \bigotimes_{T \in \mathcal{T}} C^i_T$. For any LOCC operation performing a split-transfer of the state $(\psi^T \mathcal{A}^{ABR})^\otimes n$ with error $\epsilon$, we have

$$\sum_k q_k F^2(\psi^k_{T^1 A_T^1 A^n_T^1 B_T^1 B^n_T R_{T^1 R^n_T}}, \psi_{A_T^1 A_R^1 B_T^1 B_R^1 R_{T^1 R^n_T}}^\otimes n \otimes \Phi^L \otimes \Gamma^N) \geq (1 - \epsilon/2)^2. \quad (134)$$

This follows from the definition of a split-transfer (eq. (112)) and the fact that $F^2$ is linear when one argument is pure. Using Lemma 21, we can rewrite this as

$$\sum_k q_k \left\| \psi^k_{T^1 A_T^1 A^n_T^1 B_T^1 B^n_T R_{T^1 R^n_T}} - \psi_{A_T^1 A_R^1 B_T^1 B_R^1 R_{T^1 R^n_T}}^\otimes n \otimes \Phi^L \otimes \Gamma^N \right\| \leq 2 \sqrt{\epsilon(1 - \epsilon/2)}. \quad (135)$$

By monotonicity of the trace norm under partial tracing, we get

$$\sum_k q_k \left\| \psi^k_{T^1 A_T^1 A^n_T^1} - \psi_{A_T^1}^\otimes n \otimes \bigotimes_{i \in \mathcal{X}} \Phi^L_i \right\| \leq 2 \sqrt{\epsilon(1 - \epsilon/2)}. \quad (136)$$

Using the Fannes inequality (Lemma 22) and the concavity of the $\eta$-function, we have

$$\sum_k q_k \left| S(\mathcal{X}^i A_T^1 A^n_T^1)_{\psi^k} - \sum_{i \in \mathcal{X}} \log L_i - nS(\mathcal{X}^i A_T^1)_{\psi} \right| \leq (2 \sum_{i \in \mathcal{T}} \log L_i + n \log d_A + n \log d_{A_T^1}) \eta(2 \sqrt{\epsilon(1 - \epsilon/2)})$$

$$\leq O(n) \eta(2 \sqrt{\epsilon(1 - \epsilon/2)}). \quad (137)$$

Finally, using eq. (132), we have

$$\sum_{i \in \mathcal{X'}} 1 \left( \log K_i - \log L_i \right) \geq S(\mathcal{X'|A}_T)_{\psi} - O(1) \eta(2 \sqrt{\epsilon(1 - \epsilon/2)}) \quad (138)$$

for any non empty subset $\mathcal{X'} \subseteq \mathcal{T}$. Using a similar argumentation, we can show that

$$\sum_{i \in \mathcal{Y'}} 1 \left( \log M_i - \log N_i \right) \geq S(\mathcal{Y'|B}_T)_{\psi} - O(1) \eta(2 \sqrt{\epsilon(1 - \epsilon/2)}) \quad (139)$$

holds for any non-empty subset $\mathcal{Y'} \subseteq \mathcal{T}$. By letting $n \to \infty$ and $\epsilon \to 0$, we get the converse. \qed

If only a single copy of $\psi^T \mathcal{A}^{ABR}$ is available to the involved parties, we can adapt the argument of Theorem [11] and prove the following result concerning the existence of split-transfer protocols with error $\epsilon$:

**Proposition 18** Given a partition $\mathcal{T} \subseteq \{1, 2, \ldots, m\}$ of the helpers $C_1, C_2, \ldots, C_m$, let $\psi^T \mathcal{T}^{ABR}$ be a $(m+3)$-partite pure state and fix $\epsilon_1, \epsilon_2 > 0$. Then, for any entanglement cost $E_T = \bigoplus_{i \in \mathcal{T}} (\log K_i - \log L_i)$ and $E_T' = \bigoplus_{i \in \mathcal{T}} (\log M_i - \log N_i)$ satisfying

$$\log K_S - \log L_S := \sum_{i \in S} (\log K_i - \log L_i) \geq -H_{\min}(\psi^S \mathcal{T}^{BR} | \psi^T \mathcal{T}^{BR}) + 4 \log \left( \frac{1}{\epsilon_1} \right) + 2|\mathcal{T}| + 8$$

$$\log M_{S'} - \log L_{S'} := \sum_{i \in S'} (\log M_i - \log N_i) \geq -H_{\min}(\psi^{S'} \mathcal{A}_T^{AR} | \psi^{A_T} \mathcal{A}^{AR}) + 4 \log \left( \frac{1}{\epsilon_2} \right) + 2|\mathcal{T}| + 8 \quad (140)$$

for all non-empty subsets $S \subseteq \mathcal{T}$ and $S' \subseteq \mathcal{T}$, there exists a split-transfer protocol acting on $\psi^T \mathcal{T}^{ABR}$ with error $\epsilon_1 + \epsilon_2$. 


Lemma 10 to the state $\psi$. We proceed in a manner analogous to the proof of Theorem 11 and get the following bound on the $C$ for each helper $Q_j$.

For each helper $Q_j$ is a random Haar unitary on $C_j$, $T_{\psi}$. If $F_i, L_i < d_C, K_i$, we also have a partial isometry of rank $L_i' < L_i$. Similarly, for each helper $Q_i$ in $\mathcal{T}$, we have $G_i = \frac{d_C M_i}{N_i}$ random partial isometries $Q_i^j U_i$ of rank $N_i$, and one of rank $N_i'$ if $G_i N_i < d_C M_i$. For a measurement outcome $J := (j_1, j_2, \ldots, j_m)$, let $J_T = \bigoplus_{i \in T} J_i$ be the vector of length $t = |T|$ whose components correspond to the measurement outcomes for the helpers belonging to the cut $T$. The $i$-th element of $J_T$ will be denoted by $j_{T,i}$.

Define

$$\omega_{J_T}^{T \mathcal{R}^T} := (Q_j^T U_T \otimes I_{\mathcal{T}BR}) \psi^{T \mathcal{R}^T} (Q_j^T U_T \otimes I_{\mathcal{T}BR})^\dagger,$$

where $Q_j^T := \bigotimes_{i \in T} Q_j^i$, and the shorthand $U_T$ denotes the tensor product $\bigotimes_{i \in T} U_i$. If we apply Lemma 10 to the state $\psi^{T \mathcal{R}^T} \otimes \tau^K$, where $\tau^K := \bigotimes_{i \in T} \tau^K_i$ and $\mathcal{R}^i := \mathcal{T} \otimes B \otimes R$, we get

$$E \left[ \sum_{j_{T,1}=0}^{F_1} \cdots \sum_{j_{T,t}=0}^{F_t} \left\| \omega_{J_T}^{T \mathcal{R}^T} - \frac{L}{d_C} \tau^{T \mathcal{R}^T} \otimes \psi \right\|_1 \right] \leq \sum_{S \subseteq T} 2^{-H_{\text{min}}(\psi^{S \mathcal{R}^T} | \psi^R) + \log K_S - \log L_S},$$

where $K_S = \prod_{i \in S} K_i$.

Using the hypothesis that $\log K_S - \log L_S \geq -H_{\text{min}}(\psi^{S \mathcal{R}^T} | \psi^R) + 4 \log \left( \frac{1}{c_1} \right) + 2|T| + 8$, we can proceed in a manner analogous to the proof of Theorem 11 and get the following bound on the expectation of the quantum error $Q_i^T (\psi^{T \mathcal{R}^T} \otimes \Phi^K)$:

$$E \left[ \sum_{j_{T,1}=0}^{F_1} \cdots \sum_{j_{T,t}=0}^{F_t} p_{J_T} \left\| \psi_{J_T}^{T \mathcal{R}^T} \otimes \tau^{T \mathcal{R}^T} \psi \right\|_1 \right] \leq 2 \sum_{S \subseteq T} \frac{L_i}{d_C} K_i + E \left[ \sum_{j_{T,1}=0}^{F_1} \cdots \sum_{j_{T,t}=0}^{F_t} p_{J_T} \left\| \psi_{J_T}^{T \mathcal{R}^T} \otimes \tau^{T \mathcal{R}^T} \psi \right\|_1 \right]$$

$$\leq \frac{2e^A H_{\text{min}}(\psi^S)}{22 + 8d_C} + \frac{\epsilon_1^2}{8},$$

where $t = |T|$, $p_{J_T} = \text{Tr}(\omega_{J_T}^{T \mathcal{R}^T} \otimes \psi_{J_T}^{T \mathcal{R}^T} \otimes \psi \psi) = \frac{1}{p_{J_T}} \omega_{J_T}^{T \mathcal{R}^T}$. In a similar way, we can bound the
expected value of the quantum error $Q^2_T$ as follows:

\[
\mathbb{E} \left[ \sum_{j_1=0}^{G_1} \sum_{j_2=0}^{G_2} \cdots \sum_{j_{m-t}=0}^{G_{m-t}} p_j \left\| \psi_A \tau^A_T \psi_B - \tau^B_T \otimes \psi_A AR \right\|_1 \right] \leq \sum_{S' \subseteq S} \frac{2e_2H_{\text{min}}(\psi^{S'})}{2^{2(m-t)+8}d_{C_{S'}}} + \frac{\epsilon_2}{8}
\]

\[
\leq \frac{e_2}{2^{m-t+7}} + \frac{\epsilon_2}{8} \leq \frac{\epsilon_2}{4}.
\]

From Proposition 15 we can conclude that there exists a split-transfer protocol of error $\epsilon_1 + \epsilon_2$. □

With these results in hand, we can now return to our initial motivation, which was that of proving that the min-cut entanglement of the state $\psi^{C_1 C_2 \ldots C_m AB}$ can be preserved by letting all the helpers $C_1, C_2, \ldots, C_m$ perform simultaneous random measurements on their typical subspaces. To prove this fact, we will need the following corollary to Theorem 17.

**Corollary 19** For a pure state $\psi^{C_1 C_2 \ldots C_m AB}$, we denote by $T_{\text{min}}$ a cut of the smallest possible size with the following property:

\[
\forall T \subseteq \{C_1, C_2, \ldots, C_m\} : S(AT_{\text{min}}) \psi \leq S(AT) \psi.
\]

Then, for the state $\psi^{T_{\text{min}} T_{\text{min}} AB}$, the right hand side of eq. (127) will be negative for all nonempty sets $X \subseteq T_{\text{min}}$, while the right hand side of eq. (128) will be non-positive for all nonempty sets $Y \subseteq T_{\text{min}}$.

Furthermore, if we have arbitrarily many copies of the state $\psi^{C_1 C_2 \ldots C_m AB}$ at our disposal, we can perform a split-transfer of the state $\psi^{C_1 C_2 \ldots C_m AB}$ using only local operations and classical communication.

**Proof** For any non-empty subset $X \subseteq T_{\text{min}}$, where $T_{\text{min}}$ is not the empty set, we have

\[
S(X|\overline{T})_\psi = S(T_{\text{min}}A)_\psi - S(XT_{\text{min}}B)_\psi \\
< S(\overline{T}A)_\psi - S(XT_{\text{min}}B)_\psi \\
= S(XT_{\text{min}}B)_\psi - S(XT_{\text{min}}B)_\psi \\
= 0,
\]

where in the second line we have used the fact that $S(AT_{\text{min}}) \psi < S(AT)$ when $T$ is a cut of size smaller than $|T_{\text{min}}|$.

Similarly, for any non-empty subset $Y \subseteq \overline{T}_{\text{min}}$, where $T_{\text{min}}$ is not the whole set $\{C_1, C_2, \ldots, C_m\}$, we have

\[
S(Y|\overline{B})_\psi = S(T_{\text{min}}B)_\psi - S(YT_{\text{min}}A)_\psi \\
= S(T_{\text{min}}A)_\psi - S(YT_{\text{min}}A)_\psi \\
\leq 0
\]

This proves the first part of the corollary.

To get the second part, apply Theorem 17 by setting the Schmidt ranks of the pre-shared maximally entangled states to be $K_i = 1$ for all $i \in T$ and $N_j = 1$ for all $j \in \overline{T}$. Then, for these particular values, eqs. (129) and (130) give us bounds on the ranks $L_i$ and $M_j$ of projectors corresponding to measurements performed by $C_i \in T$ and $C_j \in \overline{T}$ respectively. Since $L_i \geq 1$ and $M_j \geq 1$ must be satisfied for all $i \in T$ and $j \in \overline{T}$, we need the conditional entropies $S(X|\overline{T}A)_\psi$ and $S(Y|\overline{B})_\psi$ appearing in the upper bounds to $\prod_{i \in T} L_i$ and $\prod_{j \in \overline{T}} M_j$ to be negative. Otherwise, the helpers
Let \( m \) \( \geq 2 \) be two recipients: Alice and Bob. Given many copies of \( \psi^{C_1 C_2 \ldots C_m B A} \), if we allow LOCC operations between the helpers and the recipients, the optimal "assisted" EPR rate is given by

\[
E_{\mathcal{A}}^\infty(\psi, A : B) = \min_{\mathcal{T}} \{ S(\mathcal{AT})_\psi \}
\]  

(148)

**Theorem 20 (Multipartite Entanglement of Assistance [4])** Let \( \psi^{C_1 C_2 \ldots C_m AB} \) be a state shared between \( m \) helpers and two recipients: Alice and Bob. Given many copies of \( \psi^{C_1 C_2 \ldots C_m AB} \), if we allow LOCC operations between the helpers and the recipients, the optimal "assisted" EPR rate is given by

\[
E_{\mathcal{A}}^\infty(\psi, A : B) = \min_{\mathcal{T}} \{ S(\mathcal{AT})_\psi \}
\]  

(148)

**Proof** Let \( \mathcal{T}_{\min} \) be a cut of the smallest size attaining the minimization in eq. (148) and fix some \( \epsilon > 0 \). Then, according to Corollary [19] if \( n \) is large enough, we can perform a split-transfer protocol of the state \( \psi^{\mathcal{T}_{\min} \mathcal{T}_{\min} AB} \) with error \( \epsilon \). This will produce a state \( \varphi \) \( \in \mathcal{A}^n \mathcal{T}_{\min}^n \mathcal{B}^n \mathcal{T}_{\min}^n \) such that

\[
\| \varphi^{A^n \mathcal{T}_{\min}^n} - (\psi^{A A \mathcal{T}_{\min}}) \otimes n \|_1 \leq \| \varphi^{A^n \mathcal{T}_{\min}^n} \mathcal{B}^n \mathcal{T}_{\min}^n - (\psi^{A A \mathcal{T}_{\min}} \mathcal{B}^n \mathcal{T}_{\min}^n) \otimes n \|_1 \leq \epsilon,
\]  

(149)

where \( \psi^{A A \mathcal{T}_{\min} \mathcal{B}^n \mathcal{T}_{\min}^n} \) is the original state \( \psi^{\mathcal{T}_{\min} \mathcal{T}_{\min} AB} \) with the systems \( A \mathcal{T}_{\min} \) and \( \mathcal{B} \mathcal{T}_{\min} \) substituted for the systems \( \mathcal{T}_{\min} \) and \( \mathcal{T}_{\min} \). Applying the Fannes inequality to eq. (149), we get

\[
S(A^n \mathcal{T}_{\min}^n)_\varphi - n(S(A A \mathcal{T}_{\min})_\psi) \leq n \log(d_A d_A \mathcal{T}_{\min}) \eta(\epsilon)
\]  

(150)

which implies that

\[
S(A^n \mathcal{T}_{\min}^n)_\varphi = n(S(A A \mathcal{T}_{\min})_\psi \pm \delta) = n(S(A \mathcal{T}_{\min})_\psi \pm \delta),
\]  

(151)

where \( \delta \) can be made arbitrarily small by letting \( \epsilon \to 0 \). Thus, the min-cut entanglement \( E_{\mathcal{A}}^\infty(\psi) \) is arbitrarily well preserved after the split-transfer is performed, and so Alice and Bob can distill at this rate by applying a standard purification protocol on \( \varphi^{A^n \mathcal{T}_{\min}^n} \mathcal{B}^n \mathcal{T}_{\min}^n \). \( \square \)

### IX. DISCUSSION

We have studied the problem of multiparty state merging with an emphasis on how to accomplish merging when the participants have access only to a single copy of a quantum state. In the easier asymptotic i.i.d. setting, the rate region was characterized by a set of "entropic" inequalities which any rate-tuple \( (R_1, R_2, \ldots, R_m) \) must satisfy in order to be achievable for merging. These inequalities define a convex region \( S \) in an \( m \)-dimensional space, whose axes are the individual rates \( R_i \), and where merging can be achieved if the parties have access to many copies of \( \psi^{C_1 C_2 \ldots C_m BR} \). Our protocol for multiparty state merging distinguishes itself in that any point in the rate region can be achieved without the need for time-sharing. The main technical challenge
for showing this was to adapt the decoupling lemma of [4] and the upper bound to the quantum merging error (Proposition 4 in [4]) to the multiparty setting.

The one-shot analysis of the entanglement cost necessary to perform merging presented more difficulties than in the asymptotic setting but as compensation yielded greater rewards. Most notably, because time-sharing is impossible with only a single copy of a quantum state, our intrinsically multiparty protocol provides the first method to interpolate between achievable costs in the multiparty setting. The technical challenge was to derive an upper bound on the quantum error $Q_I(\psi)$ for a random coding strategy in terms of the min-entropies. We suspect that it might possible to further improve our bound by replacing the min-entropies with their smooth variations, but it is unclear how to proceed in order to show this. We leave it as an open problem. To illustrate the advantages of intrinsic multiparty merging over iterated two-party merging, we also performed a detailed analysis of the costs incurred by the two strategies for variants of the embezzling states.

Lastly, we have introduced the split-transfer problem, a variation on the state merging task, and applied it in the context of multiparty assisted distillation. The main technical difficulty here was to prove that the helpers in the cut $T$ do not have to wait for the helpers in $\overline{T}$ to complete their merging with the decoder $A$ before they can proceed with the transfer of their shares to the $B$ decoder. The essential ingredients for showing this were the commutativity of the Kraus operators $P^T_j$ and $P^T_{\overline{T}}$, and the triangle inequality. The rate region for a split-transfer is composed of two sub-regions, each corresponding to rates which would be achievable for a merging operation from $T$ (resp. $\overline{T}$) to $A$ (resp. $B$) with reference $TBR$ (resp. $TAR$).

In the context of assisted distillation, the existence of a split-transfer protocol which redistributes the initial pure state $\psi_{C_1C_2...C_mAB}$ to the decoders $A$ and $B$ was used to give a non-recursive proof that the optimal achievable EPR rate under assistance is given by the min-cut entanglement $\min_{T} \{S(\rho_{AT})\}$. It would be interesting to come up with other potential applications of the split-transfer protocol. State merging was used as a building block for solving various communication tasks, and we believe split-transfer could be useful in other multipartite scenarios than the assisted distillation context. Alternatively, it could also simplify some of the existing protocols which rely on multiple applications of the state merging primitive.

Acknowledgments

The authors would like to thank Jürg Wullschleger for an interesting discussion on the subject of time-sharing and Andreas Winter for discussions on multiparty state transfer. This research was supported by the Canada Research Chairs program, CIFAR, FQRNT, INTRIQ, MITACS, NSERC, ONR grant No. N000140811249 and QuantumWorks.

Appendix A: Miscellaneous Facts

For an operator $X$, the trace norm is defined as:

$$\|X\|_1 := \text{Tr} \sqrt{X^\dagger X}$$
and the trace distance of two states $\rho$ and $\sigma$ is given by $D(\rho, \sigma) = \frac{1}{2} \|\rho - \sigma\|_1$. An alternative measure of closeness of two states is given by the fidelity:

$$F(\rho, \sigma) := \left( \text{Tr} \sqrt{\sqrt{\rho} \sigma \sqrt{\rho}} \right).$$

If the state $\rho := |\psi\rangle \langle \psi|$ is pure, the fidelity between $\rho$ and $\sigma$ becomes equal to:

$$F(|\psi\rangle \langle \psi|, \sigma) = \sqrt{\langle \psi|\sigma|\psi \rangle} = \sqrt{\text{Tr}(\rho|\psi\rangle \langle \psi|)}$$

These two measures of closeness are related as follows:

**Lemma 21** [41] For states $\rho$ and $\sigma$, the trace distance is bounded by

$$1 - F(\rho, \sigma) \leq D(\rho, \sigma) \leq \sqrt{1 - F^2(\rho, \sigma)}.$$

**Lemma 22 (Fannes Inequality)** [42] Let $\rho$ and $\sigma$ be states on a $d$-dimensional Hilbert space, with $\|\rho - \sigma\|_1 \leq \epsilon$. Then $|H(\rho) - H(\sigma)| \leq \eta(\epsilon) \log d$, where $\eta(x) = x - x \log x$ for $x \leq \frac{1}{e}$. When $x > \frac{1}{e}$, we set $\eta(x) = x + \frac{\log x}{e}$.

**Lemma 23 (Gentle Measurement Lemma)** [43] Let $\rho$ be a subnormalized state (i.e. $\rho \geq 0$ and $\text{Tr}[\rho] \leq 1$). For any operator $0 \leq X \leq I$ such that $\text{Tr}[X\rho] \geq 1 - \epsilon$, we have

$$\left\| \sqrt{X}\rho\sqrt{X} - \rho \right\|_1 \leq 2\epsilon$$

**Appendix B: Proof of eq. 38**

**Lemma 24** For $n$ copies of a state $\psi^{C_1C_2...C_mBR}$, let $\Pi_{\tilde{B}}$, $\Pi_{\tilde{C}_1}$, $\Pi_{\tilde{C}_2}$, ..., $\Pi_{\tilde{C}_m}$, $\Pi_{\tilde{R}}$ be the projectors onto the typical subspaces $\tilde{B}, \tilde{C}_1, \tilde{C}_2, \ldots, \tilde{C}_m$ and $\tilde{R}$ respectively. Then, we have

$$\Pi_{\tilde{B}\tilde{C}_m\tilde{R}} := \Pi_{\tilde{B}} \otimes \Pi_{\tilde{C}_1} \otimes \cdots \otimes \Pi_{\tilde{C}_m} \otimes \Pi_{\tilde{R}} \geq \Pi_{\tilde{B}} + \Pi_{\tilde{C}_1} + \cdots + \Pi_{\tilde{C}_m} + \Pi_{\tilde{R}} - (m + 1)I_{\tilde{B}\tilde{C}_m\tilde{R}} \tag{B1}$$

where $\Pi_{\tilde{B}}$ is a shorthand for $\Pi_{\tilde{B}} \otimes I^{C_1C_2...C_mR}$, and similarly for $\Pi_{\tilde{C}_1}, \Pi_{\tilde{C}_2}, \ldots, \Pi_{\tilde{C}_m}$ and $\Pi_{\tilde{R}}$.

**Proof** The projection operators involved in the proof statement pairwise commute, and thus, are simultaneously diagonalizable. Let $\{ |e_i\rangle \}$ be a common eigenbasis for these projectors. Then any eigenvector $|e_i\rangle$ with $\Pi_{\tilde{B}\tilde{C}_m\tilde{R}}|e_i\rangle = |e_i\rangle$ satisfies

$$\left( \Pi_{\tilde{B}} + \Pi_{\tilde{C}_1} + \cdots + \Pi_{\tilde{C}_m} + \Pi_{\tilde{R}} - (m + 1)I \right)|e_i\rangle = |e_i\rangle.$$

If $|e_i\rangle$ is any eigenvector with $\Pi_{\tilde{B}\tilde{C}_m\tilde{R}}|e_i\rangle = 0$, then it must be in the kernel of at least one of the projection operators $\Pi_{\tilde{B}}, \Pi_{\tilde{C}_1}, \Pi_{\tilde{C}_2}, \ldots, \Pi_{\tilde{C}_m}$ and $\Pi_{\tilde{R}}$, which implies that

$$\left( \Pi_{\tilde{B}} + \Pi_{\tilde{C}_1} + \cdots + \Pi_{\tilde{C}_m} + \Pi_{\tilde{R}} - (m + 1)I_{\tilde{B}\tilde{C}_m\tilde{R}} \right)|e_i\rangle = \lambda_i|e_i\rangle,$$

where $\lambda_i \leq 0$. Using both of these observations, we have

$$\Pi_{\tilde{B}\tilde{C}_m\tilde{R}} = \sum_{\Pi_{\tilde{B}\tilde{C}_m\tilde{R}}|e_i\rangle = |e_i\rangle} |e_i\rangle \langle e_i| \geq \sum_{\Pi_{\tilde{B}\tilde{C}_m\tilde{R}}|e_i\rangle = |e_i\rangle} |e_i\rangle \langle e_i| + \sum_{\Pi_{\tilde{B}\tilde{C}_m\tilde{R}}|e_i\rangle = 0} \lambda_i|e_i\rangle \langle e_i| \tag{B2}$$

\[\square\]
Appendix C: Smoothing $H_{\text{max}}$

Lemma 25 Suppose the density operator $\rho$ has eigenvalues $r = (r_1, \ldots, r_d)$ with $r_j \geq r_{j+1}$. Then

$$H_{\text{max}}^\epsilon(\rho) \geq 2\log \left\{ \frac{k}{\epsilon^2} \sum_{j=1}^{k-1} \sqrt{r_j} : k \text{ such that } \sum_{j=k+1}^{d} r_j \leq \frac{\epsilon^2}{2} \right\}. \quad (C1)$$

Proof By Lemma 16 of [18], $H_{\text{max}}^\epsilon(\rho)$ is equal to the minimum of $H_{\text{max}}(\overline{\rho})$ over all positive semidefinite operators $\overline{\rho}$ no more than $\epsilon$ away from $\rho$ as measured by the purified distance. This measure is a bit awkward to work with for our purposes, but it is bounded above by $\sqrt{2}\|\rho - \overline{\rho}\|_1^{1/2}$ by eq. (49). Therefore,

$$H_{\text{max}}^\epsilon(\rho) \geq \min \{ H_{\text{max}}(\overline{\rho}) : \|\rho - \overline{\rho}\|_1 \leq \delta \} \quad (C2)$$

for $\delta = \epsilon^2/2$ and we will try to estimate the right hand side of the inequality.

Let $\overline{\rho}$ be a positive semidefinite operator such that $\|\overline{\rho} - \rho\|_1 \leq \delta$ and let $\overline{\rho} = (\overline{\tau}_1, \ldots, \overline{\tau}_d)$ be the eigenvalues of $\overline{\rho}$, ordered such that $\overline{\tau}_j \geq \overline{\tau}_{j+1}$. We will identify $r$ and $\overline{\tau}$ with their associated diagonal matrices. Then (see [24])

$$\|\overline{\tau} - r\|_1 \leq \|\overline{\rho} - \rho\|_1, \quad (C3)$$

but $H_{\text{max}}(r) = H_{\text{max}}(\rho)$ and $H_{\text{max}}(\overline{\tau}) = H_{\text{max}}(\overline{\rho})$ so we may assume without loss of generality that $\overline{\tau}$ and $\rho$ are simultaneously diagonal with diagonal entries in non-increasing order. We can therefore dispense with $\rho$ and $\overline{\tau}$, discussing only $r$ and $\overline{\rho}$ from now on.

By Theorem 3 of [31], $H_{\text{max}}(r) = 2\log \sum_j \sqrt{r_j}$, which is monotonically decreasing in each $r_j$. This implies that a minimizing $\overline{\tau}$ must satisfy $r_j \geq \overline{\tau}_j$. If not, redefining $\overline{\tau}_j = r_j$ decreases $\|\overline{\tau} - \rho\|_1$ and $H_{\text{max}}(\overline{\tau})$ at the same time.

We will now argue that there is a minimizing $\overline{\tau}$ such that there is a $j_0$ for which $r_j = \overline{\tau}_j$ for all $j < j_0$ and $\overline{\tau}_j = 0$ for all $j > j_0$. Let $s = (s_1, \ldots, s_d)$ be any vector such that $s_j \geq s_{j+1} \geq 0$ and $s_j \leq r_j$, that is, a vector that is a possible candidate for a minimizer. Suppose that $s$ does not have the prescribed form, that is, there is a $j_0$ such that $s_{j_0} < r_{j_0}$ but $s_{j_0+1} \neq 0$. Consider the family of vectors $t(\gamma)$ that arise by transferring $\gamma$ from $s_{j_0+1}$ to $s_{j_0}$ defined by $t(\gamma)_{j_0} = s_{j_0} + \delta$, $t(\gamma)_{j_0+1} = s_{j_0+1} - \delta$ and $t_j = s_j$ for $j \notin \{j_0, j_0 + 1\}$.

It is easy to check that for sufficiently small $\gamma$, it will be the case that $\|r - t(\gamma)\|_1 \leq \|r - s\|_1$. Moreover, defining $f(\gamma) = \sum_j \sqrt{t(\gamma)_j}$, we have that

$$\left. \frac{df}{d\gamma} \right|_{\gamma=0} = \frac{1}{2\sqrt{s_{j_0}}} - \frac{1}{2\sqrt{s_{j_0+1}}}, \quad (C4)$$

which is nonpositive since $s_{j_0} \geq s_{j_0+1}$. For sufficiently small $\gamma$ then, $H_{\text{max}}(t(\gamma)) \leq H_{\text{max}}(s)$. (If $s_{j_0} = 0$, the derivative does not exist but the conclusion can be confirmed by looking at finite differences.) So, if $s$ were a minimizer, it is possible to either construct a new minimizer of the prescribed form or reach a contradiction by further decreasing $H_{\text{max}}$.

The statement of the lemma follows by evaluating $H_{\text{max}}$ on a minimizer of the prescribed form.

\[\Box\]

[1] C. H. Bennett, G. Brassard, C. Crépeau, and R. Jozsa et al. Teleporting an unknown quantum state via dual classical and Einstein-Podolsky-Rosen channels. Physical Review Letters, 70(13):1895–1899, 1993.
[2] C. H. Bennett and S. J. Wiesner. Communication via one- and two-particle operators on Einstein-Podolsky-Rosen states. Physical Review Letters, 69(20):2881–2884, 1992.

[3] J. A. Smolin, F. Verstraete, and A. Winter. Entanglement of assistance and multipartite state distillation. Physical Review A, 72(5):052317, 2005. arXiv:quant-ph/0505038v1.

[4] M. Horodecki, J. Oppenheim, and A. Winter. Quantum state merging and negative information. Communication in Mathematical Physics, 269(1):107–136, 2007. arXiv:quant-ph/0512247.

[5] A. Abeyesinghe, I. Devetak, P. Hayden, and A. Winter. The mother of all protocols: restructuring quantum information’s family tree. Proceedings of the Royal Society A, 465:2537–2563, 2009.

[6] I. Devetak and A. Winter. Distillation of secret key and entanglement from quantum states. Proceedings of the Royal Society A, 461:207–235, 2005. arXiv:quant-ph/0306078v1.

[7] I. Devetak and J. Yard. The operational meaning of quantum conditional information. 2006. arXiv:quant-ph/0612050v1.

[8] J. Yard and I. Devetak. Optimal quantum source coding with quantum side information at the encoder and decoder. IEEE Transactions on Information Theory, 55(11):5339–5351, 2009.

[9] G. Smith and J. Yard. Quantum communication with zero-capacity channels. Science, 321(5897):1812–1815, 2008. arXiv:quant-ph/0807.4935v2.

[10] M. B. Hastings. Superadditivity of communication capacity using entangled inputs. Nature Physics, 5(4):255–257, 2009.

[11] B. Schumacher and M. A. Nielsen. Quantum data processing and error correction. Physical Review A, 54(4):2629–2635, 1996.

[12] N. J. Cerf and C. Adami. Negative entropy and information in quantum mechanics. Physical Review Letters, 79(26):5194–5197, 1997.

[13] M. Horodecki, J. Oppenheim, and A. Winter. Quantum information can be negative. Nature, 436:673–676, 2005. arXiv:quant-ph/0505062.

[14] C. Ahn, A. Doherty, P. Hayden, and A. Winter. On the distributed compression of quantum information. IEEE Transactions on Information Theory, 52(10):4349–4357, 2006.

[15] D. P. DiVincenzo, C. A. Fuchs, H. Mabuchi, and J. A. Smolin et al. Entanglement of assistance. In Quantum Computing and Quantum Communications First NASA International Conference, QCQC98 Palm Springs, California, USA February 17-20, 1998 Selected Papers, volume 1509 of Lecture Notes in Computer Science, pages 247–257. Springer Berlin, 1999. arXiv:quant-ph/9803033v1, 1998.

[16] M. Berta. Single-shot quantum state merging. Master’s thesis, ETH Zürich, 2009. arXiv:quant-ph/0912.4495v1.

[17] R. Renner. Security of quantum key distribution. PhD thesis, ETH Zürich, 2005. arXiv:quant-ph/0512258.

[18] M. Tomamichel, R. Colbeck, and R. Renner. Duality between smooth min- and max-entropies. IEEE Transactions on Information Theory, 56(9):4674–4681, 2010. arXiv:quant-ph/0907.5238v2.

[19] M. Berta, M. Christandl, and R. Renner. A conceptually simple proof of the quantum reverse shannon theorem. 2009. arXiv:quant-ph/0912.3805v1.

[20] F. Buscemi and N. Datta. How many singlets are needed to create a bipartite state using locc? 2009. arXiv:0906.3698v2.

[21] F. Buscemi and N. Datta. General theory of assisted entanglement distillation. 2010. arXiv:quant-ph/1009.4464v1.

[22] J. M. Renes and R. Renner. One-shot classical data compression with quantum side information and the distillation of common randomness or secret keys. 2010. arXiv:1008.0452v2.

[23] A. Uhlmann. The ‘transition probability’ in the state space of a *-algebra. Reports in Mathematical Physics, 9:273, 1976.

[24] M. Nielsen and I. L. Chuang. Quantum computation and quantum information. Cambridge University Press, 2001.

[25] A. Winter. Coding theorems of quantum information theory. PhD thesis, University of Bielefeld, 1999. arXiv:quant-ph/9907077v1.

[26] A. Rényi. On measures of entropy and information. Proceedings of the 4th Berkeley Symposium on Mathematical Statistics and Probability, 1:547–561, 1960.

[27] Christian Cachin. Smooth entropy and Rényi entropy. In Advances In Cryptology - EUROCRYPT '97, Lecture Notes in Computer Science, pages 193–208. SpringerVerlag, 1997.

[28] Renato Renner and Stefan Wolf. Simple and tight bounds for information reconciliation and privacy amplification. In Advances in Cryptology - ASIACRYPT 2005, Lecture Notes in Computer Science, pages
41

... SpringerVerlag, 2005.

[29] S. Baratpour, J. Ahmadi, and N. R. Arghami. Characterizations based on Rényi entropy of order statistics and record values. Journal of Statistical Planning and Inference, 138(8):2544–2551, 2008.

[30] M. M. Mayoral. Renyi’s entropy as an index of diversity in simple-stage cluster sampling. Information Sciences, 105:101–114, 1998.

[31] R. Koenig, R. Renner, and C. Schaffner. The operational meaning of conditional min- and max-entropy. IEEE Transactions on Information Theory, 55(9):4337–4347, 2009.

[32] W. van Dam and P. Hayden. Universal entanglement transformations without communication. Physical Review A, 67(6):060302, 2003. arXiv:quant-ph/0201041v1.

[33] A. W. Harrow. Entanglement spread and clean resource inequalities. In XVITH Internation Congress on Mathematical Physics, pages 536–540. World Scientific, 2009. arXiv:quant-ph/0909.1557.

[34] C. H. Bennett, I. Devetak, A. W. Harrow, P. W. Shor, and A. Winter. Quantum reverse shannon theorem. 2009. arXiv:quant-ph/0912.5537v1.

[35] R. A. Horn and C. R. Johnson. Matrix Analysis. Cambridge University Press, 1990.

[36] Wikipedia. Gershgorin circle theorem— Wikipedia, the free encyclopedia, 2010. [Online; accessed 28-October-2010].

[37] P. Hayden and A. Winter. Communication cost of entanglement transformations. Physical Review A, 67(1):012326, 2003. arXiv:quant-ph/0204092v3.

[38] Charles H. Bennett, Herbert J. Bernstein, Sandu Popescu, and Benjamin Schumacher. Concentrating partial entanglement by local operations. Phys. Rev. A, 53(4):2046–2052, 1996.

[39] C. H. Bennett, D. P. DiVincenzo, J. A. Smolin, and W. K. Wootters. Mixed-state entanglement and quantum error correction. Physical Review A, 54(5):3824–3851, 1996. arXiv:quant-ph/9604024v2.

[40] J. A. Smolin and A. V. Thapliyal. The power of loccq state transformations. 2002. arXiv:quant-ph/0212098.

[41] C. A. Fuchs and J. van de Graaf. Cryptographic distinguishability measures for quantum mechanical states. IEEE Transactions on Information Theory, 45:1216–1227, 1999. arXiv:quant-ph/9712042v2.

[42] M. Fannes. A continuity property of the entropy density for spin lattice systems. Communication in Mathematical Physics, 31:291–294, 1973.

[43] A. Winter. Coding theorem and strong converse for quantum channels. IEEE Transactions on Information Theory, 45(7):2481–2485, 1999.