Unification Of Quantum Information Theory

Thesis by

Anura Abeyesinghe

In Partial Fulfillment of the Requirements

for the Degree of

Doctor of Philosophy

California Institute of Technology

Pasadena, California

2006

(Defended May 23, 2006)

Acknowledgements

The work in this dissertation is the result of many fruitful collaborations. The first two chapters are the end result of many research projects done under the mentorship of Patrick Hayden, to whom I owe many thanks. The final chapter is the result of work done under the mentorship of Igor Devetak, also to whom I owe many thanks. I am also very grateful to have had the pleasure to work with my other two co-authors – Andreas Winter and Graeme Smith. Last but not least I owe many thanks to my advisor, John Preskill, for allowing me the flexibility to work at my own pace and choose projects that excite me.

Abstract

We present the unification of many previously disparate results in noisy quantum Shannon theory and the unification of all of noiseless quantum Shannon theory. More specifically we deal here with bipartite, unidirectional, and memoryless quantum Shannon theory. We find all the optimal protocols and quantify the relationship between the resources used, both for the one-shot and for the ensemble case, for what is arguably the most fundamental task in quantum information theory: sharing entangled states between a sender and a receiver. We find that all of these protocols are derived from our one-shot superdense coding protocol and relate nicely to each other. We then move on to noisy quantum information theory and give a simple, direct proof of the "mother" protocol, or rather her generalization to the Fully Quantum Slepian-Wolf protocol(FQSW). FQSW simultaneously accomplishes two goals: quantum communication-assisted entanglement distillation, and state transfer from the sender to the receiver. As a result, in addition to her other "children," the mother protocol generates the state merging primitive of Horodecki, Oppenheim, and Winter as well as a new class of distributed compression protocols for correlated quantum sources, which are optimal for sources described by separable density operators. Moreover, the mother protocol described here is easily transformed into the so-called "father" protocol, demonstrating that the division of single-sender/single-receiver protocols into two families was unnecessary: all protocols in the family are children of the mother.

Table of Contents

List of Figures

Unification of noiseless quantum information theory

Quantum information theory can be described as the effort to identify and quantify the basic resources required to communicate or, more generally, process information in a quantum mechanical setting. The dual goals of identifying new protocols and demonstrating their optimality have, respectively, helped to expose the surprising range of information processing tasks facilitated by quantum mechanics and highlighted the subtle ways in which physics dictates limitations on the transmission and processing of information.

Part of the appeal of the information theoretic paradigm is that it emphasizes the notions of interconvertibility and simulation. Identifying basic resources and evaluating their interconvertibility provides a general strategy for systematically charting the capabilities of quantum mechanical systems. Some early successes of this approach include Schumacher's quantum noiseless coding theorem [37, 38], which demonstrated that a single number quantifies the compressibility of memoryless sources of quantum states, and the theory of pure state bipartite entanglement, where a single number, likewise, determines the asymptotic interconvertibility of entanglement [39]. The last ten years have seen major advances in the area, including, among many other discoveries, the determination of the classical capacity of a quantum channel [62, 63], the capacities

of entanglement-assisted channels [64, 65], the quantum capacity of a quantum channel [66, 67, 68], and the best ways to use noisy entanglement to extract pure entanglement [69] or send classical information [70].

From the point of view of communication theory, these results identify three basic and inequivalent resources: noiseless classical channels, noiseless quantum channels and maximally entangled states. Other inequivalent resources exist, of course. One such, classically correlated bits, will prove useless for the problems we investigate. Noisy versions of the basic list of three resources identified above adds many others but we do not study them here. They will be the subject of the next section. Those caveats aside, the three basic resources serve as formalized versions of abstract "classicality," "quantumness" and "nonlocality," quantifiable in units of classical bits (cbits), quantum bits (qubits) and maximally entangled qubits (ebits). While the three basic resources are inequivalent, relationships exist between them. Because cbits can be encoded in qubits and ebits can be established by sending qubits, the noiseless quantum channel is (in this narrow sense) the strongest of the three. Because it is impossible to establish entanglement using classical communication or to communicate using only entanglement, ebits and cbits are simply incomparable; neither is truly stronger than the other.

In this section, we quantify the relationship between the three resources for a basic task in quantum information theory: communicating quantum states from a sender to a receiver (and, more generally, sharing entangled states between them). There are at least two variations on the task, depending on whether or not the sender has knowledge of the states she is required to communicate. If she is only given a copy of the quantum state and not a description, we describe the source as hidden and the encoding as oblivious (or blind). At the other extreme, if she is told which state she is required to transmit, we describe the source as visible and the encoding as non-oblivious. (Sometimes in the quantum information literature the adjective "visible" is also applied, somewhat

nonsensically, to the encoding.) While the distinction makes no difference in classical information theory, quantum mechanical restrictions on the sender's ability to measure without causing a disturbance lead to very different results for the two tasks in the quantum case. (Compare, for example, the results of Refs. [42, 43] and [44].) Our emphasis here is on the visible scenario since there is generically only a trivial trade-off for the blind encoder case: using teleportation, two cbits and one ebit can be used to simulate a noiseless one-qubit channel but no other interesting trade-offs are possible.

In the visible scenario, the relationship between the three resources becomes much more varied. When no quantum channel is permitted, we recover the problem known as remote state preparation, which was solved for the ensemble case in [45, 46], while forbidding use of the classical channel leads to superdense coding of quantum states, which was solved for the ensemble case in [47, 48]. Likewise, if entanglement is not permitted, we recover the trade-off between classical and quantum communication solved in Ref. [44]. This section completely solves the problem of trading all three resources against each other, finding that optimal protocols for any combination of resources can be constructed by appropriate combinations of the protocols representing the extremes identified above, which in turn are derived from one unifying protocol: the one-shot superdense coding protocol presented in the first chapter.

Chapter 1

Optimal superdense coding of entangled states

1.1 Introduction

A sender's power to communicate with a receiver is frequently enhanced if the two parties share entanglement. The most well-known example of this phenomenon is perhaps superdense coding [47], the communication of two classical bits of information by the transmission of one quantum bit and consumption of one ebit. If the sender knows the identity of the state to be sent, superdense coding of *quantum states* also becomes possible, with the result that, asymptotically, two qubits can be communicated by physically transmitting one qubit and consuming one bit of entanglement [48, 18]. In [48] it was furthermore shown that a sender (Alice) can asymptotically share a two-qubit entangled state with a receiver (Bob) at the same qubit and ebit rate, along with the consumption of some shared randomness. That result, however, failed to exploit one of the most basic observations about superdense coding: highly entangled states are much *easier* to prepare than non-entangled states. Indeed, maximally entangled states can be prepared with no communication from the sender at all.

In this chapter, we construct a family of protocols that take advantage of this effect, finding that even partial entanglement in the state to be shared translates directly into a reduction in the amount of communication required. Recall that every bipartite pure state can be written in the form $|\varphi_{AB}\rangle = \sum_i \sqrt{\lambda_i}|e_i\rangle|f_i\rangle$, where $\langle e_i|e_j\rangle = \langle f_i|f_j\rangle = \delta_{ij}$ and $\lambda_i \geq 0$ [25]. Since the numbers $\sqrt{\lambda_i}$, known as *Schmidt coefficients*, are the only local invariants of $|\varphi_{AB}\rangle$, they entirely determine the nonlocal features of the state. In the case of one-shot superdense coding, we find that it is the largest Schmidt coefficient that plays a crucial role. More specifically, we show how Alice can share with Bob any pure state that has reduction on Bob's system of dimension d_S and maximum Schmidt coefficient $\sqrt{\lambda_{\max}}$ by transmitting roughly $\frac{1}{2}\log d_S + \frac{1}{2}\log \lambda_{\max}$ qubits and consuming $\frac{1}{2}\log d_S - \frac{1}{2}\log \lambda_{\max}$ ebits. We also show that these rates are essentially optimal.

In the spirit of [11], this new protocol can be viewed as the "father" of the noiseless, visible state communication protocols. Composing it with teleportation generates an optimal *remote state preparation* [21, 46] protocol. Applying it to the preparation of states drawn from a memoryless source generates all the optimal rate points of the triple cbit-qubit-ebit trade-off studied in [1], when combined with quantum-classical trade-off coding [56, 44]. An inspiration for the present work was Harrow's alternative construction of optimal protocols in this memoryless setting that made use of coherent classical communication [71] and pre-existing remote state preparation protocols [50]. Harrow's techniques provided strong circumstantial evidence that the protocol we present here should exist.

The rest of this chapter is structured as follows: We begin, in Section 3.2, by presenting the universal protocol for superdense coding of entangled states and then prove its optimality, along with that of the associated remote state preparation protocol, in Section 1.3. Section 1.4 contains an easy application of typical subspace techniques to the task of developing an optimal protocol for preparing states generated by a memoryless

source. Section 1.5 provides another application of the protocol, this time to the theory of identification [29, 2]. Specifically, we show that the quantum identification capacity of an ebit is two qubits.

Notation: We use the following conventions throughout the paper: log and exp are always taken base 2. Unless otherwise stated, a "state" can be pure or mixed. The density operator $|\varphi\rangle\langle\varphi|$ of the pure state $|\varphi\rangle$ will frequently be written simply as φ. If φ_{AB} is a state on $A \otimes B$, we refer to the reduced state on A as φ_A. Sometimes we omit subscripts labelling subsystems, in which case the largest subsystem on which the state has been defined should be assumed: $\varphi = \varphi_{AB}$ in the bipartite system $A \otimes B$, for example. A system we call A will have a Hilbert space also called A with a dimension d_A. $\mathbb{U}(d)$ denotes the unitary group on $\mathbb{C}^d$, and $\mathcal{B}(\mathbb{C}^d)$ the set of linear transformations from $\mathbb{C}^d$ to itself. We write the fidelity between two states ρ and σ as $F(\rho, \sigma) = \|\sqrt{\rho}\sqrt{\sigma}\|_1^2$ and the von Neumann entropy of a state ρ as $S(\rho) = -\operatorname{Tr} \rho \log \rho$.

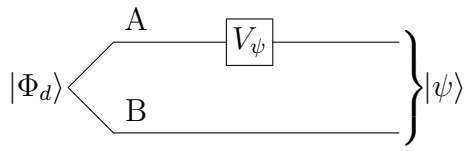
1.2 The universal protocol

To begin, suppose that Alice would like to share a maximally entangled state with Bob. Clearly, this can be accomplished without any communication – Alice need only perform operations on her half of a fixed maximally entangled state shared between them. In particular, if $|\psi\rangle$ is an arbitrary maximally entangled state and we denote by $|\Phi_d\rangle = \frac{1}{\sqrt{d}} \sum_{i=1}^{d} |i\rangle|i\rangle$ a fixed maximally entangled state, then $|\psi\rangle$ can be expressed as

$$|\psi\rangle = V_\psi \otimes \mathbb{1}_B |\Phi_d\rangle, \tag{1.1}$$

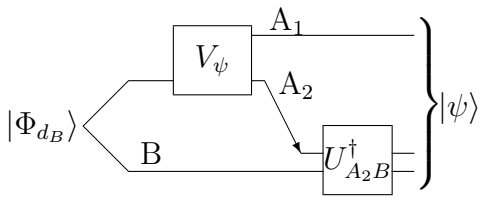
where V_ψ is a unitary transformation of Alice's system that depends on ψ. This identity is equivalent to the circuit diagram (1.2), in which time runs from left to right:

Of course, in general, we would like to prepare an arbitrary state $|\psi_{AB}\rangle$ that may *not* be maximally entangled, and to do so by using as few resources as possible. Our



(1.2)

Figure 1.1:



(1.3)

Figure 1.2:

general method is as follows: Alice and Bob initially share a fixed maximally entangled state $|\Phi_{d_B}\rangle$, to which Alice applies an isometry V_ψ. She then sends a subsystem A_2 of dimension d_{A_2} to Bob, who will apply a fixed unitary $U^\dagger_{A_2B}$. Alice's goal is to make d_{A_2} as small as possible while still reliably preparing $|\psi_{AB}\rangle$. The procedure can again be summarized with a circuit diagram, although this time it is much less clear whether there exist choices of the operations V_ψ and U_{A_2B} that will do the job:

Figure (1.3) does provide a method for preparing the state $|\psi\rangle$ as long as $\mathbb{1} \otimes U_{A_2B}|\psi\rangle$ is maximally entangled across the $A_1A_2|B$ cut. (All such states are related by an operation on Alice's system alone.) We will now use this observation, together with the fact that high-dimensional states are generically highly entangled [22, 20, 27, 13, 30, 31], to construct a protocol that prepares an arbitrary state with high fidelity. The precise statement about the entanglement of generic states that we will need is the following lemma.

Lemma 1.2.1 *Let φ be a state on $A \otimes B$ proportional to a projector of rank r and let*

$U_{AB} \in \mathbb{U}(d_A d_B)$ be chosen according to the Haar measure. Then, if $d_B \leq d_A$,

$$\begin{aligned} &\Pr\left(S(\mathrm{Tr}_A\, U_{AB}\varphi U_{AB}^\dagger) < \log d_B - \alpha - \beta\right) \\ &\qquad \leq 12r \exp\left(-r d_A d_B \frac{\alpha^2 C}{(\log d_B)^2}\right), \end{aligned} \tag{1.4}$$

where we may choose $C = (8\pi^2 \ln 2)^{-1}$, and $\beta = \frac{1}{2\ln 2}\frac{d_B}{r d_A}$.

It generalizes the following lemma for rank-one φ, which was proved in [18].

Lemma 1.2.2 *Let $|\varphi\rangle$ be chosen according to the Haar measure on $A \otimes B$. Then, if $3 \leq d_B \leq d_A$,*

$$\begin{aligned} &\Pr\left(S(\varphi_B) < \log d_B - \alpha - \beta\right) \\ &\qquad \leq \exp\left(-\left(d_B d_A - 1\right)\frac{\alpha^2 C}{(\log d_B)^2}\right), \end{aligned} \tag{1.5}$$

where $C = (8\pi^2 \ln 2)^{-1}$ as before and $\beta = \frac{1}{\ln 2}\frac{d_B}{d_A}$. □

Proof (of Lemma 1.2.1) If we let R be a space of dimension r and $|\tau_{ABR}\rangle$ be a uniformly distributed state on $A \otimes B \otimes R$, then $\frac{1}{r}\Pi_{\tau_{AB}}$ is equal in distribution to $U_{AB}\varphi U_{AB}^\dagger$, where $\Pi_{\tau_{AB}}$ is the projector onto the support of τ_{AB}. Let σ denote the unitary transformation $\sum_{j=0}^{r-1} |e_{(j+1 \bmod r)}\rangle\langle e_j| + \mathbb{1} - \Pi_{\tau_{AB}}$ that implements a cyclic permutation on the eigenvectors $\{|e_j\rangle\}$ of τ_{AB} corresponding to non-zero eigenvalues. (There are r such eigenvalues except on a set of measure zero, which we will ignore.) We then have

$$\frac{1}{r}\Pi_{\tau_{AB}} = \frac{1}{r}\sum_{k=0}^{r-1} \sigma^k \tau_{AB} \sigma^{-k}. \tag{1.6}$$

Eq. (1.6), together with the concavity of entropy, implies

$$S\left(\frac{1}{r}\,\mathrm{Tr}_A\, \Pi_{\tau_{AB}}\right) \geq \frac{1}{r}\sum_{k=1}^{r} S(\mathrm{Tr}_A\, \sigma^k \tau_{AB} \sigma^{-k}), \tag{1.7}$$

which in turn gives

$$\begin{aligned}
&\Pr\Big(S(\mathrm{Tr}_A\, U_{AB}\varphi U_{AB}^\dagger) < \log d_B - \alpha - \beta\Big) \\
&\leq \Pr\left(\frac{1}{r}\sum_{k=1}^{r} S(\mathrm{Tr}_A\, \sigma^k \tau_{AB}\sigma^{-k}) < \log d_B - \alpha - \beta\right) \\
&\leq r\Pr\big(S(\mathrm{Tr}_A\, \sigma^k \tau_{AB}\sigma^{-k}) < \log d_B - \alpha - \beta\big) \\
&= r\Pr\big(S(\tau_B) < \log d_B - \alpha - \beta\big), \qquad (1.8)
\end{aligned}$$

where the final step is a result of the unitary invariance of τ_{ABR}. Applying Lemma 1.2.2 to Eq. (1.8) with $A \to AR$ and $B \to B$ reveals that

$$\begin{aligned}
&\Pr\Big(S(\mathrm{Tr}_A\, U_{AB}\varphi U_{AB}^\dagger) < \log d_B - \alpha - \beta\Big) \\
&\qquad \leq 12r\exp\Big(-rd_Bd_A\frac{\alpha^2 C}{4(\log d_B)^2}\Big). \qquad (1.9)
\end{aligned}$$

□

The idea behind the protocol is then simple: We will show that there exists a single unitary U_{A_2B} such that $\mathbb{1}_{A_1} \otimes U_{A_2B}|\psi_{A_1A_2B}\rangle$ is almost maximally entangled across the $A_1A_2|B$ cut for *all* states $|\psi_{A_1A_2B}\rangle$ satisfying a bound on their Schmidt coefficients and whose support on $A_2 \otimes B$ lies in a large subspace $S \subset A_2 \otimes B$. Since any such $\mathbb{1}_{A_1} \otimes U_{A_2B}|\psi_{A_1A_2B}\rangle$ is almost maximally entangled, we can then find an *exactly* maximally entangled state that closely approximates it. This state, in turn, can be prepared by the method of Eq. (**??**). More formally, the following general prescription can be made to succeed:

Protocol: To send an arbitrary pure state with maximal Schmidt coefficient $\leq \sqrt{\lambda_{\max}}$ and reduction of Bob's system to dimension d_S.

1. Alice and Bob share a maximally entangled state of $\log d_B = \frac{1}{2}(\log d_S - \log\lambda_{\max}) + o(\log d_S)$ ebits on their joint system AB.

2. Alice applies a local partial isometry V_ψ with output on two subsystems A_1 and A_2. The size of A_2 is $\log d_{A_2} = \frac{1}{2}(\log d_S + \log\lambda_{\max}) + o(\log d_S)$.

3. Alice sends A_2 to Bob.

4. Bob applies $U^\dagger_{A_2B}$ followed by a projection onto S, which is embedded as a subspace of A_2B.

Proposition 1.2.3 *Let $0 < \kappa \leq 1$. For sufficiently large d_S,and for d_{A_2} and d_B as defined in the protocol, there exists choices of V_ψ that depend on the input state $|\psi_{A_1S}\rangle$ and U_{A_2B} such that for all input states $|\psi_{A_1S}\rangle$ with largest Schmidt coefficient $\leq \sqrt{\lambda_{\max}}$, the output of the protocol has fidelity at least $1-\kappa$ with $|\psi_{A_1S}\rangle$.*

Proof Our method will be to show that if U_{A_2B} is chosen according to the Haar measure, then the corresponding protocol has a non-zero probability over choices of U_{A_2B} of achieving high fidelity for all states that satisfy the restriction on their Schmidt coefficients, establishing the existence of a particular U_{A_2B} for which this is true.

Now, to ensure that the protocol succeeds on a given $|\psi_{A_1S}\rangle$, we only need to ensure that $\mathbb{1}_{A_1} \otimes U_{A_2B}|\psi_{A_1S}\rangle$ is highly entangled across the $A_1A_2|B$ cut, which amounts to showing that $S(\mathrm{Tr}_{A_2} U_{A_2B}\psi_S U^\dagger_{A_2B})$ is close to $\log d_B$. This is exactly what Lemma 1.2.1 tells us is overwhelmingly likely for an individual random state $|\varphi_{A_1S}\rangle$ maximally entangled with a subspace A_1' of A_1. By standard arguments, this will ensure that there exists a unitary U_{A_2B} such that $S(\mathrm{Tr}_{A_2} U_{A_2B}\varphi_S U^\dagger_{A_2B})$ is close to $\log d_B$ for *all* the states on S maximally entangled with A_1'. Majorization can then be used to extend the argument to general states $|\psi_{A_1S}\rangle$ with bounded largest Schmidt coefficient.

We begin by restricting to the case of states $|\varphi_{A_1'S}\rangle$ maximally entangled between S and a fixed subspace $A_1' \subseteq A_1$, with $d_{A_1'} = \lfloor 1/\lambda_{\max} \rfloor$. Now, let $\mathcal{N}^\gamma_{A_1'S}$ be a trace norm γ-net for such states. It is possible to choose $|\mathcal{N}^\gamma_{A_1'S}| \leq (5/\gamma)^{2d_{A_1'}d_S}$. (See, for example, [17]. We will fix γ later.) By the definition of the net and the contractivity of the trace norm under the partial trace, for every maximally entangled state $|\varphi\rangle$ on $A_1' \otimes S$ there is

a state $|\tilde{\varphi}\rangle \in \mathcal{N}^{\gamma}_{A_1'S}$ such that

$$\left\| \mathrm{Tr}_{A_2}(U_{A_2B}\varphi_S U^\dagger_{A_2B}) - \mathrm{Tr}_{A_2}(U_{A_2B}\tilde{\varphi}_S U^\dagger_{A_2B}) \right\|_1 \leq \|\varphi - \tilde{\varphi}\|_1 \leq \gamma, \tag{1.10}$$

which, by the Fannes inequality [59], implies that

$$\left| S\big(\mathrm{Tr}_{A_2}(U_{A_2B}\varphi_S U^\dagger_{A_2B})\big) - S\big(\mathrm{Tr}_{A_2}(U_{A_2B}\tilde{\varphi}_S U^\dagger_{A_2B})\big) \right| \leq \delta + \eta(\gamma), \tag{1.11}$$

where $\delta = \gamma \log d_B$ and $\eta(t) = -t \log t$ for $\gamma \leq 1/4$. Noting that all the states $|\psi_{A_1S}\rangle$ have the same reduction on Bob, we have

$$\begin{aligned} &\Pr\Big(\inf_{|\varphi_{A_1'S}\rangle} S(\mathrm{Tr}_{A_2}\, U_{A_2B}\varphi_S U^\dagger_{A_2B}) \\ &\qquad\qquad < \log d_B - \alpha - \beta - \delta - \eta(\gamma) \Big) \qquad (1.12) \\ &\leq \ |\mathcal{N}^{\gamma}_{A_1'S}| \Pr\Big(S(\mathrm{Tr}_{A_2}\, U_{A_2B}\varphi_S U^\dagger_{A_2B}) < \log d_B - \alpha - \beta \Big) \\ &\leq \ \left(\frac{5}{\gamma}\right)^{2d_{A_1'}d_S} 4d_{A_1'} \exp\left(-d_{A_1'}d_{A_2}d_B \frac{\alpha^2 C}{4(\log d_B)^2} \right), \qquad (1.13) \end{aligned}$$

where $\beta = d_B/(2\ln 2 d_{A_1'}d_{A_2})$. Choosing $\alpha = \beta =: \epsilon/4 \leq 1/4$, $\gamma = \alpha^2/(4\log d_B)$ and

$$d_S < d_{A_2}d_B \frac{\alpha^2 C}{8(\log d_B)^2 \log(20 \log d_B/\alpha^2)} - 1,$$

we find that the probability bound (1.13) is less than 1. For our choice of parameters, we have furthermore $\alpha + \beta + \delta + \eta(\gamma) \leq 4\alpha = \epsilon$, using $\eta(x) \leq 2\sqrt{x}$ for $x \leq 1/4$.We have chosen parameters such that $d_{A_2} = d_B/(2\ln 2\, \alpha d_{A_1'})$.

Moreover, relaxing the restriction on the input states now, suppose that $|\psi\rangle$ is any state on $A_1 \otimes S$ satisfying the condition $\|\psi_S\|_\infty \leq \lambda_{\max}$. Then any such ψ_S is majorized by any φ_S maximally entangled with A_1', so that ψ_S can be written as a convex combination $\sum_j p_j W_j \varphi_S W_j^\dagger$, where each W_j is unitary [4]. It then follows from the concavity of the

entropy that

$$S\big(\operatorname{Tr}_{A_2} U_{A_2B}\psi_S U_{A_2B}^\dagger\big) \\ \geq \min_j S\big(\operatorname{Tr}_{A_2} U_{A_2B} W_j \varphi_S W_j^\dagger U_{A_2B}^\dagger\big). \tag{1.14}$$

Therefore, the probability of Eq. (1.12) is actually an upper bound for

$$\Pr\Big(\inf_{|\psi_{A_1S}\rangle} S(\operatorname{Tr}_{A_2} U_{A_2B}\psi_S U_{A_2B}^\dagger) < \log d_B - \epsilon\Big). \tag{1.15}$$

Thus, with our choice of parameters, there is a unitary U_{A_2B} such that for all states $|\psi\rangle$ on $A_1 \otimes S$ satisfying the requirement that $\operatorname{Tr}_{A_1}\psi$ have eigenvalues $\leq \lambda_{\max}$, we have

$$S(\psi'_B) \geq \log d_B - \epsilon, \tag{1.16}$$

introducing $|\psi'\rangle = (\mathbb{1} \otimes U_{A_2B})|\psi\rangle$. Since this can be rewritten as $S(\psi'_B \| \mathbb{1}_B/d_B) = \log d_B - S(\psi'_B) \leq \epsilon$, it in turn implies [26] that, for such states,

$$\|\psi'_B - \mathbb{1}_B/d_B\|_1 \leq \sqrt{2\ln 2\epsilon} =: \kappa, \tag{1.17}$$

and, therefore, that $F(\psi'_B, \mathbb{1}_B/d_B) \geq 1-\kappa$. By Uhlmann's theorem [32, 19], there exists a purification Φ_ψ of $\mathbb{1}/d_B$ such that $|\langle\psi'|\Phi_\psi\rangle|^2 \geq 1-\kappa$. Starting from a fixed maximally entangled state $|\Phi_0\rangle$, $|\Phi_\psi\rangle$ can be prepared by Alice using a local operation V_ψ on A_1A_2 alone. Sending the system A_2 to Bob and having him perform $U_{A_2B}^\dagger$ completes the protocol. The final state has fidelity at least $1-\kappa$ with $|\psi\rangle$.

We end with the accounting: The foregoing discussion implies that we may choose

$$\begin{aligned} \log d_{A_2} &= \frac{1}{2}(\log d_S + \log \lambda_{\max}) \\ &\qquad - O(\log\kappa) + O(\log\log d_S) \\ \log d_B &= \frac{1}{2}(\log d_S - \log \lambda_{\max}) \\ &\qquad - O(\log\kappa) + O(\log\log d_S). \end{aligned}$$

□

The main idea behind the proof, combining an exponential concentration bound with discretization, has been used a number of times recently in quantum information theory [17, 50, 48]. (It is, of course, much older; see [23].) If there is a twist in the present application, it is illustrated in Eq. (1.13). Since d_S is comparable in size to $d_{A_2}d_B$, any prefactor significantly larger than $(5/\gamma)^{2d_{A_1'}d_S}$ would have caused the probability bound to fail. Therefore, it was crucial to restrict first to states maximally entangled between A_1' and S, giving the manageable prefactor, and then extend to general states and larger A_1 using majorization.

1.3 Optimality of the protocol

The communication and entanglement resources of Proposition 1.2.3 are optimal up to terms of lower order than $\log d_S$ or $\log \lambda_{\max}$: the amount of quantum communication cannot be reduced, neither can the sum of the entanglement and quantum communication. (Entanglement alone can be reduced at the cost of increasing the quantum communication.) We will demonstrate the result in two steps. First we prove an optimality result for the task of remotely preparing entangled quantum states using entanglement and *classical* communication. We then show that by teleporting the quantum communication of our superdense coding protocol for entangled states, we generate the optimal remote state preparation protocol, meaning the original superdense coding protocol must have been optimal.

Proposition 1.3.1 *A remote state preparation protocol of fidelity $F \geq 1/2$ for all d_S-dimensional states with maximum Schmidt coefficient $\leq \sqrt{\lambda_{\max}}$ must make use of at least $\log d_S + \log \lambda_{\max} + \log F - 2$ cbits and $\log d_S - 18\sqrt{1-F}\log d_S - 2\eta\left(2\sqrt{1-F}\right)$ ebits, where $\eta(t) = -t\log t$.*

Proof Consider a remote state preparation protocol involving the transmission of exactly $\log K$ cbits that can, with fidelity F, prepare all d_S dimensional states having maximum Schmidt coefficient $\sqrt{\lambda_{\max}}$. We will show that causality essentially implies that K must be roughly as large as $d_S \lambda_{\max} F$.

In particular, suppose Alice wants to send Bob a message $i \in \left\{1, \ldots, \lfloor \frac{d_S}{a} \rfloor\right\}$, with $a = \lceil \frac{1}{\lambda_{\max}} \rceil$. One way she can accomplish this is by preparing (a purification of) the state $\sigma_i = \frac{1}{a} \sum_{k=1+a(i-1)}^{ai} |k\rangle\langle k|$ on Bob's system, with some fixed basis $\{|k\rangle\}$. The remote state preparation protocol will produce a state ρ_i for Bob that will have a fidelity F with the intended state, σ_i. In order to decode the message, Bob simply measures $\Pi_i = \sum_{k=1+a(i-1)}^{ai} |k\rangle\langle k|$. His probability of decoding the message Alice intended is $\mathrm{Tr}(\rho_i \Pi_i) \geq F$.

Now, imagine that Alice and Bob use the same protocol, with the modification that rather than Alice sending cbits, Bob simply guesses which $j \in \{1, \ldots, K\}$ Alice would have sent. The probability of Bob correctly identifying i in this case is thus at least $\frac{F}{K}$ — he has a probability $\frac{1}{K}$ of correctly guessing j and, given a correct guess, a conditional probability F of correctly identifying i. However, since this protocol involves no forward communication from Alice to Bob, it can succeed with probability no greater than $\lfloor \frac{d_S}{a} \rfloor^{-1}$ (by causality), implying $K \geq F \lfloor \frac{d_S}{a} \rfloor$, which implies that $K \geq \log d_S + \log \lambda_{\max} + \log F - 2$.

The entanglement lower bound follows easily from conservation of entanglement under local operations and classical communication (LOCC): let Alice and Bob prepare a maximally entangled state $|\Phi_0\rangle$ of Schmidt rank d_S. If they were able to do this exactly, by the non-increase of entanglement under LOCC, they would need to start with at least $\log d_S$ ebits. However, the protocol only succeeds in creating a state ρ of fidelity $\geq F$ with $|\Phi_0\rangle$. By a result of Nielsen [24], this implies that for the entanglement of formation,

$$E_F(\rho) \geq \log d_S - 18\sqrt{1-F} \log d_S - 2\eta\left(2\sqrt{1-F}\right).$$

Since E_F cannot increase under LOCC, the right hand side is also a lower bound on the number of ebits Alice and Bob started with. □

Corollary 1.3.2 *A superdense coding protocol of fidelity $F \geq 1/2$ for all d_S-dimensional states with maximum Schmidt coefficient $\leq \sqrt{\lambda_{\max}}$ must make use of at least $\frac{1}{2}\log d_S + \frac{1}{2}\log \lambda_{\max} + \frac{1}{2}\log F - 1$ qubits of communication. The sum of qubit and ebit resources must be at least $\log d_S - 18\sqrt{1-F}\log d_S - 2\eta\left(2\sqrt{1-F}\right)$.*

Proof Suppose there exists an superdense coding protocol that can prepare all d_S dimensional states with maximum Schmidt coefficient $\leq \sqrt{\lambda_{\max}}$ and that uses only Q qubits and E ebits. Use teleportation to transmit the qubits, turning it into an remote state preparation protocol.

The qubit cost translates directly to a cbit cost of $2Q$. From Proposition 1.3.1 we infer the lower bound on Q. The protocol including teleportation requires $Q + E$ ebits, thus the lower bound on $Q + E$ follows from Proposition 1.3.1 as well. □

Thus, when $F \to 1$ and ignoring terms of order $o(\log d_S)$, the upper resource bounds from our protocol, and the above lower bound coincides.

1.4 Protocol for a memoryless source

The universal protocol of Proposition 1.2.3 is easily adapted to the task of sending states produced by a memoryless source. A standard application of typical subspace techniques gives control of the value of $\lambda_{\max}$ and the effective size of the states received by Bob, the two parameters determining the resources consumed by the universal protocol. We model the source $\mathcal{E}_{A_1S} = \{p_i, |\varphi_i^{A_1S}\rangle\}_{i=1}^m$ as a sequence of independent, identically distributed states:

$$|\varphi_{i^n}^{A_1S}\rangle = |\varphi_{i_1}^{A_1S}\rangle \otimes \cdots \otimes |\varphi_{i_n}^{A_1S}\rangle \tag{1.18}$$

occurs with probability $p_{i^n} = p_{i_1} p_{i_2} \ldots p_{i_n}$, where $i^n = i_1 i_2 \cdots i_n$. If we define $S(\mathcal{E}_S) = S\left(\sum_i p_i \operatorname{Tr}_{A_1} |\varphi_i\rangle\langle\varphi_i|\right)$ and $\bar{S}(\mathcal{E}_S) = \sum_i p_i S(\operatorname{Tr}_{A_1} |\varphi_i\rangle\langle\varphi_i|)$, Harrow combined coherent classical communication and a remote state preparation protocol to demonstrate that a qubit rate of $\frac{1}{2}(S(\mathcal{E}_S) - \bar{S}(\mathcal{E}_S))$ and ebit rate of $\frac{1}{2}(S(\mathcal{E}_S) + \bar{S}(\mathcal{E}_S))$ are simultaneously achievable [71], an optimal result [1] that hinted at the existence of the universal protocol. Here we show how the universal protocol provides an alternate, perhaps more direct, route to Harrow's rate pair.

Proposition 1.4.1 *There exist protocols for superdense coding of entangled states with mean fidelity approaching one and asymptotically achieving the rate pair of* $\frac{1}{2}(S(\mathcal{E}_S) - \bar{S}(\mathcal{E}_S))$ *qubits and* $\frac{1}{2}(S(\mathcal{E}_S) + \bar{S}(\mathcal{E}_S))$ *ebits.*

Proof With probability p_{i^n}, Alice needs to prepare the state $|\varphi_{i^n}^{A_1 S}\rangle$. Instead, for typical i^n, she prepares a state $|\sigma_{i^n}^{A_1 S}\rangle$ obtained by applying a typical projector and a conditional typical projector to $\varphi_{i^n}^S$. When i^n is atypical, the protocol fails.

Given a probability distribution q on a finite set χ, define the set of *typical sequences*, with $\delta > 0$, as

$$\mathcal{T}_{q,\delta}^n = \left\{ x^n : \forall x | N(x|x^n) - nq_x| \leq \delta \sqrt{n} \sqrt{q_x(1-q_x)} \right\}, \tag{1.19}$$

where $N(x|x^n)$ counts the numbers of occurences of x in the string $x^n = x_1 x_2 \cdots x_n$. If $\rho = \sum_i p_i \varphi_i^S$ has spectral decomposition $\sum_{j=1}^{d_S} R(j) \Pi_j$, we then define the typical projector to be

$$\Pi_{\rho,\delta}^n = \sum_{j^n \in \mathcal{T}_{R,\delta}^n} \Pi_{j_1} \otimes \cdots \otimes \Pi_{j_n} \tag{1.20}$$

and the conditional typical projector to be

$$\Pi_{\varphi^S,\delta}^n(i^n) = \bigotimes_{i=1}^m \Pi_{\varphi_i,\delta}^{I_i}, \tag{1.21}$$

where $I_i = \{j \in [n] : i_j = i\}$and $\Pi_{\varphi_i,\delta}^{I_i}$ refers to the typical projector in the tensor product of the systems $j \in I_i$. In terms of these definitions, $\sigma_{i^n}^{A_1 S}$, the state Alice prepares instead

of $\varphi_{i^n}^{A_1 S}$, is proportional to

$$(\mathbb{1}_{A_1} \otimes \Pi_{\rho,\delta}^n \Pi_{\varphi^S,\delta}^n(i^n)) \varphi_{i^n}^{A_1 S} (\mathbb{1}_{A_1} \otimes \Pi_{\varphi^S,\delta}^n(i^n) \Pi_{\rho,\delta}^n), \tag{1.22}$$

With respect to approximation, the relevant property of these operators is that, defining

$$\xi_{i^n} = \Pi_{\varphi^S,\delta}^n(i^n) \varphi_{i^n}^S \Pi_{\varphi^S,\delta}^n(i^n), \tag{1.23}$$

we have

$$\begin{aligned} &\mathrm{Tr}[\Pi_{\rho,\delta}^n \Pi_{\varphi^S,\delta}^n(i^n) \varphi_{i^n}^S \Pi_{\varphi^S,\delta}^n(i^n) \Pi_{\rho,\delta}^n] \\ &\quad = \mathrm{Tr}[\xi_{i^n}] - \mathrm{Tr}[(\mathbb{1} - \Pi_{\rho,\delta}^n)\xi_{i^n}] \geq 1 - \epsilon, \end{aligned} \tag{1.24}$$

if $\delta = m\sqrt{2d_S/\epsilon}$ (by Lemmas 3 and 6 in [35]). The Gentle Measurement Lemma, referred to as the tender operator inequality in [35], together with a simple application of the triangle inequality, implies that $\|\varphi_{i^n}^{A_1 S} - \sigma_{i^n}^{A_1 S}\|_1 \leq \sqrt{8\epsilon} + 2\epsilon$. For a more detailed proof of these facts and further information about typical projectors, see [35]. If i^n is typical, meaning it is in the set $\mathcal{T}_{p,\delta}^n$ (which occurs with probability at least $1 - m/\delta^2$), then it is also true that

$$\Pi_{\varphi^S,\delta}^n(i^n) \varphi_{i^n}^S \Pi_{\varphi^S,\delta}^n(i^n) \leq \Pi_{\varphi^S,\delta}^n(i^n) 2^{-n\bar{S}(\mathcal{E}_S) + c\delta\sqrt{n}}, \tag{1.25}$$

$$\mathrm{Rank}\, \Pi_{\rho,\delta}^n \leq 2^{nS(\mathcal{E}_S) + c\delta\sqrt{n}}, \tag{1.26}$$

where $c > 0$ is independent of n and δ. Equation (1.25) implies that $(1-\epsilon)\sigma_{i^n}^S \leq 2^{-n\bar{S}(\mathcal{E}_S)+c\delta\sqrt{n}} \Pi_{\rho,\delta}^n$, which in turn leads to the conclusion that $\lambda_{\max}(\sigma_{i^n}^S) \leq \frac{1}{1-\epsilon} 2^{-n\bar{S}(\mathcal{E}_S)+c\delta\sqrt{n}} =: \lambda_{\max}$; Eq. (1.26) provides a bound on the effective dimension of the system S since $\sigma_{i^n}^S \leq \Pi_{\rho,\delta}^n$ for all i^n.

Applying the universal superdense coding protocol to $\sigma_{i^n}^{A_1 S}$, we find that the number of qubits that must be sent is

$$\begin{aligned} &\tfrac{1}{2}[\log \mathrm{Rank}\, \Pi_{\rho,\delta}^n + \log \lambda_{\max}] + o(n) \\ &\quad \leq \tfrac{n}{2}[S(\mathcal{E}_S) - \bar{S}(\mathcal{E}_S)] + c\delta\sqrt{n} - \log(1-\epsilon) + o(n) \end{aligned} \tag{1.27}$$

while the number of ebits used is

$$\begin{aligned} &\tfrac{1}{2}[\log \operatorname{Rank} \Pi^n_{\rho,\delta} - \log \lambda_{\max}] + o(n) \\ &\quad \leq \tfrac{n}{2}[S(\mathcal{E}_S) + \bar{S}(\mathcal{E}_S)] + \log(1-\epsilon) + o(n), \end{aligned} \tag{1.28}$$

matching the rates of the proposition. □

We've already used, in Section 1.3, the fact that teleporting the qubits of a superdense coding protocol leads to a remote state preparation protocol. When applied to Proposition 1.4.1, we get an alternative proof of Proposition 15 of [50]:

Corollary 1.4.2 *There exist protocols for remote state preparation of entangled states with mean fidelity approaching one and asymptotically achieving the rate pair of* $S(\mathcal{E}_S) - \bar{S}(\mathcal{E}_S)$ *cbits and* $S(\mathcal{E}_S)$ *ebits.* □

1.5 Identification

Quantum message identification, a generalization of hypothesis testing to the quantum setting, has been explored recently in a series of papers [3, 34, 33]. As opposed to transmission, where the goal is to communicate a message over a channel reliably, identification only allows the receiver to answer a single binary question: Is the message x or is it not? A surprising aspect of the theory of identification is that the number of questions that can be answered grows as a doubly exponential function of the number of uses of the channel, as opposed to the well-known singly exponential behavior for transmission [29, 2]. In the quantum setting, a number of versions of the identification (ID) capacity have been defined; these divide broadly into the capacities for quantum resources to identify classical messages and the capacities for those quantum resources to identify quantum messages. In the former case, doubly exponential growth of the number of messages was found, with the most important result to date that the ID capacity of an ebit, supplemented

with negligible rate of forward classical communication, is two [33]. It follows, of course, that the ID capacity of a qubit is also two [34].

In this section, we will instead be focusing on the capacity of an ebit to identify quantum messages, that is, quantum states. We will consider the model with a visible encoder and ID-visible decoder, according to the terminology introduced in [34].

Specifically, we say that we have a quantum-ID code on $\mathcal{B}(\mathbb{C}^d)$ of error $0 < \lambda < 1$ and dimension d_C if there exists an encoding map $\varepsilon : \mathcal{B}(\mathbb{C}^{d_C}) \to \mathcal{B}(\mathbb{C}^d)$ and a decoding map $D : \mathbb{C}^{d_C} \to \mathcal{B}(\mathbb{C}^d)$ such that for all pure states $|\varphi\rangle$ and $|\psi\rangle$ on $\mathbb{C}^{d_C}$

$$\left|\mathrm{Tr}(\varphi\psi) - \mathrm{Tr}(\varepsilon(\varphi)D_\psi)\right| \leq \frac{\lambda}{2}. \tag{1.29}$$

This condition ensures that the measurement $(D_\psi, \mathbb{1} - D_\psi)$ can be used on the states $\varepsilon(\varphi)$ to simulate the test $(\psi, \mathbb{1} - \psi)$ applied to the states φ. In the blind encoder, ID-visible decoder case, ε must be a quantum channel and D can be an arbitrary assignment to operators $0 \leq D_\psi \leq \mathbb{1}$. It was shown in [34] that for all $0 < \lambda < 1$ there exists on $\mathbb{C}^d$ such a quantum-ID code of error λ and $d_C = \left\lfloor d^{2\frac{(\lambda/100)^4}{4\log(\lambda/100)}} \right\rfloor$. Since, for fixed λ, $\log d_C = 2\log d + \text{const}$, this shows that, asymptotically, one qubit of communication can identify two qubits. We claim that, again asymptotically, but now using a visible encoding map, one ebit plus a negligible (rate of) quantum communication can be used to identify two qubits. Rather than introducing another cumbersome definition, we simply state the method: The states $\varepsilon(\varphi)$ that are output by the blind encoding can be prepared visibly using superdense coding. Because they are extremely mixed, their purifications are highly entangled and Proposition 1.2.3 demonstrates that negligible communication is sufficient.

The negligible communication cost is encountered frequently in the theory of identification: The classical identification capacity of a bit of shared randomness supplemented by negligible communication is a bit. In [33], it was found that the classical identification capacity of an ebit supplemented by negligible communication is two bits. Our finding

here that the quantum identification capacity of an ebit and negligible communication is two qubits, provides, in fact, an alternative proof of this result.

Proposition 1.5.1 *If $d_C = d^2(\lambda/100)^2/(\log d)^4$, then for all states $|\varphi\rangle \in \mathbb{C}^{d_C}$, approximations $|\Phi'_\varphi\rangle$ of the* purifications *of the states $\varepsilon(\varphi)$ can be prepared on $\mathbb{C}^a \otimes \mathbb{C}^d$ using $\log d + o(\log d)$ ebits and $o(\log d)$ qubits of communication, in such a way that*

$$\left|\mathrm{Tr}[\varphi\psi] - \mathrm{Tr}[(\mathrm{Tr}_{\mathbb{C}^a}\,\Phi'_\varphi)D_\psi]\right| \leq \frac{\lambda}{2}. \tag{1.30}$$

Proof From the proof of Proposition 19 in [34], if we choose $a = \epsilon d/2$, $\eta = \lambda/8$ and $\epsilon = (\eta/6)^2$, let $\varepsilon(\varphi) = \mathrm{Tr}_{\mathbb{C}^a}(V\varphi V^\dagger)$ with $V : C^{d_C} \to \mathbb{C}^d \otimes \mathbb{C}^a$ a Haar distributed isometry and choose $D_\psi = \operatorname{supp}\varepsilon(\psi)$,

$$\begin{aligned}\Pr\left(\exists\psi,\varphi \text{ such that } |\,\mathrm{Tr}(\varphi\psi) - \mathrm{Tr}(\varepsilon(\varphi)D_\psi)| > \frac{\lambda}{4}\right)\\ \leq \left(\frac{10}{\eta}\right)^{4d_C} 3\exp(-d^2\epsilon^2/16).\end{aligned} \tag{1.31}$$

with absolute constants c_0 and c_1. (Note that this statement is trivial for $a = 0$ or $a = 1$.) To be precise, in [34] the above probability bound is derived for states in the net, but it is also explained how to use triangle inequality to lift this to all states.

Therefore, the states $\varepsilon(\varphi)$ form a good quantum-ID code. We will demonstrate how to make them using superdense coding. Arguing along the lines of Eq. (1.13), we find that for all $\alpha > 0$, $d_C = ad/(\log a)^4$ and sufficiently large d,

$$\Pr\left(\inf_\varphi S\big(\varepsilon(\varphi)\big) < \log a - \alpha\right) < \frac{1}{2}. \tag{1.32}$$

(This is also a special case of Theorem IV.1 from [18].) From here, we observe that by the same reasoning given after Eq. (1.17), there exists a maximally entangled state $|\Phi_\varphi\rangle$ such that $|\langle\Phi_\varphi|V|\varphi\rangle|^2 \geq 1 - \sqrt{2\alpha\ln 2}$. We can, therefore, invoke Proposition 1.2.3 with $d_S = d$ and $\lambda_{\max} = 1/a$ to conclude that for sufficiently large d, states $|\Phi'_\varphi\rangle$ approximating $|\Phi_\varphi\rangle$ to within fidelity $\sqrt{2\alpha\ln 2}$ can be prepared using $\log d + o(\log d)$ ebits and $o(\log d)$

qubits of communication. By an appropriate choice of α, we can therefore ensure that $|\langle\Phi_\varphi|V|\varphi\rangle|^2 \geq \lambda/4$. Using the triangle inequality, we then find that

$$|\operatorname{Tr}(\varphi\psi) - \operatorname{Tr}(\Phi'_\varphi D_\psi)| \leq \lambda/2 \tag{1.33}$$

for all pure states $|\varphi\rangle \in \mathbb{C}^{d_C}$. □

There is a little subtlety in the proof that is worth considering briefly. The states to be prepared, $|\Phi_\varphi\rangle$, are maximally entangled, so one might think that they can be prepared without any communication at all. The party holding $\mathbb{C}^d$ can, indeed, create them without communication. The party holding the smaller $\mathbb{C}^a$, however, cannot; local unitary transformations on $\mathbb{C}^a$ will not change the support of the reduction to $\mathbb{C}^d$, for example. Nonetheless, by appealing to Proposition 1.2.3, we see that the asymmetry disappears in the asymptotic limit if negligible communication is allowed.

1.6 Discussion

We have proved the existence of protocols that allow a sender to share entangled states with a receiver while using as little quantum communication as is possible. These protocols interpolate between requiring no communication at all for maximally entangled states and a rate of two remote qubits per sent qubit for product states. An immediate application of the result was a proof that the identification capacity of an ebit is two qubits when visible encoding is permitted.

The question of efficient constructions remains – we would like to have protocols with the same ebit and qubit rates that are implementable in polynomial time (as has been demonstrated for state randomization [17] by Ambainis and Smith [5]). It would also be interesting to know whether stronger success criteria can be satisfied while still achieving the same rates. Specifically, the universal remote state preparation protocol of [50] produces an *exact* copy of the desired state when the protocol succeeds, not just a

high fidelity copy. Is such a probabilistic-exact protocol possible in the superdense coding setting? (One could even ask questions about perfectly faithful superdense coding, in analogy to what has been done for remote state preparation in [28, 36, 9].) Another natural question is the quantum identification capacity of an ebit in the *blind* scenario. We have shown that it is possible to achieve the identification rate of two qubits per ebit in the case when the identity of the encoded qubits is known, but it is not at all clear whether this rate is achievable when the identity of the qubits is unknown.

In the next chapter we show how the optimal superdense coding protocol for a memoryless source leads to the optimal trade-off of all three resources – qubits, cbits and ebits in noiseless quantum communication.

Chapter 2

Generalized remote state preparation: Trading cbits, qubits and ebits in quantum communication

2.1 Introduction

We consider the problem of communicating quantum states by simultaneously making use of a noiseless classical channel, a noiseless quantum channel and shared entanglement. We specifically study the version of the problem in which the sender is given knowledge of the state to be communicated. In this setting, a trade-off arises between the three resources, some portions of which have been investigated previously in the contexts of the quantum-classical trade-off in data compression, remote state preparation and superdense coding of quantum states, each of which amounts to allowing just two out of these three resources. We present a formula for the triple resource trade-off that reduces its calculation to evaluating the data compression trade-off formula. In the process, we also construct protocols achieving all the optimal points. These turn out to be achievable by trade-off coding and suitable time-sharing between optimal protocols for cases involving two resources out of the three mentioned above.

The rest of this chapter is structured as follows. Section 2.2 defines the problem

rigorously and describes previous results for the cases when one of the three resources is not used, along with some minor extensions. Section 2.3 studies the relationship between the trade-off between qubits and cbits in quantum data compression (QCT) and the trade-off between ebits and cbits in remote state preparation (RSP). In Section 2.4 these connections and the results described in section 2.2 are used to obtain optimal protocols and optimal resource trade-offs for communicating quantum states when all three resources are used simultaneously: the full "triple trade-off."

We use the following conventions throughout this chapter. If $\mathcal{E}_{AB} = \{\varphi_i^{AB}, p_i\}$ is an ensemble of bipartite states then we write $\mathcal{E}_A$ for the ensemble $\{\varphi_i^A, p_i\}$ of reduced states on system A. Sometimes we omit subscripts (or superscripts) labelling subsystems, in which case the largest subsystem on which the ensemble (or state) has been defined should be assumed: $\mathcal{E} = \mathcal{E}_{AB}$ and $\varphi_i = \varphi_i^{AB}$. We identify states with their density operators and if $|\varphi\rangle$ is a pure state, we use the notation $\varphi = |\varphi\rangle\langle\varphi|$ for its density operator. The function $S(\rho)$ is the von Neumann entropy $S(\rho) = -\operatorname{Tr}\rho\log\rho$ and $S(\mathcal{E})$ the von Neumann entropy of the average state of the ensemble $\mathcal{E}$. Functions like $S(A|B)_\rho$ and $S(A:B|C)_\rho$ are defined in the same way as their classical counterparts:

$$S(A:B|C)_\rho = S(\rho^{AC}) + S(\rho^{BC}) - S(\rho^{ABC}) - S(\rho^C), \tag{2.1}$$

for example. $\chi(\mathcal{E})$ is the Holevo χ quantity of $\mathcal{E}$ [49]. Given a bipartite ensemble $\mathcal{E}_{AB} = \{\varphi_i^{AB}, p_i\}$, we also make use the abbreviations $S = S(\mathcal{E}_B)$, $\bar{S} = \sum_i p_i\varphi_i^B$, $\chi = \chi(\mathcal{E}_B)$ and $H = H(p_i)$. Throughout, log and exp are taken base 2.

2.2 Definition of the problem and previous results

We now give a more formal definition of the task to be completed by the sender and receiver, henceforth, respectively, Alice and Bob. The reader can also refer to Figure 2.1, which illustrates the definition. We consider an ensemble of bipartite quantum states

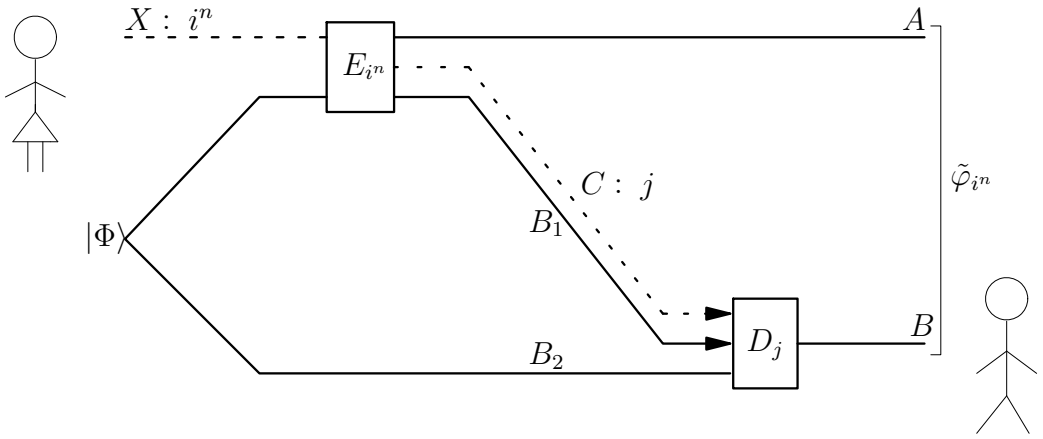


Figure 2.1: In the above quantum circuit diagram for generalized remote state preparation time goes from left to right, solid lines represent quantum registers and dashed lines represent classical registers. The registers connected in the left represent a maximally entangled state of $\log d_E$ ebits initially shared between Alice and Bob. The $\log d_Q$-qubit quantum register B_1 is sent from Alice to Bob, as is the $\log d_C$ cbit classical message m. Alice's encoding operation is denoted by E_{i^n} and Bob's decoding operation, which is conditioned on m, by D_m.

$\mathcal{E} = \{|\varphi_i\rangle^{AB}, p_i\}$ on a finite-dimensional Hilbert space $\mathcal{H}_{AB} = \mathcal{H}_A \otimes \mathcal{H}_B$ and the product ensembles $\mathcal{E}^{\otimes n} = \{|\varphi_{i^n}\rangle^{AB}, p_{i^n}\}$ on $\mathcal{H}_{AB}^{\otimes n}$, where

$$\begin{aligned} i^n &= i_1 i_2 \dots i_n, \\ p_{i^n} &= p_{i_1} p_{i_2} \dots p_{i_n} \quad \text{and} \\ |\varphi_{i^n}\rangle &= |\varphi_{i_1}\rangle \otimes |\varphi_{i_2}\rangle \otimes \dots \otimes |\varphi_{i_n}\rangle. \end{aligned}$$

At the end of the protocol, Alice and Bob are to reproduce the states of the bipartite ensemble with high fidelity. (Regardless of whether pure states are *prepared* in Bob's system, or entangled states are *shared* between Alice and Bob, we will always refer to the task simply as *communicating* from Alice to Bob.) We imagine that there is a noiseless classical channel from Alice to Bob capable of sending one of d_C messages, a noiseless quantum channel capable of sending a d_Q-dimensional quantum system and a maximally entangled state $|\Phi\rangle = d_E^{-1/2} \sum_{i=1}^{d_E} |i\rangle|i\rangle$ of Schmidt rank d_E. A source provides Alice with i^n, drawn with probability p_{i^n}, at which point Alice applies a quantum operation E_{i^n} to her half of $|\Phi\rangle$ that without loss of generality has output of the form

$$\sum_{j=1}^{d_C} \rho_{i^n,j}^{AB_1B_2} \otimes q(j|i^n)|j\rangle\langle j|^C, \tag{2.2}$$

where B_1 is a d_Q-dimensional quantum system, B_2 is the quantum system supporting Bob's half of $|\Phi\rangle$, the states $\{|j\rangle\}$ are orthonormal (*i.e* classical) and $q(\cdot|i)$ is a probability distribution. Alice then sends register B_1 to Bob over her noiseless quantum channel and C to Bob over the noiseless classical channel. The protocol is completed by Bob performing a quantum operation D_j on registers B_1 and B_2. Write $\tilde{\varphi}_{i^n}$ for the joint Alice-Bob output state averaged over different values of j. We say that the protocol has fidelity $1 - \epsilon$ if

$$\sum_{i^n} p_{i^n} \langle\varphi_{i^n}|\tilde{\varphi}_{i^n}|\varphi_{i^n}\rangle \geq 1 - \epsilon. \tag{2.3}$$

Likewise, (R, Q, E) is an achievable rate triple for the ensemble $\mathcal{E}$ if for all $\delta, \epsilon > 0$ there exists N such that for all $n > N$ there is a protocol for $\mathcal{E}^{\otimes n}$ with fidelity $1 - \epsilon$ and

$$\frac{1}{n} \log d_C \leq R + \delta \quad \frac{1}{n} \log d_Q \leq Q + \delta \quad \frac{1}{n} \log d_E \leq E + \delta. \tag{2.4}$$

Our goal will be to identify these achievable triples. In particular, we will find a formula for the function

$$E^*(R, Q) = \inf\{E : (R, Q, E) \text{ is achievable}\}. \tag{2.5}$$

We refer to rate triples of the form $(R, Q, E^*(R, Q))$ as optimal rate triples and the protocols that achieve them as optimal protocols. We will indicate that a rate triple (R, Q, E) is optimal by writing it as $(R, Q, E)^*$. Throughout the paper, unless otherwise stated, all entropic quantities will be taken with respect to 4-partite states ω of the following form:

$$\omega = \sum_i p_i |i\rangle\langle i|^X \otimes \varphi_i^{AB} \otimes \sum_{j=1}^{m+1} p(j|i) |j\rangle\langle j|^C, \tag{2.6}$$

where m is the number of states in $\mathcal{E}_{AB}$ (if that number is finite), and $p(\cdot|\cdot)$ is a classical noisy channel. Note that for all such states

$$S(X : B|C) = S(B|C) - \bar{S}, \quad \text{where } \bar{S} = \sum_i p_i S(\varphi_i^B), \tag{2.7}$$

a fact that will be useful later.

Before moving on to the general problem, we consider the special cases given by setting one of the three rates to zero.

2.2.1 $Q = 0$: Remote state preparation (RSP)

This problem was studied extensively in Ref. [50]. It is impossible to achieve an entanglement rate of less than $\sum_i p_i \varphi_i^B$, essentially because that is the amount of entanglement shared between Alice and Bob at the end of any successful protocol. The optimal cbit

rate when the entanglement is minimal is just $H(p_i)$, meaning that the simple protocol consisting of Alice communicating i^n to Bob and then the pair performing entanglement dilution is optimal. At the other extreme, the cbit rate is minimized (at least for irreducible sources) by a protocol achieving the rate $(\chi(\mathcal{E}_B), 0, S(\mathcal{E}_B))$. In general, we introduce the function

$$E^*(R) = \inf\{E : (R, 0, E) \text{ is achievable}\}. \tag{2.8}$$

This choice, a slight abuse of notation given our earlier definition of a function E^* with two arguments, is chosen for consistency with the remote state preparation paper. Note that $E^*(R) = E^*(R, 0)$. We have the following theorem from Ref. [50]:

Theorem 2.2.1 *For the ensemble $\mathcal{E} = \{|\varphi_i\rangle^{AB}, p_i\}$ of pure bipartite states and $R \geq 0$,*

$$E^*(R) = \min\{S(B|C) : S(X : BC) \leq R\}, \tag{2.9}$$

where the entropic quantities are with respect to the state ω, minimization is over all 4-partite states ω of the form of Eq. (2.6) with classical channels $p(j|i)$, and m the number of states in $\mathcal{E}$. E^* is convex, continuous and strictly decreasing in the interval in which it takes positive values.

We will also use the simple fact that the inequality in Eq. (2.9) can be replaced by equality.

2.2.2 $E = 0$: Quantum-classical trade-off (QCT)

The case where the ensemble $\mathcal{E}$ consists only of product states $|\varphi_i\rangle^{AB} = |0\rangle^A|\varphi_i\rangle^B$ was the focus of Ref. [44]. At the extreme when $R = 0$, only quantum communication is permitted so the problem of finding achievable rates is answered by the quantum noiseless coding theorem: $(0, S(\mathcal{E}_B), 0)$ is an *optimal point*, in the sense that none of the three rates can be

reduced. Likewise, the optimal point when $Q = 0$ is given by $(H(p_i), 0, 0)$, meaning that Alice has no better strategy than to communicate the label i^n to Bob. More generally, when the ensemble is allowed to contain entangled states, the techniques of Refs. [44, 50] are easily adapted to yield a formula for

$$Q^*(R) = \inf\{Q : (R, Q, 0) \text{ is achievable}\}. \tag{2.10}$$

In particular, we have the following analog of Theorem 2.2.1:

Theorem 2.2.2 *For the ensemble $\mathcal{E} = \{|\varphi_i\rangle^{AB}, p_i\}$ of pure bipartite states and $R \geq 0$,*

$$Q^*(R) = \min\{S(B|C) : S(X:C) \leq R\}, \tag{2.11}$$

where the entropic quantities are with respect to the state ω, minimization is over all 4-partite states ω of the form of Eq. (2.6) with classical channels $p(j|i)$, and m the number of states in $\mathcal{E}$. Q^* is convex, continuous and strictly decreasing in the interval in which it takes positive values. There exists a critical value of R, hereafter referred to as H_c such that $R + Q^*(R) = S(B)$ for $R \leq H_c$ and $R + Q^*(R) > S(B)$ otherwise.

As before, the inequality in Eq. (2.11) can be replaced by equality.

2.2.3 $R = 0$: Superdense coding of quantum states (SDC)

In the last chapter we showed that

$$\left(0, \tfrac{1}{2}\chi(\mathcal{E}_B), S(\mathcal{E}_B) - \tfrac{1}{2}\chi(\mathcal{E}_B)\right) \tag{2.12}$$

is an achievable rate triple. Using this construction, we can easily find the $R = 0$ trade-off curve:

Theorem 2.2.3 *For the ensemble $\mathcal{E} = \{|\varphi_i\rangle^{AB}, p_i\}$ of pure bipartite states and $Q \geq 0$,*

$$E^*(0, Q) = \begin{cases} S(\mathcal{E}_B) - Q & \text{if } Q \geq \chi(\mathcal{E}_B)/2 \\ +\infty & \text{otherwise.} \end{cases} \tag{2.13}$$

Proof Since $(0, S, 0)$ and $(0, \chi/2, S - \chi/2)$ (S and χ are defined in the introduction) are both achievable rate triples, any convex combination of the two is an achievable rate triple corresponding to a time-shared protocol. Thus, if $0 \leq \lambda \leq 1$,

$$(0, \lambda S + (1-\lambda)\chi/2, (1-\lambda)(S - \chi/2)) \tag{2.14}$$

is achievable. Suppose these points are not optimal. Then there exists $\epsilon > 0$ such that

$$(0, \lambda S + (1-\lambda)\chi/2, (1-\lambda)(S - \chi/2) - \epsilon) \tag{2.15}$$

is optimal. By using quantum communication to establish entanglement, however, protocols achieving this rate can be converted into protocols with the rate triple

$$(0, \lambda S + (1-\lambda)\chi/2 + (1-\lambda)(S - \chi/2) - \epsilon, 0) = (0, S - \epsilon, 0), \tag{2.16}$$

contradicting the optimality of Schumacher compression. We conclude that $E^*(0, Q) = S - Q$ when this conversion is possible, that is, when $Q \geq \chi/2$. This condition is required by causality. (For a detailed proof, see Section 2.4.3.) □

The simple argument used in the proof of Theorem 2.2.3 is characteristic of what will follow. Our evaluation of $E^*(R, Q)$ will be accomplished via operational reductions to the three extremal cases we have now completed, just as Theorem 2.2.3 was demonstrated using a reduction from the unknown $E^*(0, Q)$ curve to the known Schumacher compression point.

Later we will also have occasion to make use of the following analog of the QCT and RSP constructions. Given a state ω of the form of Eq. (2.6), the trade-off coding technique from Ref. [44] then gives protocols achieving all the rate triples of the form

$$\left(S(X:C), \tfrac{1}{2}S(X:B|C), S(B|C) - \tfrac{1}{2}S(X:B|C)\right). \tag{2.17}$$

Briefly, once an optimal channel $p(j|i)$ is chosen, Alice and Bob can share (typical) $j^n = j_1 \ldots j_n$ at a cost of $nS(X:C) + o(n)$ bits of communication plus shared random

bits using the reverse Shannon theorem [65]. Harrow's protocol is then used on the induced "conditional" ensembles

$$\begin{aligned} \{|\varphi_{i^n}\rangle^{AB}, q(i^n|j^n) &= q(i_1|j_1)\dots q(i_n|j_n)\}, \quad \text{where} \\ q(i|j) &= \left(\sum_{i'} p_{i'} p(j|i')\right)^{-1} p(j|i) p_i. \end{aligned} \tag{2.18}$$

The shared random bits are then seen to be unnecessary because we only require high fidelity on average (so that some particular value of the shared random bits can be used). Evaluation of the rates for the approach gives exactly Eq. (2.17).

Given any $(R, Q^*(R), 0)$ there is a state ω of the form Eq. (2.6) for which $(S(X:C), S(B|C), 0) = (R, Q^*(R), 0)$. For this state, we therefore find a new achievable rate triple:

$$\left(S(X:C), \tfrac{1}{2}S(X:B|C), S(B|C) - \tfrac{1}{2}S(X:B|C)\right) = \left(R, \tfrac{1}{2}(Q^*(R) - \bar{S}), \tfrac{1}{2}(Q^*(R) + \bar{S})\right), \tag{2.19}$$

where we have used Eq. (2.7) to arrive at the expression on the right hand side.

2.3 Relating optimal QCT and optimal RSP

Any protocol for quantum-classical compression can be converted into an RSP protocol by using RSP to send the compressed qubits. One might hope that if the original QCT point was optimal that the resulting RSP point would also be optimal. For classical rates above H_c this is indeed the case but otherwise it need not be. Consider, for example, the ensemble consisting of the orthonormal states $|0\rangle$ and $|1\rangle$, each occurring with probability $\frac{1}{2}$. In this case, $Q^*(0) = 1$ but the corresponding RSP protocol would wastefully consume 1 cbit and 1 ebit per signal when 1 cbit and no entanglement are sufficient.

As an aside, while we have described a natural way to convert optimal QCT protocols into optimal RSP protocols (that works when $R \geq H_c$), there is no known way to do the

opposite. An appendix to Ref. [50], however, demonstrates the existence of just such an operational reduction but only under the assumption that the mixed state compression conjecture is true. (See Refs. [53, 54, 55] for more details on the conjecture.)

The following two lemmas formally express the relationship between optimal QCT and optimal RSP:

Lemma 2.3.1 *When* $R \geq H_c$, $E^*(R+Q^*(R)-\bar{S}) = Q^*(R)$. *Otherwise,* $E^*(R+Q^*(R)-\bar{S}) = Q^*(H_c)$.

Proof We begin by showing that $E^*(R + Q^*(R) - \bar{S}) \leq Q^*(R)$. We know that $(S(X:BC), 0, S(B|C))$ is an achievable rate triple for any ω of the form of Eq. (2.6). In particular, it is achievable when $(S(X:C), S(B|C), 0) = (R, Q^*(R), 0)$, in which case

$$\begin{aligned} (S(X:BC), 0, S(B|C)) &= \left(S(X:C) + S(B|C) - \bar{S}, 0, S(B|C)\right) && (2.20) \\ &= \left(R + Q^*(R) - \bar{S}, 0, Q^*(R)\right). && (2.21) \end{aligned}$$

This proves the claim. Note that this inequality is true regardless of whether R is greater or less than H_c.

We now prove the opposite inequality: $E^*(R + Q^*(R) - \bar{S}) \geq Q^*(R)$ when $R \geq H_c$. Substituting our expressions for $E^*(R)$ and $Q^*(R)$ shows that what we need to prove is that

$$\begin{aligned} & \min\{S(B|C) : S(X:C) + S(B|C) = R + Q^*(R)\} && (2.22) \\ \geq\ & \min\{S(B|C) : S(X:C) = R\}. && (2.23) \end{aligned}$$

Let ω be the state that minimizes the first expression for fixed R. If $S(X:C)_\omega \leq R$ then we're done so we may suppose not: $S(X:C)_\omega = R + \Delta$ for some $\Delta > 0$. By convexity and the definition of H_c, for any $R \geq H_c$,

$$\frac{Q^*(R+\Delta) - Q^*(R)}{\Delta} > -1. \qquad (2.24)$$

Rearranging this inequality yields

$$(R + \Delta) + Q^*(R + \Delta) > R + Q^*(R). \tag{2.25}$$

Using the hypothesis $S(X : C)_\omega = R + \Delta$ and the fact that the right hand side of the above inequality is $S(X : C)_\omega + S(B|C)_\omega$, we find that $S(B|C)_\omega < Q^*(R + \Delta)$. But, again by hypothesis, $S(X : C)_\omega = R + \Delta$ so we have a contradiction of the definition of $Q^*(R + \Delta)$. We conclude that $S(X : C)_\omega \leq R$.

Finally, $R + Q^*(R) - \bar{S} = \chi$ when $R < H_c$ so $E^*(R) = E^*(\chi)$ is constant. Using the first half of the lemma, we then find $E^*(\chi) = E^*(H_c + Q^*(H_c) - \bar{S}) = Q^*(H_c)$. □

Lemma 2.3.2 *$Q^*(R - E^*(R) + \bar{S}) = E^*(R)$ when $R \geq \chi$. Otherwise $E^*(R) = +\infty$.*

Proof Let $H_c \leq R_1$ and consider $R = R_1 + Q^*(R_1) - \bar{S}$. R is a strictly increasing function of R_1 by the definition of H_c, taking all values $\chi \leq R$. Substituting into Lemma 2.3.1 gives

$$\begin{aligned} Q^*(R - E^*(R) + \bar{S}) &= Q^*(R_1 + Q^*(R_1) - \bar{S} - Q^*(R_1) + \bar{S}) && (2.26)\\ &= Q^*(R_1) && (2.27)\\ &= E^*(R_1 + Q^*(R_1) - \bar{S}) && (2.28)\\ &= E^*(R). && (2.29) \end{aligned}$$

Also, $R < \chi$ is not achievable (by causality, see section 2.4.3), yielding the second half of the lemma. □

2.4 The triple trade-off

The following theorem is the main result of the paper: a prescription for calculating the minimal amount of entanglement required given any cbit and qubit rate.

Theorem 2.4.1

$$E^*(R,Q) \;=\; \begin{cases} 0 & \text{if} \quad Q^*(R) < Q \\ Q^*(R) - Q & \text{if} \;\; \frac{1}{2}(Q^*(R) - \bar{S}) \leq Q \leq Q^*(R) \\ E^*(R+2Q) - Q & \text{if} \;\; \frac{1}{2}(\chi - R) \leq Q < \frac{1}{2}(Q^*(R) - \bar{S}) \\ +\infty & \text{if} \, Q < \frac{1}{2}(\chi - R) \end{cases}$$

We discuss each of the four ranges for Q separately, referring to them, in order, as the *QCT region*, the *low-entanglement region*, the *high-entanglement region* and the *forbidden region*. The names of the first and last regions should be self-explanatory. (QCT is optimal by definition in the QCT region and no amount of entanglement is sufficient in the forbidden region.) In the low-entanglement region we'll find that optimal protocols can be found by time-sharing between QCT and SDC (the first of which does not use entanglement) while the optimal protocols for the high-entanglement region are found by time-sharing between RSP and SDC, *both* of which rely on entanglement.

While H_c does not appear explicitly in our formula, it once again delineates the boundary between two qualitatively different regimes: for $R < H_c$ we have that $\frac{1}{2}(Q^*(R) - \bar{S}) = \frac{1}{2}(\chi - R)$ so there is no high-entanglement region in this case. The region defined by $R < H_c$ and $Q \geq \frac{1}{2}(\chi - R)$ is entirely contained in low-entanglement region.

Before giving a proof of Theorem 2.4.1, we consider the standard example: $\mathcal{E}_{AB}$ being the uniform (unitarily invariant) ensemble over qubit states on B. Devetak and Berger gave an explicit parametrization [56] of the function identified as $Q^*(R)$ for this ensemble in Ref. [44] and the corresponding RSP curve appeared in Ref. [50]. We present the full trade-off surface $E^*(R,Q)$ in figure 2.3. (In the case of an infinite ensemble, theorems 2.2.1 and 2.2.2 need to be slightly modified: the min should be replaced by an inf as explained in Theorem 10.1 of Ref. [44]. The only modification required to the argument of this paper is in the second half of Lemma 2.3.1, where a sequence of ω_n needs to be considered instead of a fixed minimizing ω.)

Figure 2.2: Achievable rate triples and conversions

Rate triple	Description
$(R, Q^*(R), 0)$	QCT
$(R, 0, E^*(R))$	RSP
$(R, \frac{1}{2}(Q^*(R) - \bar{S}), \frac{1}{2}(Q^*(R) + \bar{S}))$	SDC on QCT: Eq. (2.19)
$(R + Q^*(R) - \bar{S}, 0, Q^*(R))$ for $R \geq H_c$	QCT to RSP: lemma 2.3.1
$(R - E^*(R) + \bar{S}, E^*(R), 0)$	RSP to QCT: lemma 2.3.2
$(R, Q, E) \longrightarrow (R + 2Q, 0, E + Q)$	Teleportation (of qubits)
$(R, Q, E) \longrightarrow (0, Q + \frac{1}{2}R + Q, \frac{1}{2}R + E)$	Superdense coding (of cbits)
(R_1, Q_1, E_1) & (R_2, Q_2, E_2) $\longrightarrow \lambda(R_1, Q_1, E_1) + (1 - \lambda)(R_2, Q_2, E_2)$	Time-sharing
$(R, Q, E) \longrightarrow (R, Q + E, 0)$	Sending entanglement using qubits

We also summarize for convenience in Table 2.2 all the rate triples and conversions between them that we will use in the proof. We use the notation $(R, Q, E) \longrightarrow (R', Q', E')$ to indicate that if the rate triple (R, Q, E) is achievable then so is the rate triple (R', Q', E'); *i.e.* (R, Q, E) can be *converted* into (R', Q', E'). Similarly, if we write $(R, Q, E)^* \longrightarrow (R', Q', E')$ then the conversion is possible conditional on (R, Q, E) being optimal.

2.4.1 The low-entanglement region: $\frac{1}{2}(Q^*(R) - \bar{S}) \leq Q \leq Q^*(R)$

Define $\lambda = 2(Q^*(R) - Q)/(Q^*(R) + \bar{S})$. By the definition of the low-entanglement region, $0 \leq \lambda \leq 1$. Both $(R, Q^*(R), 0)$ and $(R, \frac{1}{2}(Q^*(R) - \bar{S}), \frac{1}{2}(Q^*(R) + \bar{S}))$ are achievable so the convex combination

$$(R, Q, Q^*(R) - Q) = \lambda(R, Q^*(R), 0) \; + \; (1 - \lambda)\left(R, \frac{1}{2}(Q^*(R) - \bar{S}), \frac{1}{2}(Q^*(R) + \bar{S})\right) \quad (2.30)$$

is achievable by time-sharing.

The proof that these points are optimal is very simple. Suppose they are not. Then there would exist an ϵ such that $(R, Q, Q^*(R) - Q - \epsilon)$ were optimal. Now, using the conversion $(R, Q, E) \to (R, Q+E, 0)$, it follows that $(R, Q^*(R)-\epsilon, 0)$ is achievable, which is a contradiction of the definition of Q^*.

2.4.2 The high-entanglement region: $\frac{1}{2}(\chi - R) \leq Q < \frac{1}{2}(Q^*(R) - \bar{S})$

We first define two new variables R_1 and R_2, which are functions of R and Q but much easier to work with:

$$R_1 = R + 2Q - E^*(R + 2Q) + \bar{S}, \tag{2.31}$$

$$R_2 = R - R_1 + \bar{S} = E^*(R + 2Q) - 2Q. \tag{2.32}$$

We collect for future use some simple facts about R_1 and R_2:

1. $R_1 \geq H_c$:

   The function $R' - E^*(R') + \bar{S}$ is a monotonically increasing function of R'. By causality, therefore, the minimum of this function over achievable R' occurs when $R' = \chi$. From Lemma 2.3.1, $E^*(\chi) = Q^*(H_c) = S - H_c$, so $R' - E^*(R') + \bar{S} \geq H_c$. Since $R + 2Q \geq \chi$ in the high-entanglement region, we conclude that $R_1 \geq H_c$.

2. $Q = \frac{1}{2}(Q^*(R_1) - R_2)$:

   This follows by Lemma 2.3.2: $Q^*(R_1) = E^*(R + 2Q) = R_2 + 2Q$.

3. $E^*(R + 2Q) - Q = R_2 + Q = \frac{1}{2}(Q^*(R_1) + R_2)$:

   This follows by the definition of R_2 and the previous fact.

4. $R_2 \leq Q^*(R_1)$:

   By fact 1, $R_2 = Q^*(R_1) - 2Q$.

5. $Q^*(R_1) \geq \bar{S}$:

   $Q^*(R_1) - \bar{S} \;=\; S(B|C) - \bar{S} \;=\; S(X:B|C) \geq 0$ (for optimal ω).

6. $R_2 \geq \bar{S}$ (for $Q \leq \frac{1}{2}(Q^*(R) - \bar{S})$) :

   This is equivalent to $E^*(R+2Q) \geq 2Q + \bar{S}$. Since $2Q \leq Q^*(R) - \bar{S}$ in this region, we have by the monotonicity of E^* and by Lemma 2.3.1 that

$$\begin{aligned} E^*(R+2Q) &\geq E^*(R + Q^*(R) - \bar{S}) && (2.33)\\ &= Q^*(R) && (2.34)\\ &\geq 2Q + \bar{S}. && (2.35) \end{aligned}$$

Equipped with these observations we can now proceed to the proof of Theorem 2.4.1 in the high-entanglement region. That is, we will prove that $E^*(R,Q) = E^*(R+2Q) - Q$ when $\frac{1}{2}(\chi - R) \leq Q < \frac{1}{2}(Q^*(R) - \bar{S})$. Note that

$$(R, Q, E^*(R+2Q) - Q) = \left(R_1 + R_2 - \bar{S}, \tfrac{1}{2}(Q^*(R_1) - R_2), \tfrac{1}{2}(Q^*(R_1) + R_2)\right) \quad (2.36)$$

in terms of the new variables, by the definition of R_1 and R_2 as well as facts 2 and 3.

Proof of achievability

$(R_1, \frac{1}{2}(Q^*(R_1) - \bar{S}), \frac{1}{2}(Q^*(R_1) + \bar{S}))$ is achievable by Eq. (2.19) and $(R_1 + Q^*(R_1) - \bar{S}, 0, Q^*(R_1))$ is achievable by Lemma 2.3.1. By facts 4,5,and 6, $\lambda = (Q^*(R_1) - R_2)/(Q^*(R_1) - \bar{S})$ is between 0 and 1. Therefore, the convex combination

$$\begin{aligned} &\left(R_1 + R_2 - \bar{S}, \tfrac{1}{2}(Q^*(R_1) - R_2), \tfrac{1}{2}(Q^*(R_1) + R_2)\right) && (2.37)\\ = \;& \lambda \left(R_1, \tfrac{1}{2}(Q^*(R_1) - \bar{S}), \tfrac{1}{2}(Q^*(R_1) + \bar{S})\right) + (1-\lambda)\left(R_1 + Q^*(R_1) - \bar{S}, 0, Q^*(R_1)\right) && (2.38) \end{aligned}$$

is also achievable by time-sharing.

Proof of optimality

Suppose these points were not optimal. Then there would exist an ϵ such that $\left(R_1 + R_2 - \bar{S}, \frac{1}{2}(Q^*(R_1) - R_2), \frac{1}{2}(Q^*(R_1) + R_2) - \epsilon\right)$ were optimal. Performing teleportation yields the conversion

$$\left(R_1 + R_2 - \bar{S}, \tfrac{1}{2}(Q^*(R_1) - R_2), \tfrac{1}{2}(Q^*(R_1) + R_2) - \epsilon\right) \longrightarrow (R_1 + Q^*(R_1) - \bar{S}, 0, Q^*(R_1) - \epsilon) \quad (2.39)$$

(Note that teleportation is appropriate here instead of RSP because the encoding map corresponding to the first triple will generally produce complicated entangled states between Alice and Bob, conditioned on the classical bits being communicated. Teleportation will preserve this entanglement.)

This is a contradiction by Lemma 2.3.1.

2.4.3 The forbidden region: $Q < \frac{1}{2}(\chi - R)$

In keeping with the operational spirit of the other arguments in this paper, we argue that achievability in this region would lead to a violation of causality. A classical channel of dimension d_C and a quantum channel of dimension d_Q can be used to transmit at most $\log d_C + 2 \log d_Q$ bits of classical information by the optimality of superdense coding [47, 49]. Success in the ensemble communication task, however, results in Bob holding a high-fidelity copy of $\mathcal{E}_B$. By using coding, Alice could then communicate approximately $\chi(\mathcal{E}_B)$ classical bits to Bob per usage of the protocol [63, 58], a violation of causality (for sufficiently high fidelity and small δ in the notation of section 2.2) if $\chi(\mathcal{E}_B) > R + 2Q$.

A simple entropic argument is also possible. Consider the state

$$\rho = \sum_{i^n, j} p_{i^n} |i^n\rangle\langle i^n|^X \otimes \rho_{i^n, j}^{AB_1B_2} \otimes q(j|i^n)|j\rangle\langle j|^C, \quad (2.40)$$

which represents the output of Alice's encoding operation for a given (unspecified) pro-

tocol of the form of figure 2.1. We can estimate

$$\begin{aligned} \tfrac{1}{n}\chi(\{\tilde{\varphi}^B_{i^n}, p_{i^n}\}) &\leq S(X : B_1\, B_2\, C) \quad \text{(by monotonicity of } \chi) && (2.41)\\ &= S(X : B_2) + S(X : C|B_2) + S(X : B_1|B_2C) && (2.42)\\ &\leq \log d_C + 2\log d_Q, && (2.43) \end{aligned}$$

using Lemma 2.4.2 (see below) twice and the fact that $S(X : B_2) = 0$ since B_2 is maximally mixed for all i^n. On the other hand, applying the Fannes inequality [59] and the fidelity condition implies that

$$\tfrac{1}{n}\chi(\{\tilde{\varphi}^B_{i^n}, p_{i^n}\}) \xrightarrow{\epsilon\to 0} \chi, \tag{2.44}$$

giving the constraint $\chi \leq R + 2Q$.

Lemma 2.4.2 *Let ρ be a tripartite density operator of the form*

$$\rho = \sum_i p_i |i\rangle\langle i|^X \otimes \rho_i^{AB}, \tag{2.45}$$

where the states $\{|i\rangle^X\}$ are orthonormal and the p_i are probabilities. Then

$$S(X : A|B) \leq \min(\log \dim X, 2\log \dim A). \tag{2.46}$$

Proof We can expand $S(X : A|B) = S(X|B) - S(X|AB)$. By subadditivity of the von Neumann entropy, the first term is less than or equal to $S(X)$, which is in turn no more than $\log \dim X$. Moreover, because ρ is separable across the X/AB cut, $S(X|AB) \geq 0$. (This follows immediately from concavity of the entropy [60, 61].)

To prove the second inequality, we expand the definition of $S(X : A|B)$ differently:

$$S(X : A|B) = S(A|B)_{\rho^{AB}} + \sum_i p_i S(A|B)_{\rho_i^{AB}}. \tag{2.47}$$

Using subadditivity of the von Neumann entropy again, $S(A|B) \leq S(A)$ for any density operator. $S(A)$, in turn, is always less than or equal to $\log \dim A$. □

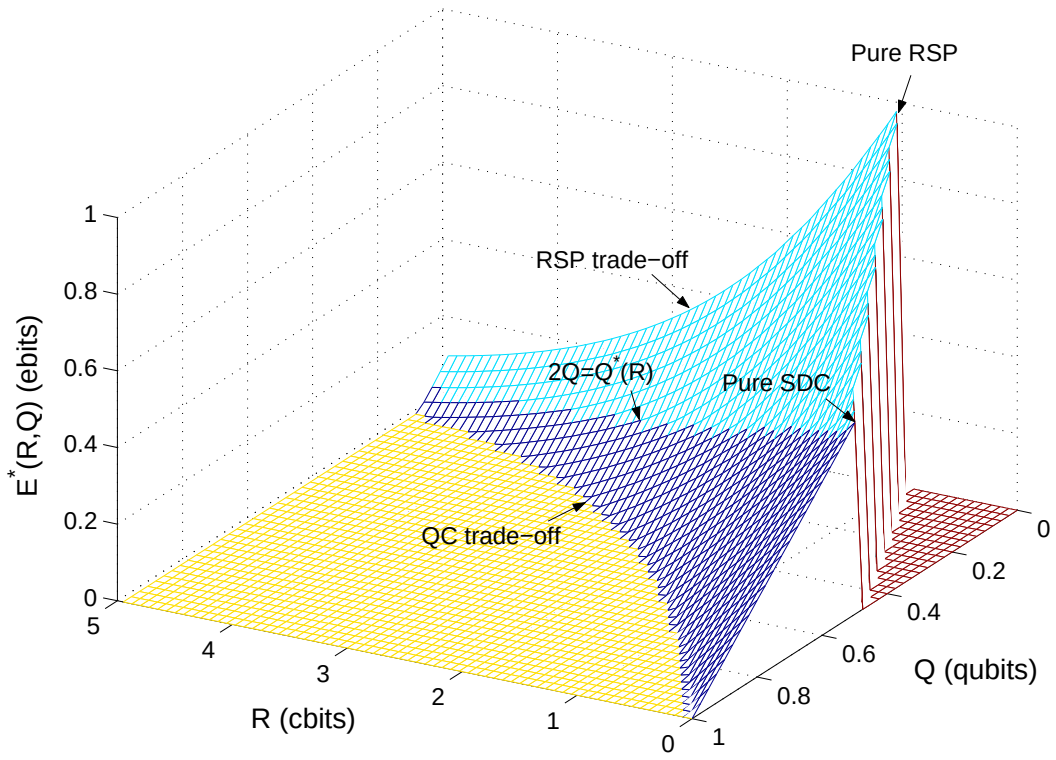


Figure 2.3: Trade-off surface for the uniform qubit ensemble. The region on the left for which $E^*(R,Q) = 0$ is the QCT region, whose boundary with the low-entanglement region is given by the curve $(R, Q^*(R), 0)$. The transition to the high-entanglement region then occurs when $2Q = Q^*(R)$; note that the surface is not smooth at the transition. Finally the points corresponding to pure RSP, $(1, 0, 1)$, and pure SDC, $(0, 1/2, 1/2)$, define the boundary of the forbidden region. In the low-entanglement region, the trade-off is a ruled surface, linear for constant R.

Unification of noisy quantum information theory

One of the major goals of quantum information theory is to find the optimal ways to make use of noisy quantum states or channels for communication or establishing entanglement. Quantum Shannon theory attacks the problem in the limit of many copies of the state or channel in question, in which situation the answers often simplify to the point where they can be expressed by relatively compact formulae. The last ten years have seen major advances in the area, including, among many other discoveries, the determination of the classical capacity of a quantum channel [62, 63], the capacities of entanglement-assisted channels [64, 65], the quantum capacity of a quantum channel [66, 67, 68], and the best ways to use noisy entanglement to extract pure entanglement [69] or send classical information [70]. Until recently, however, each new problem was solved essentially from scratch and no higher-level structure was known connecting the different results. Harrow's discovery of the *cobit* [71] and its subsequent application to the construction of the so-called "mother" and "father" protocols provided that missing structure. All the problems listed above and others were shown to fall into two families, first the mother and her descendants, and second the father and his [72]. Appending or prepending simple transformations like teleportation and superdense coding sufficed to transform the parents into their children.

In this paper, we provide a direct proof of the mother protocol or, more precisely, of the existence of a protocol performing the same task as the mother. In contrast to most proofs in information theory, instead of showing how to establish perfect correlation of some kind between the sender (Alice) and the receiver (Bob), our proof proceeds by showing that the protocol *destroys* all correlation between the sender and a reference system. Since destroying correlation is a relatively straightforward task, the resulting proof is correspondingly simple. This approach also makes it clear that the mother actually accomplishes even more than originally thought. In particular, in addition to distilling entanglement between Alice and Bob, the protocol transfers all of Alice's entanglement with a reference system to Bob. This side effect is extremely important in its own right, and a major focus of our paper. To start, it places the state merging protocol of Horodecki, Oppenheim and Winter [73] squarely within the mother's brood. In addition, it makes it possible to use the mother as a building block for distributed compression. We analyze the resulting protocols, finding they are optimal for sources described by separable density operators, as well as inner and outer bounds on the achievable rate region in general.

Finally, the new approach to the mother solves a major problem left unanswered in the original family paper. No operational relationship between the mother and father could be identified but they were nonetheless connected by a symmetry called *source-channel duality* [74]. This new mother protocol can be directly transformed into the father, resolving the mystery of the two parents' formal similarity and collapsing the two families into one.

We use the following conventions throughout this chapter. For a quantum system A, let $d_A = \dim A$. For two quantum systems A and A', let F_A be the operator that swaps the two systems. An operator acting on a subsystem is freely identified with its extension (via tensor product with the identity) to larger systems. Π_A^+ denotes the projector onto

the symmetric subspace of $A \otimes A'$ and Π_A^- the projector onto the antisymmetric subspace of $A \otimes A'$. Let $\mathbb{U}(A)$ be the unitary group on A. $H(A)_\varphi$ is the von Neumann entropy of φ^A, $I(A;B)_\varphi = H(A)_\varphi + H(B)_\varphi - H(AB)_\varphi$ is the mutual information between the A and B parts of φ and $H(A|B)_\varphi = H(AB)_\varphi - H(B)_\varphi$ the conditional entropy. The symbol $|\Phi\rangle^{AB}$ will be used to represent a maximally entangled state between A and B.

Chapter 3

Restructuring quantum information's family tree

3.1 The family of quantum protocols

The mother protocol is a transformation of a tensor power quantum state $(|\varphi\rangle^{ABR})^{\otimes n}$. At the start, Alice holds the A shares and Bob the B shares. R is a reference system purifying the AB systems and does not participate actively in the protocol. In the original formulation, the mother protocol accomplished a type of entanglement distillation between Alice and Bob in which the only communication permitted was the ability to send *qubits* from Alice to Bob. The transformation can be expressed concisely in the resource inequality formalism as

$$\langle \varphi^{AB} \rangle + \frac{1}{2} I(A;R)_{\varphi} [q \to q] \geq \frac{1}{2} I(A;B)_{\varphi} [qq]. \tag{3.1}$$

We will informally explain the resource inequalities used here, but the reader is directed to Ref. [75] for a rigorous treatment. $[q \to q]$ represents one qubit of communication from Alice to Bob and $[qq]$ represents an ebit shared between them. In words, n copies of the

state φ shared between Alice and Bob can be converted into $\frac{1}{2}I(A;B)_\varphi$ EPR pairs per copy provided Alice is allowed to communicate with Bob by sending him qubits at rate $\frac{1}{2}I(A;R)_\varphi$ per copy. Small imperfections in the final state are permitted provided they vanish as n goes to infinity.

In this chapter, we prove a stronger resource inequality that we call the *fully quantum Slepian-Wolf* (FQSW) inequality. The justification for this name will become apparent in Section 3.6, where we study its applicability to distributed compression. The inequality states that starting from state $(|\varphi\rangle^{ABR})^{\otimes n}$ and using only quantum communication at the rate $\frac{1}{2}I(A;R)_\varphi$ from Alice to Bob, they can distill EPR pairs at the rate $\frac{1}{2}I(A;B)_\varphi$ and produce a state approximating $(|\psi\rangle^{R\hat{B}})^{\otimes n}$, where $\hat{B}$ is held by Bob and $\varphi^R = \psi^R$. That is, Alice can *transfer* her entanglement with the reference system R to Bob while simultaneously distilling ebits with him. This can be expressed as a resource inequality in the following way:

$$\langle \mathcal{U}^{S\to AB} : \varphi^S \rangle + \frac{1}{2}I(A;R)_\varphi \, [q \to q] \geq \frac{1}{2}I(A;B)_\varphi \, [qq] + \langle \mathrm{id}^{S\to\hat{B}} : \varphi^S \rangle. \tag{3.2}$$

This inequality makes use of the concept of a relative resource. A resource of the form $\langle \mathcal{N} : \rho^S \rangle$ is a channel with input system S that is guaranteed to behave like the channel $\mathcal{N}$ provided the reduced density operator of the input state on S is ρ^S. In the inequality, $\mathcal{U}^{S\to AB}$ is an isometry taking the S system to AB. Thus, on the left hand side of the inequality, a state is distributed to Alice and Bob while on the right hand side, that same state is given to Bob alone. Transforming the first situation into the second means that Alice transfers her portion of the state to Bob.

Since the relationship of the mother to entanglement distillation and communication supplemented using noisy entanglement is explained at length in the original family paper, we will not describe the connections here. The FQSW inequality is stronger than the mother, however, and leads to more children. In particular, if the entanglement produced

at the end of the protocol is then re-used to perform teleportation, we get the following resource inequality:

$$\langle \mathcal{U}^{S\to AB} : \varphi^S \rangle + H(A|B)_\varphi \, [q \to q] + I(A;B)_\varphi \, [c \to c] \geq \langle \mathrm{id}^{S\to \hat{B}} : \varphi^S \rangle, \tag{3.3}$$

which is known as the *state merging* primitive, discovered by Oppenheim and Winter. It is of note both because it is a useful building block for multiparty protocols and because it provides an operational interpretation of the conditional entropy $H(A|B)_\varphi$ as the number of qubits Alice must send Bob in order to transfer her state to him, ignoring the classical communication cost.

On the other side of the family there is the father protocol. In contrast to the mother, in which Alice and Bob share a mixed state $(\varphi^{AB})^{\otimes n}$, for the father protocol they share a noisy channel $\mathcal{N}^{A'\to B}$. Let $U^{A'\to BE}$ be a Stinespring dilation of $\mathcal{N}$ with environment system E, such that $\mathcal{N}(\rho) = \mathrm{Tr}_E \, U\rho U^\dagger$, and define $|\varphi\rangle^{ABE} = U^{A'\to BE}|\varphi\rangle^{AA'}$ for a pure state $|\varphi\rangle^{AA'}$. The resource inequality is

$$\langle \mathcal{N}^{A'\to B} \rangle + \frac{1}{2} I(A;E)_\varphi \, [qq] \geq \frac{1}{2} I(A;B)_\varphi \, [q \to q]. \tag{3.4}$$

Thus, Alice and Bob use pre-existing shared entanglement and the noisy channel to produce noiseless quantum communication. Comparing Eq. (3.4) to the mother, Eq. (3.1), reveals the two to be strikingly similar: To go from one to the other it suffices to replace channels by states and vice-versa, as well as replace the reference R by the environment E. This is known as source-channel duality [74].

3.2 The fully quantum Slepian-Wolf protocol

The input to the fully quantum Slepian-Wolf protocol is a quantum state, $(|\varphi\rangle^{RAB})^{\otimes n}$, and the output is also a quantum state, $|\Phi\rangle^{A_2\tilde{B}}(|\varphi\rangle^{R\hat{B}})^{\otimes n}$. A_2 is a quantum system

held by Alice while both $\tilde{B}$ and $\hat{B}$ are held by Bob. $|\Phi\rangle^{A_2\tilde{B}}$ therefore represents a maximally entangled state shared between Alice and Bob. The size of the A_2 system is $n\frac{1}{2}I(A;B)_\varphi - o(n)$ qubits. The steps in the protocol that transform the input state to the output state are as follows:

1. Alice performs Schumacher compression on her system A^n. The output space A_S factors into two subsystems A_1 and A_2 with $\log d_{A_1} = n\frac{1}{2}I(A;R) + o(n)$.

2. Alice applies a unitary transformation U_A to A_S and then sends A_1 to Bob.

3. Bob applies an isometry V_B taking A_1B^n to $\hat{B}\tilde{B}$.

It remains to specify which transformations U_A and V_B Alice and Bob should apply, as well as a more precise bound on d_{A_1}. Observe that each step in the protocol is essentially non-dissipative. Since no information is leaked to the environment at any step, Bob will hold a purification of the A_2R^n system after step 2, regardless of the choice of U_A. Because all purifications are equivalent up to local isometric transformations of the purifying space, it therefore suffices to ensure that the reduced state on A_2R^n approximates $\Phi^{A_2} \otimes (\varphi^R)^{\otimes n}$ after step 2. Bob's isometry U_B will be the one taking the purification he holds upon receiving A_2 to the one approximating $|\Phi\rangle^{A_2\tilde{B}}(|\varphi\rangle^{R\hat{B}})^{\otimes n}$.

From this perspective, the operation $\rho \to \mathrm{Tr}_{A_1}(U_A\,\rho\,U_A^\dagger)$ should be designed to *destroy* the correlation between A_2 and R^n: the mother will succeed provided the state on $A_2 \otimes R^n$ is a product state and A_2 is maximally mixed. The operation U_A does not itself destroy the correlation; the partial trace over A_1 does that. U_A should therefore be chosen in order to ensure that tracing over A_1 should be maximally effective. Because one qubit can carry at most two bits of information, tracing over a qubit can reduce mutual information by at most two bits. The starting state $(\varphi^{AR})^{\otimes n}$ has $nI(A;B)_\varphi$ bits of mutual information, which means that A_1 must consist of at least $\frac{n}{2}I(A;B)_\varphi$ qubits. We will see that choosing

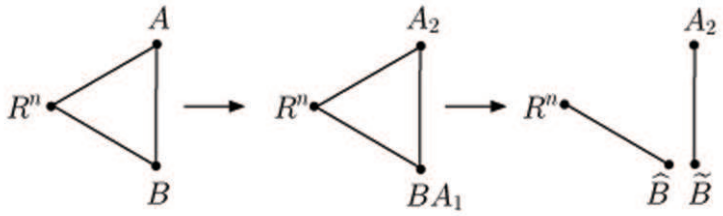


Figure 3.1: The transformations of the input state to the output state in the FQSW protocol

U_A randomly according to the Haar measure will come close to achieving this rate. The result is similar in spirit to a recent result of Groisman et al. that demonstrated that in order to destroy correlation in the state φ by discarding *classical* information instead of quantum, Alice must discard twice as large a system as she does here: $I(A;B)_\varphi$ bits per copy [76].

3.3 Fully quantum Slepian-Wolf: one-shot version

While the tensor power structure of $(|\varphi\rangle^{ABR})^{\otimes n}$ allows the fully quantum Slepian-Wolf inequality (3.2) to be expressed conveniently in terms of mutual information quantities, our approach allows us to treat arbitrary input states without such structure as well. In this section, we will prove a general "one-shot" version of the fully quantum Slepian-Wolf result that leads quickly to Inequality (3.2) in the special case where the input state is a tensor power.

For this section, we will therefore dispense with $|\varphi\rangle^{\otimes n}$ and instead study a general state $|\psi\rangle^{ABR}$ shared between Alice, Bob and the reference system. We also eliminate the Schumacher compression step: assume that A has been decomposed into subsystems A_1 and A_2 satisfying $d_A = d_{A_1} d_{A_2}$. Finally, let $\sigma^{A_2 R}(U) = \mathrm{Tr}_{A_1}[(U \otimes I_R)\psi^{AR}(U^\dagger \otimes I_R)]$

be the state remaining on A_2R after the unitary transformation U has been applied to $A = A_1A_2$.

The following inequality is essentially the one-shot version of fully quantum Slepian-Wolf:

Theorem 3.3.1 (One-shot, fully quantum Slepian-Wolf bound) *Let ψ and σ be defined as in the previous paragraph. Then*

$$\int_{\mathbb{U}(A)} \|\sigma^{A_2R}(U) - \sigma^{A_2}(U) \otimes \sigma^R(U)\|_1^2 \, dU \leq \frac{d_A d_R}{d_{A_1}^2} \Big\{ \operatorname{Tr}[(\psi^{AR})^2] + \operatorname{Tr}[(\psi^A)^2] \operatorname{Tr}[(\psi^R)^2] \Big\}. \tag{3.5}$$

The theorem quantifies how distinguishable $\sigma^{A_2R}(U)$ will be from its completely decoupled counterpart $\sigma^{A_2}(U) \otimes \sigma^R(U)$ if U is chosen at random according to the Haar measure. As a first observation, note that as d_{A_1} grows, the two states become progressively more indistinguishable. Also, the upper bound on the right hand side is expressed entirely in terms of the dimensions of the spaces involved and the purities $\operatorname{Tr}[(\psi^{AR})^2]$, $\operatorname{Tr}[(\psi^A)^2]$ and $\operatorname{Tr}[(\psi^R)^2]$. In the tensor power source setting, both dimensions and purities can be replaced by functions of the corresponding entropies, but in the one-shot setting they must be distinguished.

Focusing on the first term in the upper bound reveals that in order for the bound to be small, it is necessary that

$$\log d_{A_1} \gg \frac{1}{2} \Big[\log d_A + \log d_R + \log \operatorname{Tr}[(\psi^{AR})^2] \Big]. \tag{3.6}$$

This expression plays the role of $\frac{1}{2} I(A;R) = \frac{1}{2}[H(A) + H(R) - H(AR)]$ in the one-shot setting.

According to the proof strategy outlined in the previous section, if $\sigma^{A_2R}(U)$ is close to $\sigma^{A_2}(U) \otimes \sigma^R(U)$, then $\sigma^{A^2R}(U)$ has a purification which is itself close to a product

state. This argument can be made quantitative using the Uhlmann fidelity $F(\tau, \omega) = (\mathrm{Tr}\sqrt{\tau^{1/2}\omega\tau^{1/2}})^2$. If $\left\|\sigma^{A_2R}(U) - \sigma^{A_2}(U) \otimes \sigma^R(U)\right\|_1 \leq \epsilon$, then standard inequalities imply that $F\big(\sigma^{A_2R}(U), \sigma^{A_2}(U) \otimes \sigma^R(U)\big) \geq 1 - \epsilon$ [77]. But if two mixed states are close in fidelity, then they have equally close purifications [78, 79]. That is, σ^{A_2R} can be purified to a state $|\xi\rangle^{A_2\tilde{B}\hat{B}R}$ and $\sigma^{A_2}(U) \otimes \sigma^R(U)$ to a state $|\xi_1\rangle^{A_2\tilde{B}}|\xi_2\rangle^{R\hat{B}}$ such that $|\langle\xi|\,|\xi_1\rangle|\xi_2\rangle|^2 \geq 1 - \epsilon$. Since all purifications are locally equivalent, Bob could apply a transformation V to A_1B taking $|\psi\rangle^{A_1A_2BR}$ to $|\xi\rangle^{A_2\tilde{B}\hat{B}R}$. The net result would be that $\hat{B}$ would hold nearly all of the original Alice-Bob entanglement with the reference system R and that $A_2\tilde{B}$ would contain nearly pure entanglement between Alice and Bob.

The proof of the Theorem 3.3.1 is quite straightforward. We will evaluate the corresponding average over the unitary group exactly for the Hilbert-Schmidt norm and then use simple inequalities to extract Inequality (3.5). Before starting in earnest, we perform a calculation that will prove to be useful:

Lemma 3.3.2

$$\int_{\mathbb{U}(A)} (U^\dagger \otimes U^\dagger \otimes I_{RR'}) F_{A_2R} (U \otimes U \otimes I_{RR'})\, dU = [p\Pi_A^+ + q\Pi_A^-] \otimes F_R, \tag{3.7}$$

where

$$p = \frac{d_{A_1} + d_{A_2}}{d_A + 1} \quad \textit{and} \quad q = \frac{d_{A_1} - d_{A_2}}{d_A - 1}. \tag{3.8}$$

Proof Let X be Hermitian. By Schur's lemma,

$$\int_{\mathbb{U}(A)} (U^\dagger \otimes U^\dagger) X (U \otimes U)\, dU = \alpha_+(X)\Pi_A^+ + \alpha_-(X)\Pi_A^-, \tag{3.9}$$

with the coefficients $\alpha_\pm(X) = \mathrm{Tr}(X\Pi_A^\pm)/\,\mathrm{Rank}(\Pi_A^\pm)$.

Recall that $\Pi_A^{\pm} = \frac{1}{2}(I_{AA'} \pm F_A)$.

$$\begin{aligned}
\operatorname{Rank}(\Pi_A^{\pm})\,\alpha_{\pm}(F_{A_2}) &= \operatorname{Tr}(\Pi_A^{\pm} F_{A_2}) && (3.10)\\
&= \frac{1}{2}\operatorname{Tr}\left[\left(I_{AA'} \pm F_{A_1}\otimes F_{A_2}\right)F_{A_2}\right] && (3.11)\\
&= \frac{1}{2}\left[\operatorname{Tr}(I_{A_1A_1'}\otimes F_{A_2}) \pm \operatorname{Tr}(F_{A_1}\otimes I_{A_2A_2'})\right] && (3.12)\\
&= \frac{1}{2}[d_{A_1}^2 d_{A_2} \pm d_{A_1} d_{A_2}^2]. && (3.13)
\end{aligned}$$

The second line uses the identity $F_A = F_{A_1} \otimes F_{A_2}$. The third follows from $F^2 = I$ and the explicit inclusion of previously implicit identity operators to help in the evaluation of the trace in line four. The formula then follows after a little algebra, using that $F_{A_2R} = F_{A_2} \otimes F_R$ and $\operatorname{Rank}(\Pi_A^{\pm}) = d_A(d_A \pm 1)/2$. □

The next step is an exact evaluation of the Hilbert-Schmidt analogue of the one-shot, fully quantum Slepian-Wolf inequality.

Lemma 3.3.3

$$\begin{aligned}
&\int_{\mathbb{U}(A)} \|\sigma^{A_2R}(U) - \sigma^{A_2}(U)\otimes\sigma^R(U)\|_2^2\, dU = \\
&\frac{d_{A_1}d_{A_2}^2 - d_{A_1}}{d_A^2 - 1}\Big\{\operatorname{Tr}[(\psi^{AR})^2] - 2\operatorname{Tr}[\psi^{AR}(\psi^A\otimes\psi^R)] + \operatorname{Tr}[(\psi^A)^2]\operatorname{Tr}[(\psi^R)^2]\Big\}. \qquad (3.14)
\end{aligned}$$

Proof Note that

$$\|\sigma^{A_2R} - \sigma^{A_2}\otimes\sigma^R\|_2^2 = \operatorname{Tr}[(\sigma^{A_2R})^2] - 2\operatorname{Tr}[\sigma^{A_2R}(\sigma^{A_2}\otimes\sigma^R)] + \operatorname{Tr}[(\sigma^{A_2})^2]\operatorname{Tr}[(\sigma^R)^2]. \quad (3.15)$$

Starting with the first term,

$$\begin{aligned}
\int_{\mathbb{U}(A)} \operatorname{Tr}[(\sigma^{A_2R}(U))^2]\, dU &= \int \operatorname{Tr}\left[\left(\sigma^{A_2R}(U)\otimes\sigma^{A_2'R'}(U)\right)F_{A_2R}\right] dU && (3.16)\\
&= \int \operatorname{Tr}\left[\left(\operatorname{Tr}_{A_1}(U\psi^{AR}U^\dagger)\otimes\operatorname{Tr}_{A_1'}(U\psi^{A'R'}U)^\dagger\right)F_{A_2R}\right] dU && (3.17)\\
&= \operatorname{Tr}\left[(\psi^{AR}\otimes\psi^{A'R'})\cdot\int (U^\dagger\otimes U^\dagger)(I_{A_1A_1'}\otimes F_{A_2R})(U\otimes U)\, dU\right] && (3.18)\\
&= \operatorname{Tr}\left[(\psi^{AR}\otimes\psi^{A'R'})\cdot(p\Pi_A^+ + q\Pi_A^-)\otimes F_R\right] && (3.19)\\
&= \frac{p+q}{2}\operatorname{Tr}[(\psi^R)^2] + \frac{p-q}{2}\operatorname{Tr}[(\psi^{AR})^2], && (3.20)
\end{aligned}$$

where p and q are defined as in Eq. (3.8). In the fourth line we've used the result of Lemma 3.3.2, and in the fifth the identity $\Pi_A^{\pm} = \frac{1}{2}(I_{AA'} \pm F_A)$. The third term in Eq. (3.15) can also be evaluated using this formula and the observation that $\sigma^R(U) = \psi^R$, giving

$$\int_{\mathbb{U}(A)} \operatorname{Tr}[(\sigma^{A_2})^2] \operatorname{Tr}[(\sigma^R)^2]\, dU = \left\{ \frac{p+q}{2} + \frac{p-q}{2} \operatorname{Tr}[(\psi^A)^2] \right\} \operatorname{Tr}[(\psi^R)^2]. \tag{3.21}$$

That leaves the second term of Eq. (3.15), which can be calculated in the same way as Eq. (3.16), with the result that

$$\int_{\mathbb{U}(A)} \operatorname{Tr}[\sigma^{A_2 R}(\sigma^{A_2} \otimes \sigma^R)]\, dU = \frac{p+q}{2} \operatorname{Tr}[(\psi^R)^2] + \frac{p-q}{2} \operatorname{Tr}[\psi^{AR}(\psi^A \otimes \psi^R)]. \tag{3.22}$$

Substituting back into Eq. (3.15) shows that $\int_{\mathbb{U}(A)} \|\sigma^{A_2 R}(U) - \sigma^{A_2}(U) \otimes \sigma^R(U)\|_2^2\, dU$ is equal to

$$\frac{p-q}{2} \Big\{ \operatorname{Tr}[(\psi^{AR})^2] - 2 \operatorname{Tr}[\psi^{AR}(\psi^A \otimes \psi^R)] + \operatorname{Tr}[(\psi^A)^2] \operatorname{Tr}[(\psi^R)^2] \Big\}, \tag{3.23}$$

which, after substitution for p and q, yields (3.14). □

The one-shot fully quantum Slepian-Wolf theorem is then an easy corollary:

Proof (of Theorem 3.3.1) The Cauchy-Schwarz inequality can be used to relate the two norms: $\|\cdot\|_1^2 \leq d_{A_2} d_R \|\cdot\|_2^2$. Also, $\operatorname{Tr}[\psi^{AR}(\psi^A \otimes \psi^R)]$ is non-negative. Finally,

$$\frac{d_{A_1} d_{A_2}^2 - d_{A_1}}{d_A^2 - 1} \leq \frac{1}{d_{A_1}} \tag{3.24}$$

holds for all $d_{A_1} \geq 1$. □

For good measure, we can check how entangled the state Bob shares with Alice will

be. In general, the decoupling can occur without producing a maximally entangled state.

$$\begin{aligned}
\int_{\mathbb{U}(A)} \left\| \sigma^{A_2}(U) - \frac{I}{d_{A_2}} \right\|_1^2 dU &\leq d_{A_1} \int_{\mathbb{U}(A)} \left\| \sigma^{A_2}(U) - \frac{I}{d_{A_2}} \right\|_2^2 dU && (3.25)\\
&= d_{A_2} \int_{\mathbb{U}(A)} \operatorname{Tr}[(\sigma^{A_2}(U))^2] dU - 1 && (3.26)\\
&= d_{A_2} \left\{ \frac{p+q}{2} + \frac{p-q}{2} \operatorname{Tr}[(\psi^A)^2] \right\} - 1 && (3.27)\\
&\leq d_{A_2} \frac{p-q}{2} \operatorname{Tr}[(\psi^A)^2] && (3.28)\\
&\leq \frac{d_A}{d_{A_1}^2} \operatorname{Tr}[(\psi^A)^2]. && (3.29)
\end{aligned}$$

The first line is Cauchy-Schwarz, the second is an integral that we already performed in proving Lemma 3.3.3, and the last is an application of Eq. (3.24).

3.4 Fully quantum Slepian-Wolf: i.i.d. version

We return now to the setting where Alice, Bob and the reference system share the state $|\psi'\rangle = (|\varphi\rangle^{ABR})^{\otimes n}$. Combining the one-shot, fully quantum Slepian-Wolf result with Schumacher compression will lead to the FQSW resource inequality. Let Π_A^{typ}, Π_R^{typ} and Π_{AR}^{typ} be projectors onto the typical subspaces of the systems indicated by the subscripts. If we define

$$|\psi\rangle = \Pi_A^{typ} \Pi_R^{typ} \Pi_{AR}^{typ} |\psi'\rangle, \tag{3.30}$$

then we can choose the projectors such that $\operatorname{Rank} \Pi_F^{typ} \leq 2^{nH(F)+o(n)}$ and $\Pi_F^{typ} \psi \Pi_F^{typ} \leq 2^{-nH(F)-o(n)}$ for any combination of subsystems F and such that $\langle \psi | \psi' | \psi \rangle \geq 1 - o(1)$. Therefore, while we are concerned with the output of the protocol when it is applied to the state $(|\varphi\rangle^{ABR})^{\otimes n}$, we will analyze its effect on $|\psi\rangle$ and be able to conclude that the outputs are indistinguishable in the limit of many copies. In particular, we can *assume* that Schumacher compression has been performed on the AR and R systems even though the definition of the protocol only has it performed on A. This is an important point

because R is the reference system and is therefore not permitted to participate actively in the protocol.

Thanks to the Schumacher compression, the various dimensions in Theorem 3.3.1 then get replaced by the ranks of the corresponding projectors and the purities by the corresponding eigenvalue bounds. For an arbitrary combination of subsystems F, let F^{typ} denote the support of Π_F^{typ} and assume $A^{typ} = A_1 \otimes A_2$. Let $\sigma^{A_2R^{typ}}(U) = \mathrm{Tr}_{A_1}(U\psi U^\dagger)/\langle\psi|\psi\rangle$. (Note that U now acts only on the typical subspace of A.) By Theorem 3.3.1, we find that

$$\int_{\mathbb{U}(A^{typ})} \left\| \sigma^{A_2R^{typ}}(U) - \sigma^{A_2}(U) \otimes \sigma^{R^{typ}}(U) \right\|_1^2 dU \leq \frac{1}{d_{A_1}^2} \left\{ 2^{-nI(A;R)+o(n)} + 2^{o(n)} \right\}. \quad (3.31)$$

Therefore, if $\log d_{A_1} \geq n[I(A;R)/2 + \delta]$ for any $\delta > 0$, $\sigma^{A_2R^{typ}}$ and Alice sends A_1 to Bob, he will end up holding a close approximation to a purification of $\varphi^{R^{\otimes n}}$ for sufficiently large n. To a good approximation, he will also hold a purification of the state remaining on Alice's system. We can use Eq. (3.29) to test how entangled we expect the state to be:

$$\int_{\mathbb{U}(A^{typ})} \left\| \sigma^{A_2}(U) - \frac{I}{d_{A_2}} \right\|_1^2 dU \leq \frac{1}{d_{A_1}^2} 2^{o(n)}. \quad (3.32)$$

Therefore, Alice and Bob will share *maximal entanglement* provided Alice sends qubits to Bob at a positive rate. As discussed after Theorem 3.3.1, Bob could therefore apply a transformation V taking A_1B to $\tilde{B}\hat{B}$ with final state $|\xi\rangle^{A_2\tilde{B}\hat{B}R}$ satisfying $\left\| \langle\xi||\Phi\rangle^{A_2\tilde{B}} |\psi'\rangle^{\hat{B}R} \right\|^2 \geq 1 - o(1)$.

3.5 Father from FQSW

We begin with a third version the FQSW theorem that we will need to invoke in order to prove the existence of the father protocol. Let $|\psi\rangle$ and $|\psi'\rangle$ be as before. Now define Π_A^t

to be the projector onto a particular *typical* type t and consider the normalized states

$$|\psi'_t\rangle = \frac{1}{\sqrt{P'_t}}\Pi^t_A|\psi'\rangle \quad \text{and} \quad |\psi_t\rangle = \frac{1}{\sqrt{P_t}}\Pi^t_A|\psi\rangle. \tag{3.33}$$

By standard arguments using the method of types, we know that $P_t \geq 2^{-o(n)}$ for each typical type t. Since $\Pi^{typ}_A|\psi'\rangle = \sum_t \sqrt{P'_t}|\psi'_t\rangle$, and $|\psi\rangle = \sum_t \sqrt{P_t}|\psi_t\rangle$, we have

$$\sum_t \sqrt{P_t P'_t}|\langle\psi_t|\psi'_t\rangle| \geq |\langle\psi|\psi'\rangle| \geq 1 - o(1). \tag{3.34}$$

Again by standard methods, there exists a t for which both

$$|\langle\psi_t|\psi'_t\rangle| \geq 1 - o(1) \tag{3.35}$$

and $P'_t \leq 3P_t$. Choose t to be such. Then $P_t \geq 2^{-o(n)}$. Thus, from (3.33)

$$\|\psi^A_t\|_\infty \leq 2^{o(n)}\|\psi^A\|_\infty \leq 2^{-nH(A)+o(n)} \tag{3.36}$$

and similarly for R and AR. Decomposing $A_t = A_1A_2$, by Theorem 3.3.1 and Eq. (3.35) there exists a unitary U on A_t such that

$$\left\|\sigma_t^{A_2R}(U) - \sigma_t^{A_2}(U) \otimes \sigma_t^R(U)\right\|_1 \leq \frac{1}{d_{A_1}}\{2^{nI(A;R)/2+o(n)} + 2^{o(n)}\} + o(1), \tag{3.37}$$

where $\sigma_t^{A_2R}(U) = \mathrm{Tr}_{A_1}[(U \otimes I^R)\psi_t^{A_tR}(U^\dagger \otimes I^R)]$. Choosing $d_{A_1} \geq 2^{nI(A;R)/2+o(n)}$, there is thus an isometry $V^{A_1B\to\tilde{B}\hat{B}}$ such that

$$\left|(V \circ U)\langle\psi_t|^{A_tBR}|\Phi_2\rangle^{A_2\tilde{B}}|\Gamma\rangle^{R\hat{B}}\right| \geq 1 - o(n),$$

where $|\Phi_2\rangle$ is a maximally entangled state of rate $I(A;B)/2 - o(1)$ ebits. This is the required variant of the fully quantum Slepian-Wolf theorem.

Given a channel $\mathcal{N}^{A'\to B}$, choose a Stinespring dilation $U_{\mathcal{N}}{}^{A'\to BE}$ such that $\mathcal{N}(\rho) = \mathrm{Tr}_E\, U\rho U^\dagger$ and define $|\varphi\rangle^{ABE} = U_{\mathcal{N}}|\varphi\rangle^{AA'}$. If t is a type, then $|\psi_t\rangle^{A_tBE}$ is the result of sending a maximally entangled state proportional to $|\Phi\rangle^{A_tA'_t} = (\Pi^A_t \otimes \Pi^{A'}_t)(|\varphi\rangle^{AA'})^{\otimes n}$ through $U_{\mathcal{N}}{}^{\otimes n}$. Recall that the father protocol is a type of entanglement-assisted quantum

Figure 3.2:

Father	FQSW
B_3	A_1
R	A_2
B	B
E	R

communication. Let Alice and Bob initially share a maximally entangled state $|\Phi_3\rangle^{A_3B_3}$ of $nI(A;E)_\varphi/2 + o(n)$ ebits. In order to verify that the quantum communication performed by the protocol is correct, we will introduce a reference system R and a maximally entangled state $|\Phi_0\rangle^{RA_0}$. Alice performs an isometry identifying A_0A_3 with A_t'. There is a corresponding isometry identifying RB_3 with A_t. Thus, we have identified $|\Phi\rangle^{A_tA_t'}$ with $|\Phi_0\rangle^{RA_0}|\Phi_3\rangle^{B_3A_3}$. The situation is illustrated in Figure 3.2. We will now invoke the type class variant FQSW result but now with the subsystems playing different roles: If the FQSW unitary U were applied to B_3R, we see that the effect would be to decouple the registers corresponding in the FQSW picture to A_2 and R, meaning R and E, respectively. This is precisely what is desired because there will exist a decoding protocol for Bob to complete the father provided that, after the application of the channel $\mathcal{N}$, he holds the entire purification of R. Of course, applying U to B_3R is not possible because, as always, the reference system cannot participate in the protocol. Because the states $|\Phi_0\rangle^{RA_0}$ and $|\Phi_3\rangle^{A_3B_3}$ are maximally entangled, however, it would be equivalent for Alice to apply U^T to A_0A_3. This, in fact, will be her encoding operation. After sending A_t' through the channel, the state is precisely $U|\psi_t'\rangle^{A_tBE}$. Bob performs $V^{B_3B\to\tilde{B}\hat{B}}$, obtaining a state in which the $\tilde{B}$ system approximately purifies the reference system, which plays the role of A_2 in the FQSW picture.

3.6 Correlated source coding: distributed compression

One of the major applications of the state merging inequality (3.3) is to the problem of distributed compression with free forward classical communication. For this problem, Horodecki, Oppenheim and Winter demonstrated that the resulting region of achievable rates has the same form as the classical Slepian-Wolf problem [80, 73]. In this section, we consider the application of the fully quantum Slepian-Wolf inequality to distributed compression without classical communication.

Because distributed compression studies multiple senders, it no longer fits into the resource inequality framework. We therefore begin with some definitions describing the task to be performed. A source provides Alice and Bob with the A and B parts of a quantum state $|\psi\rangle = (|\varphi\rangle^{ABR})^{\otimes n}$ purified by a reference system R. They must independently compress their shares and transmit them to a receiver Charlie. That is, they will perform encoding operations E_A and E_B described by completely positive, trace-preserving (CPTP) maps with outputs on systems C_A and C_B of dimensions 2^{nQ_A} and 2^{nQ_B}, respectively. The receiver, Charlie, will then perform a decoding operation, again described by a CPTP map, this time with output systems $\hat{A}$ and $\hat{B}$ isomorphic to A and B. A rate pair (Q_A, Q_B) will be said to be achievable if for all $\epsilon > 0$ there exists an $N(\epsilon) > 0$ and a sequence of increasing n_k with corresponding (E_A, E_B, D) such that

$$\langle\psi|^{\hat{A}\hat{B}R}(D \circ (E_A \otimes E_B))(\psi^{ABR})|\psi\rangle^{\hat{A}\hat{B}R} \geq 1 - \epsilon \tag{3.38}$$

for all $n_k > N$. The achievable rate region $\mathcal{SW}(\varphi)$ for a given $|\varphi\rangle$ is the closure of the set of achievable rates.

The fully quantum Slepian-Wolf inequality provides a natural class of protocols for this task. One party, say Bob, first Schumacher compresses his share and sends it to

Charlie. This is possible provided $Q_B \geq H(B)_\varphi$. The other party, in this case Alice, then implements the fully quantum Slepian-Wolf protocol with Charlie playing the role of Bob. This is possible provided $Q_A \geq I(A;R)/2$. Looking at the total number of qubits required gives a curious symmetrical formula:

$$Q_A + Q_B \geq \frac{1}{2} I(A;R)_\varphi + H(B)_\varphi = \frac{H(A)_\varphi + H(B)_\varphi + H(AB)_\varphi}{2}. \tag{3.39}$$

By switching the roles played by Alice and Bob and also time-sharing between the resulting two protocols, we find that the region defined by

$$\begin{aligned} Q_A &\geq \tfrac{1}{2} I(A;R)_\varphi \\ Q_B &\geq \tfrac{1}{2} I(B;R)_\varphi \\ Q_A + Q_B &\geq \tfrac{H(A)_\varphi + H(B)_\varphi + H(AB)_\varphi}{2} \end{aligned} \tag{3.40}$$

is contained in the achievable rate region $\mathcal{SW}(\varphi)$.

In fact, the region is in some cases *equal* to $\mathcal{SW}(\varphi)$, as we will see by proving a general outer bound on the achievable rate region. Assume that $(Q_A, Q_B) \in \mathcal{SW}(\varphi)$. To begin, fix $n > N(\epsilon)$ and let W_A and W_B be the environments for the Stinespring dilations of the encoding operations E_A and E_B.

To bound Q_A, assume that Charlie has received both C_B and W_B, that is, all of B^n. Let W_C be the output environment for the dilation of Charlie's D. Again, without loss of generality we can assume that the initial environment state is an unentangled pure state.

First we will make rigorous the intuition that $\widehat{A}\widehat{B}$ is almost decoupled from W_C.

By the fidelity condition,

$$\begin{aligned} \lambda_{\max}(\varphi^{R^n \widehat{A}\widehat{B}}) &\geq \langle \psi |^{A^n B^n R^n} \varphi^{R^n A^n B^n} | v \rangle^{A^n B^n R^n} \\ &\geq 1 - \epsilon, \end{aligned}$$

where $\lambda_{\max}(\varphi^{R^n \widehat{A}\widehat{B}})$ denotes the maximum eigenvalue of $\varphi^{R^n \widehat{A}\widehat{B}}$.

Therefore, $|\varphi\rangle^{R^n\widehat{A}\widehat{B}W_AW_C}$ can be Schmidt decomposed as

$$|\varphi\rangle^{R^n\widehat{A}\widehat{B}W_AW_C} = \sum_i \sqrt{\lambda_c}|i\rangle^{W_AW_C}|i\rangle^{W_AW_C}, \tag{3.41}$$

where $\lambda_{\max} \geq 1-\epsilon$ and

$$\begin{aligned}
& \langle\varphi|^{R^n\widehat{A}\widehat{B}W_AW_C}(\varphi^{R^n\widehat{A}\widehat{B}}\otimes\varphi^{W_AW_C})|\varphi\rangle^{R_n\widehat{A}\widehat{B}W_AW_C} \\
= & \left(\sum_i \sqrt{\lambda_i}\langle i|^{R^n\widehat{A}\widehat{B}}\langle i|^{W_AW_C}\right)\left(\sum_j \lambda_j|j\rangle\langle j|^{R^n\widehat{A}\widehat{B}}\otimes\sum_k \lambda_k|kX_k|^{W_AW_C}\right) \\
& \qquad\qquad \left(\sum_\ell \sqrt{\lambda_\ell}|\ell\rangle^{R^n\widehat{A}\widehat{B}}|\ell\rangle^{W_AW_B}\right) \\
= & \sum_i \lambda_i^3 \\
\geq & (1-\epsilon)^3 \\
\geq & 1-3\epsilon^3.
\end{aligned}$$

So,

$$F(|\varphi\rangle^{R^n\widehat{A}\widehat{B}W_AW_C}, \varphi^{R^n\widehat{A}\widehat{B}}\oplus\varphi^{W_AW_C}) \geq 1-\frac{3}{2}\,\epsilon^2 \tag{3.42}$$

and

$$D(|\varphi\rangle^{R^n\widehat{A}\widehat{B}W_AW_C}, \varphi^{R^n\widehat{A}\widehat{B}}\oplus\varphi^{W_AW_C}) \leq \sqrt{3}\epsilon. \tag{3.43}$$

by the contractivity of distance we have

$$D(\varphi^{\widehat{A}\widehat{B}W_C}, \varphi^{\widehat{A}\widehat{B}}\otimes\varphi^{W_C}) \leq \sqrt{3}\epsilon$$

We can now apply the Fannes inequality to yield

$$\left|H(\varphi^{\widehat{A}\widehat{B}W_C}) - H(\varphi^{\widehat{A}\widehat{B}}\otimes\varphi^{W_C})\right| \leq \sqrt{3}\epsilon\log(d_A^n d_B^n d_{W_C}) + \eta(\sqrt{3}\epsilon) \tag{3.44}$$

for $\epsilon \leq \frac{1}{\sqrt{3}e}$.

Since the environment for any quantum operation $\rho \mapsto \epsilon(\rho)$ can be modelled as a Hilbert Space of less than d^2 dimensions, where $d = \dim(\rho)$ we have that

$$\dim W_C \leq d_A^{2n} d_B^{2n}.$$

Therefore,

$$|H(\varphi^{\widehat{A}\widehat{B}W_C}) - H(\varphi^{\widehat{A}\widehat{B}} \otimes \varphi^{W_C})| \leq n3\sqrt{3}\epsilon \log(d_A d_B) + \eta(\sqrt{3}\epsilon) \tag{3.45}$$

for $\epsilon \leq \frac{1}{\sqrt{3}\, e}$.

Next we will make rigorous the intuition that W_A nearly purifies W_C.

Since

$$F(|\psi\rangle^{A^n B^n R^n}, \varphi^{R^n \widehat{A}\widehat{B}}) \geq 1 - \frac{\epsilon}{2}, \tag{3.46}$$

$$D(|\varphi\rangle^{A^n B^n R^n}, \varphi^{R^n \widehat{A}\widehat{B}}) \leq \sqrt{\epsilon} \tag{3.47}$$

and by Fannes inequality,

$$H(\varphi^{R^n \widehat{A}\widehat{B}}) \leq n\sqrt{\epsilon} \log(d_A d_B d_R) + \eta(\sqrt{\epsilon}), \tag{3.48}$$

(for $\sqrt{\epsilon} \leq \frac{1}{e}$). Therefore,

$$n\sqrt{\epsilon} \log(d_n d_B d_A) + \eta(\sqrt{\epsilon}) \geq H(\varphi^{R^n \widehat{A}\widehat{B}}) = H(W_A W_C) \geq |H(W_A) - H(W_C)| \tag{3.49}$$

The first inequality is because $R^n \widehat{A}\widehat{B} W_A W_C$ is in a pure state, and the second inequality is the Aracki-Lieb inequality.

Finally we will make rigorous the intuition that R^n nearly purifies $\widehat{A}\widehat{B}$.

By the contractiivity of distance,

$$D(\psi^{A^n B^n}, \varphi^{\widehat{A}\widehat{B}}) \leq D(|\psi\rangle^{A^n B^n R^n}, \varphi^{R^n \widehat{A}\widehat{B}}) \leq \sqrt{\epsilon} \tag{3.50}$$

and

$$D(\psi^{R^n}, \varphi^{R^n}) \leq D(|\psi\rangle^{A^n B^n R^n}, \varphi^{R^n \widehat{A}\widehat{B}}) \leq \sqrt{\epsilon} \tag{3.51}$$

By application of the Fannes inequality,

$$|H(\varphi^{\widehat{A}\widehat{B}}) - H(\psi^{A^n B^n})| \leq n\sqrt{\epsilon} \log(d_A d_B) + \eta(\sqrt{\epsilon}) \tag{3.52}$$

and

$$|H(\varphi^{R^n}) - H(\psi^{R^n})| \leq n\sqrt{\epsilon}\log(d_A d_B) + \eta(\sqrt{\epsilon}). \tag{3.53}$$

Therefore,

$$\begin{aligned} &|H(\varphi^{\widehat{A}\widehat{B}}) - H(\varphi^{R^n})| \\ &\qquad \leq \; |H(\varphi^{\widehat{A}\widehat{B}}) - H(\psi^{A^n B^n})| + |H(\psi^{A^n B^n}) - H(\varphi^{A^n})| \\ &\qquad = \; |H(\varphi^{\widehat{A}\widehat{B}}) - H(\psi^{A^n B^n})| + |H(\psi^{R^n}) - H(\varphi^{R^n})| \\ &\qquad \leq \; n\sqrt{\epsilon}\log(d_A d_B d_R) + 2\eta(\sqrt{\epsilon}). \end{aligned} \tag{3.54}$$

The first inequality is an application of the triangle inequality, while the first equality is because $A^n B^n R^n$ is a pure state.

Putting (3.45),(3.49) and (3.54) together and using the subadditivity of the Von Neumann entropy and the fact that the overall state is pure we have

$$\begin{aligned} &|H(B^n) + H(C^n)| \geq H(B^n C_A) = H(w_c \widehat{A}\widehat{B}) \\ \geq \; & H(W_c) + H(\widehat{A}\widehat{B}) - n3\sqrt{3}\epsilon\log(d_A d_B) - \eta(\sqrt{3}\epsilon) \\ \geq \; & H(W_A) - n\sqrt{\epsilon}\log(d_A d_B d_R) - \eta(\sqrt{\epsilon}) + H(R^n) \\ & -n\sqrt{\epsilon}\log(d_A d_B d_R) - 2\eta(\sqrt{\epsilon}) - n3\sqrt{3}\log(d_A d_B) - \eta(\sqrt{3}\epsilon) \\ = \; & H(W_A) + H(R^n) - n\sqrt{\epsilon}\log(d_A^2 d_B^2 D_R^2) - n3\sqrt{3}\epsilon\log(d_A d_B) - 4\eta(\sqrt{\epsilon}) \\ \geq \; & H(A^n) - H(C_A) + H(R^n) - n\sqrt{\epsilon}\log(d_A^2 d_B^2 d_R^2) - n3\sqrt{3}\epsilon\log(d_A d_B) - 4\eta(\sqrt{\epsilon}). \end{aligned}$$

Therefore,

$$\begin{aligned} 2nQ_A \; &\geq \; 2nH(C_A) \\ &\geq \; H(A^n) - H(B^n) + H(R^n) - n\sqrt{\epsilon}\log(d_A^2 d_B^2 d_n^2) - n3\sqrt{3}\epsilon\log(d_A d_B) - 4\eta(\sqrt{\epsilon}) \\ &= \; nI(A;R) - n\sqrt{\epsilon}\log(d_A^2 d_B^2 d_R^2) - n3\sqrt{3}\log(d_A d_B) - 4\eta(\sqrt{\epsilon}) \end{aligned}$$

and so,

$$Q_A \geq \frac{1}{2}I(A;R) - \frac{1}{2}\sqrt{\epsilon}\log(d_A^2 d_B^2 d_R^2) - \frac{3\sqrt{3}}{2}\epsilon\log(d_A d_B) - \frac{2\eta(\sqrt{\epsilon})}{n}. \tag{3.55}$$

Since this is true for all $\epsilon > 0$ we have that

$$Q_A \geq \frac{1}{2} I(A;R). \tag{3.56}$$

Switching the roles of Alice and Bob gives the corresponding inequality

$$Q_B \geq \frac{1}{2} I(B;R). \tag{3.57}$$

To bound $Q_A + Q_B$ let us return to the situation where Alive and Bob perform their original encoding. Then,

$$H(A^n) = H(C_A W_A) \leq H(W_A) + H(C_A) \leq H(W_A) + nQ_A. \tag{3.58}$$

The first equality follows from the fact that the initial environment is a pure unentangled state and from the unitary invariance of the Von Neumann entropy.

Combining with the analogous inequality for B leads to

$$n(Q_A + Q_B) \geq n[H(A) + H(B)] - H(W_A) - H(W_B). \tag{3.59}$$

By similar arguments as before,

$$|H(W_A W_B R^n) - H(W_A W_B) - H(R^n)| \leq n\sqrt{3}\epsilon \log(d_A^2 d_B^2 d_R) + \eta(\sqrt{3}\epsilon), \tag{3.60}$$

for $\epsilon \leq \frac{1}{\sqrt{3}e}$.

So,

$$\begin{aligned} & H(C_A C_B) = H(W_A W_B R^n) \\ \geq \; & H(W_A) + H(W_B) - I(W_A; W_B) + H(R^n) \\ & -n\sqrt{3}\epsilon \log(d_A^2 d_B^2 d_R) - \eta(\sqrt{3}\epsilon). \end{aligned}$$

Using the purity of overall the overall state, however, gives $H(R^n) = nH(AB)$, which combined with the bound $H(C_A C_B) \leq n(Q_A + Q_B)$, leads to the inequality

$$\begin{aligned} H(W_A) + H(W_B) \; \leq \; & n(\varphi_A + \varphi_B) - nH(AB) + I(W_A; W_B) \\ & + n\sqrt{3}\epsilon \log(d_A^2 d_B^2 d_R) + n(\sqrt{3}\epsilon). \end{aligned} \tag{3.61}$$

Adding equations (3.59) and (3.61),

$$2n(Q_A+Q_B) \geq nH(A)+nH(B)+nH(AB)-I(W_A;W_B)-n\sqrt{3}\epsilon\log(d_A^2d_B^2d_R)-\eta(\sqrt{3}\epsilon). \tag{3.62}$$

Thus,

$$Q_A+Q_B \geq \frac{1}{2}\left[H(A)+H(B)+H(AB)-\frac{I(W_A;W_B)}{n}-\sqrt{3}\epsilon\log(d_A^2d_B^2d_R)-\frac{\eta(\sqrt{3}\epsilon)}{n}\right]. \tag{3.63}$$

Now, let $T : R^n \to R'$ be any CPTP map on R^n

$$\begin{aligned}
& I(W_A;W_B)-I(W_A;W_B|R') \\
= \;& (H(W_A)+H(W_B)-H(W_AW_B))-(H(W_A|R')+H(W_B|R')-H(W_AW_B|R')) \\
= \;& H(W_A)+H(W_B)-H(W_AW_B)-H(W_AR')+H(R') \\
& -H(W_BR')+H(R')+H(W_AW_BR')-H(R') \\
= \;& H(W_A)-H(W_AR')+H(W_B)-H(W_BR')-H(W_AW_B)+H(W_AW_BR') \\
\leq \;& -H(R')+n\sqrt{3}\epsilon\log(d)A^2d_R)+\eta(\sqrt{3}\epsilon)-H(B') \\
& +n\sqrt{3}\epsilon\log(d_B^2d_R)+\eta(\sqrt{3}\epsilon)+H(R')-H(W_AW_B)+H(W_AW_BR') \\
= \;& H(W_AW_BR')-H(W_AW_B)-H(R')+n\sqrt{3}\epsilon\log(d_A^2d_R) \\
& +n\sqrt{3}\epsilon\log(d_B^2d_R)+2\eta(\sqrt{3}\epsilon) \\
\leq \;& n\sqrt{3}\epsilon\log(d_A^2d_B^2d_R)+n\sqrt{3}\epsilon\log(d_A^2d_R)+n\sqrt{3}\epsilon\log(d_B^2d_R)+3\eta(\sqrt{3}\epsilon) \\
= \;& n\sqrt{3}\epsilon\log(d_A^4d_B^4d_R^2)+3\eta(\sqrt{3}\epsilon),
\end{aligned}$$

here we have used similar arguments as before to make rigorous the intuitions that W_AR', W_BR' and W_AW_BR' are almost uncorrelated with R', followed by the Fannes inequality.

We have

$$I(W_A;W_B) \leq I(W_A;W_B|R')+n\sqrt{3}\epsilon\log(d_A^4d_B^4d_R^2)+3\eta(\sqrt{3}\epsilon) \tag{3.64}$$

By the monotonicity of mutual information under local operations,

$$I(W_A; W_B) \leq I(A^n; B^n|R') + n\sqrt{3}\epsilon \log(d_A^4 d_B^4 d_R^2) + 3\eta(\sqrt{3}\epsilon). \tag{3.65}$$

Therefore,

$$\begin{aligned} Q_A + Q_B &\geq \frac{1}{2}[H(A) + H(B) + H(AB)] - \frac{1}{2}I(A;B|R) \\ &\quad -3\sqrt{3}\epsilon \log(d_A^4 d_B^4 d_R^2) - \frac{4\eta(\sqrt{3}\epsilon)}{n} - \sqrt{3}\epsilon \log(d_A^2 d_B^2 d_R) \\ &= \frac{1}{2}[H(A) + H(B) + H(AB)] - \frac{1}{2}I(A;B|R) \\ &\quad -\sqrt{3}\epsilon \log(d_A^6 d_B^6 d_R^3) - \frac{4\eta(\sqrt{3}\epsilon)}{n} \\ &\geq \frac{1}{2}[H(A) + H(B) + H(AB)] - E_{sq}(\varphi^{AB}), \end{aligned} \tag{3.66}$$

where $E_{sq}(\varphi^{AB})$ is the squashed entanglement of φ^{AB}, defined as the infimum of $\frac{1}{2}I(A;B|E)$ over extensions φ^{ABE} of φ^{AB} [81]. We have used explicitly the fact, proved in the cited paper, that $E_{sq}(\varphi^{\otimes n}) = nE_{sq}(\varphi)$.

We have therefore proved the following outer bound on the achievable rate region $\mathcal{SW}(\varphi)$:

$$\begin{aligned} Q_A &\geq \tfrac{1}{2}I(A;R)_\varphi \\ Q_B &\geq \tfrac{1}{2}I(B;R)_\varphi \\ Q_A + Q_B &\geq \tfrac{H(A)_\varphi + H(B)_\varphi + H(AB)_\varphi}{2} - E_{sq}(\varphi). \end{aligned} \tag{3.67}$$

In the special case where φ^{AB} is separable, $E_{sq}(\varphi) = 0$, which implies that the region defined by Eq. (3.40) is optimal. Under certain further technical assumptions, namely that φ^{AB} be the density operator of an ensemble of product pure states satisfying a condition called irreducibility, the same conclusion could be found in Ref. [82]. That paper, however, was unable to show that the bound was achievable.

The appearance of the squashed entanglement in (3.67) may seem somewhat mysterious, but a slight modification of the protocols based on fully quantum Slepian-Wolf will

lead to an inner bound on the achievable region that is of a similar form. Specifically, let $D_0(\varphi^{AB})$ be the amount of pure state entanglement that Alice and Bob can distill from φ^{AB} without engaging in any communication. Since this pure state entanglement is decoupled from the reference system R, they could actually perform this distillation process and discard the resulting entanglement before beginning one of their FQSW-based compression protocols. While neither $I(A;R)$ nor $I(B;R)$ would change, each of $H(A)$ and $H(B)$ would decrease by $D_0(\varphi^{AB})$. The corresponding inner bound on the achievable rate region $\mathcal{SW}(\varphi)$ would therefore be defined by the inequalities

$$\begin{aligned} Q_A &\geq \tfrac{1}{2} I(A;R)_\varphi \\ Q_B &\geq \tfrac{1}{2} I(B;R)_\varphi \\ Q_A + Q_B &\geq \tfrac{H(A)_\varphi + H(B)_\varphi + H(AB)_\varphi}{2} - D_0(\varphi). \end{aligned} \tag{3.68}$$

The only gap between the inner and outer bounds, therefore, is a gap between different measures of entanglement.

3.7 On efficiency

While the protocols described so far make use of a unitary transformation drawn at random according to the Haar measure, that is not essential. In fact, the only place the Haar measure was used was in the proof of Lemma 3.3.2. Therefore, the full unitary group could be replaced by any subset yielding the same average as in the lemma. (We thank Debbie Leung for alerting us to this possibility.) In fact, DiVincenzo, Leung and Terhal have shown that

$$\int_{\mathbb{U}(\mathbb{C}^{2^n})} (U \otimes U) X (U^\dagger \otimes U^\dagger)\, dU = \frac{1}{|G_n|} \sum_{g \in G} (g \otimes g) X (g^\dagger \otimes g^\dagger), \tag{3.69}$$

where G_n is the Clifford group on n qubits [83]. They also demonstrate in that paper that choosing an element of G_n from the uniform distribution can be done in time polynomial

in n. More specifically, they show that a random walk on a particular set of generators for G_n mixes in $O(n^8)$ time, leading to an associated quantum circuit for the selected element that is of size $O(n^2)$ gates.

Since the Schumacher compression portion of the fully quantum Slepian-Wolf protocol can also be done in polynomial time [84], we conclude that the encoding portion of the mother can be done efficiently. Since her immediate children, including entanglement distillation and state merging, are built by composing the mother with efficient protocols, namely superdense coding and teleportation, their encodings can also be found and implemented efficiently.

The transformation from FQSW to the father, however, included another non-constructive step, namely the choice of a good type class. Since the number of type classes is polynomial in the number of qubits in the input, however, that step could also be implemented efficiently. The corresponding isometries mapping the shared maximally entangled state and the input space into A_t can also be performed efficiently. Finally, while the proof presented here implies that the transpose of a random Clifford group element can be used as the encoding operation, there is in fact no need for the transpose because the Clifford group is closed under transposition. Thus, the encoding for the father can be found and implemented in polynomial time, as can those of his children, entanglement-assisted classical communication and quantum communication over a noisy channel.

Bibliography

[1] A. Abeyesinghe and P. Hayden. Generalized remote state preparation: Trading cbits, qubits, and ebits in quantum communication. *Phys. Rev. A*, 68(6):062319, 2003. arXiv:quant-ph/0308143.

[2] R. Ahlswede and G. Dueck. Identification via channels. *IEEE Trans. Inf. Theory*, 35(1):15–29, 1989.

[3] R. Ahlswede and A. Winter. Strong converse for identification via quantum channels. *IEEE Trans. Inf. Theory*, 48(3):569–579, 2002. arXiv:quant-ph/0012127.

[4] P. M. Alberti and A. Uhlmann. *Stochasticity and partial order*. Dordrecht, Boston, 1982.

[5] A. Ambainis and A. Smith. Small pseudo-random families of matrices: Derandomizing approximate quantum encryption. arXiv:quant-ph/0404075.

[6] C. Bennett and S.J. Wiesner. Communication via one- and two-particle operators on Einstein-Podolsky-Rosen states. *Phys. Rev. Lett.*, 69:2881, 1992.

[7] C. H. Bennett, D. P. DiVincenzo, P. W. Shor, J. A. Smolin, B. M. Terhal, and W. K. Wootters. Remote state preparation. *Phys. Rev. Lett.*, 87:077902, 2001. arXiv:quant-ph/0006044.

[8] C. H. Bennett, P. Hayden, D. Leung, P. W. Shor, and A. Winter. Remote preparation of quantum states. arXiv:quant-ph/0307100.

[9] D.W. Berry. Resources required for exact remote state preparation. arXiv:quant-ph/0404004.

[10] I. Devetak and T. Berger. Low-entanglement remote state preparation. *Phys. Rev. Lett.*, 87(9):197901, 2001. arXiv:quant-ph/0102123.

[11] I. Devetak, A. W. Harrow, and A. Winter. A family of quantum protocols. arXiv:quant-ph/0308044.

[12] M. Fannes. A continuity property of the entropy density for spin lattice systems. *Comm. Math. Phys.*, 31:291–294, 1973.

[13] S. K. Foong and S. Kanno. Proof of Page's conjecture on the average entropy of a subsystem. *Phys. Rev. Lett.*, 72:1148–1151, 1994.

[14] A. Harrow. Coherent communication of classical messages. *Phys. Rev. Lett.*, 92(9):097902, 2004. arXiv:quant-ph/0307091.

[15] A. Harrow, P. Hayden, and D. Leung. Superdense coding of quantum states. *Phys. Rev. Lett.*, 92:187901, 2004. arXiv:quant-ph/0307221.

[16] P. Hayden, R. Jozsa, and A. Winter. Trading quantum for classical resources in quantum data compression. *J. Math. Phys.*, 43(9):4404–4444, 2002. arXiv:quant-ph/0204038.

[17] P. Hayden, D. Leung, P.W. Shor, and A. Winter. Randomizing quantum states: Constructions and applications. arXiv:quant-ph/0307104, 2003. *Commun. Math. Phys.*, to appear.

[18] P. Hayden, D. Leung, and A. Winter. Aspects of generic entanglement. arXiv:quant-ph/0407049.

[19] R. Jozsa. Fidelity for mixed quantum states. *J. Mod. Opt.*, 41:2315–2323, 1994.

[20] S. Lloyd and H. Pagels. Complexity as thermodynamic depth. *Ann. Phys.*, 188(1):186–213, 1988.

[21] H.-K. Lo. Classical communication cost in distributed quantum information processing – A generalization of quantum communication complexity. *Phys. Rev. A*, 62:012313, 2000. arXiv:quant-ph/9912009.

[22] E. Lubkin. Entropy of an n-system from its correlation with a k-reservoir. *J. Math. Phys.*, 19:1028–1031, 1978.

[23] V. D. Milman and G. Schechtman. *Asymptotic theory of finite dimensional normed spaces*, volume 1200 of *Lecture Notes in Mathematics*. Springer-Verlag, 1986.

[24] M. Nielsen. Continuity bounds for entanglement. *Phys. Rev. A*, 61:064301, 2000.

[25] M. A. Nielsen and I. L. Chuang. *Quantum computation and quantum information*. Cambridge University Press, Cambridge, 2000.

[26] M. Ohya and D. Petz. *Quantum entropy and its use*. Texts and monographs in physics. Springer-Verlag, Berlin, 1993.

[27] D. Page. Average entropy of a subsystem. *Phys. Rev. Lett.*, 71:1291–1294, 1993.

[28] A.K. Pati. Minimum classical bit for remote preparation and measurement of a qubit. *Phys. Rev. A*, 63:14302, 2001. arXiv:quant-ph/9907022.

[29] M. O. Rabin and A. C.-C. Yao. Unpublished manuscript. Cited in A. C.-C. Yao, "Some questions on the complexity of distributive computing," in : *Proc. 11th ACM Symposium on the Theory of Computing*, 209–213, 1979.

[30] J. Sanchez-Ruiz. Simple proof of Page's conjecture on the average entropy of a subsystem. *Phys. Rev. E*, 52(5):5653, 1995.

[31] S. Sen. Average entropy of a quantum subsystem. *Phys. Rev. Lett.*, 77(1):1–3, 1996.

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

[33] A. Winter. Identification via quantum channels in the presence of prior correlation and feedback. arXiv:quant-ph/0403203.

[34] A. Winter. Quantum and classical message identification via quantum channels. arXiv:quant-ph/0401060.

[35] A. Winter. Coding theorem and strong converse for quantum channels. *IEEE Trans. Inf. Theory*, 45:2481–2485, 1999.

[36] M.-Y. Ye, Y.-S. Zhang, and G.-C. Guo. Faithful remote state preparation using finite classical bits and a nonmaximally entangled state. *Phys. Rev. A*, 69:022310, 2004. arxiv:quant-ph/0307027.

[37] B. Schumacher. Quantum coding. *Phys. Rev. A*, 51:2738–2747, 1995.

[38] M. Ohya and D. Petz. *Quantum entropy and its uses.* Texts and monographs in physics. Springer-Verlag, Berlin, 1993.

[39] C. H. Bennett, H. J. Bernstein, S. Popescu, and B. Schumacher. Concentrating partial entanglement by local operations. *Phys. Rev. A*, 53(4):2046–2052, 1996.

[40] G. Kuperberg. The capacity of hybrid quantum memory. arXiv quant-ph/0203105, 2002.

[41] C. H. Bennett, I. Devetak, A. Harrow, P. W. Shor, and Winter A. The quantum reverse Shannon theorem. In preparation.

[42] H. Barnum, P. Hayden, R. Jozsa, and A. Winter. On the reversible extraction of classical information from a quantum source. *Proc. R. Soc. London Ser. A*, 457(2012):2019–2039, 2001. arXiv quant-ph/0011072.

[43] M. Koashi and N. Imoto. Operations that do not disturb partially known quantum states. *Phys. Rev. A*, 66(2):022318, 2002. arXiv quant-ph/0101144.

[44] P. Hayden, R. Jozsa, and A. Winter. Trading quantum for classical resources in quantum data compression. *J. Math. Phys.*, 43(9):4404–4444, 2002. arXiv quant-ph/0204038.

[45] H.-K. Lo. Classical communication cost in distributed quantum information processing – A generalization of quantum communication complexity. *Phys. Rev. A*, 62:012313, 2000. arXiv quant-ph/9912009.

[46] C. H. Bennett, D. P. DiVincenzo, P. W. Shor, J. A. Smolin, B. M. Terhal, and W. K. Wootters. Remote state preparation. *Phys. Rev. Lett.*, 87:077902, 2001. arXiv quant-ph/0006044.

[47] C. H. Bennett and S. Wiesner. Communication via one- and two-particle operators on Einstein-Podolsky-Rosen states. *Phys. Rev. Lett.*, 69(20):2881–2884, 1992.

[48] A. Harrow, P. Hayden, and D. Leung. Superdense coding of quantum states. arXiv quant-ph/0307221, 2003.

[49] A. S. Holevo. Bounds for the quantity of information transmitted by a quantum communication channel. *Problemy Peredachi Informatsii*, 9(3):3–11, 1973.

[50] C. H. Bennett, P. Hayden, D. Leung, P. W. Shor, and A. Winter. Remote preparation of quantum states. arXiv quant-ph/0307100, 2003.

[51] A. Harrow. Coherent classical communication. arXiv quant-ph/0307091, 2003.

[52] C. H. Bennett, P. W. Shor, J. A. Smolin, and A. V. Thapliyal. Entanglement-assisted classical cpacity of noisy quantum channels. *IEEE Trans. Inf. Theory*, 48(10):2637–2655, 2002. arXiv quant-ph/0106052.

[53] M. Horodecki. Optimal compression for mixed signal states. *Phys. Rev. A*, 61:052309, 2000. arXiv quant-ph/9905058.

[54] H. Barnum, C. M. Caves, C. A. Fuchs, R. Jozsa, and B. W. Schumacher. On quantum coding for ensembles of mixed states. *J. Phys. A: Math. and Gen.*, 34(35):6767–6785, 2001. arXiv quant-ph/0008024.

[55] A. Winter. Compression of sources of probability distributions and density operators. arXiv quant-ph/0208131, 2002.

[56] I. Devetak and T. Berger. Low-entanglement remote state preparation. *Phys. Rev. Lett.*, 87(9):197901, 2001. arXiv quant-ph/0102123.

[57] B. Schumacher and M. D. Westmoreland. Sending classical information via noisy quantum channels. *Phys. Rev. A*, 56:131–138, 1997.

[58] A. S. Holevo. The capacity of the quantum channel with general signal states. *IEEE Trans. Inf. Theory*, 44:269–273, 1998.

[59] M. Fannes. A continuity property of the entropy density for spin lattice systems. *Communications of mathematical physics*, 31:291–294, 1973.

[60] N. J. Cerf and C. Adami. Quantum conditional operator and a criterion for separability. *Phys. Rev. A*, 60:893–898, 1999. arXiv quant-ph/9710001.

[61] P. Horodecki, R. Horodecki, and M. Horodecki. Entanglement and thermodynamical analogies. *Acta Phys. Solv.*, 48:141, 1998. arXiv quant-ph/9805072.

[62] A. S. Holevo. The capacity of the quantum channel with general signal states. *IEEE Trans. Inf. Theory*, 44:269–273, 1998.

[63] B. Schumacher and M. D. Westmoreland. Sending classical information via noisy quantum channels. *Phys. Rev. A*, 56:131–138, 1997.

[64] C. H. Bennett, P. W. Shor, J. A. Smolin, and A. V. Thapliyal. Entanglement-assisted classical capacity of noisy quantum channels. *Phys. Rev. Lett.*, 83:3081, 1999. arXiv.org:quant-ph/9904023.

[65] C. H. Bennett, P. W. Shor, J. A. Smolin, and A. V. Thapliyal. Entanglement-assisted capacity of a quantum channel and the reverse Shannon theorem. *IEEE Trans. Inf. Th.*, 48(10):2637, 2002. arXiv.org:quant-ph/0106052.

[66] S. Lloyd. Capacity of the noisy quantum channel. *Phys. Rev. A*, 55:1613, 1996.

[67] P. W. Shor. The quantum channel capacity and coherent information. Lecture notes, MSRI workshop on quantum computation, 2002. Available at http://www.msri.org/publications/ln/msri/2002/quantumcrypto/shor/1/.

[68] I. Devetak. The private classical capacity and quantum capacity of a quantum channel. *IEEE Trans. Inf. Th.*, 51(1):44, 2005. arXiv.org/0304127.

[69] I. Devetak and A. Winter. Distillation of secret key and entanglement from quantum states. *Proc. R. Soc. Lond. A*, 461:207–237, 2005. arXiv.org:quant-ph/0306078.

[70] M. Horodecki, P. Horodecki, R. Horodecki, D. W. Leung, and B. M. Terhal. Classical capacity of a noiseless quantum channel assisted by noisy entanglement. *Quant. Inf. Comp.*, 1:70–78, 2001. arXiv.org:quant-ph/0106080.

[71] A. W. Harrow. Coherent communication of classical messages. *Phys. Rev. Lett.*, 92:097902, 2004. arXiv.org:quant-ph/0307091.

[72] I. Devetak, A. W. Harrow, and A. Winter. A family of quantum protocols. *Phys. Rev. Lett.*, 93:230504, 2004. arXiv.org:quant-ph/0308044.

[73] M. Horodecki, J. Oppenheim, and A. Winter. Partial quantum information. *Nature*, 436:673–676, 2005. arXiv.org:quant-ph/0505062.

[74] I. Devetak. A triangle of dualities: reversibly decomposable quantum channels, source-channel duality, and time reversal. arXiv.org:quant-ph/0505138, 2005.

[75] I. Devetak, A. W. Harrow, and A. Winter. A resource framework for quantum Shannon theory. arXiv.org:quant-ph/0512015, 2005.

[76] B. Groisman, S. Popescu, and A. Winter. On the quantum, classical and total amount of correlations in a quantum state. *Phys. Rev. A*, 72:032317, 2005. arXiv.org:quant-ph/0410091.

[77] C. A. Fuchs and J. van de Graaf. Cryptographic distinguishability measures for quantum mechanical states. *IEEE Trans. Inf. Theory*, 45:1216–1227, 1999.

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

[79] R. Jozsa. Fidelity for mixed quantum states. *J. Mod. Opt.*, 41:2315–2323, 1994.

[80] D. Slepian and J. K. Wolf. Noiseless coding of correlated information sources. *IEEE Trans. Inf. Theory*, 19:461–480, 1971.

[81] M. Christandl and A. Winter. Squashed entanglement – An additive entanglement measure. *J. Math. Phys.*, 45(3):829–840, 2004. arXiv.org:quant-ph/0308088.

[82] C. Ahn, A. Doherty, P. Hayden, and A. Winter. On the distributed compression of quantum information. arXiv.org:quant-ph/0403042, 2004.

[83] D.P. DiVincenzo, D. W. Leung, and B. M. Terhal. Quantum data hiding. *IEEE Trans. Inf. Theory*, 48(3):580–598, 2002. arXiv.org:quant-ph/0103098.

[84] R. Cleve and D. P. DiVincenzo. Schumacher's quantum data compression as a quantum computation. *Phys. Rev. A*, 54:2636–2650, 1996. arXiv.org:quant-ph/9603009.