PROCEEDINGS
OF
THE ROYAL
SOCIETY A

Proc. R. Soc. A (2009) **465**, 2537–2563
doi:10.1098/rspa.2009.0202
Published online 5 June 2009

The mother of all protocols: restructuring quantum information's family tree

By Anura Abeyesinghe[1], Igor Devetak[2], Patrick Hayden[3] and Andreas Winter[4,*]

[1] *Physics Department, Institute for Quantum Information, Caltech 103-33, Pasadena, CA 91125, USA*
[2] *Electrical Engineering Department, University of Southern California, Los Angeles, CA 90089, USA*
[3] *School of Computer Science, McGill University, Montreal, Quebec H3A 2A7, Canada*
[4] *Department of Mathematics, University of Bristol, University Walk, Bristol BS8 1TW, UK*

We give a simple, direct proof of the 'mother' protocol of quantum information theory. In this new formulation, it is easy to see that the mother, or rather her generalization to the fully quantum Slepian–Wolf protocol, 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, a fully quantum reverse Shannon theorem, and 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 whose children provide the quantum capacity and the entanglement-assisted capacity of a quantum channel, 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.

Keywords: quantum information; coding; entanglement

1. Introduction

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 10 years have seen major advances in the area, including, among many other discoveries, the determination of the classical capacity of a quantum channel (Schumacher & Westmoreland 1997; Holevo 1998),

*Author for correspondence (a.j.winter@bris.ac.uk).

Received 16 April 2009
Accepted 1 May 2009

the capacities of entanglement-assisted channels (Bennett *et al.* 1999, 2002), the quantum capacity of a quantum channel (Lloyd 1996; Shor 2002; Devetak 2005) and the best ways to use noisy entanglement to extract pure entanglement (Devetak & Winter 2005) or to help send classical information (Horodecki *et al.* 2001). Until recently, however, each new problem was solved essentially from scratch, and no higher-level structure connecting the different results was known. Harrow's (2004) introduction of the *cobit* and its subsequent application to the construction of the so-called 'mother' and 'father' protocols provided that missing structure. Almost all the problems listed above were shown to fall into two families: first the mother and her descendants, and second the father and his (Devetak *et al.* 2004). Appending or prepending simple transformations such as teleportation and superdense coding are suffice 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. As destruction is a relatively indiscriminate goal, the resulting proof is correspondingly simple. This approach also makes it clear that the mother actually accomplishes 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 very important in its own right and also the major focus of our paper. To start with, it places the state-merging protocol of Horodecki *et al.* (2005*a*, 2007) squarely within the mother's brood. In addition, it makes it possible to use the mother as a building block for distributed compression. We analyse 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.

We also emphasize a further connection, first identified in Devetak (2006), that requires both the state transfer and entanglement-distillation capabilities of the mother: the entire protocol allows for a time-reversed interpretation as a quantum reverse Shannon theorem, i.e. an efficient simulation of a noisy quantum channel using a noiseless quantum channel along with entanglement.

Finally, the new approach to the mother solves a major problem left unanswered in the original family paper. There, no operational relationship between the mother and father protocols could be identified, but the two were nonetheless connected by a formal symmetry called *source-channel duality* (Devetak 2006). This new mother protocol can directly be transformed into the father, resolving the mystery of the two parents' formal similarity and merging the two families into one.

The structure of the paper is as follows. After reviewing the family of quantum protocols in §2 and providing in §3 a high-level description of the improved mother, henceforth the fully quantum Slepian–Wolf (FQSW) protocol, we go straight to the statement and proof of the central result of this paper in §4: a one-shot version of FQSW. The middle section of the paper is devoted to a number of applications of one-shot FQSW. Sections 5 and 6 describe one-shot versions of the 'father' and the fully quantum 'quantum reverse Shannon' (FQRS) protocol, respectively. The one-shot theorems quickly yield memoryless forms for all three: FQSW in §7, the father in §8 and FQRS in §9. Then we turn to the other highlight of this paper, a treatment of the fully quantum version of the distributed compression problem, which we can solve completely for a large class of sources by providing general inner and outer bounds on the rate region, in §10. In §11, we point out that the FQSW protocol allows for efficient encoding via Clifford operations, after which we conclude. An appendix collects useful facts on typical subspaces.

Notation: 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. Logarithms are taken base 2 throughout.

2. 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

$$\langle \varphi^{AB} \rangle + \frac{1}{2} I(A;R)_\varphi \, [q \to q] \geq \frac{1}{2} I(A;B)_\varphi \, [qq]. \tag{2.1}$$

We will informally explain the resource inequalities used here, but the reader is directed to Devetak *et al.* (2008) for a rigorous treatment. Here $[q \to q]$ 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 $I(A;B)_\varphi$ EPR pairs per copy, provided Alice is allowed to communicate with Bob by sending him qubits at the rate $I(A;R)_\varphi$ per copy. Small imperfections in the final state are permitted provided they vanish as n goes to infinity.

In this paper, we prove a stronger resource inequality which we call the FQSW inequality. The justification for this name will become apparent in §10, where we study its applicability to distributed compression, solved classically by Slepian & Wolf (1971). 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, the two parties 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,

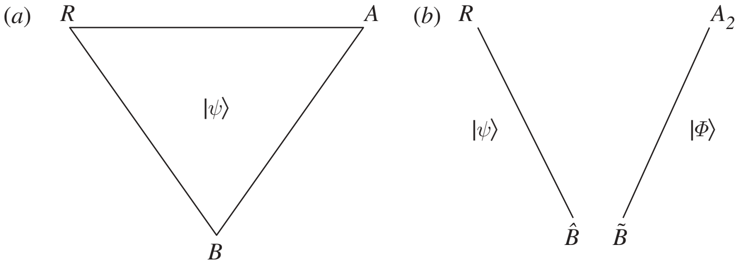

Figure 1. (*a*) The starting point for the FQSW protocol, a pure tripartite entangled state $|\psi\rangle = (|\varphi\rangle^{ABR})^{\otimes n}$. (*b*) After execution of the protocol, Alice's portion of the original tripartite state has been transferred to Bob, so that Bob holds a purification of the unchanged reference system in his register $\hat{B}$. He also shares pure state entanglement with Alice in the form of the state $|\Phi\rangle$.

while simultaneously distilling ebits with him. A graphical depiction of this transformation is given in figure 1. The process can also be expressed as a resource inequality in the following way:

$$\langle W^{S\to AB} : \varphi^S\rangle + \frac{1}{2}I(A;R)_\varphi\,[q\to q] \geq \frac{1}{2}I(A;B)_\varphi\,[qq] + \langle \mathrm{id}^{S\to\hat{B}} : \varphi^S\rangle. \tag{2.2}$$

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, $W^{S\to AB}$ is an isometry taking the system S to AB. Thus, on the left-hand side of the inequality, a state is distributed to Alice and Bob, whereas 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.

As 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 obtain the following resource inequality:

$$\langle W^{S\to AB} : \varphi^S\rangle + H(A|B)_\varphi\,[q\to q] + I(A;B)_\varphi\,[c\to c] \geq \langle \mathrm{id}^{S\to\hat{B}} : \varphi^S\rangle, \tag{2.3}$$

which is known as the *state-merging* primitive (Horodecki *et al.* 2005*a*). It is of note both because it is a useful building block for multiparty protocols (Horodecki *et al.* 2005*a*, 2007; Yard *et al.* 2007) 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 are connected by 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 $|\varphi\rangle^{AA'}$.

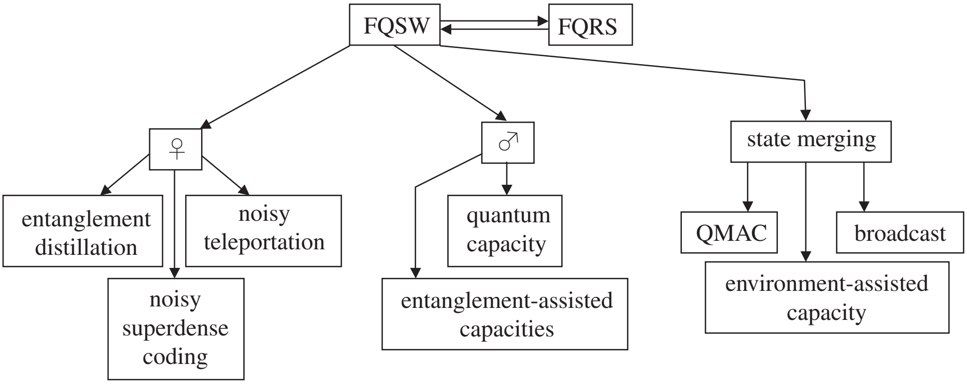

Figure 2. Partial quantum information theory family tree. The symbols ♀ and ♂ represent the ‘old’ mother and father protocols from Devetak *et al.* (2004) and arrows indicate that a protocol accomplishing the task at the start of the arrow can be transformed into a protocol accomplishing the task at the end. The relationships between ♀, ♂ and their children are discussed in detail in Devetak *et al.* (2004, 2008). ‘QMAC’ refers to the task of sending quantum data through a quantum multiple access channel (Horodecki *et al.* 2005*a*; Yard *et al.* 2008), ‘broadcast’ the task of sending quantum data through a quantum broadcast channel Yard *et al.* (2006) and the environment-assisted quantum capacity is discussed in Smolin *et al.* (2005).

The resource inequality is

$$\langle \mathcal{N}^{A'\to B}\rangle + \frac{1}{2}I(A;E)_{\varphi}\,[qq] \geq \frac{1}{2}I(A;B)_{\varphi}\,[q\to q]. \tag{2.4}$$

Thus, Alice and Bob use pre-existing shared entanglement and the noisy channel to produce noiseless quantum communication. Comparison of equation (2.4) to the mother, equation (2.1) 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 formal symmetry is known as source-channel duality (Devetak 2006). Just as the mother can be strengthened to the FQSW protocol, there is a fully coherent version of the father known as the *feedback father* (Devetak 2006).

The relationships between different protocols are sketched as a family tree in figure 2.

3. The fully quantum Slepian–Wolf protocol

The input to the FQSW 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, whereas 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 $nI(A;B)_{\varphi} - o(n)$ qubits. The steps in the protocol that transform the input state to the output state are as follows:

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

(ii) Alice applies a unitary transformation U_A to A_S and then sends A_1 to Bob.
(iii) Bob applies an isometry V_B taking $A_1 B^n$ 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. As essentially no information is leaked to the environment at any step, Bob will hold a purification of the $A_2 R^n$ system after step (ii), 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 $A_2 R^n$ approximates $\Phi^{A_2} \otimes (\varphi^R)^{\otimes n}$ after step (ii). Bob's isometry V_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, whereas 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 $nI(A;R)_\varphi$ bits of mutual information, which means that A_1 must consist of at least $n/2I(A;R)_\varphi$ qubits. We will see that by choosing U_A randomly according to the Haar measure, we will come close to achieving this rate.

The result is similar in spirit to a recent result of Groisman *et al.* (2005) 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: $I(A;R)_\varphi$ cbits per copy. In fact, it is clear that we can derive that result from ours: after Alice's unitary, the state remaining between A_2 and R is almost a product because Alice's entanglement with the reference gets transferred to Bob, so Alice only needs to discard the system A_1 of roughly $n/2I(A;R)$ qubits, which she can do by erasing it entirely via random Pauli operations, at randomness cost amounting to $I(A;R)$ cbits per copy.

4. Fully quantum Slepian–Wolf: one-shot version

Although the tensor power structure of $(|\varphi\rangle^{ABR})^{\otimes n}$ allows the FQSW inequality (2.2) 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 FQSW result that leads quickly to inequality (2.2) 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 $d_A = d_{A_1} d_{A_2}$.

The following inequality is the one-shot version of FQSW.

Theorem 4.1 (One-shot, FQSW bound). *There exist isometries $U^{A\to A_1A_2}$ and $V^{A_1B\to\hat{B}\tilde{B}}$ such that*

$$\left\|(V\circ U)\psi^{RAB}(V\circ U)^{\dagger}-\psi^{R\hat{B}}\otimes\Phi^{A_2\tilde{B}}\right\|_1\leq 2\left[\frac{d_Ad_R}{d_{A_1}^2}\mathrm{Tr}[(\psi^{AR})^2]\right]^{1/4},$$

where $W^{\hat{B}\to AB}|\psi\rangle^{R\hat{B}}=|\psi\rangle^{RAB}$ for some isometry W.

The protocol corresponding to the above theorem consists of Alice performing U, sending the system A_1 to Bob, and Bob performing V. The number of qubit channels used up is $\log d_{A_1}$, whereas the number of ebits distilled is $\log d_{A_2}=\log d_A-\log d_{A_1}$.

The main ingredient is the following decoupling theorem.

Theorem 4.2 (Decoupling). *Let $\sigma^{A_2R}(U)=\mathrm{Tr}_{A_1}[(U\otimes I^R)\psi^{AR}(U^{\dagger}\otimes I^R)]$ be the state remaining on A_2R after the unitary transformation U has been applied to $A=A_1A_2$. Then*

$$\int_{\mathbb{U}(A)}\left\|\sigma^{A_2R}(U)-\frac{I^{A_2}}{d_{A_2}}\otimes\sigma^R(U)\right\|_1^2\mathrm{d}U\leq\frac{d_Ad_R}{d_{A_1}^2}\mathrm{Tr}[(\psi^{AR})^2]. \tag{4.1}$$

The theorem quantifies how distinguishable $\sigma^{A_2R}(U)$ will be from the completely decoupled state $I^{A_2}/d_{A_2}\otimes\sigma^R(U)$ 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 purity $\mathrm{Tr}[(\psi^{AR})^2]$. To assure good decoupling, it suffices that

$$\log d_{A_1}\gg\frac{1}{2}\left[\log d_A+\log d_R+\log\mathrm{Tr}[(\psi^{AR})^2]\right]. \tag{4.2}$$

In the tensor power source setting, both dimensions and purities can be tightly bounded by functions of the corresponding entropies. When that is the case, the expression on the right-hand side of equation (4.2) plays the role of the $\frac{1}{2}I(A;R)=\frac{1}{2}[H(A)+H(R)-H(AR)]$ from the FQSW resource inequality (2.2).

According to the proof strategy outlined in the previous section, if $\sigma^{A_2R}(U)$ is close to $I^{A_2}(U)/d_{A_2}\otimes\sigma^R(U)$, then $\sigma^{A_2R}(U)$ has a purification that is itself close to a product state. This argument will be made quantitative in the proof of theorem 4.1.

The proof of theorem 4.2 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 (4.1). The evaluations of the relevant averages are mechanical but slightly lengthy calculations.

Before starting in earnest, we perform a calculation whose result will allow us to evaluate all necessary averages over the unitary group. Recall from the notation summary from §1 that F^{A_2R} is the operator that swaps the composite system A_2R with a duplicate composite system $A_2'R'$ and that $\Pi_{+(-)}^{A}$ is the projector onto the (anti-)symmetric subspace of AA'.

Lemma 4.3. *For* $A = A_1A_2$,

$$\int_{\mathbb{U}(A)} (U^A \otimes U^{A'} \otimes I^{RR'})^\dagger (I^{A_1} \otimes F^{A_2R})(U^A \otimes U^{A'} \otimes I^{RR'})\,\mathrm{d}U$$
$$= [p\Pi_+^A + q\Pi_-^A] \otimes F^R, \tag{4.3}$$

where

$$p = \frac{d_{A_1} + d_{A_2}}{d_A + 1} \quad \text{and} \quad q = \frac{d_{A_1} - d_{A_2}}{d_A - 1}. \tag{4.4}$$

Proof. Let X be Hermitian. By Schur–Weyl duality,

$$\int_{\mathbb{U}(A)} (U^\dagger \otimes U^\dagger) X (U \otimes U)\,\mathrm{d}U = \alpha_+(X)\Pi_+^A + \alpha_-(X)\Pi_-^A \tag{4.5}$$

with the coefficients $\alpha_\pm(X) = \mathrm{Tr}(X\Pi_\pm^A)/\mathrm{rank}(\Pi_\pm^A)$. Recall that $\Pi_\pm^A = \frac{1}{2}(I^{AA'} \pm F^A)$.

$$\begin{aligned}\mathrm{rank}(\Pi_\pm^A)\,\alpha_\pm(F^{A_2}) &= \mathrm{Tr}(\Pi_\pm^A F^{A_2})\\ &= \frac{1}{2}\mathrm{Tr}[(I^{AA'} \pm F^{A_1} \otimes F^{A_2})F^{A_2}]\\ &= \frac{1}{2}[\mathrm{Tr}(I^{A_1A_1'} \otimes F^{A_2}) \pm \mathrm{Tr}(F^{A_1} \otimes I^{A_2A_2'})]\\ &= \frac{1}{2}[d_{A_1}^2 d_{A_2} \pm d_{A_1} d_{A_2}^2].\end{aligned} \tag{4.6}$$

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 $\mathrm{rank}(\Pi_\pm^A) = d_A(d_A \pm 1)/2$. ■

The decoupling theorem is a direct application of this formula.

Proof of theorem 4.2. First note that

$$\left\| \sigma^{A_2R}(U) - \frac{I^{A_2}}{d_{A_2}} \otimes \sigma^R(U) \right\|_2^2 = \mathrm{Tr}[(\sigma^{A_2R}(U))^2] - \frac{1}{d_{A_2}}\mathrm{Tr}[(\sigma^R(U))^2]. \tag{4.7}$$

As $\sigma^R = \psi^R$, the second term on the right-hand side is independent of U and it is sufficient to calculate

$$\begin{aligned}&\int_{\mathbb{U}(A)} \mathrm{Tr}[(\sigma^{A_2R}(U))^2]\,\mathrm{d}U\\ &\quad = \int \mathrm{Tr}[(\sigma^{A_2R}(U) \otimes \sigma^{A_2'R'}(U))F^{A_2R}]\,\mathrm{d}U\\ &\quad = \int \mathrm{Tr}[(\mathrm{Tr}_{A_1}(U\psi^{AR}U^\dagger) \otimes \mathrm{Tr}_{A_1'}(U\psi^{A'R'}U^\dagger))F^{A_2R}]\,\mathrm{d}U\end{aligned}$$

$$\begin{aligned}
&= \mathrm{Tr}[(\psi^{AR} \otimes \psi^{A'R'}) \cdot \int (U^\dagger \otimes U^\dagger)(I^{A_1A_1'} \otimes F^{A_2R})(U \otimes U)\,\mathrm{d}U] \\
&= \mathrm{Tr}[(\psi^{AR} \otimes \psi^{A'R'}) \cdot (p\Pi_+^A + q\Pi_-^A) \otimes F^R] \\
&= \frac{p-q}{2}\mathrm{Tr}[(\psi^{AR})^2] + \frac{p+q}{2}\mathrm{Tr}[(\psi^R)^2], \qquad (4.8)
\end{aligned}$$

where p and q are defined as in equation (4.4). In the fourth line, we have used the result of lemma 4.3, and in the fifth the identity $\Pi_\pm^A = \frac{1}{2}(I^{AA'} \pm F^A)$. But

$$\frac{p-q}{2} = \frac{d_{A_1}d_{A_2}^2 - d_{A_1}}{d_A^2 - 1} \le \frac{1}{d_{A_1}} \qquad (4.9)$$

holds for all $d_{A_1}, d_{A_2} \ge 1$. Likewise, $(p+q)/2 \le 1/d_{A_2}$ under the same assumption. This when compared with equation (4.7) shows that we can drop the $\mathrm{Tr}[(\psi^R)^2]$ term to obtain

$$\int_{\mathbb{U}(A)} \left\| \sigma^{A_2R}(U) - \frac{I^{A_2}}{d_{A_1}} \otimes \sigma^R(U) \right\|_2^2 \mathrm{d}U \le \frac{1}{d_{A_1}}\mathrm{Tr}[(\psi^{AR})^2]. \qquad (4.10)$$

The decoupling theorem then follows by the Cauchy–Schwarz inequality: $\|\cdot\|_1^2 \le d_{A_2}d_R\|\cdot\|_2^2$. ■

Theorem 4.1 is then an easy corollary exploiting the fact that a product mixed state has a product purification.

Proof of theorem 4.1. Observe that there exists a particular U such that $\|\sigma^{A_2R} - I^{A_2}/d_{A_2} \otimes \sigma^R\|_1^2$ is bounded as in equation (4.1). The final ingredient is Uhlmann's theorem (Uhlmann 1976), in the version of lemma 2.2 of Devetak *et al.* (2008): If $\|\rho^C - \sigma^C\|_1 \le \epsilon$, where ρ^{BC} is a purification of ρ^C, and σ^{DC} is a purification of σ^C, then there exists an isometry $V^{B\to D}$ such that $\|(V^B \otimes I^C)\rho^{BC}(V^B \otimes I^C)^\dagger - \sigma^{BC}\|_1 \le 2\sqrt{\epsilon}$. As $\Phi^{A_2\tilde{B}} \otimes \psi^{R\hat{B}}$ is a purification of $I^{A_2}/d_{A_2} \otimes \sigma^R$, and $U\psi^{RAB}U^\dagger$ is a purification of σ^{A_2R}, there is an isometry $V^{A_1B\to\tilde{B}\hat{B}}$ such that the statement of the theorem holds. ■

5. Father from FQSW: one-shot version

A few simple observations will allow us to transform the one-shot FQSW protocol into a one-shot father protocol. The father implements entanglement-assisted noiseless quantum communication over a noisy channel $\mathcal{N}^{A\to B}$. The protocol consumes (maximal) entanglement initally shared between Alice and Bob, and in registers we will call A_3 and B_3. Mathematically, we verify that the protocol implements noiseless quantum communication by applying it to one-half of a maximally entangled state, the other half of which is held by a reference system R. This is equivalent to verifying that after the application of $\mathcal{N}^{A\to B}$, the reference system R is decoupled from the channel's environment E. In the one-shot FQSW protocol, the objective was to decouple R and A_2.

Precisely, we apply the FQSW protocol to the following state:

$$|\psi\rangle^{B_3RBE} = U_{\mathcal{N}}^{A\to BE} \circ W_1^{A_0A_3\to A}(|\Phi_0\rangle^{RA_0}|\Phi_3\rangle^{B_3A_3}) \tag{5.1}$$

for a Stinespring dilation $U_{\mathcal{N}}^{A\to BE}$ of the noisy channel $\mathcal{N}^{A\to B}$ and some fixed isometry $W_1^{A_0A_3\to A}$. (The latter will allow us to choose which part of the input system we want to use for encoding.) Note that as a product of two maximally entangled states, ${\Phi_0}^{RA_0} \otimes {\Phi_3}^{B_3A_3}$ really is a single maximally entangled state between B_3R and A_3A_0.

Now we make the corresponding replacements in theorem 4.1:

Father	FQSW
B_3	A_1
R	A_2
B_3R	A
E	R

Thus, there exist isometries $U^{B_3R\to B_3R}$ and $V^{B_3B\to\hat{B}\tilde{B}}$ such that

$$\left\|(V\circ U)\psi^{B_3RBE}(V\circ U)^\dagger - \psi^{\hat{B}E}\otimes\Phi_0^{R\tilde{B}}\right\|_1 \le 2\left[\frac{d_{B_3R}\,d_E}{d_{B_3}^2}\mathrm{Tr}[(\psi^{B_3RE})^2]\right]^{1/4},$$

where an appropriate unitary equivalence between $\hat{B}$ and B_3RB allows us to identify $|\psi\rangle^{B_3RBE}$ with $|\psi\rangle^{\hat{B}E}$.

As $V^{B_3B\to\hat{B}\tilde{B}}$ acts entirely on systems held by Bob, it could be performed by him as a decoding operation. The isometry $U^{B_3R\to B_3R}$, on the other hand, acts on the reference system, which is not allowed to participate actively in the protocol. The situation up to this point is depicted in figure 3. However, because $\Phi_0^{RA_0}\otimes\Phi_3^{B_3A_3}$ is maximally entangled between A_3A_0 and B_3R,

$$U^{B_3R\to B_3R}(|\Phi_0\rangle^{RA_0}|\Phi_3\rangle^{B_3A_3}) = (U^\top)^{A_3A_0\to A_3A_0}(|\Phi_0\rangle^{RA_0}|\Phi_3\rangle^{B_3A_3}),$$

where $\top$ denotes transposition. Thus, the effect of U can be achieved by acting instead with $U^\top$ on A_3A_0, systems held by Alice. Defining $W_2^{A_0A_3\to A} := W_1\circ U^\top$, we obtain

$$\begin{aligned}&\left\|(V\circ U_{\mathcal{N}}\circ W_2)(\Phi_0^{RA_0}\otimes\Phi_3^{B_3A_3})(V\circ U_{\mathcal{N}}\circ W_2)^\dagger - \psi^{\hat{B}E}\otimes\Phi_0^{R\tilde{B}}\right\|_1\\ &\quad\le 2\left[\frac{d_{A_0A_3}\,d_E}{d_{A_3}^2}\mathrm{Tr}[(\psi^B)^2]\right]^{1/4}.\end{aligned} \tag{5.2}$$

This is precisely the setting of the father protocol, as illustrated in figure 4.

Alice needs to transfer the purification of some maximally mixed state Φ_0^R to Bob. The resources at their disposal are the channel $\mathcal{N}^{A\to B}$ and a maximally entangled state $\Phi_3^{B_3A_3}$. Alice performs the encoding W_2, sends the resulting state

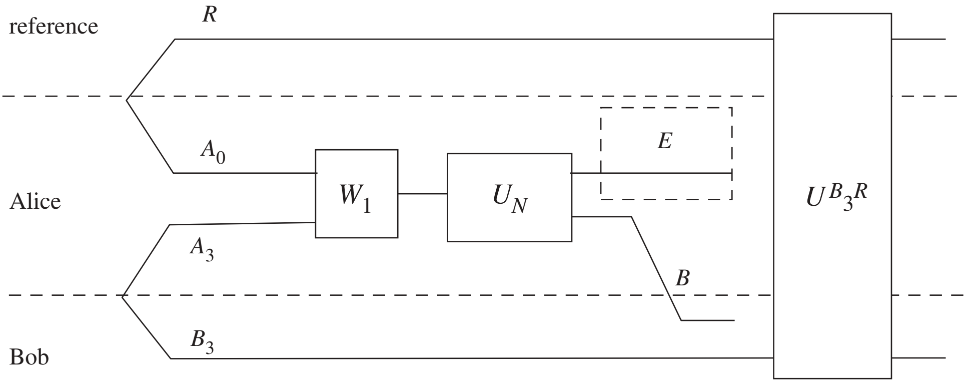

Figure 3. Partial reduction from the father to the mother. Dotted lines are used to demarcate domains controlled by the different partipicants and solid lines represent quantum information. Note that Alice starts the protocol sharing one maximally entangled state with the reference, $|\Phi_0\rangle^{A_0R}$, and another with Bob, $|\Phi_3\rangle^{A_3B_3}$. The unitary transformation U^{B_3R} comes from an application of the FQSW theorem with B_3R replacing A_1A_2. After the application of the unitary, the registers R and E are nearly decoupled, as desired, but unfortunately, because it requires acting on the reference system R, U^{B_3R} cannot be used in this way.

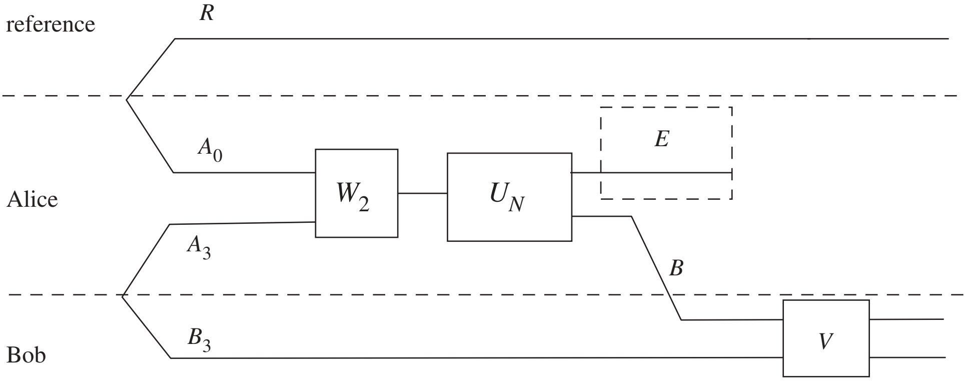

Figure 4. Final version of the father protocol generated from FQSW. As figure 3 makes clear, U^{B_3B} was required to act on one-half of a maximally entangled state, the other half of which is found in A_3A_0, register held by Alice. Thus, Alice can instead implement the encoding operation $W_2 = W_1 \circ U^T$. Bob performs the decoding operation V mandated by FQSW, resulting in the one-shot father.

through the channel $\mathcal{N}$ and Bob decodes with V. The number of ebits used up is $\log d_{A_3}$, whereas the number of qubits transmitted in the process is $\log d_{A_0}$.

6. Fully quantum reverse Shannon theorem: one-shot version

The quantum reverse Shannon theorem was conceived of in Bennett *et al.* (1999, 2002), and is proved in full in Bennett *et al.* (in preparation). It asserts that in

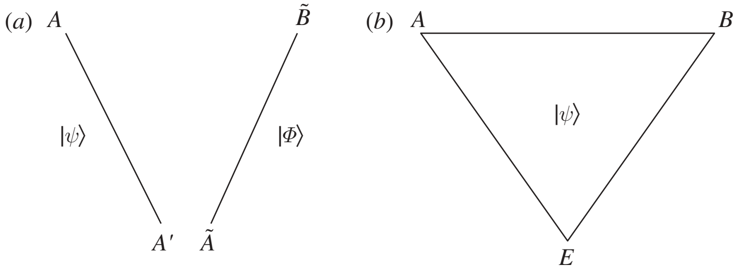

Figure 5. (*a*) The starting point for FQRS, a pair of pure entangled states. The system A is a purification of Alice's input system A', while $\tilde{A}\tilde{B}$ holds the entanglement that Alice–Bob will consume to execute the protocol. (*b*) After execution of the protocol, the reference system A is unchanged, while Alice receives the environment feedback system E and Bob receives his share B of the state $|\psi\rangle^{ABE} = U_{\mathcal{N}}^{A' \to BE}|\psi\rangle^{AA'}$.

the presence of entanglement, a noisy quantum channel $\mathcal{N}$ can be simulated by $C_E(\mathcal{N})$ cbits of forward classical communication per copy of the channel, where C_E is the entanglement-assisted capacity of the channel.

Here, following Devetak (2006), we demonstrate how, by running the mother protocol backwards, one obtains a simple proof of a fully quantum version of this result (however, unlike Bennett *et al.* (in preparation) we do not obtain a simulation of the channel on arbitrary inputs, but only relative to a fixed source). The Stinespring dilation $U_{\mathcal{N}} : A' \to BE$ of $\mathcal{N}^{A' \to B}$ is simulated in such a way that E ends up with Alice. For that reason, we say that the protocol simulates the *feedback channel* associated to $\mathcal{N}^{A' \to B}$.

Ultimately, in §9, we will show the FQRS resource inequality

$$\frac{1}{2}I(A;B)_{\varphi}[q \to q] + \frac{1}{2}I(B;E)_{\varphi}[qq] \geq \left\langle U_{\mathcal{N}}^{A' \to BE} : \rho^{A'} \right\rangle, \tag{6.1}$$

where $|\varphi\rangle^{ABE} = U_{\mathcal{N}}^{A' \to BE}|\varphi\rangle^{AA'}$ and $|\varphi\rangle^{AA'}$ is a purification of $\rho^{A'}$. In this section, we will actually prove a one-shot version of this resource inequality, by a simple re-interpretation of the systems of the mother, and running her backwards in time. The task is to simulate with high fidelity the feedback channel $U_{\mathcal{N}} : A' \to BE$ on a source $\psi^{AA'}$, using some maximal entanglement $\Phi^{\tilde{A}\tilde{B}}$ and quantum communication of a system A_1 of dimension d_{A_1}. From a mathematical point of view, the state $|\psi\rangle^{ABE} = U_{\mathcal{N}}^{A' \to BE}|\psi\rangle^{AA'}$ has to be created from $|\psi\rangle^{AA'} \otimes |\Phi\rangle^{\tilde{A}\tilde{B}}$, as illustrated in figure 5.

Recall that the one-shot FQSW protocol created a product state starting from an arbitrary pure tripartite entangled state, whereas here the goal is to do the reverse. Hence the need to run the protocol backwards in time. To help see the appropriate choice of relabellings, note that in the FQSW case, Bob holds purifications of the R and A_2 systems, called $\hat{B}$ and $\tilde{B}$, respectively. In the present setting, Alice starts holding purifications A' and $\tilde{A}$ of A and $\tilde{B}$, respectively. Matching the corresponding systems suggests the following replacements in the one-shot mother:

FQRS	FQSW
A'	$\hat{B}$
A	R
B	A
E	B
$\tilde{A}$	$\tilde{B}$
$\tilde{B}$	A_2

A comparison of figure 5 with the FQRS analogue, figure 1 is also very helpful in clarifying the role of the substitutions. We can interpret theorem 4.1 as saying that there exist isometries $U^{B\rightarrow A_1\tilde{B}}$ and $V^{A_1E\rightarrow A'\tilde{A}}$ such that

$$\begin{aligned}&\left\|\psi^{ABE}-(U^\dagger\circ V^\dagger)(\psi^{AA'}\otimes\Phi^{\tilde{A}\tilde{B}})(U^\dagger\circ V^\dagger)^\dagger\right\|_1\\&=\left\|(V\circ U)\psi^{ABE}(V\circ U)^\dagger-\psi^{AA'}\otimes\Phi^{\tilde{A}\tilde{B}}\right\|_1\\&\leq 2\left[\frac{d_Bd_A}{d_{A_1}^2}\mathrm{Tr}[(\psi^{AB})^2]\right]^{1/4}.\end{aligned}$$

In other words, Alice performs $V^\dagger: A'\tilde{A}\rightarrow A_1E$ on her part of the system and sends A_1 to Bob; she keeps E which will be the environment of the channel. (Note that $V^\dagger$ can be implemented as an isometry for this particular input state.) Bob can perform the isometry $U^\dagger: A_1\tilde{B}\rightarrow B$ to obtain the channel output in B.

7. Fully quantum Slepian–Wolf: i.i.d. version

We now return to the setting where Alice, Bob and the reference system share the state $|\psi'\rangle=(|\varphi\rangle^{ABR})^{\otimes n}$. This is often called the i.i.d. case because each copy of the state is identical and independently distributed. Combining the one-shot, FQSW result with Schumacher compression will lead to the FQSW resource inequality (2.2). In appendix A, we show the following: For any $\epsilon,\delta>0$ and sufficiently large n, we can define projectors Π_A,Π_B and Π_R onto the δ-typical subspaces of the systems indicated by the subscripts such that the following properties hold for any subsystem $F=A,B,R$:

(i) $\|\mathcal{E}(\psi')-\psi'\|_1\leq\epsilon$,

(ii) $\|\psi-\psi'\|_1\leq\epsilon$,

(iii) $2^{n[H(F)-\delta]}\leq\mathrm{rank}\ \Pi_F\leq 2^{n[H(F)+\delta]}$, and

(iv) $\mathrm{Tr}[(\psi^F)^2]\leq 2^{-n[H(F)-\delta]}$.

Here $\mathcal{E}^{A\rightarrow A^{\mathrm{typ}}}$ is the Schumacher compression operation (one of whose Kraus elements is Π_A) and $|\psi\rangle$ the normalized version of the state

$$(\Pi_A\otimes\Pi_B\otimes\Pi_R)|\psi'\rangle. \tag{7.1}$$

Although we are concerned with the output of the protocol when it is applied to the state $|\psi'\rangle = (|\varphi\rangle^{ABR})^{\otimes n}$, by properties (i) and (ii), we can analyse its effect on the nearly indistinguishable $|\psi\rangle$ instead.

Thanks to the properties of the typical projectors, namely properties (iii) and (iv), the various quantities appearing in the upper bound of theorem 4.2 get replaced by entropic formulas in the i.i.d. case. For an arbitrary subsystem F, let F^{typ} denote the support of Π_F and assume $A^{typ} = A_1 \otimes A_2$. By theorem 4.1, there exist isometries $U^{A^{typ}\to A_1A_2}$ and $V^{A_1B\to\hat{B}\tilde{B}}$ such that

$$\left\|(V\circ U)\psi^{R^{typ}A^{typ}B}(V\circ U)^\dagger - \psi^{R^{typ}\hat{B}}\otimes\Phi^{A_2\tilde{B}}\right\|_1 \leq 2\left[\frac{d_{A^{typ}}d_{R^{typ}}}{d_{A_1}^2}\mathrm{Tr}[(\psi^{A^{typ}R^{typ}})^2]\right]^{1/4}$$
$$\leq 2[2^{n[I(A;R)+3\delta]}/d_{A_1}^2]^{1/4}. \tag{7.2}$$

Choosing $\log d_{A_1} = n[I(A;R)/2 + 2\delta]$, the bound of equation (7.2) becomes less than or equal to $\sqrt{8}2^{-n\delta/4}$.

As ψ, $\mathcal{E}(\psi')$ and ψ' are close, performing the protocol on the Schumacher compressed state $\mathcal{E}(\psi')$ will also do well. More precisely, a double application of the triangle inequality and properties (i) and (ii) give

$$\left\|(V\circ U)\mathcal{E}(\psi'^{RAB})(V\circ U)^\dagger - \psi'^{R\hat{B}}\otimes\Phi^{A_2\tilde{B}}\right\|_1 \leq 2\epsilon + \sqrt{8}2^{-n\delta/4}.$$

The number of qubit channels used up is thus $n[I(A;R)/2 + 2\delta]$, whereas the number of ebits distilled is $\log d_{A_2} = \log d_{A^{typ}} - \log d_{A_1} \geq n[I(A;B)/2 - 3\delta]$.

8. Father: i.i.d. version

In the i.i.d. father setting described by the resource inequality (2.4), Alice and Bob are given a channel of the form $(\mathcal{N}^{A'\to B})^{\otimes n}$. Choose a Stinespring dilation $U_{\mathcal{N}}{}^{A'\to BE}$ such that $\mathcal{N}(\rho) = \mathrm{Tr}_E\, U_\rho U^\dagger$ and define $|\varphi\rangle^{ABE} = U_{\mathcal{N}}|\varphi\rangle^{AA'}$. Let $|\psi\rangle$ and $|\psi'\rangle$ be as in the previous section, only with R replaced by E. Now define Π_A^t to be the projector onto a particular *typical* type t and define $|\psi_t'\rangle$ and $|\psi_t\rangle$ to be the normalized versions of the states $\Pi_A^t|\psi'\rangle$ and $\Pi_A^t|\psi\rangle$, respectively. In appendix A, it is shown that there exists a particular Π_A^t such that the following properties hold:

(i) $\psi_t^A = I/(\mathrm{rank}\,\Pi_A^t)$,

(ii) $\|\psi_t - \psi_t'\|_1 \leq \epsilon$,

(iii) $2^{n[H(F)-\delta]} \leq \mathrm{rank}\,\Pi_F \leq 2^{n[H(F)+\delta]}$,

(iv) $\mathrm{Tr}[(\psi_t^F)^2] \leq 2^{-n[H(F)-\delta]}$, and

(v) $2^{n[H(A)-\delta]} \leq \mathrm{rank}\,\Pi_A^t \leq 2^{n[H(A)+\delta]}$.

Let A_t denote the support of Π_A^t. By property (i), $|\psi_t'\rangle^{A_tBE}$ is the result of sending a maximally entangled state proportional to $|\Phi\rangle^{A_tA_t'} = (\Pi_t^A \otimes \Pi_t^{A'})(|\varphi\rangle^{AA'})^{\otimes n}$ through $U_{\mathcal{N}}{}^{\otimes n}$. Similarly, $|\psi_t\rangle^{A_tB^{typ}E^{typ}}$ arises from the modified channel $(\Pi_B \otimes \Pi_E) \circ U_{\mathcal{N}}^{\otimes n}$. Thus, $|\psi_t\rangle^{A_t BE}$ is of the form (5.1) and we can apply the results of §5. Proceeding as in the previous section and using the above properties, we conclude that there exist isometries $W_2^{A_0 A_3 \rightarrow A}$ and $V^{B_3 B \rightarrow \hat{B}\tilde{B}}$ such that

$$\begin{aligned}&\|(V \circ U_{\mathcal{N}}^{\otimes n} \circ W_2)(\Phi_0^{RA_0} \otimes \Phi_3^{B_3 A_3})(V \circ U_{\mathcal{N}}^{\otimes n} \circ W_2)^{\dagger} - \psi_t^{\hat{B}E} \otimes \Phi_0^{R\tilde{B}}\|_1\\ &\quad \leq 2\epsilon + \sqrt{8}2^{-n\delta/4}.\end{aligned}$$

The number of ebits used up is $\log d_{A_3} = n[I(A;E)/2 + 2\delta]$ and the number of qubits transmitted is $\log d_{A_0} = \log d_{A_t} - \log d_{A_3} \geq n[I(A;B)/2 - 3\delta]$, leading to the asymptotic rates required by the father resource inequality.

9. Fully quantum reverse Shannon theorem: i.i.d. version

As in the previous two sections, we can consider the special case in which Alice and Bob want to simulate many realizations of the channel $\mathcal{N}: A \rightarrow B$, or rather its feedback isometry $U_{\mathcal{N}}: A \rightarrow BE$, relative to a source ρ^A. The FQRS resource inequality (6.1) was described in §6. Just as in §7, the resource inequality is achieved by mentally truncating the state $\left(|\varphi\rangle^{ABE}\right)^{\otimes n}$ to its typical part, introducing small disturbances and then running the one-shot protocol on the truncated state. We omit the details.

10. Correlated source coding: distributed compression

One of the major applications of the state merging inequality (2.3) is to the problem of distributed compression with free forward (or indeed completely unrestricted) 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 (Slepian and Wolf 1971; Horodecki *et al.* 2005*a*). In this section, we consider the application of the FQSW inequality to distributed compression without classical communication.

Because distributed compression studies multiple senders, it no longer fits into the resource inequality framework as laid out in Devetak *et al.* (2008). 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^n and B^n. A rate pair (Q_A, Q_B) will be said to be achievable if for all $\epsilon > 0$ there exists an $N(\epsilon) > 0$, such that for all $n \geq N(\epsilon)$ there exists a corresponding (E_A, E_B, D) such that

$$\langle\psi|^{R^n\hat{A}\hat{B}}(D \circ (E_A \otimes E_B))(\psi^{R^n A^n B^n})|\psi\rangle^{R^n\hat{A}\hat{B}} \geq 1 - \epsilon. \tag{10.1}$$

The achievable rate region $\mathcal{SW}(\varphi)$ for a given $|\varphi\rangle$ is the closure of the set of achievable rates. By time-sharing, it is a convex set.

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

$$Q_A + Q_B > \frac{1}{2} I(A;R)_\varphi + H(B)_\varphi = \frac{1}{2}[H(A)_\varphi + H(B)_\varphi + H(AB)_\varphi] =: \frac{1}{2} J(A;B)_\varphi, \tag{10.2}$$

introducing a new symbol $J(A;B) = H(A) + H(B) + H(AB)$ for the characteristic rate sum above, a kind of quasi-mutual information with a plus sign instead of minus.

By switching the roles played by Alice and Bob and also time-sharing between the resulting two protocols, we find the following theorem.

Theorem 10.1. *The region defined by*

$$\left.\begin{aligned} Q_A &\geq \frac{1}{2} I(A;R)_\varphi, \\ Q_B &\geq \frac{1}{2} I(B;R)_\varphi \\ \text{and} \qquad\qquad & \\ Q_A + Q_B &\geq \frac{1}{2} J(A;B)_\varphi \end{aligned}\right\} \tag{10.3}$$

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

In fact, the region of theorem 10.1 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. We may, without loss of generality, assume that their dimensions d_{W_A} and d_{W_B} are bounded above by d_A^{2n} and d_B^{2n}, respectively, because every CPTP map from a space of dimension d to a space of dimension at most d can be written using at most d^2 Kraus operators.

To bound Q_A, assume that Charlie has received both C_B and W_B, i.e. all of B^n. Let W_C be the environment for the dilation of Charlie's D. Again, without loss of generality, we can assume that the Stinespring dilations are implemented by preparing the environment systems in pure unentangled states and then applying unitary transformations. Because at the end of the protocol Charlie must have essentially $A^n B^n$, which purifies R^n, the registers $W_A W_C$ have to be in a pure state of their own, product with R^n and Charlie's output $\hat{A}\hat{B}$. Of course, this is not exactly true, only with high fidelity, so we proceed to make these statements rigorous.

Let $|\xi\rangle^{R^n \hat{A}\hat{B} W_A W_C}$ be the final state after the application of the Stinespring dilations of the encoding and decoding. By the fidelity condition,

$$\lambda_{\max}(\xi^{R^n \hat{A}\hat{B}}) \geq \mathrm{Tr}[\xi^{R^n \hat{A}\hat{B}} |\psi\rangle\langle\psi|^{R^n \hat{A}\hat{B}}] \geq 1 - \epsilon,$$

where $\lambda_{\max}(\xi^{R^n\hat{A}\hat{B}})$ denotes the maximum eigenvalue of $\xi^{R^n\hat{A}\hat{B}}$. Therefore, $|\xi\rangle^{R^n\hat{A}\hat{B}W_AW_C}$ has Schmidt decomposition

$$|\xi\rangle^{R^n\hat{A}\hat{B}W_AW_C} = \sum_i \sqrt{\lambda_i}|v_i\rangle^{R^n\hat{A}\hat{B}}|w_i\rangle^{W_AW_C}, \tag{10.4}$$

where $\lambda_1 = \lambda_{\max} \geq 1-\epsilon$, and consequently,

$$\mathrm{Tr}[|\xi\rangle\langle\xi|^{R_n\hat{A}\hat{B}W_AW_C}(\xi^{R^n\hat{A}\hat{B}} \otimes \xi^{W_AW_C})] \geq \sqrt{1-\epsilon}^2(1-\epsilon)^2 \geq 1-3\epsilon.$$

So, as the above is the fidelity between states,

$$\left\| |\xi\rangle^{R^n\hat{A}\hat{B}W_AW_C} - \xi^{R^n\hat{A}\hat{B}} \otimes \xi^{W_AW_C} \right\|_1 \leq 2\sqrt{3\epsilon},$$

by Fuchs & van de Graaf (1999), and with the contractivity of the trace distance, we now have

$$\left\|\xi^{R^nW_A} - \xi^{R^n} \otimes \xi^{W_A}\right\|_1 \leq 2\sqrt{3\epsilon}. \tag{10.5}$$

We can now apply the Fannes inequality (Fannes 1973) to yield

$$\begin{aligned}\left|H(\xi^{R^nW_A}) - H(\xi^{R^n} \otimes \xi^{W_A})\right| &\leq 2\sqrt{3\epsilon}\ \log(d_A^n d_B^n d_{W_A}) + \eta(2\sqrt{3\epsilon})\\ &\leq 2\sqrt{3\epsilon}n\ \log(d_A^3 d_B) + \eta(2\sqrt{3\epsilon}),\end{aligned} \tag{10.6}$$

for $\epsilon \leq \frac{1}{12e^2}$, $\eta(x) = -x\log x$ and using $d_{W_A} \leq d_A^{2n}$.

Now, using the subadditivity of the von Neumann entropy and the fact that the overall state is pure we have

$$\begin{aligned}H(B^n) + H(C_A) &\geq H(B^nC_A) = H(W_C\hat{A}\hat{B}) = H(W_AR^n)\\ &\geq H(W_A) + H(R^n) - 2\sqrt{3\epsilon}n\log(d_A^3d_B) - \eta(2\sqrt{3\epsilon})\\ &\geq H(A^n) - H(C_A) + H(R^n) - 2\sqrt{3\epsilon}n\log(d_A^3d_B) - \eta(2\sqrt{3\epsilon}).\end{aligned}$$

Therefore,

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

Dividing by n and letting $\epsilon \to 0$, we obtain

$$Q_A \geq \frac{1}{2}I(A;R). \tag{10.7}$$

Switching the roles of Alice and Bob gives the corresponding inequality,

$$Q_B \geq \frac{1}{2}I(B;R). \tag{10.8}$$

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

$$H(A^n) = H(C_AW_A) \leq H(W_A) + H(C_A) \leq H(W_A) + nQ_A. \tag{10.9}$$

The first equality follows from the fact that the environment system is initiated as a pure unentangled state and from the unitary invariance of the von Neumann entropy.

Combining with the analogous inequality for B leads to,

$$n(Q_A + Q_B) \geq n[H(A) + H(B)] - H(W_A) - H(W_B). \tag{10.10}$$

By similar arguments as before

$$|H(W_A W_B R^n) - H(W_A W_B) - H(R^n)| \leq 2\sqrt{3\epsilon} n \log(d_A^2 d_B^2 d_R) + \eta(2\sqrt{3\epsilon}) \tag{10.11}$$

for ϵ small enough. So,

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

Using the purity of 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

$$\begin{aligned} H(W_A) + H(W_B) \leq n(Q_A + Q_B) - nH(AB) + I(W_A; W_B) \\ + 2\sqrt{3\epsilon} n \log(d_A^2 d_B^2 d_R) + \eta(2\sqrt{3\epsilon}). \end{aligned} \tag{10.12}$$

Adding equations (10.10) and (10.12), we obtain

$$\begin{aligned} 2n(Q_A + Q_B) \geq n\big(H(A) + H(B) + H(AB)\big) - I(W_A; W_B) \\ - n\sqrt{3\epsilon} \log(d_A^2 d_B^2 d_R) - \eta(\sqrt{3\epsilon}). \end{aligned} \tag{10.13}$$

Thus,

$$Q_A + Q_B \geq \frac{1}{2} J(A; B) - \frac{1}{n} I(W_A; W_B) - 2\sqrt{3\epsilon} \log(d_A^2 d_B^2 d_R) - \frac{\eta(2\sqrt{3\epsilon})}{n}. \tag{10.14}$$

Now, let $T : R^n \to R'$ be any CPTP map on R^n. Then we can bound the mutual information $I(W_A; W_B)$ as follows:

$$\begin{aligned} & I(W_A; W_B) - I(W_A; W_B|R') \\ &= \big(H(W_A) - H(W_A R')\big) + \big(H(W_B) - H(W_B R')\big) \\ &\quad - \big(H(W_A W_B) - H(W_A W_B R')\big) - H(R') \\ &\leq 8\sqrt{3\epsilon}\big(\log d_{W_A} + \log d_{W_B} + \log d_{W_A} d_{W_B}\big) + 6H_2(2\sqrt{3\epsilon}) \\ &\leq 8\sqrt{3\epsilon} n \log(d_A^4 d_B^4) + 6H_2(2\sqrt{3\epsilon}), \end{aligned}$$

where we have used that $W_A W_B$ is almost uncorrelated with R' (via the contractivity of the trace distance under CPTP maps):

$$\|\xi^{W_A W_B R'} - \xi^{W_A W_B} \otimes \xi^{R'}\|_1 \leq 2\sqrt{3\epsilon}$$

followed by the Alicki–Fannes inequality (Alicki & Fannes 2004). The function $H_2(x)$ is the binary entropy $H_2(x) = -x\log x - (1-x)\log(1-x)$. Note that in this way the dimension of R' does not enter, which is desirable as we do not wish to constrain it in any way.

In particular, for small ϵ,

$$\begin{aligned} I(W_A; W_B) &\leq I(W_A; W_B|R')\, 8\sqrt{3\epsilon}n\, \log(d_A^4 d_B^4) + 6H_2(2\sqrt{3\epsilon}) \\ &\leq I(A^n; B^n|R') + 8\sqrt{3\epsilon}n\, \log(d_A^4 d_B^4) + 6H_2(2\sqrt{3\epsilon}), \end{aligned} \tag{10.15}$$

where in the second line we have invoked the monotonicity of mutual information under local operations. Therefore,

$$\begin{aligned} Q_A + Q_B \geq \frac{1}{2}J(A;B) - \frac{1}{2n}I(A^n; B^n|R') - 2\sqrt{3\epsilon}\, \log(d_A^2 d_B^2 d_R) - \frac{\eta(2\sqrt{3\epsilon})}{n} \\ - 8\sqrt{3\epsilon}\, \log(d_A^4 d_B^4) - \frac{6H_2(2\sqrt{3\epsilon})}{n}. \end{aligned}$$

By optimizing over the CPTP map T, we thus obtain

$$\begin{aligned} Q_A + Q_B &\geq \frac{1}{2}J(A;B) - \frac{1}{n}E_{\text{sq}}\big((\varphi^{AB})^{\otimes n}\big) - 10\sqrt{3\epsilon}\, \log(d_A^4 d_B^4) - \frac{7H_2(2\sqrt{3\epsilon})}{n} \\ &= \frac{1}{2}J(A;B) - E_{\text{sq}}(\varphi^{AB}) - 10\sqrt{3\epsilon}\, \log(d_A^4 d_B^4) - \frac{7H_2(2\sqrt{3\epsilon})}{n}, \end{aligned}$$

where $E_{\text{sq}}(\varphi^{AB})$ 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} (Christandl & Winter 2004). We have used explicitly the fact, proved in the cited paper, that $E_{\text{sq}}(\varphi^{\otimes n}) = nE_{\text{sq}}(\varphi)$.

As $\epsilon > 0$ was arbitrary, we have therefore proved the following outer bound on the achievable rate region.

Theorem 10.2. *The rate region $\mathcal{SW}(\varphi)$ of fully quantum distributed compression of the source φ is contained in the set defined by the inequalities*

$$\left.\begin{aligned} Q_A &\geq \frac{1}{2}I(A;R)_\varphi \\ Q_B &\geq \frac{1}{2}I(B;R)_\varphi \\ \textit{and} \qquad Q_A + Q_B &\geq \frac{1}{2}J(A;B)_\varphi - E_{\text{sq}}(\varphi^{AB}). \end{aligned}\right\} \tag{10.16}$$

In the special case where φ^{AB} is separable, $E_{\text{sq}}(\varphi) = 0$, which implies that the region defined by equation (10.3) is optimal. Under a certain technical assumption the same conclusion was found in Ahn *et al.* (2006): namely, there it was required that φ^{AB} is the density operator of an ensemble of product pure states satisfying a condition called irreducibility. That paper, however, was unable to show that the bound was achievable.

The appearance of the squashed entanglement in equation (10.16) may seem somewhat mysterious, but a slight modification of the protocols based on FQSW 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. As 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. Although 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

$$\left.\begin{aligned} Q_A &\geq \frac{1}{2}I(A;R)_\varphi, \\ Q_B &\geq \frac{1}{2}I(B;R)_\varphi \\ \text{and} \qquad & \\ Q_A + Q_B &\geq \frac{1}{2}J(A;B)_\varphi - D_0(\varphi^{AB}). \end{aligned}\right\} \tag{10.17}$$

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

We close this section by exhibiting a class of example sources for which we believe that the above inner bound is not tight. It is based on the observation that to arrive at equation (10.17), we considered a case where the structure of W_C was very simple. Although, in principle, W_C could harbour arbitrary tripartite entanglement with W_A and W_B, the decoding for equation (10.17), which is just the FQSW protocol's decoding, is simply an isometry separating the entanglement with R^n from that with one, and only one, of W_A and W_B. Hence, we are motivated to try and construct a source that permits Alice and Bob to extract and discard some 'waste', such that later on Charlie can finish off by discarding exactly the purification of that waste. The purified source is one of the *twisted states* (Horodecki *et al.* 2005*b*) of the form

$$|\varphi\rangle^{RA'A''B'B''} = \sum_{i=1}^{d} \sqrt{p_i}|i\rangle^{A'}|i\rangle^{B'}(U_i^{A''B''} \otimes I^R)|\phi_0\rangle^{RA''B''}$$

arbitrary unitaries U_i on the joint system $A''B''$. (It is understood that $A = A'A''$ and $B = B'B''$.)

Now let us assume that the reduced states $\tau_i^{A''B''} = U_i\phi_0^{A''B''}U_i^\dagger$ are mutually orthogonal for $i = 1, \dots, d$. Furthermore, we restrict to the case of *non-local* unitaries U_i, i.e. U_i is not a tensor product of local unitaries. We conjecture that $D_0(\varphi^{AB}) = 0$ or, more specifically, that because of the non-local 'twist', Alice and Bob cannot extract pure states from φ^{AB} by local operations alone. This would mean that our inner bound yields an achievable rate sum of

$$R_A + R_B = \frac{1}{2}J(A;B) = H(A) + \frac{1}{2}I(B;R).$$

However, a better rate sum is attainable because neither Alice nor Bob need to send the A' and B' registers, respectively: if A'' and B'' are transmitted faithfully, Charlie can coherently measure i, use it to undo U_i, so that he is left with the state ϕ_0^{CR}. He then has $|i\rangle$ in his waste register W_C, entangled only with the contents of Alice's and Bob's waste registers $W_A = A'$ and $W_B = B'$. He finishes off by discarding the waste register, creating $\sum_i \sqrt{p_i}|i\rangle|i\rangle$ afresh and using a controlled unitary to put back the twist U_i onto ϕ_0. Instead of the rates

$$R_A = H(A) = H(A'') + H(A'|A''), \quad R_B = \frac{1}{2}I(B;R),$$

they now use strictly less qubit resources,

$$R'_A = H(A'') < H(A), \quad R'_B = \frac{1}{2}I(B'';R) \leq \frac{1}{2}I(B;R).$$

11. On encoding complexity

Although 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 4.3. 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 *et al.* (2002) have shown that

$$\int_{\mathbb{U}(\mathbb{C}^{2^n})} (U \otimes U)X(U^\dagger \otimes U^\dagger)\,\mathrm{d}U = \frac{1}{|G_n|}\sum_{g \in G_n}(g \otimes g)X(g^\dagger \otimes g^\dagger), \tag{11.1}$$

where G_n is the Clifford group on n qubits. 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.

As the Schumacher compression portion of the FQSW protocol can also be done in polynomial time (Cleve & DiVincenzo 1996), we conclude that the encoding portion of the mother can be done efficiently. As 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. As 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 (Cleve & DiVincenzo 1996). Finally, although 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.

Finally, because the quantum reverse Shannon protocol consists of running FQSW backwards in time, it is Bob's *decoding* that can be found and implemented efficiently instead of Alice's encoding.

12. Discussion

We have shown that simple representation-theoretic reasoning, specifically some quadratic averages, are sufficient to derive the powerful mother protocol: a fully quantum version of entanglement distillation with state merging. The mother, in proper mythical fashion, not only generates her children in the family tree, but also the father protocol and his offspring, the quantum reverse Shannon theorem, plus an almost complete solution to the distributed quantum compression problem. We leave it as an open problem to determine the exact rate region, which we conjecture to be given by

$$\left.\begin{aligned} Q_A &\geq \frac{1}{2} I(A;R), \\ Q_B &\geq \frac{1}{2} I(B;R) \\ \text{and} \qquad & \\ Q_A + Q_B &\geq \frac{1}{2} J(A;B) - F(\varphi_{AB}), \end{aligned}\right\}$$

with some functional $F(\varphi_{AB})$ of the source density operator. It is tempting to speculate that F, as in our inner and outer bounds on the rate region, is an entanglement monotone; note that for separable and for pure states, our inner and outer bounds coincide, giving 0 and the entropy of entanglement, respectively, in agreement with the idea that F should be an entanglement measure.

We also note that although we have not pursued the opportunity here, the one-shot versions of the FQSW, father and reverse Shannon theorem are natural starting points for developing versions of the theorem adapted to states or channels with some internal structure more complicated than i.i.d. It would be interesting to compare the results of such an effort with the insights of Bowen & Mancini (2004) and Kretschmann & Werner (2005).

We close by highlighting a peculiar feature of the FQSW protocol. Let $|\psi\rangle$ be a pure state and suppose that Alice–Bob and Alice–Rebecca both share n copies of $|\psi\rangle$, so that the global pure state is $|\varphi\rangle^{\otimes n} = (|\psi\rangle^{A_1 R} |\psi\rangle^{A_2 B})^{\otimes n}$. This is a 'trivial' situation for FQSW. Instead of using our protocol, Alice can simply transfer her entanglement with Rebecca to Bob by compressing and sending him her A_1 registers, requiring a rate of $H(A_1) = I(A;R)/2$. As Alice and Bob already share $H(A_2) = I(A;B)/2$ ebits of pure state entanglement, that completes the FQSW protocol. Because of the symmetry of the situation, the roles of Rebecca and Bob could also be reversed. Thus, Alice could transfer her Bob entanglement to Rebecca by Schumacher compressing and sending A_2 to her, requiring a rate $H(A_2) = I(A;B)/2$. It is quite clear that Alice's system decomposes into an A_1 part, which contains her entanglement with Rebecca, and an A_2, which contains her entanglement with Bob. Note that the entanglement structure of the final state is very different in the two cases (figure 6). Here is the weirdness: if they use the general FQSW protocol instead, then because $H(A_1) = H(A_2)$, the *same* unitary will work in both cases with high probability. In other words, Alice could *first* apply the unitary and *then* decide whether to transfer her Rebecca entanglement to Bob or her Bob entanglement to Eve. The only difference in Alice's part of the protocol is whether she sends the

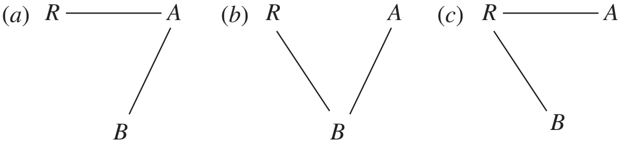

Figure 6. (*a*) A trivial starting configuration for FQSW. Solid lines represent pure state entanglement between two parties. (*b*) The result of Alice sending her Rebecca entanglement to Bob. (*c*) The result of Alice sending her Bob entanglement to Rebecca.

qubits (at rate arbitrarily small above $H(A_1)$) to Bob or to Rebecca. Thus, the localization of the entanglement so evident in the trivial implementation of the protocol disappears in the general implementation. The same subsystem can be made to carry both forms of entanglement simultaneously, compatible with either recipient!

The authors thank Debbie Leung for bringing to their attention the possibility of replacing Haar measure unitaries with random Clifford group elements. They also thank Isaac Chuang, Ignacio Cirac, Frédéric Dupuis, Renato Renner and Jürg Wullschleger for their helpful comments. A.A. appreciates the support of the US National Science Foundation through grant no. EIA-0086038. I.D. was partially supported by the NSF under grant no. CCF-0524811. P.H. was supported by the Canada Research Chairs program, the Canadian Institute for Advanced Research, and Canada's NSERC. He is also grateful to the Benasque Centre for Science and CQC Cambridge for their hospitality. A.W. was supported by the UK Engineering and Physical Sciences Research Council's 'IRC QIP', and by the EC projects RESQ (contract IST-2001-37759) and QAP (contract IST-2005-15848), as well as by a University of Bristol Research Fellowship.

Appendix A. Properties of typical and type projectors

We present here a number of consequences of the method of type classes. Denote by x^n a sequence $x_1x_2\cdots x_n$, where each x_i belongs to the finite set $\mathcal{X}$. Denote by $|\mathcal{X}|$ the cardinality of $\mathcal{X}$. Denote by $N(x|x^n)$ the number of occurrences of the symbol x in the sequence x^n. The *type* t^{x^n} of a sequence x^n is a probability vector with elements $t_x^{x^n} = N(x_i|x^n)/n$. Denote the set of sequences of type t by

$$\mathcal{T}_t^n = \{x^n \in \mathcal{X}^n : t^{x^n} = t\}.$$

For the probability distribution p on the set $\mathcal{X}$ and $\delta > 0$, let $\tau_\delta = \{t : \forall x \in \mathcal{X}, |t_x - p_x| \leq \delta\}$. $|\tau_\delta| = a$. Define the set of δ-typical sequences of length n as $\mathcal{T}_{p,\delta}^n$,

$$\mathcal{T}_{p,\delta}^n = \bigcup_{t\in\tau_\delta} \mathcal{T}_t^n = \{x^n : \forall x \in \mathcal{X},\ |t_x^{x^n} - p_x| \leq \delta\}. \tag{A 1}$$

Define the probability distribution p^n on $\mathcal{X}^n$ to be the n-fold product of p. The sequence x^n is drawn from p^n if and only if each letter x_i is drawn independently from p. Typical sequences enjoy many useful properties (Csiszar & Körner 1981; Cover & Thomas 1991). Let $H(p) = -\sum_x p_x \log p_x$ be the Shannon entropy of p. For any $\epsilon, \delta > 0$, and all sufficiently large n for which

$$p^n(\mathcal{T}_{p,\delta}^n) \geq 1 - \epsilon, \tag{A 2}$$

$$2^{-n[H(p)+c\delta]} \leq p^n(x^n) \leq 2^{-n[H(p)-c\delta]} \quad \forall x^n \in \mathcal{T}_{p,\delta}^n \tag{A 3}$$

and

$$(1-\epsilon)^{-1}2^{n[H(p)-c\delta]} \le |\mathcal{T}^n_{p,\delta}| \le 2^{n[H(p)+c\delta]}, \tag{A 4}$$

for some constant c. For $t \in \tau_\delta$ and for sufficiently large n, the cardinality D_t of $\mathcal{T}_t^n$ is bounded as (Winter 1999)

$$D_t \ge 2^{n[H(p)-\iota(\delta)]} \tag{A 5}$$

and the function $\iota(\delta) \to 0$ as $\delta \to 0$.

The above concepts generalize to the quantum setting by virtue of the spectral theorem. Let $\rho = \sum_{x\in\mathcal{X}} p_x |x\rangle\langle x|$ be the spectral decomposition of a given density matrix ρ. In other words, $|x\rangle$ is the eigenstate of ρ corresponding to eigenvalue p_x. The von Neumann entropy of the density matrix ρ is

$$H(\rho) = -\mathrm{Tr}\,\rho\log\rho = H(p).$$

The type projector is defined as

$$\Pi^n_t = \sum_{x^n\in\mathcal{T}_t^n} |x^n\rangle\langle x^n|.$$

The typical subspace associated with the density matrix ρ is defined as

$$\Pi^n_{\rho,\delta} = \sum_{x^n\in\mathcal{T}^n_{p,\delta}} |x^n\rangle\langle x^n| = \sum_{t\in\tau_\delta} \Pi^n_t.$$

Properties analogous to equations (A 2)–(A 5) hold. For any $\epsilon, \delta > 0$, and all sufficiently large n for which

$$\mathrm{Tr}\rho^{\otimes n}\Pi^n_{\rho,\delta} \ge 1-\epsilon, \tag{A 6}$$

$$2^{-n[H(\rho)+c\delta]}\Pi^n_{\rho,\delta} \le \Pi^n_{\rho,\delta}\rho^{\otimes n}\Pi^n_{\rho,\delta} \le 2^{-n[H(\rho)-c\delta]}\Pi^n_{\rho,\delta} \tag{A 7}$$

and

$$(1-\epsilon)^{-1}2^{n[H(\rho)-c\delta]} \le \mathrm{Tr}\,\Pi^n_{\rho,\delta} \le 2^{n[H(\rho)+c\delta]} \tag{A 8}$$

for some constant c. For $t \in \tau_\delta$ and for sufficiently large n, the support dimension of the type projector Π^n_t is bounded as

$$\mathrm{Tr}\,\Pi^n_t \ge 2^{n[H(\rho)-\iota(\delta)]}. \tag{A 9}$$

Henceforth, we shall drop the n and δ indices. In dealing with a multiparty system such as $|\psi'\rangle = (|\varphi\rangle^{ABR})^{\otimes n}$, we shall label the typical projectors corresponding to the various subsystems by Π_A, etc. A variant of the gentle measurement lemma (Winter 1999) states that if $\mathrm{Tr}\,\Pi\rho \ge 1-\epsilon$, then $\|\rho-\hat{\sigma}\|_1 \le 2\sqrt{\epsilon}$, where $\hat{\sigma} = \sigma/\mathrm{Tr}\,\sigma$ and $\sigma = \Pi\rho\Pi$. Applying it together with (A 6) gives

$$\left\|\psi' - \frac{\Pi_A\psi'\Pi_A}{\mathrm{Tr}\,\psi'\Pi_A}\right\|_1 \le 2\sqrt{\epsilon}.$$

The Schumacher compression operation $\mathcal{E}$ projects onto Π_A with probability $\mathrm{Tr}\,\psi'\Pi_A \ge 1-\epsilon$. Thus,

$$\left\|\mathcal{E}(\psi') - \frac{\Pi_A\psi'\Pi_A}{(\mathrm{Tr}\psi'\Pi_A)}\right\|_1 \le 2\epsilon.$$

The triangle inequality now gives

$$\|\mathcal{E}(\psi') - \psi'\|_1 \le 2\epsilon + 2\sqrt{\epsilon}.$$

Define $|\psi\rangle$ to be the normalized version of the state

$$(\Pi_A \otimes \Pi_B \otimes \Pi_R)|\psi'\rangle. \tag{A 10}$$

As Π_R, Π_A and Π_B commute, they satisfy a sort of union bound,

$$\Pi_A \otimes \Pi_B \otimes \Pi_R \ge \Pi_A + \Pi_B + \Pi_R - 2I. \tag{A 11}$$

Combining this with the same variant of the gentle measurement lemma as before and (A 6) gives

$$\|\psi' - \psi\|_1 \le 2\sqrt{3\epsilon}.$$

Observe

$$\Pi_A {\psi'}^{ABR} \Pi_A \ge \Pi_A {\psi'}^{ABR} (\Pi_B \otimes \Pi_R) {\psi'}^{ABR} \Pi_A.$$

Then,

$$\begin{aligned} \Pi_A {\psi'}^{A} \Pi_A &= \mathrm{Tr}_{BR}[\Pi_A {\psi'}^{ABR} \Pi_A] \\ &\ge \mathrm{Tr}_{BR}[\Pi_A {\psi'}^{ABR} (\Pi_B \otimes \Pi_R) {\psi'}^{ABR} \Pi_A] \\ &= \mathrm{Tr}_{BR}[(\Pi_A \otimes \Pi_B \otimes \Pi_R) {\psi'}^{ABR} (\Pi_A \otimes \Pi_B \otimes \Pi_R)] \\ &\ge (1-3\epsilon)\psi^A. \end{aligned} \tag{A 12}$$

Combining with inequalities (A 7) and (A 8) gives

$$\mathrm{Tr}\,[(\psi^A)^2] \le (1-3\epsilon)^{-1} 2^{-n[H(A)-c\delta]}.$$

Define $P'_t = \mathrm{Tr}\,\psi' \Pi_A^t$ and $P_t = \mathrm{Tr}\,\psi \Pi_A^t$. By equations (A 7) and (A 9), $P'_t \ge 2^{-n[c\delta+\iota(\delta)]}$ for all $t \in \tau_\delta$. Define $|\psi'_t\rangle$ and $|\psi_t\rangle$ to be the normalized versions of the states $\Pi_A^t|\psi'\rangle$ and $\Pi_A^t|\psi\rangle$, respectively. As $\Pi_A|\psi'\rangle = \sum_{t\in\tau_\delta} \sqrt{P'_t}|\psi'_t\rangle$ and $|\psi\rangle = \sum_{t\in\tau_\delta} \sqrt{P_t}|\psi_t\rangle$, we have

$$\sum_{t\in\tau_\delta} \sqrt{P_t P'_t}|\langle\psi_t|\psi'_t\rangle| \ge |\langle\psi|\psi'\rangle| \ge 1 - 3\epsilon. \tag{A 13}$$

We now claim that there exists a t for which both

$$|\langle\psi_t|\psi'_t\rangle| \ge 1 - 18\epsilon \quad \text{and} \quad P_t \ge \tfrac{1}{3}P'_t \ge \tfrac{1}{3}2^{-n[c\delta+\iota(\delta)]}. \tag{A 14}$$

First, by Cauchy–Schwarz,

$$\sum_t \frac{1}{2}(P_t + P'_t)|\langle\psi_t|\psi'_t\rangle| \ge 1 - 3\epsilon,$$

so that

$$\sum_t P'_t|\langle\psi_t|\psi'_t\rangle| \ge 1 - 6\epsilon.$$

Thinking of P'_t as a probability distribution over t, the probability that $P'_t > 3P_t$ is upper bounded by $\frac{1}{3}$, as is the probability that $|\langle\psi_t|\psi'_t\rangle| \le 1 - 18\epsilon$. Hence, there exists a t for which both events are false, yielding the claim. Choose t to be one that satisfies the claim. Then

$$\|\psi_t - \psi'_t\|_1 \leq 12\sqrt{\epsilon}.$$

From

$$\begin{aligned}&\mathrm{Tr}_{AR}[(\Pi_A^t \otimes \Pi_B \otimes \Pi_R)\psi'^{ABR}(\Pi_A^t \otimes \Pi_B \otimes \Pi_R)]\\ &\quad \leq \mathrm{Tr}_{AR}[(\Pi_A \otimes \Pi_B \otimes \Pi_R)\psi'^{ABR}(\Pi_A \otimes \Pi_B \otimes \Pi_R)]\end{aligned}$$

and $\mathrm{Tr}\,\Pi_A^t\psi \geq \frac{1}{3}2^{-n[c\delta+\iota(\delta)]}$, it follows that

$$\mathrm{Tr}\,[(\psi_t^B)^2] \leq 3 \cdot 2^{n[c\delta+\iota(\delta)]}\mathrm{Tr}\,[(\psi^B)^2] \leq 3 \cdot (1-3\epsilon)^{-1}2^{-n[H(B)-2c\delta-\iota(\delta)]}.$$

A similar bound holds for $\mathrm{Tr}[(\psi_t^R)^2]$.

Thus we have shown properties (i)–(iv) of §7 and (i)–(v) of §8.

References

Ahn, C., Doherty, A., Hayden, P. & Winter, A. 2006 On the distributed compression of quantum information. *IEEE Trans. Inf. Theory* **52**, 4349–4357. (doi:10.1109/TIT.2006.881734)

Alicki, R. & Fannes, M. 2004 Continuity of quantum conditional information. *J. Phys. A* **37**, L55–L57. (doi:10.1088/0305-4470/37/5/L01)

Bennett, C. H., Shor, P. W., Smolin, J. A. & Thapliyal, A. V. 1999 Entanglement-assisted classical capacity of noisy quantum channels. *Phys. Rev. Lett.* **83,** 3081. (doi:10.1103/PhysRevLett.83.3081)

Bennett, C. H., Shor, P. W., Smolin, J. A. & Thapliyal, A. V. 2002 Entanglement-assisted capacity of a quantum channel and the reverse Shannon theorem. *IEEE Trans. Inf. Theory* **48,** 2637. (http://arxiv.org/abs/quant-ph/0106052)

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

Bowen, G. & Mancini, S. 2004 Quantum channels with a finite memory. *Phys. Rev. A* **69**, 12306. (doi:10.1103/PhysRevA.69.012306)

Christandl, M. & Winter, A. 2004 “Squashed entanglement”: an additive entanglement measure. *J. Math. Phys.* **45**, 829–840. (doi:10.1063/1.1643788)

Cleve, R. & DiVincenzo, D. P. 1996 Schumacher’s quantum data compression as a quantum computation. *Phys. Rev. A* **54**, 2636–2650. (doi:10.1103/PhysRevA.54.2636)

Cover, T. M. & Thomas, J. A. 1991 *Elements of information theory.* Hoboken, NJ: Wiley.

Csiszar, I. & Körner, J. 1981 *Information theory: coding theorems for discrete memoryless systems.* New York, NY: Academic Press.

Devetak, I. 2005 The private classical capacity and quantum capacity of a quantum channel. *IEEE Trans. Inf. Theory* **51,** 44. (doi:10.1109/TIT.2004.839515)

Devetak, I. 2006 A triangle of dualities: reversibly decomposable quantum channels, source-channel duality, and time reversal. *Phys. Rev. Lett.* **97**, 140503. (doi:10.1103/PhysRevLett.97.140503)

Devetak, I. & Winter, A. 2005 Distillation of secret key and entanglement from quantum states. *Proc. R. Soc. A* **461,** 207–237. (doi:10.1098/rspa.2004.1372)

Devetak, I., Harrow, A. W. & Winter, A. 2004 A family of quantum protocols. *Phys. Rev. Lett.* **93**, 230504. (doi:10.1103/PhysRevLett.93.230504)

Devetak, I., Harrow, A. W. & Winter, A. 2008 A resource framework for quantum Shannon theory. *IEEE Trans. Inf. Theory* **54**, 4587–4618. (doi:10.1109/TIT.2008.928980)

DiVincenzo, D. P., Leung, D. W. & Terhal, B. M. 2002 Quantum data hiding. *IEEE Trans. Inf. Theory* **48**, 580–598. (doi:10.1109/18.985948)

Fannes, M. 1973 A continuity property of the entropy density for spin lattice systems. *Commun. Math. Phys.* **31**, 291–294. (doi:10.1007/BF01646490)

Fuchs, C. A. & van de Graaf, J. 1999 Cryptographic distinguishability measures for quantum mechanical states. *IEEE Trans. Inf. Theory* **45**, 1216–1227. (doi:10.1109/18.761271)

Groisman, B., Popescu, S. & Winter, A. 2005 On the quantum, classical and total amount of correlations in a quantum state. *Phys. Rev. A* **72**, 032317. (doi:10.1103/PhysRevA.72.032317)

Harrow, A. W. 2004 Coherent communication of classical messages. *Phys. Rev. Lett.* **92,** 097902. (doi:10.1103/PhysRevLett.92.097902)

Holevo, A. S. 1998 The capacity of the quantum channel with general signal states. *IEEE Trans. Inf. Theory* **44,** 269–273. (doi:10.1109/18.651037)

Horodecki, M., Horodecki, P., Horodecki, R., Leung, D. W. & Terhal, B. M. 2001 Classical capacity of a noiseless quantum channel assisted by noisy entanglement. *Quant. Inf. Comp.* **1,** 70–78. (http://arxiv.org/abs/quant-ph/0106080)

Horodecki, M., Oppenheim, J. & Winter, A. 2005*a* Partial quantum information. *Nature* **436**, 673–676. (doi:10.1038/nature03909)

Horodecki, K., Horodecki, M., Horodecki, P. & Oppenheim, J. 2005*b* Secure key from bound entanglement. *Phys. Rev. Lett.* **94**, 160502. (doi:10.1103/PhysRevLett.94.160502)

Horodecki, M., Oppenheim, J. & Winter, A. 2007 Quantum state merging and negative information. *Commun. Math. Phys.* **269**, 107–136. (doi:10.1007/s00220-006-0118-x)

Kretschmann, D. & Werner, R. F. 2005 Quantum channels with memory. *Phys. Rev. A* **72**, 062323. (doi:10.1103/PhysRevA.72.062323)

Lloyd, S. 1996 Capacity of the noisy quantum channel. *Phys. Rev. A* **55,** 1613–1622. (doi:10.1103/PhysRevA.55.1613)

Schumacher, B. & Westmoreland, M. D. 1997 Sending classical information via noisy quantum channels. *Phys. Rev. A* **56,** 131–138. (doi:10.1103/PhysrevA.56.131)

Shor, P. W. 2002 The quantum channel capacity and coherent information. Lecture Notes in MSRI Workshop on Quantum Computation. See http://www.msri.org/publications/ln/msri/2002/quantumcrypto/shor/1/.

Slepian, D. & Wolf, J. K. 1971 Noiseless coding of correlated information sources. *IEEE Trans. Inf. Theory* **19**, 461–480.

Smolin, J. A., Verstraete, F. & Winter, A. 2005 Entanglement of assistance and multipartite state distillation. *Phys. Rev. A* **72**, 052317. (doi:10.1103/PhysRevA.72.052317)

Uhlmann, A. 1976 The 'transition probability' in the state space of a *-algebra. *Rep. Math. Phys.* **9**, 273. (doi:10.1016/0034-4877(76)90060-4)

Winter, A. 1999 Coding theorem and strong converse for quantum channels. *IEEE Trans. Inf. Theory* **45**, 2481–2485. (doi:10.1109/18.796385)

Yard, J., Hayden, P. & Devetak, I. 2006 Quantum broadcast channels. (http://arxiv.org/abs/quant-ph/0603098)

Yard, J., Devetak, I. & Hayden, P. 2008 Capacity theorems for quantum multiple access channels—part I: Classical–quantum and quantum–quantum capacity regions. *IEEE Trans. Inf. Theory* **54**, 3091–3113. (doi:10.1109/TIT.2008.924665)