A father protocol for quantum broadcast channels

Frédéric Dupuis and Patrick Hayden

Abstract—A new protocol for quantum broadcast channels based on the fully quantum Slepian-Wolf protocol is presented. The protocol yields an achievable rate region for entanglement-assisted transmission of quantum information through a quantum broadcast channel that can be considered the quantum analogue of Marton’s region for classical broadcast channels. The protocol can be adapted to yield achievable rate regions for unassisted quantum communication and for entanglement-assisted classical communication. Regularized versions of all three rate regions are provably optimal.

Index Terms—quantum information, broadcast channels

I. INTRODUCTION

DISCRETE memoryless broadcast channels are channels with one sender and multiple receivers, modelled using a probability transition matrix \( p(y_1, \ldots, y_n|x) \). There are many natural tasks that one may want to perform using these channels, such as sending common messages to all the users, sending separate information to each user, sending data to each user privately, or some combination of these tasks. Here we shall focus only on sending separate data, and most of our discussions will only involve channels with two receivers.

These channels were first introduced by Tom Cover in [1], where he suggested that it may be possible to use them more efficiently than by timesharing between the different users. Since then, several results concerning broadcast channels have been found, such as the capacity of degraded broadcast channels (see, for example, [2]).

The best known achievable rate region for general classical broadcast channels is due to Marton [3]: given a probability distribution \( p(x, u_1, u_2) = p(u_1, u_2)p(x|u_1, u_2) \), the following rate region is achievable for the general two-user broadcast channel \( p(y_1, y_2|x) \):

\[
\begin{align*}
0 & \leq R_1 \leq I(U_1; Y_1) \\
0 & \leq R_2 \leq I(U_2; Y_2) \\
R_1 + R_2 & \leq I(U_1; Y_1) + I(U_2; Y_2) - I(U_1; U_2)
\end{align*}
\]

(1)

It is conjectured that this characterizes the capacity region of general broadcast channels, but despite considerable efforts, no one has been able to prove a converse theorem.

The quantum generalization of broadcast channels was first studied in [4] as part of a recent effort to develop a network quantum information theory [5]. In [6], [7]. [8], [9], [10], [11], [12], [13]. In [4], the authors derived three classes of results, the first one about channels with a classical input and quantum outputs, the second one about sending a common classical message while sending quantum information to one receiver, and the third about sending qubits to one receiver while establishing a GHZ state with the two receivers.

In this paper, we study quantum broadcast channels using a different approach. Over the past few years, several results in quantum Shannon theory have been unified and simplified by the introduction of the mother and father protocols [4] and, more recently, by the fully quantum Slepian-Wolf (FQSW) protocol [15] [16]. Thus, a whole array of results, such as the quantum reverse Shannon theorem [17], the Lloyd-Shor-Devetak (LSD) theorem [18] [19] [20], one-way entanglement distillation [21], and distributed compression [15] can be derived from the FQSW protocol in various ways. The results presented here are of the same flavour: we will derive a new coding theorem for general quantum broadcast channels using the FQSW theorem. The new protocol corresponds to a father protocol for broadcast channels: the sender transmits independent quantum information to each of the receivers using entanglement he already shares with each of them. Like the original father protocol, it can easily be transformed into a protocol for entanglement-assisted transmission of classical information via superdense coding or into a protocol for unassisted transmission of qubits by using part of the transmission capacity to send the needed entanglement.

The paper is structured as follows. After introducing our notation and giving some background on quantum information in section II as well as a quick review of the FQSW protocol in section III we present a high-level overview of the protocol in section IV. We then state and prove a one-shot version of the protocol in section V and then move on to the i.i.d. version of the protocol in section VI. Finally, we conclude in section VII.

II. BACKGROUND AND NOTATION

Quantum subsystems will be labelled by capital letters \( A, B, \) etc.; and their associated Hilbert spaces will be denoted by \( \mathcal{H}_A, \mathcal{H}_B, \) etc. When necessary, we will use superscripts to indicate which subsystems a pure or mixed state is defined on; for instance, \(|\psi\rangle^{AB} \in \mathcal{H}_{AB} \). We will abbreviate \( \dim \mathcal{H}_A \) by \( |A| \).

Quantum operations will also be written using superscripts to denote the input and output systems; for example, \( U^{A'\rightarrow B} \) is an operator which takes the quantum subsystem \( A' \) as input and yields output on subsystem \( B \). Generally, isometries will be written as \( U, V \), and so forth, whereas quantum channels (also known as superoperators, or completely positive trace-preserving maps) will be written using calligraphic letters, such as \( N^{A'\rightarrow B} \). A quantum broadcast channel is a quantum channel with one input subsystem and two or more output subsystems.

Note that a quantum channel can always be extended to an isometry by adding another output subsystem which represents
the environment of the channel (see, for example, [22]). This isometric extension implements exactly the same operation as the original channel if we trace out the environment subsystem. The isometric extension of $N^A\rightarrow B$ will be denoted by $U_N^{AB}$, where $E$ is the environment.

We denote conjugation of $B$ by $A$ using the symbol $\cdot$ in the form $A \cdot B := ABA^\dagger$. This will allow us to avoid writing symbols twice when applying several operators to a quantum state.

We will also denote a “standard” entangled pair between subsystems $S$ and $S'$ of equal size as $|\Phi^S_{SS'} = \sum_{i=0}^{S} |ii^S \rangle \otimes |ii^{S'} \rangle$, where the $|i^S \rangle$ and $|i^{S'} \rangle$ are some standard bases on $S$ and $S'$.

We will often use the trace norm of a hermitian matrix $M$, defined to be $\|M\| := \text{Tr} |M|$. It is particularly useful because it induces a statistically important metric on the space of quantum states; we call the quantity $\|\rho - \sigma\|$ the trace distance between $\rho$ and $\sigma$.

The von Neumann entropy of a density operator $\rho_A$ will be denoted $H(\rho_A) = H(\hat{A})$. The quantum mutual information of $\rho_{AB}$ is the function $I(A:B) = H(\hat{A}) + H(\hat{B}) - H(\hat{AB})$ when there exists a local unitary transformation to her entire share of the state (a random unitary selected according to the Haar measure will do), splitting her share into two subsystems $\hat{A}$ and $\hat{A}'$, and then sending $\hat{A}$ to Bob.

III. The FQSW Protocol

Before presenting our protocol, we first give a quick overview of the fully quantum Slepian-Wolf protocol [15]. Suppose Alice and Bob hold a mixed state $\rho^{\bar{A}B}$. We introduce a reference system $R$ to purify the state; the resulting state is $|\psi^{\bar{A}B}R\rangle$. Alice would like to transfer her state to Bob by sending him as few qubits as possible. The FQSW theorem states that Alice can do this by first applying a unitary transformation to her entire share of the state (a random unitary selected according to the Haar measure will do), splitting her share into two subsystems $\bar{A}$ and $\bar{A}'$, and then sending $\bar{A}$ to Bob.

Note that this scheme works provided that the subsystems $\bar{A}$ and $R$ are in a product state after applying the random unitary: since Bob holds the purifying system of $\bar{A}R$, there exists a local unitary that Bob can apply to turn his purifying system into separate purifying systems of the two subsystems. The purifying system of $R$ is exactly the original state that Alice wanted to send to Bob, and $\bar{A}'$ and its purifying system is an EPR pair shared by Alice and Bob. This last feature is an added bonus of the protocol: Alice and Bob get some free entanglement at the end.

It is possible to calculate how close $\bar{A}$ and $R$ are to being in a product state. The result of the calculation is the following (see [15] for details):

$$\int_{U(A)} \left\| \rho^{\bar{A}R}(U) - \frac{1}{|A|} \otimes |\psi^R\rangle \right\|_1^2 \, dU \leq \frac{|A||R|}{|A|^2} \text{Tr} \left[ (\psi^{\bar{A}R})^2 \right],$$

(2)

where $\rho_{\bar{A}R}(U) = \text{Tr}_{\bar{A}} U[A \cdot |\psi^{\bar{A}R}\rangle\langle\psi^{\bar{A}R}|]$. Since the inequality holds for the average over choices of $U$, there must exist at least one $U$ that satisfies it.

A special case of interest is when the initial state is an i.i.d. state of the form $(|\psi^{\bar{A}B}R\rangle)^{\otimes n}$. In this case, it can be shown
that as long as \( \log |\tilde{A}| \geq n(\frac{1}{2} I(A; R) + \delta) \), it will be true that
\[
\varphi^{AB^\otimes n} \approx (a) \frac{I^A}{|A|} \otimes \varphi^{R^\otimes n}
\]
(3)
where \( \varphi^{AB^\otimes n} \) is the result of applying the random unitary to \( \Pi_A \cdot (\psi^{ABR})^\otimes n \), where \( \Pi_A \) is the projector onto the typical subspace of the \( A \) subsystem, as defined in Appendix \[A\] and \( \delta > 0 \).

**IV. Overview of the Protocol**

Returning now to the broadcast setting, let’s suppose Alice would like to send the maximally mixed system \( A_1 \) (which is purified by \( R_1 \)) to Bob 1, and \( A_2 \) to Bob 2 using \( n \) instances of the quantum broadcast channel \( N^{A'\rightarrow B_1 B_2} \). In addition, she has shared EPR pairs with both of them, represented by systems \( A_1 B_1 \) and \( A_2 B_2 \). We represent the channel by its isometric extension \( U_{N}^{A'\rightarrow B_1 B_2 E} \). Alice encodes her information using the encoding isometry \( W : A_1 A_2 A' \rightarrow B_1 B_2 E A' \); \( A' \) is then transmitted through the channel, and \( A \) is discarded (discarding a subsystem will turn out to be useful when discussing the i.i.d. case). Thus, after using the channel, the state of the system is \( |\psi\rangle = U_{N}^{\otimes n} W |\varphi\rangle \), where \( |\varphi\rangle = |\psi\rangle_{R_1 A_1} \otimes |\tilde{A}_1 B_1 \rangle \otimes |\tilde{A}_2 B_2 \rangle \). See Figure 2 for a diagram illustrating this.

In order for Bob 1 to be able to decode, we have to make sure that \( R_1 \) is in a product state with everything else that Bob 1 doesn’t have access to, namely \( R_2 B_2 B_2 E A \). Likewise, \( R_2 \) must be in a product state with \( R_1 B_1 B_2 E A \). This is accomplished by applying an FQS scheme to each of \( R_1 B_1 \) and another on \( R_2 B_2 \), where \( R_1 \) and \( R_2 \) play the role of the system that stays behind. Of course, it is impossible to apply these unitaries directly, since no one has access to \( R_1 \) and \( R_2 \), but since they are each applied to one end of a maximally entangled state, we can have the same effect by applying their transposes to the other end.

**V. One-shot Version**

We first prove a generic “one-shot” version of our theorem which works for general states and channels; we will then use it to derive an achievable rate region for the case of many independent uses of the channel.

**Theorem 1:** For any encoding isometry \( W : A_1 A_2 A' \rightarrow A' A \), there exist unitaries \( U_{1}^{A_1 A_1} \) and \( U_{2}^{A_2 A_2} \), and decoding isometries \( V_{1}^{B_1 B_1 B_1} \) and \( V_{2}^{B_2 B_2 B_2} \) such that
\[
\left< (V_{2} V_{1} U_{N} W U_{2}^{T} U_{1}^{T}) \cdot \varphi, \psi^{B_1 B_2 E A} \right> \leq 2 \left\{ \left( \sum_{1} |\tilde{B}_1| |\tilde{B}_2| B_2 E A \right) \right\}^{\frac{1}{2}}
\]
(7)
where \( \varphi = |\psi\rangle_{R_1 A_1} \otimes |\tilde{A}_1 B_1 \rangle \otimes |\tilde{A}_2 B_2 \rangle \). See Figure 2 for a diagram illustrating this.

**Proof:** Applying formula 2 twice, once for a random unitary over \( R_1 B_1 \) and once for a random unitary over \( R_2 B_2 \), yields:
\[
\int_{U(R_1 B_1)} \left| \sigma^{R_1 R_2 B_1 B_2 E A} (U) - \frac{R_1}{|R_1|} \otimes \psi^{B_2 E A} \right|_{1}^{2} dU
\]
(5)
and
\[
\int_{U(R_2 B_2)} \left| \sigma^{R_2 R_1 B_1 B_2 E A} (U) - \frac{R_2}{|R_2|} \otimes \psi^{B_1 B_2 E A} \right|_{1}^{2} dU
\]
(6)
where \( \sigma^{R_1 R_2 B_1 B_2 E A} (U) = T_{B_1} \left( U \cdot \psi^{R_1 R_2 B_1 B_2 E A} \right) \) and \( \sigma^{R_2 R_1 B_1 B_2 E A} (U) = T_{B_2} \left( U \cdot \psi^{R_2 R_1 B_1 B_2 E A} \right) \).

This means that there exist unitaries \( U_{1}^{R_1 B_1} \) and \( U_{2}^{R_2 B_2} \) that satisfy the above inequalities. As mentioned before, since \( R_1 B_1 \) and \( R_2 B_2 \) are maximally entangled, we can achieve the same effect by applying \( U_{1}^{T} \) and \( U_{2}^{T} \) on \( A_1 A_1 \) and \( A_2 A_2 \) respectively.

Now, using Uhlmann’s theorem (see Appendix \[A\]), we get that there exist decoding unitaries \( V_{1}^{B_1 B_1 B_1} \) and \( V_{2}^{B_2 B_2 B_2} \) such that
\[
\left| \left( (V_{2} V_{1} U_{N} W U_{2}^{T} U_{1}^{T}) \cdot \varphi, \psi^{B_1 B_2 E A} \right) \right|_{1}^{2}
\]
and
where $\psi_1$ and $\psi_2$ are some pure states determined by the theorem. To finish, we need the following lemma:

**Lemma 1:** If we have

\[
\begin{align*}
\|\rho_{ABC} - \sigma^A \otimes \sigma_{BC}\| &\leq \varepsilon_1 \\
\|\rho_{ABC} - \tau_{AB} \otimes \tau_{C}\| &\leq \varepsilon_2
\end{align*}
\]

then \(\|\rho_{ABC} - \sigma^A \otimes B \otimes C\| \leq 2\varepsilon_1 + \varepsilon_2\).

**Proof:**

\[
\begin{align*}
\|\rho_{ABC} - \sigma^A \otimes B \otimes C\| &= \|\rho_{ABC} - \sigma^A \otimes B \otimes C\| \\
&+ \|\sigma^A \otimes B \otimes C - \sigma^A \otimes B \otimes C\| \\
&\leq 2\varepsilon_1 + \varepsilon_2.
\end{align*}
\]

Applying this to our system, we get equation (4).

**VI. I.I.D VERSION**

**Theorem 2:** Let $\mathcal{N}_{A^1B^1}$ be a quantum broadcast channel. Then the following rate region is achievable for $|\psi\rangle_{A^1A_1B_1B_2} = U_{N}^{A^1B_1B_2E} |\phi\rangle_{A_1A_2D}$ where $|\phi\rangle$ is any pure state:

\[
\begin{align*}
0 &\leq Q_1 \leq \frac{1}{2} \left( I(A_1; B_1) \right)_\psi \\
0 &\leq Q_2 \leq \frac{1}{2} \left( I(A_2; B_2) \right)_\psi \\
Q_1 + Q_2 &\leq \frac{1}{2} \left( I(A_1; B_1) + I(A_2; B_2) - I(A_1; A_2) \right)_\psi.
\end{align*}
\]

$Q_1$ is the rate at which Alice sends qubits to Bob 1, and likewise for $Q_2$ for Bob 2.

Note that including the $D$ subsystem is equivalent to allowing $\phi^{A_1A_2A'}$ to be a mixed state; we find this formulation more convenient for our purposes.

**Proof:** To get this rate region, we must apply the one-shot theorem to an i.i.d. state. The main challenge is that for an arbitrary i.i.d. state of the form $|\phi_n\rangle^{A_1^nA_2^nB_1^nB_2^nD^nE^n}_N = U_N^{A_1^nA_2^nB_1^nB_2^nD^nE^n}|\phi\rangle_{A_1^nA_2^nD^nE}$, the $A_1^n$ and $A_2^n$ subsystems can be correlated, and to apply the one-shot theorem, it is crucial that $A_1^n$ and $A_2^n$ be maximally mixed and decoupled in order to play the roles of $R_1B_1$ and $R_2B_2$ respectively. (We use the term decoupled to indicate that the density operator of a composite quantum system is the product of the reduced density operators of its component systems. The analogous notion in probability theory is independence.)

We can remedy this situation by using the FQSW protocol to decouple $A_1^n$ and $A_2^n$. Whether we apply it to $A_1^n$ or to $A_2^n$, it will require us to remove $n[\log(I(A_1; A_2) + \delta)]$ qubits, where $\delta > 0$ can be arbitrarily small (note that here, and throughout this proof, the mutual information is taken with respect to $|\psi\rangle$ as defined in the statement of the theorem). The removed qubits will play the role of $\hat{A}$ in the previous section.

At the end of this process, it can be shown (see equation (3)) that the $\hat{A}$ subsystem of $W_{1} \cdot \phi^n$ is asymptotically equal to the maximally mixed state. To get $A_2^n$ also to be maximally mixed, we can apply another FQSW unitary to it, and discard $n\delta$ qubits from it (where $\delta$ can be arbitrarily small); this also leaves $A_2$ asymptotically equal to the maximally mixed state.

We can remedy this situation by using the FQSW protocol to decouple $A_1^n$ and $A_2^n$. Whether we apply it to $A_1^n$ or to $A_2^n$, it will require us to remove $n[\log(I(A_1; A_2) + \delta)]$ qubits, where $\delta > 0$ can be arbitrarily small (note that here, and throughout this proof, the mutual information is taken with respect to $|\psi\rangle$ as defined in the statement of the theorem). The removed qubits will play the role of $\hat{A}$ in the previous section.

At the end of this process, it can be shown (see equation (3)) that the $\hat{A}$ subsystem of $W_{1} \cdot \phi^n$ is asymptotically equal to the maximally mixed state. To get $A_2^n$ also to be maximally mixed, we can apply another FQSW unitary to it, and discard $n\delta$ qubits from it (where $\delta$ can be arbitrarily small); this also leaves $A_2$ asymptotically equal to the maximally mixed state.

Applying this to our system, we get equation (4).

**VI. I.I.D VERSION**

**Theorem 2:** Let $\mathcal{N}_{A^1B^1}$ be a quantum broadcast channel. Then the following rate region is achievable for $|\psi\rangle_{A^1A_1B_1B_2DE} = U_{N}^{A^1B_1B_2E} |\phi\rangle_{A_1A_2D}$ where $|\phi\rangle$ is any pure state:

\[
\begin{align*}
0 &\leq Q_1 \leq \frac{1}{2} \left( I(A_1; B_1) \right)_\psi \\
0 &\leq Q_2 \leq \frac{1}{2} \left( I(A_2; B_2) \right)_\psi \\
Q_1 + Q_2 &\leq \frac{1}{2} \left( I(A_1; B_1) + I(A_2; B_2) - I(A_1; A_2) \right)_\psi.
\end{align*}
\]

$Q_1$ is the rate at which Alice sends qubits to Bob 1, and likewise for $Q_2$ for Bob 2.

Note that including the $D$ subsystem is equivalent to allowing $\phi^{A_1A_2A'}$ to be a mixed state; we find this formulation more convenient for our purposes.

**Proof:** To get this rate region, we must apply the one-shot theorem to an i.i.d. state. The main challenge is that for an arbitrary i.i.d. state of the form $|\phi_n\rangle^{A_1^nA_2^nB_1^nB_2^nD^nE^n}_N = U_N^{A_1^nA_2^nB_1^nB_2^nD^nE^n}|\phi\rangle_{A_1^nA_2^nD^nE}$, the $A_1^n$ and $A_2^n$ subsystems can be correlated, and to apply the one-shot theorem, it is crucial that $A_1^n$ and $A_2^n$ be maximally mixed and decoupled in order to play the roles of $R_1B_1$ and $R_2B_2$ respectively. (We use the term decoupled to indicate that the density operator of a composite quantum system is the product of the reduced density operators of its component systems. The analogous notion in probability theory is independence.)
Appendix (11) to choose \( \varepsilon(n) \) such that \( \lim_{n \to \infty} \varepsilon(n) = 0 \) and such that the argument relies on the transitivity of asymptotic equality. We will therefore select \( \varepsilon(n) \) such that \( \xi \approx_{(a)} \xi_{U} \approx_{(a)} \xi_{1} \approx_{(a)} \xi_{2} \).

We will now evaluate the first term on the right-hand side of (11) using \( \xi_{1,\mathcal{N}} = U_{\mathcal{N}}^{\otimes n} \xi_{1} \) (where \( \mathcal{A}_1 \) will be split into \( R_1 \) and \( B_1 \), and likewise for \( \mathcal{A}_2 \)). From basic properties of typical subspaces (see Appendix (11)), for sufficiently large \( n \) we have:

\[
|B_1| \leq 2^{|H(A_1) - 1/2 I(A_1;A_2) + \delta|} \tag{12}
\]

Since \( \mathcal{A}_1 \) is the result of taking the typical subspace of \( A_1 \) (size \( 2^n|H(A_1) + \delta| \)) and removing a random subsystem of size \( 2^n|1/2 I(A_1;A_2) + \delta| \). We also have

\[
|B_2| \leq 2^{|H(A_2 B_2) + \delta|} 2^n|1/2 I(A_1;A_2) + \delta| \tag{13}
\]

Note above that after projecting onto the typical subspace of \( A_1^{\otimes n} \), \( \mathcal{A}_1 \) can be considered to have dimension \( 2^n|1/2 I(A_1;A_2) + \delta| \) in the sense that the post-projection subnormalized density operator has support only on a subspace of that dimension. Likewise, \( A_2^{\otimes n} B_2^{\otimes n} D^{\otimes n} E^{\otimes n} \) can also be considered to have dimension \( 2^n|H(A_2 B_2) + \delta| \) because of the typical projector.

Finally, we have

\[
\begin{align*}
\text{Tr} & \left[ \left( \xi_{1,\mathcal{N}}^{\otimes n} A_2 B_2 D^{\otimes n} E^{\otimes n} \right)^2 \right] \\
& = \text{Tr} \left[ \left( W_1 W_2 U_{\mathcal{N}}^{\otimes n} \cdot \xi_{1,\mathcal{N}}^{\otimes n} A_2 B_2 D^{\otimes n} E^{\otimes n} \right)^2 \right] \\
& = \text{Tr} \left[ \left( \Pi A_2 B_2 D E \Pi A_1 A_2 B_2 D E \right)^{\otimes n} \cdot \xi_{1,\mathcal{N}}^{\otimes n} \right] \\
& \leq 2^{-n|H(A_1;A_2 B_2 D E) - \delta|} \\
& \leq 4 \left\{ \frac{2^n I(A_1;A_2 B_2 D E) + 3\delta}{|B_1|^2} \right\}^{4/5}
\end{align*}
\]

Assuming \( |B_1| \geq 2^n|I(A_1;A_2 B_2 D E)| + 2\delta \), we get

\[
\begin{align*}
4 \left\{ \frac{|B_1|^2}{2^n|I(A_1;A_2 B_2 D E)| + 2\delta} \right\}^{4/5} & \leq 4 \times 2^{-n\delta/4}
\end{align*}
\]

Likewise, we can evaluate the second term on the right-hand side of equation (13) using \( \xi_{2,\mathcal{N}} = U_{\mathcal{N}}^{\otimes n} \xi_{2} \) and obtain that we need |\( B_2 \) | ≥ 2^n|I(A_2;A_1 B_1 D E) + 2\delta| to make it vanish.

Now, since \( \xi_{1,\mathcal{N}} \approx_{(a)} \xi_{2,\mathcal{N}} \approx_{(a)} U_{\mathcal{N}}^{\otimes n} \xi_{U} \), if we had calculated the LHS of (13) using \( U_{\mathcal{N}}^{\otimes n} \xi_{U} \) instead of \( \xi_{1,\mathcal{N}} \) and \( \xi_{2,\mathcal{N}} \), by the triangle inequality, we could only have gotten a value that is larger by at most a vanishing term. Hence, by combining the two bounds, we get that

\[
|B_1| \geq 2^n|I(A_1;A_2 B_2 D E) + 2\delta|
\]

We can now easily verify that our conditions on \(|B_1|\) and \(|B_2|\) indeed correspond to the rates advertised in the statement of the theorem. First, we have

\[
\begin{align*}
nQ_1 & = \log |R_1| \\
& = \log |\bar{A}_1| - \log |B_1| \\
& \leq n \left[ H(A_1) - \frac{1}{2} I(A_1;A_2) - \frac{1}{2} I(A_1;A_2 B_2 D E) - \delta \right] \\
& = \frac{1}{2} n \left[ I(A_1;B_1) - I(A_1;A_2) - \delta \right]
\end{align*}
\]

and

\[
\begin{align*}
nQ_2 & = \log |R_2| = \log |\bar{A}_2| - \log |B_2| \\
& \leq n \left[ H(A_2) - \frac{1}{2} I(A_2;A_1 B_1 D E) - \delta \right] \\
& = \frac{1}{2} n \left[ I(A_2;B_2) - \delta \right]
\end{align*}
\]

where \( \delta \) vanishes as \( n \to \infty \). We can, of course, exchange the roles of Bob 1 and Bob 2; combining this with time-sharing gives the asymptotic rates given in (9).

We can also calculate how much entanglement is needed between Alice and the two Bobs; let \( E_1 \) be the rate at which EPR pairs between Alice and Bob 1 are used during the protocol, and define \( E_2 \) similarly for Bob 2. We have

\[
\begin{align*}
nE_1 & = \log |\bar{B}_1| \\
& \geq n \left[ \frac{1}{2} I(A_1;A_2 B_2 D E) + 2\delta \right] \\
nE_2 & = \log |\bar{B}_2| \\
& \geq n \left[ \frac{1}{2} I(A_2;A_1 B_1 D E) + 2\delta \right]
\end{align*}
\]

A. Unassisted transmission

Note that a simple modification of this protocol allows us to achieve transmission of qubits without needing preshared entanglement. We can first let Alice establish initial entanglement with Bob 1 using the LSD theorem [18], [19], [20] (ignoring Bob 2 during this phase of the protocol); likewise, she can establish initial entanglement with Bob 2. Then, they can use the entanglement-assisted protocol just shown for the rest of the transmission, using part of the rate to maintain their stock of entanglement, and using the surplus to transmit qubits. Since we only need to use this suboptimal protocol for
the union of all rate points for some state of the form quantum broadcast channel capacity. It is understood to be only a very weak characterization of the regions defined by very different formulas can nonetheless uses of the channel. It is important to remember, however, quantum broadcast channels provided we regularize over many

\[ B. \text{ Regularized converse} \]

The rate region given in theorem[2] is indeed the capacity of quantum broadcast channels provided we regularize over many uses of the channel. It is important to remember, however, that regions defined by very different formulas can nonetheless agree after regularization, so the following theorem should be understood to be only a very weak characterization of the capacity.

**Theorem 3:** The entanglement-assisted capacity region of a quantum broadcast channel \( \mathcal{N}^{A'\rightarrow B_1 B_2} \) is the convex hull of the union of all rate points \((Q_1, Q_2)\) satisfying

\[
0 \leq Q_1 \leq \frac{1}{2n} I(A_1; B_1^{\otimes n}) \\
0 \leq Q_2 \leq \frac{1}{2n} I(A_2; B_2^{\otimes n}) \\
Q_1 + Q_2 \leq \frac{1}{2n} [I(A_1; B_1^{\otimes n}) + I(A_2; B_2^{\otimes n}) - I(A_1; A_2)]
\]

for some state of the form \( |\psi\rangle^{A_1 A_2 B_1^{\otimes n} B_2^{\otimes n} D^{\otimes n}} = U_{N^{\otimes n}}^{\otimes n} |\psi\rangle^{A_1 A_2 A'\otimes n D^{\otimes n}} \), where \(|\psi\rangle\) is a pure state.

**Proof:** It is immediate from theorem[2] that the region is achievable. We now prove the converse.

Suppose that \((Q_1, Q_2)\) is an achievable rate pair. That means that there exists a sequence of \((Q_1, Q_2, n, \epsilon_n)\) codes such that \(\epsilon_n \rightarrow 0\) as \(n \rightarrow \infty\). Consider the code of block size \(n\) in this sequence. Let \(|\varphi\rangle = |\phi\rangle^{R_1 A_1} \otimes |\phi\rangle^{A_1 B_1} \otimes |\phi\rangle^{R_1 A_1} \otimes |\phi\rangle^{B_1 A_2} \otimes \cdots \) be the input state as in theorem[1], \(W^{A'\rightarrow A_1 A_2 A_1 \rightarrow A^\otimes n D}\) be the encoding isometry, and let \(|\psi\rangle^{R_1 R_2 B_1^{\otimes n} B_2^{\otimes n} B_1 B_2 E^{\otimes n}} = U_{N^{\otimes n}}^{\otimes n} W|\varphi\rangle\). As usual, we will evaluate entropic quantities with respect to \(|\psi\rangle\).

Given that Bob 1 must be able to recover a system which purifies \(R_1\) from \(B_1^{\otimes n}\) and \(B_1\), we have by Fannes’ inequality \([23]\) that \(I(R_1; B_1^{\otimes n} B_1) \geq 2 \log |R_1| - n\delta_n\), where \(\delta_n \rightarrow 0\) as \(n \rightarrow \infty\), and likewise for Bob 2. We also have

\[
I(R_1; B_1^{\otimes n} B_1) = H(R_1) + H(B_1^{\otimes n} B_1) - H(R_1 B_1^{\otimes n} B_1) \\
\leq H(R_1) + H(B_1^{\otimes n}) + H(B_1) - H(R_1 B_1^{\otimes n} B_1) \\
= H(R_1) + H(B_1^{\otimes n}) - H(R_1 B_1^{\otimes n} B_1) \\
= I(R_1; B_1^{\otimes n} B_1)
\]

yielding, via time-sharing, the following rate region:

\[
0 \leq Q_1 \leq I(A_1; B_1^{\otimes n}) \\
0 \leq Q_2 \leq I(A_2; B_2^{\otimes n}) \\
Q_1 + Q_2 \leq \frac{1}{2n} [I(A_1; B_1^{\otimes n}) + I(A_2; B_2^{\otimes n}) - I(A_1; A_2)]
\]

(17)

where the second line follows from subadditivity, and the third line from the fact that \(R_1\) and \(B_1\) are in a product state. Hence, \(I(R_1; B_1^{\otimes n} B_1) \geq 2 \log |R_1| - n\delta_n\) and likewise, \(I(R_2; B_2^{\otimes n} B_2) \geq 2 \log |R_2| - n\delta_n\). Now, if we identify \(R_1 B_1\) as \(A_1\) and \(R_2 B_2\) as \(A_2\), we see that

\[
Q_1 \leq \frac{1}{2n} I(A_1; B_1^{\otimes n}) + \delta_n \\
Q_2 \leq \frac{1}{2n} I(A_2; B_2^{\otimes n}) + \delta_n
\]

(18)(19)

where \(\delta_n \rightarrow 0\) as \(n \rightarrow \infty\). Since \(I(A_1; A_2) = 0\), this rate point is clearly inside the region in equation (16), and it follows that this is indeed the capacity of the channel. An analogous theorem can easily be shown to hold for the unassisted capacity:

**Theorem 4:** The unassisted capacity region of a quantum broadcast channel \(\mathcal{N}^{A'\rightarrow B_1 B_2}\) is the convex hull of the union of all rate points \((Q_1, Q_2)\) satisfying

\[
0 \leq Q_1 \leq \frac{1}{2n} I(A_1; B_1^{\otimes n}) \\
0 \leq Q_2 \leq \frac{1}{2n} I(A_2; B_2^{\otimes n}) \\
Q_1 + Q_2 \leq \frac{1}{2n} [I(A_1; B_1^{\otimes n}) + I(A_2; B_2^{\otimes n}) - I(A_1; A_2)]
\]

(20)

for some state of the form \(|\phi\rangle^{A_1 A_2 B_1^{\otimes n} B_2^{\otimes n} D^{\otimes n}} = U_{N^{\otimes n}}^{\otimes n} |\phi\rangle^{A_1 A_2 A'\otimes n D^{\otimes n}}\), where \(|\phi\rangle\) is a pure state.

While one might conjecture that Theorem 3 characterizes the entanglement-assisted capacity region of a broadcast channel even with the restriction \(n = 1\), the analogous conjecture for the unassisted capacity is false. In fact, it isn’t even true for a channel with a single receiver [24].

**C. Generalization to more receivers**

It is possible to generalize the protocol to more than two receivers. Without going into details, it is straightforward to show that a one-shot version of the protocol holds if there are more receivers; we simply get equations of the form of equations (7) and (8) for each receiver, and then we put them
together in a way that is analogous to what we have done for two receivers.

To generalize this to the i.i.d. setting, the idea is to use a
multiparty version of the FQSW protocol to decouple all the
$A_1 \cdots A_n$ subsystems [25]. Thus, instead of simply having a
constraint on $Q_1 + Q_2$, we get nontrivial constraints on every
possible subset of receivers. The result is the following rate region:

$$\sum_{j \in K} Q_j \leq \frac{1}{2} \left[ \sum_{j \in K} I(A_j; B_j) - J(A_K) \right]$$

(21)

where $J(A_K) = H(A_{j_1}) + \cdots + H(A_{j_{|K|}}) - H(A_{j_1}, \cdots, A_{j_{|K|}})$,
for all $K = \{j_1, \cdots, j_{|K|}\} \subseteq \{1, \cdots, m\}$. The mutual
informations are defined on the state $|\phi^N\rangle_{A_1 \cdots A_n A_1' \cdots A_n'} = U_N^A|\phi\rangle_{A_1 \cdots A_n}$. \dashv

VII. DISCUSSION

We have shown that a new protocol for entanglement-assisted
communication of quantum information through quantum
broadcast channels can be obtained from the FQSW protocol.
Our protocol achieves the following rate region for every state $|\phi\rangle_{A_1 A_2 A'D}$:

$$0 \leq Q_1 \leq \frac{1}{2} I(A_1; B_1)_{\psi}$$

$$0 \leq Q_2 \leq \frac{1}{2} I(A_2; B_2)_{\psi}$$

$$Q_1 + Q_2 \leq \frac{1}{2} I(\psi)_{A_1 A_2 B_1 B_2 D} + I(A_2; B_2)_{\psi} - I(A_1; A_2)_{\psi}$$

(22)

where $|\psi\rangle_{A_1 A_2 B_1 B_2 D} = U_N^{A_1 A_2 B_1 B_2 D}|\phi\rangle_{A_1 A_2 A'D}$.

Note that the corresponding rate region (equation (2)) is
very similar to Marton’s region for classical broadcast channels
(equation (1)) [3]; except for the factors of 1/2, the two
expressions are identical. In fact, for classical channels, the
rates for entanglement-assisted quantum communication found
here can be achieved directly using teleportation between the
senders and the receiver, with the classical communication
required by teleportation transmitted using Marton’s protocol.
From this point of view, our results can be viewed as a direct
generalization of Marton’s region to quantum channels.

Therefore, once again, it is the entanglement-assisted version
of the quantum capacity that bears the strongest resemblance
to its classical counterpart. The same is true for both the regular
point-to-point quantum channel [26] and the quantum multiple-access
channel [27] [28]. In both those cases, the known achievable rate regions for
entanglement-assisted quantum communication are identical to their classical
counterparts. This collection of similarities suggests a
fundamental question. To what extent does the addition of
free entanglement make quantum information theory similar to
classical information theory?

Of course, the lack of a single-letter converse for Marton’s
region and, by extension, for our region, leaves open the
possibility that the analogy might break down for a new, better
broadcast region that remains to be discovered. A first step
towards eliminating that uncertainty could be to find a better
characterization of the quantum regions we have presented here.
The presence of the “discarded” system $D$ in theorem
2 is equivalent to optimizing over all mixed states $\phi^{A_1 A_2 A'}$
rather than only over pure states. This is not required for
most theorems in quantum information theory, but we have
not found a way to prove the regularized converse without
allowing for the possibility of mixed states. We leave it
d as an open problem to determine whether it is possible to
demonstrate a converse theorem that does not require allowing
mixed states.

Acknowledgments

The authors would like to thank Gilles Brassard, Igor
Devetak, Young-Han Kim, Ivan Savov, Andreas Winter and
Jon Yard for conversations that helped them in this research.
They are also grateful for support from CIAR, the Canada
Research Chairs program, FQRNT, MITACS and NSERC.

APPENDIX I

ASYMPTOTIC EQUALITIES

Here we formally define the asymptotic equalities involving
the $\approx_a$ relation. Let $\psi = \{\psi_{(1)}, \psi_{(2)}, \cdots\}$ and $\varphi = \{\varphi_{(1)}, \varphi_{(2)}, \cdots\}$ be two families of quantum states, where $\psi_{(n)}$
and $\varphi_{(n)}$ are defined on a Hilbert space $\mathcal{H}_{\otimes n}$. Then we say
that $\psi \approx_a \varphi$ if $\lim_{n \to \infty} \|\psi_{(n)} - \varphi_{(n)}\| = 0$. We then say that
$\psi$ and $\varphi$ are asymptotically equal. Note that, by the triangle
inequality, $\approx_a$ is transitive for any finite number of steps
independent of $n$.

It should be mentioned that throughout the paper, asymptotic
families of states are not always explicitly referred to
as such, but generally speaking, whenever a state depends on
the number of copies, it should be considered as a family of
states. In addition, with a slight abuse of notation, we allow
quantum operations on families of states; it should be clear
which operation is done on each member of the family.

APPENDIX II

TYPICAL SUBSPACES

Much of information theory relies on the concept of typical
sequences. Let $X$ be some alphabet and let $X$ be a random
variable defined on $X$ and distributed according to $p(x)$. Define the $\varepsilon$-typical set as follows:

$$T^{(n)}_{\varepsilon} = \left\{ x^n \in X^n \mid \frac{1}{n} \log \Pr\{X^n = x^n\} - H(X) \leq \varepsilon \right\}$$

where $X^n$ refers to $n$ independent, identically-distributed
copies of $X$. It can be shown that the two following properties hold:

1) There exists a function $\varepsilon(n)$ such that $\lim_{n \to \infty} \varepsilon(n) = 0$
   and such that $\Pr\{X^n \in T^{(n)}_{\varepsilon(n)}\} \geq 1 - \varepsilon(n)$.

2) There exists an $n_0$ such that for all $n > n_0$, $|T^{(n)}_{\varepsilon}| \leq 2^n[H(X) + \varepsilon]$.

The quantum generalization of these concepts is relatively
straightforward: let $\rho^A = \sum_{x \in X} p(x)|x\rangle\langle x|$ be the spectral
decomposition of a quantum state $\rho^A$ on a quantum system
A. Then we can define the typical projector on the quantum system $A^\otimes n$ as follows:

$$\Pi^{(n)}_\varepsilon = \sum_{x^n \in T^{(n)}_{\varepsilon}} |x^n\rangle \langle x^n|$$

We call the support of $\Pi^{(n)}_\varepsilon$ the $\varepsilon$-typical subspace of $A^\otimes n$. (For brevity, we often omit $\varepsilon$ and refer simply to the typical subspace. In this case, unless otherwise stated, $\varepsilon$ can be assumed to be a positive constant, independent of $n$.) The two properties given above generalize to the quantum case:

1) There exists a function $\varepsilon(n)$ such that $\lim_{n \to \infty} \varepsilon(n) = 0$ and such that $\text{Tr}[\Pi^{(n)}_\varepsilon \rho^{A^\otimes n}] \geq 1 - \varepsilon(n)$.

2) There exists an $n_0$ such that for all $n > n_0$, $\text{Tr}[\Pi^{(n)}_\varepsilon] \leq 2^n H(A) + \varepsilon$.

Note that the first of these two properties implies that $\Pi^{(n)}_\varepsilon \approx_{(a)} \rho^{A^\otimes n}$, via the “gentle measurement” lemma (Lemma 9 in [29]). One can also easily show that the normalized version of $\Pi^{(n)}_\varepsilon$ is also asymptotically equal to $\rho^{A^\otimes n}$ and that it also holds for i.i.d. states with more than one subsystem.

### APPENDIX III

**Uhlmann’s Theorem**

In this paper, we use Uhlmann’s theorem [30] several times, in the form first presented as Lemma 2.2 in [31]:

**Theorem 5:** Let $|\psi\rangle^{AB}$ and $|\varphi\rangle^{A'B'}$ be two quantum states such that $\|\psi^A - \varphi^A\| \leq \varepsilon$. Then there exists an isometry $U_{B' \rightarrow B}$ such that $\|\psi^{AB} - U_{B' \rightarrow B} \cdot \varphi^{AB'}\| \leq 2\sqrt{\varepsilon}$.

**REFERENCES**

[1] T. Cover, “Broadcast channels,” IEEE Transactions on Information Theory, vol. 18, pp. 2–14, 1972.
[2] T. Cover and J. Thomas, Elements of Information Theory. John-Wiley and Sons, 1991.
[3] K. Marton, “A coding theorem for the discrete memoryless broadcast channel,” IEEE Transactions on Information Theory, vol. IT-25, pp. 306–311, 1979.
[4] J. Yard, P. Hayden, and I. Devetak, “Quantum broadcast channels,” quant-ph/0603098.
[5] M. Demianowicz and P. Horodecki, “Capacity regions for multiparty quantum channels,” quant-ph/0603112.
[6] ———, “Quantum channel capacities - multiparty communication,” quant-ph/0603106.
[7] J. Yard, I. Devetak, and P. Hayden, “Capacity theorems for quantum multiple access channels: Classical-quantum and quantum-quantum capacity regions,” quant-ph/051045.
[8] D. Leung, J. Oppenheim, and A. Winter, “Quantum network communication – the butterfly and beyond,” quant-ph/0608223.
[9] M. Hayashi, K. Iwama, H. Nishimura, R. Raymond, and S. Yamashita, “Quantum network coding,” quant-ph/0601089.
[10] A. Winter, “The capacity of the quantum multiple access channel,” IEEE Trans. Info. Theory, vol. 47, pp. 3059–3065, 2001, quant-ph/9807019.
[11] G. Klímovič, “On the classical capacity of a quantum multiple access channel,” Proc. IEEE Intern. Sympos. Info. Theory, p. 278, 2001.
[12] J. A. Smolin, F. Verstraete, and A. Winter, “Entanglement of assistance and multipartite state distillation,” Phys. Rev. A, vol. 72, no. 5, pp. 052317+, Nov. 2005, quant-ph/0505038.
[13] J. H. Shapiro, Classical information capacity of the bosonic broadcast channel,” arXiv:0704.1901.
[14] I. Devetak, A. Harrow, and A. Winter, “A family of quantum protocols,” Phys. Rev. Lett., vol. 93, no. 230504, 2004, quant-ph/0308044.
[15] A. Abeyesingehe, I. Devetak, P. Hayden, and A. Winter, “The mother of all protocols: Restructuring quantum information’s family tree,” quant-ph/0606225.