Coherent Communication of Classical Messages

Aram Harrow

MIT Physics, 77 Massachusetts Ave, Cambridge, MA 02139

(Dated: January 12, 2004)

We define coherent communication in terms of a simple primitive, show it is equivalent to the ability to send a classical message with a unitary or isometric operation, and use it to relate other resources in quantum information theory. Using coherent communication, we are able to generalize super-dense coding to prepare arbitrary quantum states instead of only classical messages. We also derive single-letter formulae for the classical and quantum capacities of a bipartite unitary gate assisted by an arbitrary fixed amount of entanglement per use.

Beyond qubits and cbits

The basic units of communication in quantum information theory are usually taken to be qubits and cbits, meaning one use of a noiseless quantum or classical channel respectively. Qubits and cbits can be converted to each other and to and from other quantum communication resources, but often only irreversibly. For example, one qubit can yield at most one cbit, but to obtain a qubit from cbits requires the additional resource of entanglement. By introducing a new resource, intermediate in power between a qubit and a cbit, we will show that many resource transformations can be modified to be reversible and thus more efficient.

If \( \{ |x\rangle \}_{x=0,1} \) is a basis for \( \mathbb{C}^2 \), then a qubit channel can be described as the isometry \( |x\rangle_A \rightarrow |x\rangle_B \) and a cbit can be written as \( |x\rangle_A \rightarrow (|x\rangle_B|x\rangle_E) \). Here \( A \) is the sender Alice, \( B \) is the receiver Bob and \( E \) denotes an inaccessible environment, sometimes personified as a malicious Eve. Tracing out Eve yields the traditional definition of a cbit channel with the basis \( \{ |x\rangle \}_x \) considered the computational basis.

Define a coherent bit (or “cobit”) of communication as the ability to perform the map \( |x\rangle_A \rightarrow |x\rangle_A|x\rangle_B \). Since Alice is free to copy or destroy her channel input, 1 qubit \( \geq 1 \) cobit \( \geq 1 \) cbit, where \( X \geq Y \) means that the resource \( X \) can be used to simulate the resource \( Y \). We will also write \( X \geq Y \) (c) when \( X + Z \geq Y + Z \) for some resource \( Z \) used as a catalyst and \( X \geq Y \) (a) when the conversion is asymptotic, by which we mean that \( X^{\otimes n} \geq Y^{\otimes n} \) for \( n \) sufficiently large.

In Prop. 1, we will see that coherent bits come from any method for sending bits using a coherent procedure (a unitary or isometry on the joint Alice-Bob Hilbert space); hence their name. From their definitions, we see that cbits can be thought of as cbits where Alice controls the environment, making them like classical channels with quantum feedback. Further connections between cbits and cbits will be seen later in this Letter, where we show that irreversible resource transformations are often equivalent to performing “1 cobit \( \geq 1 \) cbit,” and in [1] where several communication protocols are “made coherent” with the effect of replacing cbits with cbits.

In the first half of this paper we will describe how to obtain coherent bits and then how to use them, allowing us to exactly describe the power of cobits in terms of conventional resources. The purpose of this paper is thus not to define a new incomparable quantum resource, but rather to introduce a technique for relating and composing communication protocols. Then we will apply coherent communication to remote state preparation[2]. This will allow us to generalize super-dense coding and determine the classical and quantum capacities of unitary gates with an arbitrary amount of entanglement assistance.

Sources of coherent communication

Qubits and cbits arise naturally from noiseless and dephasing channels respectively, and can be obtained from any noisy channel by appropriate coding [3, 4, 5, 6]. Similarly, we will show both a natural primitive yielding coherent bits and a coding theorem that can generate coherent bits from a broad class of unitary operations.

The simplest way to send a coherent message is by modifying super-dense coding[7]. In [7], Alice and Bob begin with the state \( |\Phi_2\rangle = \frac{1}{\sqrt{2}} \sum_{x=0}^{1} |x\rangle_A|x\rangle_B \). which we call an ebit. Alice encodes a two-bit message \( a_1a_2 \) by applying \( Z^{a_1}X^{a_2} \) to her half of \( |\Phi_2\rangle \) and then sending it to Bob, who decodes by applying \( (H \otimes I)\text{CNOT} \) to the state, obtaining

\[
(H \otimes I)\text{CNOT}(Z^{a_1}X^{a_2} \otimes I)|\Phi_2\rangle = |a_1\rangle|a_2\rangle
\]

Now modify this protocol so that Alice starts with a quantum state \( |a_1a_2\rangle \) and applies \( Z^{a_1}X^{a_2} \) to her half of \( |\Phi_2\rangle \) conditioned on her quantum input. After she sends her qubit and Bob decodes, they will be left with the state \( |a_1a_2\rangle_A|a_1a_2\rangle_B \). Thus,

\[
1 \text{ qubit} + 1 \text{ ebit} \geq 2 \text{ cobits} \tag{1}
\]

In fact, any unitary operation capable of classical communication is also capable of an equal amount of coherent communication, though in general this only holds asymptotically and for one-way communication.

Proposition 1 If \( U \) is a bipartite unitary or isometry such that

\[
U + e \text{ ebits} \geq C \text{ cbits} \tag{a}
\]
then

\[ U + e \text{ ebits} \geq C \text{ cbits (a)} \]  

Here we consider \( U \) a resource in the sense of [8], which considered the asymptotic capacity of nonlocal unitary operations to generate entanglement and to send ebits. By appropriate coding (as in [9]), we can reduce Proposition 1 to the following coherent analogue of Holevo-Schumacher-Westmoreland [3] (HSW) coding.

**Lemma 2 (coherent HSW)** Given a set of bipartite pure states \( S = \{ |\psi_{ij}\rangle\}_{i,j=1}^{n} \), an isometry \( U_{\psi} \) such that \( U_{\psi}|x\rangle_A = |x\rangle_A|\psi_{ij}\rangle_{AB} \), and an arbitrary probability distribution \( \{ p_{x} \} \) then

\[ U_{\psi} \geq \chi \text{ cobsit} + E \text{ ebits (a)} \]

where \( E = \sum_{x} p_{x} S(\text{Tr}_{A}|\psi_{x}\rangle) \) and \( \chi = S(\sum_{x} p_{x} \text{Tr}_{A}|\psi_{x}\rangle) - E \).

**Proof** One can show [3, 5] that for any \( \delta > 0, \epsilon > 0 \) and every \( n \) sufficiently large there exists a code \( C \subset S^{n} \) with \( |C| = \exp(n(\chi - \delta)) \), a decoding POVM \( \{ D_{x}\}_{x \in C} \) with error \( \epsilon \) and a type \( q \) with \( ||p - q||_{1} \leq \ell/n \) such that every codeword \( c := c_{1} \ldots c_{n} \in C \) has type \( q \) (i.e. \( \forall x, |\{ c_{j} = x \}| = n q_{x} \)). By error \( \epsilon \), we mean that for any \( x \in C \), \( \langle \psi_{c_{x}} \mid (I \otimes D_{c}) \mid \psi_{c_{x}} \rangle > 1 - \epsilon \).

Using Neumark’s theorem [10], Bob can make his decoding POVM into a unitary operation \( U_{D} \) defined by \( U_{D}|0\rangle = \sum_{c} |c\rangle \sqrt{D_{c}} \langle \phi |. \) Applying this to his half of a codeword \( |\psi_{c}\rangle := |\psi_{c_{1}}\rangle \ldots |\psi_{c_{n}}\rangle \) will yield a state within \( \epsilon \) of \( |\psi_{c}\rangle \), since measurements with nearly certain outcomes cause almost no disturbance.

The communication strategy begins by applying \( U_{\psi} \) to \( |c\rangle_{A} \) to obtain \( |c\rangle_{A}|\psi_{c}\rangle_{AB} \). Bob then decodes unitarily with \( U_{D} \) to yield a state within \( \epsilon \) of \( |c\rangle_{A}|\psi_{c}\rangle_{AB} \). Since \( c \) is of type \( q \), Alice and Bob can coherently permute the states of \( |\psi_{c}\rangle \) to obtain a state within \( \epsilon \) of \( |c\rangle_{A}|\psi_{c}\rangle_{AB} \). Then they can apply entanglement concentration [11] to \( |\psi_{c}\rangle_{AB} \) to obtain \( \approx nE \) cobsit with slightly disturbing the coherent message \( |c\rangle_{A}|\psi_{c}\rangle_{B} \).

There are many cases in which no ancillas are produced, so we do not need the assumptions of Lemma 2 that communication is one-way and occurs in large blocks. For example, a CNOT can transmit one coherent bit from Alice to Bob or one coherent bit from Bob to Alice. Given one ebit, a CNOT can send one coherent bit in both directions at once using the encoding

\[ (H \otimes I)\text{CNOT}(Y^{a} \otimes Z^{b})|\Phi_{2}\rangle_{AB} = |b\rangle_{A}|a\rangle_{B} \]

This can be made coherent by conditioning the encoding on a quantum register \( |a\rangle_{A}|b\rangle_{B} \), so that

\[ \text{CNOT + ebit} \geq \text{cabit} (+) + \text{cabit} (-) \]

**Uses of coherent communication**

By discarding her state after sending it, Alice can convert coherent communication into classical communication, so \( 1 \text{ cbit} \geq 1 \text{ ebit} \). Alice can also generate entanglement by inputting a superposition of messages (as in [8]), so \( 1 \text{ cbit} \geq 1 \text{ ebit} \). The true power of coherent communication comes from performing both tasks—classical communication and entanglement generation—simultaneously. This is possible whenever the classical message sent is random and nearly independent of the other states at the end of the protocol (as in the “oblivious” condition of [12]).

Teleportation [13] satisfies these conditions, and indeed a coherent version has already been proposed in [14]. Given an unknown quantum state \( |\psi_{A}\rangle \) and an EPR pair \( |\Phi_{2}\rangle_{AB} \), Alice begins coherent teleportation not by a Bell measurement on her two qubits but by unitarily rotating the Bell basis into the computational basis via a CNOT and Hadamard gate. This yields the state \( \frac{1}{2} \sum_{i,j} |ij\rangle_{A}X^{i}Z^{j}|\psi_{B}\rangle \). Using two coherent bits, Alice can send Bob a copy of her register to obtain \( \frac{1}{2} \sum_{i,j} |ij\rangle_{A}X^{i}Z^{j}|\psi_{B}\rangle \). Bob’s decoding step can now be made unitary, leaving the state \( |\Phi_{2}\rangle_{AB}|\psi_{B}\rangle \). In terms of resources, this can be summarized as: \( 2 \text{ cobsit} + 1 \text{ ebit} \geq 1 \text{ qubit} + 2 \text{ ebits} \). Since the ebit that we start with is returned at the end of the protocol, we need only use it catalytically: \( 2 \text{ cobsit} \geq 1 \text{ qubit} + 1 \text{ ebit} \). Combining this relation with Eq. (1) yields the equality

\[ 2 \text{ cobsit} = 1 \text{ qubit} + 1 \text{ ebit} \]  

Thus teleportation and super-dense coding are reversible so long as all of the classical communication is left coherent.

Another protocol that can be made coherent is Gottesman’s method [15] for simulating a distributed CNOT (i.e. \( x_{A} \otimes \{ |\rangle_{A} \} \rightarrow |x\rangle_{A}x_{A} \otimes \{ |\rangle_{B} \} \)) using one ebit and one qubit in either direction; i.e. \( 1 \text{ qubit} + 1 \text{ ebit} \geq 1 \text{ CNOT} \). At first glance, this appears completely irreversible, since a CNOT can be used to send one ebit forward or backwards, or to create one ebit, but no more than one of these at a time.

Using coherent bits as inputs, though, allows the recovery of 2 ebits at the end of the protocol, so \( 1 \text{ cbit} \geq 1 \text{ ebit} \). This allows us to generate entanglement catalytically, \( 1 \text{ cbit} + 1 \text{ ebit} \geq 1 \text{ CNOT} + 1 \text{ ebit} \). Combined with Eq. (3), this yields another equality:

\[ \text{CNOT + ebit} = \text{cabit} (-) + \text{cabit} (+) \]  

Another useful bipartite unitary gate is SWAP, which is equivalent to \( 1 \text{ qubit} + 1 \text{ ebit} \), up to catalysis.
Applying Eq. (4) then yields

$$2 \text{CNOT} = 1 \text{SWAP} \quad (c)$$

which explains the similar communication and entanglement capacities for these gates found in [8]. Previously, the most efficient methods known to transform between these gates gave $3 \text{CNOT} \geq 1 \text{SWAP} \geq 1 \text{CNOT}$.

**Remote state preparation**

In remote state preparation (RSP), Alice uses entanglement and classical communication to prepare a state of her choice in Bob’s lab [2]. Asymptotically, this relation is

$$1 \text{cbit} + 1 \text{ebit} \geq 1 \text{ remote qubit} \quad (a)$$

where “$n$ remote qubits” mean the ability of Alice to prepare an arbitrary $2^n$-dimensional pure state in Bob’s lab. While the entanglement cost of RSP is shown to be optimal in [2], this lower bound does not necessarily apply to extended communication protocols that use RSP as a subroutine. In particular, using coherent communication allows all of the entanglement to be recovered at the end of the protocol, so that asymptotically

$$1 \text{cbit} \geq 1 \text{ remote qubit} \quad (a) \quad (5)$$

However, here “$n$ remote qubits” means the slightly weaker ability to prepare $\sqrt{n}$ states, each with $2^{2^n}$ dimensions, so we cannot achieve the full range of remote states possible in [2] without $n$ ebits used as catalysts.

To show Eq. (5), we will need to examine the protocol in [2] more closely. Alice wishes to transmit a $d$-dimensional pure state $\psi = |\psi\rangle$ to Bob using the shared state $|\Phi_d\rangle = \frac{1}{\sqrt{d}} \sum_{i=1}^{d} |ii\rangle_{AB}$. For any $\epsilon > 0$, choose $n = O(d(\log d/\epsilon^2))$. Ref. [2] proved the existence of a set of $d \times d$ unitary gates $R_1, \ldots, R_n$ such that for any $\psi$,

$$\left(1 - \frac{\epsilon}{2}\right) I \leq \frac{d}{n} \sum_{k=1}^{n} R_k \psi R_k^\dagger \leq \left(1 + \frac{\epsilon}{2}\right) I \quad (6)$$

For $k = 1, \ldots, n$, define $A_k = \frac{1}{n(d+1/\epsilon^2)} (R_k \psi R_k^\dagger)^T$, where the transpose is taken in the Schmidt basis of $\Phi_d$. Due to Eq. (6), we have the operator inequalities $(1-\epsilon)I \leq \sum_k A_k \leq I$. Thus, if we define $A_{\text{fail}} = I - \sum_k A_k$, then $A_{\text{fail}} \geq 0$, $\text{Tr} A_{\text{fail}} < d\epsilon$, and $A \equiv \{A_1, \ldots, A_n, A_{\text{fail}}\}$ is a valid POVM. When $A \otimes I$ is applied to $\Phi_d$, the probability of outcome $A_{\text{fail}}$ is less than $\epsilon$. Since for any operator $O$, $(O \otimes I)|\Phi_d\rangle = (I \otimes O^T)|\Phi_d\rangle$, measurement outcome $k$ leaves Alice and Bob with the state $R_k^\dagger \psi \otimes R_k |\psi\rangle$ where $\psi^T = |\psi\rangle \langle \psi|$

We now show how to apply coherent communication to the above procedure. First we apply Neumark’s theorem[10] to convert $A$ into a unitary operation $U_A$

such that

$$U_A |\varphi\rangle_A |0\rangle_{A'} = \sum_k \sqrt{A_k} |\varphi\rangle_A |k\rangle_{A'} + \sqrt{A_{\text{fail}}} |\varphi\rangle_A |\text{fail}\rangle_{A'}$$

After applying $U_A$, Alice will perform the two-outcome measurement $\{\sum_k |k\rangle_A \langle k|_{A'}, |\text{fail}\rangle_{A'}\}$ on system $A'$. The probability of failure is less than $\epsilon$ and upon success, the resulting state is

$$\frac{1}{\sqrt{n}} \sum_{k=1}^{n} R_k^\dagger \psi |k\rangle_{A'} R_k |\psi\rangle_B$$

This can be simplified if Alice applies the unitary operation $\sum_k |k\rangle_A \otimes (R_k^T)_{A'}$ (i.e. $R_k^T$ to system $A$ conditioned on the value of system $A'$). Since $R_k^T R_k = I$, this leaves $A$ in the state $|\tilde{\psi}\rangle$, disentangled from the rest of the system so she can safely discard it. After this, Alice and Bob share the state $\frac{1}{\sqrt{n}} \sum_{k=1}^{n} |k\rangle_A R_k |\psi\rangle_B$. Alice now uses $\log n$ cbits to transmit $k$ to Bob coherently, obtaining $\frac{1}{\sqrt{n}} \sum_{k=1}^{n} |k\rangle_A |\tilde{\psi}\rangle_B$. Bob performs the unitary $\sum_k |k\rangle_B \otimes (R_k^T)_{B'}$ and ends with $|\tilde{\varphi}\rangle_{A' B'} |\tilde{\psi}\rangle_B$.

Alice and Bob have gone from $\Phi_2$ to $\Phi_n$, which is a slight increase in entanglement, though asymptotically insignificant. Thus $\log n$ cbits $\geq \log d \text{ remote qubits} + \log(n/d)$ ebits ($\epsilon$), which implies 1 coherent bit $\geq 1$ remote qubit (a). However, for the cost of preparing the initial catalyst to vanish asymptotically, we need to perform RSP in many separate blocks, say $\sqrt{n}$ blocks of $\sqrt{n}$ qubits each.

**Corollary 3 (_RSP capacity of unitary gates)** If $U \geq C \text{ cbits (a)}$ then $U \geq C \text{ remote qubits (a)}$.

**Corollary 4 (super-dense coding of quantum states)** 1 qubit $+ 1$ ebit $\geq 2 \text{ remote qubits (a)}$

Corollary 4 was first proven directly in [16] and in fact, finding an alternate proof was the original motivation for this work.

**Entangled RSP**: The states prepared in RSP need not always be completely remote; [2] also showed how Alice can prepare the ensemble $E = \{p_i, \psi_i\}$ of bipartite states $\chi = S - E \text{ ebits}$ and $S \text{ ebits}$, where $S = \sum_i p_i \text{Tr}_A \psi_i$ and $E = \sum_i p_i S(\text{Tr}_A \psi_i)$. Using coherent communication allows $\chi$ ebits to be recovered, so that

$$\chi \text{ ebits} + E \text{ ebits} \geq E \quad (a).$$

Here, by $\geq nE$, we mean $n$ uses of $U_{\varphi} : |i\rangle_A \rightarrow |i\rangle_A |\psi_i\rangle_{AB}$ where Alice’s input is restricted to the space spanned by $|i_1 \cdots i_n\rangle$ for $p$-typical sequences $i_1 \cdots i_n$.

The proof of Eq. (7) is just like the proof of Eq. (5), but even simpler since the original protocol in [2] already left Alice’s state independent of $k$ upon success. The same benefits of coherent communication do not apply
to the “low-entanglement” version of RSP in [2]; here entanglement can only be recovered from the part of the message corresponding to the measurement.]

Using super-dense coding with Eq. (7) allows super-dense coding of entangled quantum states according to \( \chi/2 \) qubits \(+E + \chi/2\) qubits \( \geq E \), a claim for which no direct proof is known.

Coherent RSP of entangled states can also help determine the communication capacity of bipartite unitary gates as introduced in [8]. For any \( e \in \mathbb{R} \), define \( C_e(U) \) to be the largest number such that \( U + e \) qubits \( \geq C_e(U) \) qubits. For negative values of \( e \), \( C_e \) corresponds to creating \(-e\) qubits per use of \( U \) in addition to communicating \( C_e \) qubits; we arbitrarily set \( C_e = -\infty \) when \( U \geq -e \) qubits.

**Proposition 5**

\[
C_e(U) = \Delta \chi_e(U) := \sup \left\{ \chi(U) - \chi(E) : E - E(U) \leq e \right\}
\]

where \( \mathcal{E} = \{p_i, \psi_i\} \) is an ensemble of bipartite states, \( U, E = \{p_i, U \psi_i U^\dagger\} \) and \( \chi \) and \( E \) are defined as above.

Thus the asymptotic capacity equals the largest increase in mutual information possible with one use of \( U \) if the average entanglement decreases by no more than \( e \). This was proven for \( e = \infty \) by [8].

**Achieving \( C_e(U) \geq \Delta \chi_e(U) \):**

We base our protocol on the one used in Section 4.3 of [8] to show \( C_{\infty} = \Delta \chi_{\infty} \). For any ensemble \( \mathcal{E} \),

\[
U + \mathcal{E} \geq U \mathcal{E} \geq (\chi(U) - \chi(E)) \text{qubits} + E(U) \text{qubits} \geq (\chi(U) - E(U)) \text{qubits} + E \text{qubits} \quad (8)
\]

Here Eq. (8) used Lemma 2 and Eq. (9) used Eq. (7).

Now we move the qubits to the left hand side and neglect the catalytic use of \( \mathcal{E} \) to get

\[
U + (E(U) - E(U)) \text{qubits} \geq (\chi(U) - (\chi(E)) \text{qubits}
\]

Taking the supremum of the right-hand side over all \( \mathcal{E} \) with \( E(U) - E(U) \leq e \) yields \( C_e(U) \geq \Delta \chi_e(U) \).

**Proving \( C_e(U) \leq \Delta \chi_e(U) \):**

Without loss of generality, we can assume Alice and Bob defer all measurements until the end of the protocol, so at all points we work with pure state ensembles. Since \( \chi(E) \) and \( E(E) \) are invariant under local unitaries and non-increasing under the final measurement, we need only consider how they are modified by \( U \). Thus, \( n \) uses of \( U \) and \( ne \) qubits can increase the mutual information by at most \( n \sum_j p_j \Delta \chi_{E_j} \) for some \( \sum_j p_j = 1 \) and \( \sum_j p_j e_j \leq e \). Our bound will follow from proving that \( \Delta \chi_e \) is a concave function of \( e \).

To prove concavity, consider a probability distribution \( \{p_j\}, \) a set of ensembles \( \mathcal{E}_j \) and the new ensemble \( \mathcal{E} = \{p_j, \chi_{E_j} \} \otimes \mathcal{E}_j \). Taking the simultaneous supremum of \( \chi(U \mathcal{E}_j) - \chi(E_j) \) over each \( \mathcal{E}_j \) subject to \( E(E_j) - E(U \mathcal{E}_j) \leq e \), completes the proof.

**Quantum capacities of unitary gates**

We can also consider the ability of unitary gates to send quantum information. Define \( Q_e \) to be the largest number such that \( U + e \) qubits \( \geq Q_e \) qubits. Then from Eq. (4), one can obtain

\[
Q_e = \frac{1}{2} C_e + Q_e \quad (10)
\]

where \( C_e \) can be determined from Proposition 5.

**Conclusions**

Coherent communication offers a new way of looking at quantum information resources in which irreversible transformations occur only when coherence is discarded and not just because we are transforming incomparable resources. Whenever classical communication is used in a quantum protocol, either as input or output, and the classical message is nearly independent of the other quantum states, there may be gains from making the communication coherent.

The case of noisy coherent communication remains to be fully solved and it would be interesting to find the coherent capacity of a noisy channel, or more generally to find their \( C_e \) and \( Q_e \) tradeoff curves. Some preliminary results in these directions have been obtained\[1\].

My thanks go to C.H. Bennett, P. Hayden, D. Leung, J. Smolin, A. Winter and especially I. Devetak for helping me come up with and refine the ideas in this paper. I am also grateful for the hospitality of the Caltech IQI, IBM T.J. Watson research center and the ERATO center. My funding is from the NSA and ARDA under ARO contract DAAD19-01-1-06.

---