Purification of Noisy Entanglement and Faithful Teleportation via Noisy Channels

Charles H. Bennett(1), Gilles Brassard(2), Sandu Popescu(3), Benjamin Schumacher(4), John A. Smolin(5), and William K. Wootters(6)

(1) IBM Research Division, Yorktown Heights, NY 10598; (2) Département IRO, Université de Montréal, C.P. 6128, Succursale centre–ville, Montréal (Québec), Canada H3C 3J7; (3) Physics Department, Tel Aviv University, Tel Aviv, Israel; (4) Physics Department, Kenyon College, Gambier, OH 43022; (5) Physics Department, University of California at Los Angeles, Los Angeles, CA 90024; (6) Physics Department, Williams College, Williamstown, MA 01267

(December 29, 1995)

quant-ph/9511027 22 Nov 1995

Two separated observers, by applying local operations to a supply of not-too-impure entangled states (*e.g.* singlets shared through a noisy channel), can prepare a smaller number of entangled pairs of arbitrarily high purity (*e.g.* near-perfect singlets). These can then be used to faithfully teleport unknown quantum states from one observer to the other, thereby achieving faithful transmission of quantum information through a noisy channel. We give upper and lower bounds on the yield $D(M)$ of pure singlets ($|\Psi^-\rangle$) distillable from mixed states M, showing $D(M) > 0$ if $\langle\Psi^-|M|\Psi^-\rangle > \frac{1}{2}$.

The techniques of quantum teleportation [1] and quantum data compression [2, 3] exemplify a new goal of quantum information theory, namely to understand the kind and quantity of channel resources needed for the transmission of intact quantum states, rather than classical information, from a sender to a receiver.

In this approach, the quantum source $\mathcal{S}$ is viewed as an ensemble of pure states ψ_i, typically not all orthogonal, emitted with known probabilities p_i. Transmission of quantum information through a channel is considered successful if the channel outputs closely approximate the inputs as quantum states. Because non-orthogonal states in principle cannot be observed without disturbing them, their faithful transmission requires that the entire transmission processes be carried out by a physical apparatus that functions obliviously, that is, without knowing or learning which ψ_i are passing through.

Just as classical data compression techniques allow data from a classical source to be faithfully transmitted using a number of bits per signal asymptotically approaching the source's Shannon entropy, $-\sum_i p_i \log_2 p_i$, quantum data compression [2, 3] allows quantum data to be transmitted, with asymptotically perfect fidelity, using a number of 2-state quantum systems or *qubits* (*e.g.* spin-$\frac{1}{2}$ particles) asymptotically approaching the source's von Neumann entropy

$$S(\rho) = -\mathrm{Tr}\rho\log_2\rho, \quad \text{where } \rho = \sum_i p_i|\psi_i\rangle\langle\psi_i|. \tag{1}$$

Quantum teleportation achieves the goal of faithful transmission in a different way, by substituting classical communication and prior entanglement for a direct quantum channel. Using teleportation, an arbitrary unknown qubit can be faithfully transmitted via a pair of maximally-entangled qubits (*e.g.* two spin-$\frac{1}{2}$ particles in a pure singlet state) previously shared between sender and receiver, and a 2-bit classical message from the sender to the receiver.

Both quantum data compression and teleportation require a noiseless quantum channel—in the former case for the direct quantum transmission and in the latter for sharing the entangled particles—yet available channels are typically noisy. Since quantum information cannot be cloned [4], it would perhaps appear impossible to use redundancy in the usual way to correct errors. Nevertheless, quantum error-correcting codes have recently been discovered [5] which operate in a subtler way, essentially by embedding the quantum information to be protected in a subspace so oriented in a larger Hilbert space as to leak little or no information to the environment, within a given noise model. We describe another approach, in which the noisy channel is not used to transmit the source states directly, but rather to share entangled pairs (*e.g.* singlets) for use in teleportation. But before they can be used to teleport reliably, the entangled pairs must be purified—converted to almost-perfectly entangled states from the mixed entangled states that result from transmission through the noisy channel. We show below how the two observers can accomplish this purification, by performing local unitary operations and measurements on the shared entangled pairs, coordinating their actions through classical messages, and sacrificing some of the entangled pairs to increase the purity of the remaining ones. Once this is done, the resulting almost perfectly pure, almost perfectly entangled pairs can be used, in conjunction with classical messages, to teleport the unknown quantum states ψ_i from sender to receiver with high fidelity. The overall result is to simulate a noiseless quantum channel by a noisy one, supplemented by local actions and classical communication.

Let M be a general mixed state of two spin-$\frac{1}{2}$ particles, from which we wish to distill some pure entanglement. The state M could result, for example, when one or both members of an initially pure singlet state $\Psi^- = (\uparrow\downarrow - \downarrow\uparrow)/\sqrt{2}$ are transmitted through a noisy channel to two separated observers, whom we shall call Alice and Bob. The purity of M can be conveniently expressed by its fidelity [2]

$$F = \langle \Psi^- | M | \Psi^- \rangle \tag{2}$$

relative to a perfect singlet. Though nonlocally defined, the purity F can be computed from the probability $P_\parallel$ of obtaining parallel outcomes if the two spins are measured locally along the same random axis: one finds that $F = 1 - 3P_\parallel/2$.

The recovery of entanglement from M is best understood in the special case that M is already a pure state of the two particles, $M = |\Upsilon\rangle\langle\Upsilon|$ for some Υ. The quantity of entanglement, $E(\Upsilon)$, in such a pure state is naturally defined by the von Neumann entropy of the reduced density matrix of either particle considered separately:

$$E(\Upsilon) = S(\rho_A) = S(\rho_B), \tag{3}$$

where $\rho_A = \mathrm{Tr}_B(|\Upsilon\rangle\langle\Upsilon|)$, and similarly for ρ_B. For pure states, this entanglement can be efficiently concentrated into singlets by the methods of [6], which use local operations and classical communication to transform n input states Υ into m singlets with a yield m/n approaching $E(\Upsilon)$ as $n \to \infty$. Conversely, given n shared singlets, local actions and classical communication suffice to prepare m arbitrarily good copies of Υ with a yield m/n approaching $1/E(\Upsilon)$ as $n \to \infty$.

Returning now to the problem of obtaining singlets from mixed states, the first step in our purification protocol is to have Alice and Bob perform a *random bilateral rotation* on each shared pair, choosing a random SU(2) rotation independently for each pair and applying it locally to both members of the pair (the same result could also be achieved by choosing from a finite set of rotations $\{B_x, B_y, B_z, I\}$ defined below). This transforms the initial general two-spin mixed state M into a rotationally symmetric mixture

$$\begin{aligned} W_F = \;& F\cdot|\Psi^-\rangle\langle\Psi^-| \;+\; \tfrac{1-F}{3}|\Psi^+\rangle\langle\Psi^+| \\ +\;& \tfrac{1-F}{3}|\Phi^+\rangle\langle\Phi^+| \;+\; \tfrac{1-F}{3}|\Phi^-\rangle\langle\Phi^-| \end{aligned} \tag{4}$$

of the singlet state Ψ^- and the three triplet states $\Psi^+ = (\uparrow\downarrow + \downarrow\uparrow)/\sqrt{2}$ and $\Phi^\pm = (\uparrow\uparrow \pm \downarrow\downarrow)/\sqrt{2}$. Because of the singlet's invariance under bilateral rotations, the symmetrized state W_F, which we shall call a Werner state [7] of purity F, has the same F as the initial mixed state M from which it was derived.

At this point it should be recalled that two mixed states having the same density matrix are physically indistinguishable, even though they may have had different preparations. Therefore, subsequent steps in the purification can be carried out without regard to any properties of the original mixed state M, or of the noisy channel(s) that may have generated it, except for the purity F.

Mixtures of the four states $\Psi^\pm$ and $\Phi^\pm$—known as the four Bell states—are particularly easy to analyze because the Bell states transform simply under several kinds of local unitary operations. Besides the random bilateral rotation already described, several other local operations will be used in entanglement purification.

• Unilateral Pauli rotations (that is, rotations by π radians about the x, y, or z axis) of *one* particle in an entangled pair. These operations map the Bell states onto one another in a 1:1 pairwise fashion, leaving no state unchanged; thus σ_x maps $\Psi^\pm \leftrightarrow \Phi^\pm$; σ_z maps $\Psi^\pm \leftrightarrow \Psi^\mp$ and $\Phi^\pm \leftrightarrow \Phi^\mp$, while σ_y maps $\Psi^\pm \leftrightarrow \Phi^\mp$. We ignore overall phase changes because they do not affect our arguments.

• Bilateral $\pi/2$ rotations B_x, B_y, and B_z of *both* particles in a pair about the x, y, or z axis respectively. Each of these operations leaves the singlet state and a different one of the triplets invariant, interchanging the other two triplets, with B_x mapping $\Phi^+ \leftrightarrow \Psi^+$, B_y mapping $\Phi^- \leftrightarrow \Psi^+$, and B_z mapping $\Phi^+ \leftrightarrow \Phi^-$. Again we omit phases.

• The quantum XOR or controlled-NOT operation [8] performed bilaterally by both observers on corresponding members of two shared pairs. The *unilateral* quantum-XOR is an operation on two qubits held by the same observer which conditionally flips the second or "target" spin if the first or "source" spin is up, and does nothing otherwise. As a unitary operator it is expressed

$$\begin{aligned} U_{XOR} \;=\;& |\uparrow_S\uparrow_T\rangle\langle\uparrow_S\downarrow_T| + |\uparrow_S\downarrow_T\rangle\langle\uparrow_S\uparrow_T| \\ +\;& |\downarrow_S\downarrow_T\rangle\langle\downarrow_S\downarrow_T| + |\downarrow_S\uparrow_T\rangle\langle\downarrow_S\uparrow_T|. \end{aligned} \tag{5}$$

The bilateral XOR (henceforth BXOR) operates in a similar fashion on corresponding members of two pairs shared between Alice and Bob: if Alice holds spins 1 and 3, and Bob holds spins 2 and 4, a BXOR, with spins 1 and 2 as source and spins 3 and 4 as target would conditionally flip spin 3 if and only if spin 1 was up, while conditionally flipping spin 4 if and only if spin 2 was up. A BXOR on two Φ^+ states leaves them both invariant. The result of applying BXOR to other combinations of Bell states is shown below, omitting phases.

Before		After (n.c. = no change)	
Source	Target	Source	Target
$\Phi^\pm$	Φ^+	n.c.	n.c.
$\Psi^\pm$	Φ^+	n.c.	Ψ^+
$\Psi^\pm$	Ψ^+	n.c.	Φ^+
$\Phi^\pm$	Ψ^+	n.c.	n.c.
$\Phi^\pm$	Φ^-	$\Phi^\mp$	n.c.
$\Psi^\pm$	Φ^-	$\Psi^\mp$	Ψ^-
$\Psi^\pm$	Ψ^-	$\Psi^\mp$	Φ^-
$\Phi^\pm$	Ψ^-	$\Phi^\mp$	n.c.

(6)

• Besides these unitary operations, Alice and Bob perform one kind of measurement: measuring both spins in a given pair along the z spin axis. This reliably distinguishes Ψ states from Φ states, but cannot distinguish $+$ from $-$ states. Of course after the measurement has been performed, the measured pair is no longer entangled.

We now show that, given two Werner pairs of fidelity $F > \frac{1}{2}$, Alice and Bob can use local operations and two-way classical communication to obtain, with probability greater than $\frac{1}{4}$, one Werner pair of fidelity $F' > F$, where the F' satisfies the recurrence relation

$$F' = \frac{F^2 + \frac{1}{9}(1-F)^2}{F^2 + \frac{2}{3}F(1-F) + \frac{5}{9}(1-F)^2}. \qquad (7)$$

To achieve this, the following protocol is used:

A1. A unilateral σ_y rotation is performed on each of the two pairs, converting them from mostly-Ψ^- Werner states to the analogous mostly-Φ^+ states, ie states with a large component $F > \frac{1}{2}$ of Φ^+ and equal components of the other three Bell states.

A2. A BXOR is performed on the two impure Φ^+ states, after which the target pair is locally measured along the z axis. If the target pair's z spins come out parallel, as they would if both inputs were true Φ^+ states, the unmeasured source pair is kept; otherwise it is discarded.

A3. If the source pair has been kept, it is converted back to a mostly-Ψ^- state by a unilateral σ_y rotation, then made rotationally symmetric by a random bilateral rotation (cf. [9]).

Because $F'(F)$ is continuous and exceeds F over the entire range $\frac{1}{2} < F < 1$, iteration of the above protocol can distill Werner states of arbitrarily high purity $F_{\rm out} < 1$ from a supply of input mixed states M of any purity $F_{\rm in} > \frac{1}{2}$. The yield (purified output pairs per impure input pair) is rather poor, and tends to zero in the limit $F_{\rm out} \to 1$; but by BXORing a variable number $k(F) \approx 1/\sqrt{1-F}$ of source pairs, rather than 1, into each target pair before measuring it, the yield can be increased and made to approach a positive limit as $F_{\rm out} \to 1$. (For this choice of k, to lowest order in $1-F$, the iteration formula for purity agrees with Eq. (7): $F'(F) = 1 - \frac{2}{3}(1-F)$. The expected fraction of the pairs discarded at each step, also to lowest order, is $\frac{2}{3}(1-F)^{\frac{1}{2}}$. One thus obtains a non-zero yield as $F_{\rm out} \to 1$.) We do not give the asymptotic yield from this method, because a higher yield can be obtained by combining it with another method to be described below, which uses a supply of previously purified Φ^+ pairs in the manner of a breeder reactor, consuming some in order to produce more than the number consumed.

The basic step (a "BXOR test") used in this method consists of bilaterally XORing a subset of the impure pairs, used as sources, into one of the pure Φ^+ states, used as a target, followed by measurement of the target. Consulting the table above, we see that each Ψ^+ or Ψ^- source pair toggles the target between Φ^+ and Ψ^+, without affecting the source. Thus a BXOR test, like a parity check on classical data, tells Alice and Bob whether there are an even or odd number of Ψ states in the tested subset. By performing a number of BXOR tests, on different subsets of the original impure pairs, all the Ψ states can be found and corrected to Φ states. A similar procedure is then used to find all the Φ^- states and correct them to the desired Φ^+. The full protocol is described below.

B1. Alice and Bob start with n impure pairs described by any Bell-diagonal density matrix W with $S(W) < 1$, and $n(S(W)+\delta)$ prepurified Φ^+ states, prepared, for example, by the variable blocksize recurrence method described above. Here δ is a positive constant that can be allowed to approach 0 in the limit of large n.

B2. Using the prepurified Φ^+ pairs as targets, Alice and Bob perform BXOR tests on sufficiently many random subsets of the impure pairs to locate all Ψ states, with high probability, without distinguishing Ψ^+ from Ψ^-. Once found, the $\Psi^\pm$ are converted respectively to $\Phi^\pm$ by applying a unilateral σ_y rotation to each of them. The impure pairs now consist only of Φ^+ and Φ^- states.

B3. Next Alice and Bob do a bilateral B_y to convert the Φ^- states into Ψ^+, while leaving the Φ^+ states invariant. This done, they perform BXOR tests on sufficiently many more random subsets to find all the new Ψ^+ states with high probability. Once found, these states are corrected to the desired Φ^+ form by unilateral σ_x rotations.

The number of BXOR tests per impure pair required to find all the errors, with arbitrarily small chance of failure, approaches the entropy of the impure pairs, $S(W) = -{\rm Tr} W \log_2 W$, in the limit of large n. This follows from these facts:

- For any two distinct n-bit strings the probability that they agree on the parities of r independent random subsets of their bits is $\leq 2^{-r}$ [10].
- The probability distribution P_X over n-bit strings x, where x represents the original sequence of Φ/Ψ values of the impure pairs, receives almost all its weight from a set of "typical" strings containing $N_1 = 2^{H(X)+O(\sqrt{n})}$ members, where $H(X)$ is the Shannon entropy of P_X. Similarly, the conditional distribution $P_{Y|X=x}$ of n-bit strings y, representing the $+/-$ values of a sequence of impure pairs whose Φ/Ψ sequence is x, receives almost all its weight from a set of typical strings containing $N_2 = 2^{H(Y|X=x)+O(\sqrt{n})}$ members.

Let r_1 BXOR tests be performed in the first round, whose goal is to find x uniquely. The expected number of "false positives"—strings in the typical set, other than the correct x, which agree with it on r_1 subset parities—is $\leq N_1 2^{-r_1}$. Thus the chance of a false positive becomes negligible when $r_1 > \log_2 N_1$. Similarly the chance of a false positive in the second round after r_2 BXOR tests is negligible when $r_2 > \log_2 N_2$. Combining these results, and recalling that $\log_2(N_1 N_2) = nS(W) + O(\sqrt{n})$, we obtain the desired result, viz. that asymptotically $S(W)$ BXOR tests per impure pair suffice to find all the errors.

The breeding method has a yield $1-S(W)$, producing more pure pairs than consumed if the mixed state's von Neumann entropy, $S(W)$, is less than 1. For Werner states the yield

$$1 - S(W_F) = 1 + F\log_2 F + (1-F)\log_2 \frac{1-F}{3} \qquad (8)$$

is positive for $F > 0.8107$.

The use of prepurified pairs as targets simplifies analysis of the protocol by avoiding back-action of the targets on the sources, but is not strictly necessary. Even without the prepurified pairs, using only impure Bell-diagonal states W as input, it is possible [11] to design a sequence of BXORs and local rotations that eliminate approximately half the candidates for x or y at each step, achieving the same asymptotic yield $1-S(W)$ as the breeding method. This non-breeding protocol requires only one-way classical communication, allowing it to be used to protect quantum information from errors during storage (cf [5, 11]) as well as during transmission.

We do not yet know the optimal asymptotic yield $D(M)$ of purified singlets distillable from general mixed states M, nor even from Werner states. Fig. 1 compares the yields of several purification methods for Werner states W_F with an upper bound $E(W_F)$ given by

$$\begin{array}{ll} E(W_F) = H_2(\frac{1}{2} + \sqrt{F(1-F)}) & \textbf{if} \;\; F>1/2 \\ E(W_F) = 0 & \textbf{if} \;\; F\leq 1/2. \end{array} \tag{9}$$

Here $H_2(x) = -x\log_2 x - (1-x)\log_2(1-x)$ is the dyadic Shannon entropy. This upper bound is based on the fact that W_F, for $F > \frac{1}{2}$, can be expressed as an equal mixture of eight pure states

$$\sqrt{F}|\Psi^-\rangle + \sqrt{\frac{1-F}{3}}(\pm|\Psi^+\rangle \pm |\Phi^-\rangle \pm i|\Phi^+\rangle), \tag{10}$$

each having entropy of entanglement equal to the right side of eq. 9a, while for $F \leq \frac{1}{2}$, W_F can be expressed as mixture of unentangled product states $\uparrow\uparrow, \downarrow\downarrow, \uparrow\downarrow$, and $\downarrow\uparrow$. In fact [11] these are the least entangled ensembles realizing W_F; therefore, $E(W_F)$ may be viewed as the Werner state's "entanglement of formation"—the asymptotic number of singlets required to prepare one W_F by local actions. Because expected entanglement cannot be increased by local actions and classical communication [11], a mixed state's distillable entanglement $D(M)$ cannot exceed its entanglement of formation $E(M)$.

We have seen that $F = \frac{1}{2}$ is a threshold below which Werner states can be made from unentangled ingredients, and above which they can be used as a starting material to make pure singlets. This is further grounds (cf also [12]) for regarding all Werner states with $F > \frac{1}{2}$ as nonlocal even though only those with $F > (2+3\sqrt{2})/8 \approx 0.78$ violate the Clauser-Horne-Shimony-Holt [13] inequality. Distillable entanglement and entanglement of formation are two alternative extensions of the definition of entanglement from pure to mixed states, but for most mixed states M we do not know the value of either quantity, nor do we know an M for which they provably differ.

We thank David DiVincenzo for valuable advice.

[1] C.H. Bennett, G. Brassard, C. Crépeau, R. Jozsa, A. Peres, and W.K. Wootters, *Phys. Rev. Lett.* **70**, 1895 (1993).
[2] B. Schumacher, *Phys. Rev. A* **51**, 2738 (1995).
[3] R. Jozsa and B. Schumacher, *J. Modern Optics* **41**, 2343 (1994).
[4] W.K. Wootters and W.H.Zurek, *Nature* **299**, 802 (1982).
[5] P. Shor *Phys. Rev. A* **52,** R2493 (1995).
[6] C.H. Bennett, H. Bernstein, S. Popescu, B. Schumacher, "Concentrating Partial Entanglement by Local Operations", submitted to *Phys. Rev. A* (1995).
[7] R.F. Werner, *Phys. Rev. A* **40**, 4277 (1989).
[8] D. Deutsch, *Proc. Roy. Soc. A* **425**, 73 (1989).
[9] C. Macchiavello (private communication 1995) found an improved recurrence using B_x, σ_y in place of our step A3.
[10] C.H. Bennett, G. Brassard, C. Crépeau, and U.M. Maurer, "Generalized Privacy Amplification", to appear Nov. 1995 in *IEEE Trans. on Information Theory.*
[11] C.H. Bennett, D. DiVincenzo, J.A. Smolin, and W.K. Wootters "Mixed State Entanglement and Quantum Error Correcting Codes", preprint (1995).
[12] S. Popescu, *Phys. Rev. Lett.* **72**, 797 (1994).
[13] J.F. Clauser, M.A. Horne, A. Shimony, R.A. Holt, *Phys. Rev. Lett.* **23**, 880 (1980).

FIG. 1. Log-log plot of entanglement distillable from Werner states of purity F by various methods, versus $F-\frac{1}{2}$. D_0 = breeding method alone (eq. 8); D_R = breeding preceded by recurrence method of eq. 7; D_M = breeding preceded by recurrence of [9]; E = entanglement of formation, eq. 9, an upper bound on entanglement yield of any method.