Purification of Noisy Entanglement, and Faithful Teleportation via Noisy Channels

Charles H. Bennett,[1] Gilles Brassard,[2] Benjamin Schumacher, [3] John Smolin,[4] and William K. Wootters[5]

[1]*IBM Research Division, T.J. Watson Research Center, Yorktown Heights, NY 10598, USA*
[2]*Département IRO, Université de Montréal, C.P. 6128, Succursale "A", Montréal (Québec), Canada H3C 3J7*
[3]*Physics Department, Kenyon College, Gambier, OH 43022, USA.*
[4]*Physics Department, University of California at Los Angeles, Los Angeles, CA 90024, USA.*
[5]*Physics Department, Williams College, Williamstown, MA 01267, USA*

Abstract

Two separated observers, by applying local operations to a supply of identical slightly-impure entangled states (for example singlets shared through a slightly 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. Some channels (*e.g.* a 50% depolarizing channel), while still nonclassical, are too noisy to be made faithful in this manner.

PACS numbers: 03.65.Bz, 42.50.Dv, 89.70.+c

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 the new 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 output, in general a mixed state W_i, closely approximates its input, as measured by the fidelity,

$$F = \sum_i p_i \langle \psi_i | W_i | \psi_i \rangle. \quad (1)$$

This is the probability that the output would pass a test for being the same as the input, conducted by someone who knows what the input was. Because non-orthogonal states in principle cannot be observed without disturbing them, their faithful transmission requires that the encoding and decoding processes be carried out by a physical apparatus that functions obliviously, 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 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

$$H(\rho) = -\mathrm{Tr}\rho \log_2 \rho, \quad \text{where } \rho = \sum_i p_i |\psi_i\rangle\langle\psi_i|. \quad (2)$$

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. Is there any way to transmit unknown quantum states faithfully through a *noisy* channel? The no-cloning theorem would appear to frustrate usual error-reduction techniques such as repeating the input signal three times. Nevertheless the goal of faithful transmission can be achieved. To do so we do not use the noisy channel 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 purification has been accomplished, 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 noisy one, supplemented by local actions and two-way classical communication.

Let W' be a general mixed state of two spin-$\frac{1}{2}$ particles, from which we wish to distill some pure entanglment. The state W' 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 quality of this transmission can be conveniently expressed as a fidelity[1] F of W' relative to a perfect singlet:

$$F = \langle \Psi^- | W' | \Psi^- \rangle. \tag{3}$$

This is the probability that the mixed state W' would pass a test for being a perfect singlet. Although the test of projecting the mixed state onto a singlet is nonlocal, F can be computed from the results of local measurements. For example, if Alice and Bob perform a random bilateral SU(2) rotation of their spins, rotating each spin by the same random amount about the same random

[1]This is a slightly different use of the notion of fidelity from that suggested earlier. The fidelity considered here is the fidelity of an abstract channel whose input is always a perfect singlet and whose output is the mixed state W' resulting from sending one or both spins in the singlet through a real noisy channel.

axis, then measure their respective spins in the σ_z basis, their measurement results will always be opposite if W' is a singlet, and will be opposite with probability $(1+2F)/3$ for a general mixed state. This random bilateral rotation requires classical communication between Alice and Bob but no exchange of quantum information.

The first step in our protocol for entanglement purification is to have Alice and Bob perform the bilateral random rotation just described. This transforms a general two-spin mixed state W' into a rotationally symmetric mixture

$$W = W_F = F \cdot |\Psi^-\rangle\langle\Psi^-| + \frac{1-F}{3}(|\Psi^+\rangle\langle\Psi^+| + |\Phi^+\rangle\langle\Phi^+| + |\Phi^-\rangle\langle\Phi^-|) \quad (4)$$

of the singlet state, which is invariant under bilateral rotations, 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 state's invariance under bilateral rotations, the above symmetrization process does not affect the fidelity; indeed, as our notation suggests, the resulting symmetrized state W_F is completely parameterized by the original mixed state's fidelity. The mixed state $W_{5/8}$, which can be prepared by sending one member of a singlet through a 50% depolarizing channel, was shown by Popescu [4] to exhibit subtle non-local properties even though Werner [5] showed that it violates no Bell inequality.

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 W', or of the noisy channel(s) that may have generated it, except for the fidelity F, which parameterizes the symmetrized state W.

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

- unilateral π-radian rotations of *one* particle in an entangled pair about the x, y, or z axis. These operations map the Bell states onto one another in a 1:1 pairwise fashion, leaving no state invariant. For example, a π rotation around the y axis turns Ψ^- into Φ^+ and vice versa, and likewise interchanges Ψ^+ and Φ^-. Here we ignore overall phase changes because they do not affect any of our arguments.

- bilateral $\pi/2$ rotations of *both* particles in a pair about the x, y, or z axis. Each of these operations leaves the singlet state and a different one of the triplets invariant, interchanging the other two triplets, as shown in the table below (again we omit irrelevant overall phases).

Bilateral	State			
$\pi/2$ Rotation	Φ^+	Φ^-	Ψ^+	Ψ^-
x	Ψ^+	Φ^-	Φ^+	Ψ^-
y	Φ^+	Ψ^+	Φ^-	Ψ^-
z	Φ^-	Φ^+	Ψ^+	Ψ^-

- the quantum XOR or controlled-NOT operation [6] performed bilaterally by both observers on corresponding members of two shared pairs. For example, if Alice holds spins 1 and 3, and Bob holds spins 2 and 4, a bilateral XOR, with spins 1 and 2 as source and spins 3 and 4 as target would conditionally flip spin 3 iff spin 1 was up, while conditionally flipping spin 4 iff spin 2 was up. The bilateral XOR operating on two Φ^+ pairs leaves them both invariant. The result of applying it to some other combinations of Bell states is shown below. Cases not included in this table (*i.e.* where the target is Ψ^- or Φ^-) are not needed in the protocol.

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.

(5)

Besides these unitary operations, Alice and Bob perform one kind of measurement, which consists of measuring both spins in a given pair along the z spin axis. This measurement 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. Its entanglement has been sacrificed to enable Bob and Alice to acquire classical information about its before-measurement Bell state, and thus indirectly about other pairs with which it may have interacted.

The purification protocol can be described more simply if we assume that Alice and Bob have access to an initial pool of Φ^+ pairs of precertified high purity. Later we show how this pool can be replenished, or even created *de novo* from the slightly-mixed entangled states W_F that are the input to the purification procedure. The precertified Φ^+ states are used as targets for XORs involving one or more other pairs, then measured to gain information about these pairs without destroying their entanglement. By an "XOR test" we mean the sequential XORing of a subset of pairs into a precertified target, followed by a measurment of the target to determine whether it is a Φ or Ψ state. Consulting the table above, we see that each bad Ψ^+ or Ψ^- source pair toggles the target between Φ^+ and Ψ^+. Thus an XOR test, like a parity check on classical data, tells Alice and Bob whether there are an even or odd number of errors (ie $\Psi^\pm$ states) in their data, without disturbing any good Φ^+ states that may be present. Φ^- states, also "bad", are not distinguished from good Φ^+ states by the XOR test, but can be detected at a later stage of testing by converting them into Ψ^+ states. The full protocol is described below.

1. By a unilateral y rotation, Alice and Bob convert the approximate singlets, described by density matrix W_F, into approximate Φ^+ states.

2. Using the precertified Φ^+ pairs as targets, they perform XOR tests on sufficiently many random subsets of the pairs to locate all $\Psi^\pm$ states with high probability, without distinguishing between Ψ^+ and Ψ^-. Once found, these states are discarded.

3. Next Alice and Bob do a bilateral y rotation to convert the so far undetectable Φ^- states into Ψ^+, while leaving the Φ^+ states invariant. This done, they perform XOR tests on sufficiently many more random subsets to find all the new Ψ^+ states with high probability. Once found, these states need not be discarded, since they can be corrected to the desired Φ^+ form by a bilateral x rotation.

The XOR test is an economical way of finding errors, in the sense that each such test reveals approximately 1 bit about the pattern of errors. How many XOR tests, each consuming a precertified Φ^+, are needed to find errors occurring at some rate η? The answer depends on how much classical decoding effort Alice and Bob are willing to expend to infer the error pattern from the results of the XOR tests. (Like classical error-correcting codes, the results of a set of XOR tests may be uniquely decodable without being efficiently decodable.) If we assume that exponential decoding effort is acceptable, then they need only compute enough random subset parities to unambiguously determine the pattern of errors. It is not hard to show [7] that for any n-bit string x the probability that there exists another string y differing from x in at most k positions, but passing ℓ XOR tests (*i.e.* agreeing with x on the parities of ℓ subsets chosen randomly and independently of x), is approximately 1 for $\ell < nH_2(k/n)$ but decreases exponentially with ℓ for $\ell > nH_2(k/n)$ (here H_2 is the dyadic entropy function $H_2(p) = -p\log_2 p - (1-p)\log_2(1-p)$). Therefore $H_2(\eta)$ XOR tests per entangled pair are asymptotically both necessary and sufficient.[2]

Note that there are two rounds of XOR testing, with different error rates η. The first round finds and discards the $\Psi^\pm$ states, while the second finds and corrects the Φ^- states. Since the symmetrized state W_F has equal probabilities of all three kinds of triplet, the error rate for the first stage will be twice that of the second. Thus the amount of entanglement sacrificed in the purification process, equal to the number of XOR tests performed plus the number of unresolved $\Psi^\pm$ pairs discarded, is given by the von Neumann entropy

$$H(W_F) = -\mathrm{Tr} W_F \log_2 W_F = -F\log_2 F - (1-F)\log_2 \frac{1-F}{3} \qquad (6)$$

because each XOR test ideally (*i.e.* allowing exponential decoding effort) reveals one bit of information about the pattern of errors in the symmetrized mixture. For $F < 0.8107$ the entropy is greater than 1, meaning that no purified entangled pairs can be salvaged. The yield of purified entangled pairs is thus

$$D(W) = \max\{0, 1 + \mathrm{Tr} W \log_2 W\}. \qquad (7)$$

The above protocol can clearly be applied to any Bell-diagonal mixed state W, rotationally symmetric or not, giving an asymptotic yield $D(W)$ of

[2]If linear decoding effort is required, more XOR tests are needed, and they should be calculated not over random subsets of the entire data, but over subsets prescribed by some efficiently-decodable post-facto error correction scheme, such as convolutional coding as in [8] or blockwise XORs as in [9]. With convolutional codes, $\log_2(1 + 2\sqrt{\eta(1-\eta)})$ XOR tests per spin suffice to find the errors efficiently.

purified pairs according to the above equation. We conjecture that this yield is optimal for states of this form.

Once all the errors have been found and removed or corrected, Alice and Bob use some of the resulting purified Φ^+ states to replenish the precertified Φ^+ pool, while converting the others into standard singlets for use in teleportation. (To establish the certified Φ^+ pool initially, a batch of untested mostly-Φ^+ pairs is subjected to repeated passes of the purification procedure using other untested pairs as targets, and the entire batch rejected unless all XOR tests are passed. The fidelity of the surviving pairs, relative to a pure Φ^+ state, can be increased arbitrarily by this procedure, provided the inintial fidelity is not too low.)

The issue of entanglement purification raises the following more general question: How does one quantify the notion of "degree of entanglement" for mixed states? For a *pure* state of two particles, it is perfectly clear how to define the entanglement E: it is the entropy of the density matrix of each particle separately. (The two particles always have equal entropies if their joint state is pure.) This definition is compelling because a large collection of n pairs, each in the same pure state having entanglement E, can be converted by local operations into a collection of nE pure singlets, and vice-versa [10]. However, for mixed states there need be no such perfect equivalence with pure singlets, so there is no unique generalization of the above definition. Instead, there are two natural generalizations:

- D, the maximum amount of pure entanglement one can *distill* from the given state.
- E, the minimum amount of pure entanglement required to *construct* the given state.

We now compare these two quantities for the state W_F defined by Eq. (4).

We have shown that one can distill from W_F an amount of entanglement given by Eq. (7), and we conjecture that one can do no better than that. To establish the amount of entanglement E required to construct W_F, we first present a specific ensemble of pure states whose collective density matrix is W_F, and we calculate their average entanglement. We then show that no other pure-state ensemble one could use to construct W_F has less entanglement.

Consider first the case in which the fidelity F is greater than or equal to one-half. Our proposed minimum-entanglement ensemble in that case consists of the eight pure states of the form

$$|s\rangle = \sqrt{F}|\Psi^-\rangle + \sqrt{\frac{1-F}{3}}(\pm|\Psi^+\rangle \pm |\Phi^-\rangle \pm i|\Phi^+\rangle), \tag{8}$$

where the three variable signs are taken in all possible combinations. One can easily show that the average of the projections on these eight states, with equal weights, is equal to W_F. These states all have the same entanglement, which turns out to be

$$E = H_2(\tfrac{1}{2} + \sqrt{F(1-F)}). \tag{9}$$

Notice that when $F = \frac{1}{2}$, the entanglement of these pure states is zero, which shows that $W_{1/2}$ can be constructed locally starting with no entanglement at all. Not surprisingly, W_F for $F < \frac{1}{2}$ can likewise be constructed with no entanglement. For example, one can mix together a number of pure states each of which is an unentangled superposition of just two Bell states.

We now show that for F greater than one-half, there is no "cheaper" way to construct W_F than by using the states given in Eq. (8). Our proof relies on the fact that in order to construct W_F, one must use pure states that, on average, have enough of the singlet component to give the whole ensemble the fidelity F. But having a large singlet component forces a pure state to have a certain degree of entanglement. In fact, one can show that if a pure state has fidelity $f \geq \frac{1}{2}$ relative to the singlet, then its entanglement must be at least $H_2(\frac{1}{2} + \sqrt{f(1-f)})$. But this minimum entanglement is a convex function of f. Therefore, the *average* entanglement of the ensemble of pure states can be no less than the minimum entanglement of a pure state having the *average* fidelity F. That is, the average entanglement can be no less than $H_2(\frac{1}{2} + \sqrt{F(1-F)})$, which is the entanglement we obtained above using the states (8). Thus Eq. (9) gives the minimum entanglement required to construct the state W_F if $F \geq \frac{1}{2}$.

In Fig. (1) we show both E and D as functions of the fidelity of the state. Fidelity $F = 1$ corresponds to the pure singlet state, which requires a full bit of entanglement for its construction ($E = 1$) and also allows the extraction of a full bit of entanglement ($D = 1$). For slightly distorted singlets, however, it appears that the entanglement one can extract falls short of the the minimum amount required to make the state. This is not an unreasonable conclusion: when we mix together the states (8) to make W_F, we are performing an irreversible operation, so it is plausible that we will not be able to recover all the entanglement that we have put in.

Both the new entanglement measures D and E are distinct from another commonly used measure of entanglement, viz. violation of the Clauser-Horn-Shimony-Holt (CHSH) [11] inequality. It is easy to verify that the mixed state W_F violates the CHSH inequlity if and only if $F > (1+\sqrt{2})/3 = 0.80474$. This is much higher than the threshold $F = \frac{1}{2}$ above which E becomes positive, and slightly less than the threshold $F = 0.81071$ above which D becomes positive. Yet another criterion of nonlocality, introduced by Popescu [4], is the ability of a mixed entangled state to improve the fidelity of teleportation of a random unknown quantum state above that ($F = \frac{2}{3}$) which could be achieved by classical communication alone. The states W_F improve the fidelity of teleportation over the same range $F > \frac{1}{2}$ for which E is nonzero.

Acknowledgments

We wish to thank Sandu Popescu and David DiVincenzo for valuable advice and criticism.

References

[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**, 2343-2349 (1994).

[4] S. Popescu, *Phys. Rev. Lett.* **72**, 797 (1994).

[5] R. F. Werner, *Phys. Rev. A* **40**, 4277 (1989).

[6] D. Deutsch, *Proc. Roy. Soc. A* **425**, 73 (1989).

[7] C. Bennett, G. Brassard, C. Crépeau, and U. Maurer, "Genrealized Privacy Amplification", submitted to *IEEE Trans. on Information Theory.*

[8] C. H. Bennett, G. Brassard and J.–M. Robert, *SIAM Journal on Computing*, Vol. 17, no. 2, April 1988, pp. 210–229.

[9] Bennett, C. H., F. Bessette, G. Brassard, L. Salvail and J. Smolin, *J. Cryptology* **5**, 3 (1992).

[10] C. Bennett, H. Bernstein, S. Popescu, B. Schumacher, "Entanglement Concentration", preprint 1994.

[11] J.F. Clauser, M.A. Horne, A. Shimony, R.A. Holt, *Phys. Rev. Lett.* **23**, 880 (1980).

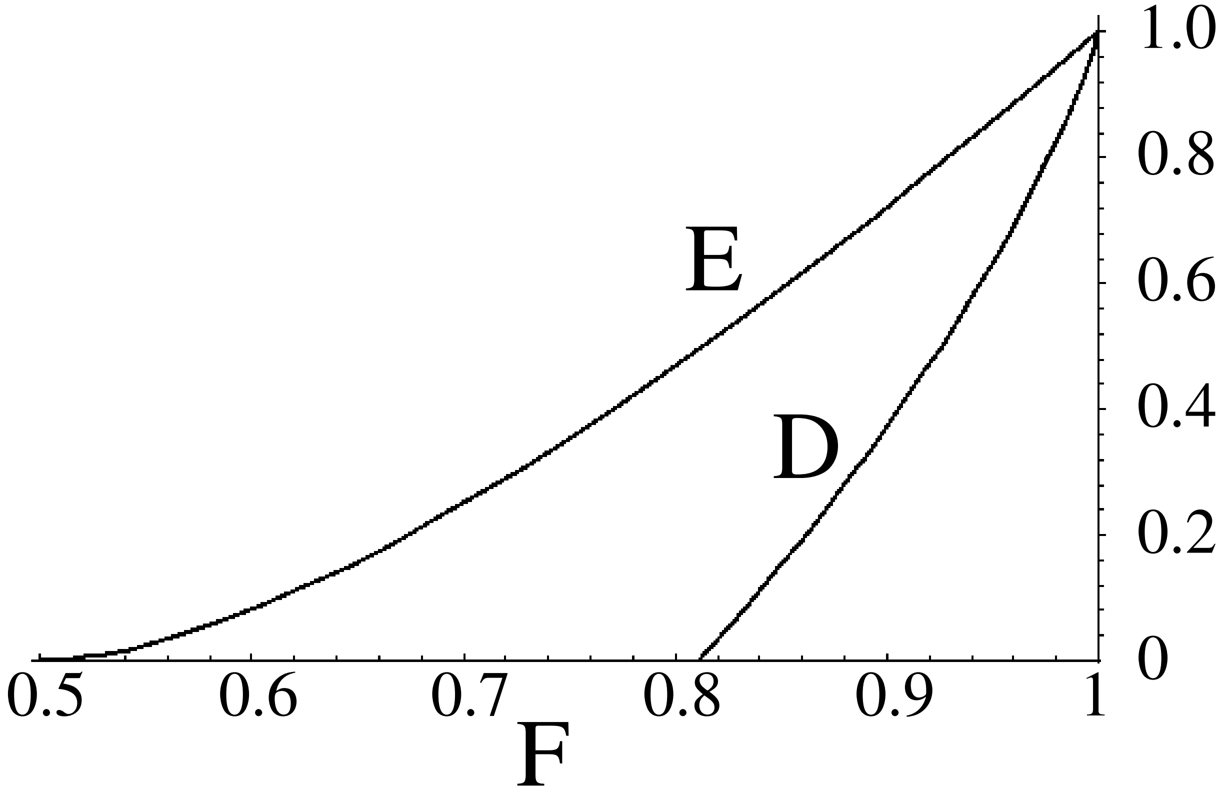

Figure 1: "Entanglement of formation" E and distillable entanglement D as functions of F, the fidelity relative to the pure singlet.