Optimal Remote State Preparation

Dominic W. Berry and Barry C. Sanders

Department of Physics and Centre for Advanced Computing—Algorithms and Cryptography, Macquarie University, Sydney, New South Wales 2109, Australia

(Received 15 September 2002; published 4 February 2003)

We prove that it is possible to remotely prepare an ensemble of noncommuting mixed states using communication equal to the Holevo information for this ensemble. This remote preparation scheme may be used to convert between different ensembles of mixed states in an asymptotically lossless way, analogous to concentration and dilution for entanglement.

DOI: 10.1103/PhysRevLett.90.057901 PACS numbers: 03.67.Hk, 03.65.Ud

In classical information theory one of the central problems is that of coding. From Shannon's noiseless coding theorem [1], if message i is given with probability p_i, then a sequence of messages may be compressed to an average number of bits per message equal to the Shannon entropy $H = -\sum_i p_i \log_2 p_i$. This result means that classical communication of H bits per message is sufficient to reconstruct the sequence of messages; conversely H bits of communication per message may be obtained. Thus the Shannon entropy may be given a definite interpretation as the classical information per message. Here we show that, in the quantum case where density ρ_i is given with probability p_i, it is possible to give a similar interpretation to the Holevo information. The Holevo information is given by $S(\sum_i p_i\rho_i) - \sum_i p_i S(\rho_i)$, where $S(\rho) = -\mathrm{Tr}(\rho\log_2\rho)$ is the von Neumann entropy. It has previously been shown [2–4] that it is possible to perform classical communication equal to the Holevo information. We show that it is possible to effectively reverse this process and remotely prepare these states using communication equal to the Holevo information. This result means that it is possible to convert from ensembles of mixed states to classical information and back again in an asymptotically lossless way, analogous to concentration and dilution of entanglement [5].

In remote state preparation [6–10], Alice (A) wishes to prepare state ρ_i with probability p_i in the laboratory of Bob (B). Because this ensemble of mixed states $\mathsf{E} = \{p_i, \rho_i\}$ may be used to perform communication equal to its Holevo information [3,4], the Holevo information provides a lower bound to the communication required for remote state preparation [8]. The problem of remote state preparation at this lower bound has hitherto been solved only for the special case that all densities to be prepared commute [11].

It is instructive to first summarize a nonoptimal scheme for preparing a single state from Ref. [8]. In order to approach this problem, it is convenient to consider preparation of a pure state $|\Phi_i\rangle$ shared between Alice and Bob, such that Bob's reduced density matrix is ρ_i. We denote this by the ensemble $\mathcal{E} = \{p_i, |\Phi_i\rangle\}$. The state $|\Phi_i\rangle$ has Schmidt decomposition

$$|\Phi_i\rangle = \sum_{j\in D}\sqrt{\lambda_i^j}|\varphi_i^j\rangle_A|\chi_i^j\rangle_B. \tag{1}$$

Both modes are of dimension d, and $D = \{1, \ldots, d\}$. Preparing this state is equivalent to remotely preparing the mixed state $\rho_i = \sum_{j\in D}\lambda_i^j|\chi_i^j\rangle_B\langle\chi_i^j|$.

In order to remotely prepare ρ_i, Alice consumes entanglement and performs classical communication to Bob. Via local operations on Alice's side, any maximally entangled state may be brought to the form

$$|\Theta_i\rangle = (1/\sqrt{d})\sum_{j\in D}|\varphi_i^j\rangle_A|\chi_i^j\rangle_B. \tag{2}$$

Alice then performs a measurement described by a positive operator-valued measure (POVM) with two elements, $\Pi_i^1 = \frac{1}{\Lambda_i}\sum_{j\in D}\lambda_i^j|\varphi_i^j\rangle_A\langle\varphi_i^j|$ and $\Pi_i^0 = \mathbb{1} - \Pi_i^1$, where $\Lambda_i = \max_j\{\lambda_i^j\}$. If the measurement result is 1, the resulting (unnormalized) state is

$$(\sqrt{\Pi_i^1}\otimes\mathbb{1})|\Theta_i\rangle = (1/\sqrt{\Lambda_i d})\sum_{j\in D}\sqrt{\lambda_i^j}|\varphi_i^j\rangle_A|\chi_i^j\rangle_B. \tag{3}$$

Normalization gives the state Alice wished to prepare, $|\Phi_i\rangle$. This result occurs with probability $1/(\Lambda_i d)$. The state resulting from a measurement result of 0 is not usable, and this measurement result counts as a failure.

The preparation of the state therefore requires the classical communication of the number of the measurement result that is a success. Let us assume that the measurement is performed a maximum of M times, where M is the smallest integer not less than $\Lambda d\ln(\Lambda d)$, for $\Lambda = \max_i\{\Lambda_i\}$. If there is a success, then the number of the success is communicated; otherwise zero is communicated. As there are no more than $M + 1$ alternative messages to communicate, the number of bits required is $\log(\Lambda d) + O[\log\log(\Lambda d)]$. Throughout this Letter we denote logarithms to base 2 by log and logarithms to base e by ln.

The probability of all the measurements being failures is $(1 - 1/\Lambda_i d)^M \leq 1/\Lambda d$. If there is a success the fidelity is equal to 1, so the average fidelity must be at least $1 - 1/\Lambda d$. Therefore, in the limit of large Λd, the communication required is $\log(\Lambda d) + O[\log\log(\Lambda d)]$, and the average fidelity is arbitrarily close to 1.

The state preparation protocol of Ref. [8] does not, in general, reach the Holevo limit. In order to reach the Holevo limit, we generalize this state preparation scheme to jointly prepare a number of states. That is, we prepare the tensor product of n states $|\Phi_u\rangle = |\Phi_{i_1}\rangle \otimes \cdots \otimes |\Phi_{i_n}\rangle$ with probability $p_u = p_{i_1} \times \cdots \times p_{i_n}$. Here we use the notation $u = (i_1, \ldots, i_n)$. This tensor product state has the Schmidt decomposition

$$|\Phi_u\rangle = \sum_{J\in D^n} \sqrt{\lambda_u^J} |\varphi_u^J\rangle_A |\chi_u^J\rangle_B, \tag{4}$$

where $\lambda_u^J = \lambda_{i_1}^{j_1} \times \cdots \times \lambda_{i_n}^{j_n}$, $|\varphi_u^J\rangle_A = |\varphi_{i_1}^{j_1}\rangle_A \otimes \cdots \otimes |\varphi_{i_n}^{j_n}\rangle_A$, and $|\chi_u^J\rangle_B = |\chi_{i_1}^{j_1}\rangle_B \otimes \cdots \otimes |\chi_{i_n}^{j_n}\rangle_B$.

We introduce the subspace of $|\Phi_u\rangle$,

$$B_u = \{J: \lambda_u^J < 2^{-n(\bar{S}-\delta)}\}, \tag{5}$$

where $\bar{S} = \sum_i p_i S(\rho_i)$. We also use the notation $\bar{\rho} = \sum_i p_i \rho_i$, so the Holevo information for the ensemble is $S(\bar{\rho}) - \bar{S}$. Given any positive ϵ and δ, there exists $n_2(\epsilon, \delta)$ such that, for all $n > n_2(\epsilon, \delta)$,

$$\left\langle \sum_{J\in B_u} \lambda_u^J \right\rangle \geq 1 - \epsilon, \tag{6}$$

where the expectation value indicates the average over u with probabilities p_u. In Ref. [4], this result is given for the subspace $B'_u = \{J: 2^{-n(\bar{S}+\delta)} < \lambda_u^J < 2^{-n(\bar{S}-\delta)}\}$. The subspace we use includes additional small values of λ_u^J, which can only increase the sum. Therefore the result (6) must hold for the subspace B_u.

Now we introduce a POVM with elements

$$\Pi_u^1 = 2^{n(\bar{S}-\delta)} \sum_{J\in B_u} \lambda_u^J |\varphi_u^J\rangle_A \langle \varphi_u^J| \tag{7}$$

and $\Pi_u^0 = \mathbb{1} - \Pi_u^1$. As above, a maximally entangled state shared between Alice and Bob may be brought, via local operations on Alice's side, to the form

$$|\Theta_u\rangle = \frac{1}{d^{n/2}} \sum_{J\in D^n} |\varphi_u^J\rangle_A |\chi_u^J\rangle_B. \tag{8}$$

Alice performs a measurement described by the above POVM on this entangled state. After a measurement which yields the result 1, the resulting unnormalized state is

$$(\sqrt{\Pi_u^1} \otimes \mathbb{1})|\Theta_u\rangle = \frac{2^{n(\bar{S}-\delta)/2}}{d^{n/2}} \sum_{J\in B_u} \sqrt{\lambda_u^J} |\varphi_u^J\rangle_A |\chi_u^J\rangle_B. \tag{9}$$

With normalization the state may be written as

$$|\Phi'_u\rangle = \frac{1}{\mathcal{N}} \sum_{J\in B_u} \sqrt{\lambda_u^J} |\varphi_u^J\rangle_A |\chi_u^J\rangle_B, \tag{10}$$

where $\mathcal{N} = \sqrt{\sum_{J\in B_u} \lambda_u^J}$ is a normalization factor. This state is not exactly equal to the state that was to be prepared (4); however, the fidelity is

$$F_u = |\langle \Phi_u | \Phi'_u \rangle|^2 = \frac{1}{\mathcal{N}^2} \left| \sum_{J\in B_u} \lambda_u^J \right|^2 = \sum_{J\in B_u} \lambda_u^J. \tag{11}$$

Using Eq. (6), we find that $\langle F_u \rangle \geq 1 - \epsilon$.

The dimension of the space used is d^n, and the maximum λ_u^J is no greater than $2^{-n(\bar{S}-\delta)}$. Therefore, from the above discussion for the preparation of a single state, the state $|\Phi'_u\rangle$ can be prepared with probability of success at least $1 - d^{-n} 2^{n(\bar{S}-\delta)}$ and communication $n(\log d - \bar{S} + \delta) + O(\log n)$. The average fidelity with the state $|\Phi_u\rangle$ must therefore be at least $1 - \epsilon - d^{-n} 2^{n(\bar{S}-\delta)}$.

The average fidelity of the reduced density matrices in Bob's mode must also be at least $1 - \epsilon - d^{-n} 2^{n(\bar{S}-\delta)}$ due to the relation [12] $F(\rho, \sigma) = \max_{|\psi\rangle, |\phi\rangle} |\langle \psi | \phi \rangle|^2$, where ρ and σ are density matrices and $|\psi\rangle$ and $|\phi\rangle$ are purifications of ρ and σ.

We therefore see that the states $|\Phi_u\rangle$, or the corresponding reduced density matrices for Bob, may be prepared with average fidelity arbitrarily close to 1 and communication per prepared state arbitrarily close to $\log d - \bar{S}$. This communication is still larger, in general, than the Holevo bound of $S(\bar{\rho}) - \bar{S}$. In order to reach this bound we combine this protocol with what is effectively Schumacher compression [13].

Using Schumacher compression, the ensemble of states to be prepared, E, may be compressed to a space of dimension $2^{S(\bar{\rho})}$. These states may therefore be prepared in the above way with communication $S(\bar{\rho}) - \bar{S}$. We now show that this compression may be applied to the preparation of entangled states, and with fidelity arbitrarily close to 1.

In order to apply Schumacher compression, we use a method similar to that of Lo [14]. Lo shows that, given an ensemble of density matrices $\mathsf{E} = \{p_i, \rho_i\}$ to be transmitted, for any $\epsilon, \delta > 0$, there exists an n such that the sequence of density matrices $\rho_u = \rho_{i_1} \otimes \cdots \otimes \rho_{i_n}$ may be compressed to $S(\bar{\rho}) + \delta$ qubits with average "distortion" less than ϵ.

It is straightforward to modify Lo's derivation so that it deals with ensembles of entangled states, and the fidelity is used rather than the distortion. For simplicity we consider preparation of the state

$$|\Phi_u\rangle = \sum_{J\in D^n} \sqrt{\lambda_u^J} |\chi_u^J\rangle_A |\chi_u^J\rangle_B. \tag{12}$$

This state may be brought to the form (4) via unitary operations on Alice's mode. As explained in Ref. [4], for all $\epsilon, \delta > 0$ there is an $n_1(\epsilon, \delta)$ such that, for all $n > n_1(\epsilon, \delta)$, $\mathrm{Tr}(\bar{\rho}^{\otimes n} P) > 1 - \epsilon$, where P is a projector onto a space of dimension $2^{n[S(\bar{\rho})+\delta]}$. We denote this space by Ξ and write $|\chi_u^J\rangle$ in the form

$$|\chi_u^J\rangle = \alpha_u^J |l_u^J\rangle + \beta_u^J |m_u^J\rangle, \tag{13}$$

where $\alpha_u^J, \beta_u^J \geq 0$, $(\alpha_u^J)^2 + (\beta_u^J)^2 = 1$, and the states $|l_u^J\rangle$ and $|m_u^J\rangle$ are in the spaces Ξ and $\Xi^\perp$, respectively.

Now we replace each state (12) with

$$|\tilde{\Phi}_u\rangle = \sum_{J\in D^n} \sqrt{\lambda_u^J}(\alpha_u^J)^2|l_u^J\rangle_A|l_u^J\rangle_B + \mathcal{B}_u|l_u^0\rangle_A|l_u^0\rangle_B. \quad (14)$$

The state $|l_u^0\rangle$ may be chosen arbitrarily. The value of the coefficient $\mathcal{B}_u$ is chosen such that $\mathcal{B}_u\langle\Phi_u|l_u^0\rangle_A|l_u^0\rangle_B$ is a positive real number and normalization is preserved. We find that

$$\begin{aligned}\langle\Phi_u|\tilde{\Phi}_u\rangle &= \sum_{J,J'} \sqrt{\lambda_u^J\lambda_u^{J'}}(\alpha_u^J)^2(\alpha_u^{J'})^2\langle l_u^J|l_u^{J'}\rangle^2 + \mathcal{B}_u\langle\Phi_u|l_u^0\rangle_A|l_u^0\rangle_B \geq \sum_J \lambda_u^J(\alpha_u^J)^4 + \sum_{J\neq J'} \sqrt{\lambda_u^J\lambda_u^{J'}}(\alpha_u^J)^2(\alpha_u^{J'})^2\langle l_u^J|l_u^{J'}\rangle^2 \\ &= \sum_J \lambda_u^J[1-2(\beta_u^J)^2+(\beta_u^J)^4] + \sum_{J\neq J'} \sqrt{\lambda_u^J\lambda_u^{J'}}(\beta_u^J)^2(\beta_u^{J'})^2\langle m_u^J|m_u^{J'}\rangle^2 \\ &= 1-2\sum_J \lambda_u^J(\beta_u^J)^2 + \sum_{J,J'} \sqrt{\lambda_u^J\lambda_u^{J'}}(\beta_u^J)^2(\beta_u^{J'})^2\langle m_u^J|m_u^{J'}\rangle^2 \geq 1-2\sum_J \lambda_u^J(\beta_u^J)^2. \end{aligned} \quad (15)$$

The average fidelity is therefore

$$\begin{aligned} F &\geq \sum_u p_u \left|1-2\sum_J \lambda_u^J(\beta_u^J)^2\right|^2 \\ &\geq \sum_u p_u\left[1-4\sum_J \lambda_u^J(\beta_u^J)^2\right] \geq 1-4\epsilon. \end{aligned} \quad (16)$$

In the last line we have used $\mathrm{Tr}(\bar{\rho}^{\otimes n}P) > 1-\epsilon$.

In addition to $n > n_1(\epsilon,\delta)$, we take $n > n_2(\epsilon,\delta)$ and $n > 1/\delta$, so the Schmidt coefficients for $|\Phi_u\rangle$ satisfy Eq. (6). Now the modified states $|\tilde{\Phi}_u\rangle$ have the Schmidt decompositions $|\tilde{\Phi}_u\rangle = \sum_{J\in D^n}\sqrt{\tilde{\lambda}_u^J}|\tilde{\chi}_u^J\rangle_A|\tilde{\chi}_u^J\rangle_B$. We introduce the subspace

$$\tilde{B}_u = \{J:\tilde{\lambda}_u^J < 2^{-n(\bar{S}-2\delta)}\}. \quad (17)$$

Now we may place limits on the sum over the $\tilde{\lambda}_u^J$ in the following way:

$$\begin{aligned}\sum_{J\in\tilde{B}_u}\tilde{\lambda}_u^J &= 1-\sum_{J\in\tilde{B}_u^\perp}\tilde{\lambda}_u^J = 1-\sum_{J\in\tilde{B}_u^\perp}\langle\tilde{\chi}_u^J|\left[\tilde{\lambda}_u^J|\tilde{\chi}_u^J\rangle\langle\tilde{\chi}_u^J| - \sum_{J'\in D^n}\lambda_u^{J'}|\chi_u^{J'}\rangle\langle\chi_u^{J'}|\right]|\tilde{\chi}_u^J\rangle - \sum_{J\in\tilde{B}_u^\perp}\sum_{J'\in D^n}\lambda_u^{J'}|\langle\chi_u^{J'}|\tilde{\chi}_u^J\rangle|^2 \\ &= 1-\mathrm{Tr}[\tilde{P}(\tilde{\rho}_u-\rho_u)] - \sum_{J\in\tilde{B}_u^\perp}\sum_{J'\in D^n}\lambda_u^{J'}|\langle\chi_u^{J'}|\tilde{\chi}_u^J\rangle|^2. \end{aligned} \quad (18)$$

Here $\tilde{P}$ is the projector $\sum_{J\in\tilde{B}_u^\perp}|\tilde{\chi}_u^J\rangle\langle\tilde{\chi}_u^J|$, $\tilde{\rho}_u$ is Bob's reduced density operator for state $|\tilde{\Phi}_u\rangle$ and $\tilde{B}_u^\perp = D^n\backslash\tilde{B}_u$. Using results from Ref. [15] we have the inequalities

$$\mathrm{Tr}[\tilde{P}(\tilde{\rho}_u-\rho_u)] \leq D(\tilde{\rho}_u,\rho_u) \leq \sqrt{1-F(\tilde{\rho}_u,\rho_u)} \leq \sqrt{1-F(|\tilde{\Phi}_u\rangle,|\Phi_u\rangle)}, \quad (19)$$

where $D(\tilde{\rho}_u,\rho_u) = \frac{1}{2}\mathrm{Tr}|\tilde{\rho}_u-\rho_u|$ is the trace distance. Note that the fidelity defined in Ref. [15] is the square root of the fidelity defined here. The third term on the right-hand side of Eq. (18) may be evaluated as

$$\begin{aligned}\sum_{J\in\tilde{B}_u^\perp}\sum_{J'\in D^n}\lambda_u^{J'}|\langle\chi_u^{J'}|\tilde{\chi}_u^J\rangle|^2 &= \sum_{J\in\tilde{B}_u^\perp}\sum_{J'\in B_u}\lambda_u^{J'}|\langle\chi_u^{J'}|\tilde{\chi}_u^J\rangle|^2 + \sum_{J'\in B_u^\perp}\lambda_u^{J'}\sum_{J\in\tilde{B}_u^\perp}|\langle\chi_u^{J'}|\tilde{\chi}_u^J\rangle|^2 \\ &\leq \sum_{J\in\tilde{B}_u^\perp}2^{-n(\bar{S}-\delta)}\sum_{J'\in B_u}|\langle\chi_u^{J'}|\tilde{\chi}_u^J\rangle|^2 + \sum_{J'\in B_u^\perp}\lambda_u^{J'} \leq \sum_{J\in\tilde{B}_u^\perp}2^{-n(\bar{S}-\delta)} + \sum_{J\in B_u^\perp}\lambda_u^J \\ &= 2^{-n\delta}\sum_{J\in\tilde{B}_u^\perp}2^{-n(\bar{S}-2\delta)} + \sum_{J\in B_u^\perp}\lambda_u^J \leq \frac{1}{2}\sum_{J\in\tilde{B}_u^\perp}\tilde{\lambda}_u^J + \sum_{J\in B_u^\perp}\lambda_u^J. \end{aligned} \quad (20)$$

Combining this with Eq. (18) gives

$$\begin{aligned}\sum_{J\in\tilde{B}_u}\tilde{\lambda}_u^J &\geq 1-\sqrt{1-F(|\tilde{\Phi}_u\rangle,|\Phi_u\rangle)} - \frac{1}{2}\sum_{J\in\tilde{B}_u^\perp}\tilde{\lambda}_u^J - \sum_{J\in B_u^\perp}\lambda_u^J \\ &\geq 1-2\sqrt{1-F(|\tilde{\Phi}_u\rangle,|\Phi_u\rangle)} - 2\sum_{J\in B_u^\perp}\lambda_u^J. \end{aligned} \quad (21)$$

Determining the expectation value over u gives

$$\left\langle\sum_{J\in\tilde{B}_u}\tilde{\lambda}_u^J\right\rangle \geq 1-4\sqrt{\epsilon}-2\epsilon. \quad (22)$$

The dimension of the space used is $2^{n[S(\bar{\rho})+\delta]}$, and the maximum $\tilde{\lambda}_u^J$ is no larger than $2^{-n(\bar{S}-2\delta)}$. Therefore it is possible to remotely prepare the state $|\tilde{\Phi}_u\rangle$ with classical communication $n[S(\bar{\rho})-\bar{S}+3\delta]+O(\log n)$ bits and probability of success at least $1-2^{-n[S(\bar{\rho})-\bar{S}+3\delta]}$. Because $\langle\sum_{J\in\tilde{B}_u}\tilde{\lambda}_u^J\rangle \geq 1-4\sqrt{\epsilon}-2\epsilon$, the average fidelity for a success is at least $1-4\sqrt{\epsilon}-2\epsilon$. The average fidelity including failures must be at least $1-4\sqrt{\epsilon}-2\epsilon-2^{-n[S(\bar{\rho})-\bar{S}+3\delta]}$.

Because the average fidelity between the states $|\tilde{\Phi}_u\rangle$ and $|\Phi_u\rangle$ is at least $1-4\epsilon$, it is easy to see from the triangle inequality for fidelities that the states $|\Phi_u\rangle$ may be prepared with average fidelity arbitrarily close to 1. Therefore we see that the states $|\Phi_u\rangle$ may be prepared with fidelity arbitrarily close to 1 and communication per

TABLE I. Three analogous processes for classical communication vs entanglement. In boldface is shown the optimal remote state preparation discussed in this Letter, as well as the conversion between ensembles, discussed in Ref. [16], for which optimal remote state preparation is required.

Classical communication	Entanglement
Communication equal to the Holevo information	Entanglement concentration
Optimal remote state preparation	Entanglement dilution
Conversion between ensembles	Conversion between entangled states

prepared state arbitrarily close to the Holevo information $S(\bar{\rho}) - \bar{S}$.

The situation that we have considered, where classical communication is the resource of interest and entanglement is a free resource, is analogous to that of entanglement concentration and dilution [5], where entanglement is the resource under consideration and classical communication is a free resource.

The results of Refs. [2–4] show that it is possible to perform classical communication equal to the Holevo information (analogous to entanglement concentration). Here we have shown that it is possible to remotely prepare ensembles of mixed states using communication equal to the Holevo information (analogous to entanglement dilution).

One consequence of our proof is that it is possible to convert between multiple copies of different ensembles with the same total Holevo information [16], in an analogous way as it is possible to convert between different pure entangled states via entanglement concentration and dilution. These analogies are summarized in Table I.

An important application of our optimal remote state preparation scheme is given in Ref. [16]. Reference [16] shows that, provided it is possible to efficiently prepare ensembles, the classical communication capacity of a unitary operation in a single direction is equal to the maximum by which the operation may increase the Holevo information of an ensemble. This result is important because it makes the evaluation of the communication capacity of an operation tractable.

It is interesting to speculate whether the same is true for bidirectional communication. In this case we would generalize to a bidirectional ensemble $\{p_i, q_j, |\Phi_{ij}\rangle\}$, where i is chosen by Alice and j is chosen by Bob. As discussed in Refs. [17,18], the same is true in the bidirectional case if it is possible to create bidirectional ensembles using as much communication as can be performed using these ensembles. This problem is a topic for future research.

The authors thank Stephen Bartlett for suggestions on the manuscript. This project has been supported by the Australian Research Council.

Note added.—Bennett *et al.* [16] refer to a private communication from P.W. Shor claiming a proof similar to that shown here for optimal remote state preparation.

[1] C. E. Shannon, Bell Syst. Tech. J. **27**, 379 (1948).

[2] A. S. Kholevo, Probl. Peredachi Inf. **9**, No. 3, 3 (1973) [Probl. Inf. Transm. (Engl. Transl.) **9**, 177 (1973)].

[3] B. Schumacher and M. D. Westmoreland, Phys. Rev. A **56**, 131 (1997).

[4] A. S. Holevo, IEEE Trans. Inf. Theory **44**, 269 (1998).

[5] C. H. Bennett, H. J. Bernstein, S. Popescu, and B. Schumacher, Phys. Rev. A **53**, 2046 (1996).

[6] H.-K. Lo, Phys. Rev. A **62**, 012313 (2000).

[7] A. K. Pati, Phys. Rev. A **63**, 014302 (2001).

[8] C. H. Bennett, D. P. DiVincenzo, P.W. Shor, J. A. Smolin, B. M. Terhal, and W. K. Wootters, Phys. Rev. Lett. **87**, 077902 (2001).

[9] I. Devetak and T. Berger, Phys. Rev. Lett. **87**, 197901 (2001).

[10] D.W. Leung and P.W. Shor, quant-ph/0201008.

[11] W. Dür, G. Vidal, and J. I. Cirac, Phys. Rev. A **64**, 022308 (2001).

[12] R. Jozsa, J. Mod. Opt. **41**, 2315 (1994).

[13] B. Schumacher, Phys. Rev. A **51**, 2738 (1995).

[14] H.-K. Lo, Opt. Commun. **119**, 552 (1995).

[15] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, Cambridge, 2000).

[16] C. H. Bennett, A. Harrow, D.W. Leung, and J. A. Smolin, quant-ph/0205057.

[17] D.W. Berry and B. C. Sanders, quant-ph/0205181.

[18] D.W. Berry and B. C. Sanders, quant-ph/0207065.