Parity bit in quantum cryptography

Charles H. Bennett,[1] Tal Mor,[2] and John A. Smolin[3]
[1]*IBM Research Division, Yorktown Heights, New York 10598*
[2]*Physics Department, Technion, Israel*
[3]*Physics Department, University of California at Los Angeles, Los Angeles, California 90024*

(Received 12 March 1996)

An n-bit string is encoded as a sequence of nonorthogonal quantum states. The parity bit of that n-bit string is described by one of two density matrices, $\rho_0^{(n)}$ and $\rho_1^{(n)}$, both in a Hilbert space of dimension 2^n. In order to derive the parity bit the receiver must distinguish between the two density matrices, e.g., in terms of optimal mutual information. In this paper we find the measurement which provides the optimal mutual information about the parity bit and calculate that information. We prove that this information decreases exponentially with the length of the string in the case where the single bit states are almost fully overlapping. We believe this result will be useful in proving the ultimate security of quantum cryptography in the presence of noise. [S1050-2947(96)10809-X]

PACS number(s): 03.65.Bz, 89.70.+c

I. INTRODUCTION

A major question in quantum information theory [1–6] is ''how well can two quantum states, or more generally, two density matrices ρ_0 and ρ_1, be distinguished?'' In terms of a communication scheme this question is translated to an identification task: A sender (Alice) sends a bit $b=i$ $(i=0;1)$ to the receiver (Bob) by sending the quantum state ρ_i, and the receiver does his best to identify the value of that bit, i.e., the quantum state. The two-dimensional Hilbert space $\mathcal{H}_2$ is usually used to implement such a binary channel, so the transmitted signals can be polarization states of photons, spin-states of spin-half particles, etc. The transmitted states may be pure states or density matrices, and need not be orthogonal. Usually, the mutual information I is used to describe distinguishability, such that $I=0$ means indistinguishable, and $I=1$ (for a binary channel) means perfect distinguishability. The ensemble of signals is agreed on in advance, and the main aim of Alice and Bob is to optimize the average mutual information over the different possible measurements at the receiving end. For a (simple) example, two orthogonal pure states transmitted through an error-free channel are perfectly distinguishable; the optimal mutual information $(I=1)$ is obtained if Bob measures in an appropriate basis. Finding the optimal mutual information is still an open question for most ensembles. Some cases with known analytic solutions are the case of two pure states and the case of two density matrices in two dimensions with equal determinants [5,6]. There are no known analytic solutions for two nontrivial density matrices in dimensions higher than 2. In this paper we find a solvable case which has very important implications to quantum cryptography.

Suppose that a source produces binary string x of length n with equal and independent probabilities for all the digits. Let the string be encoded into a quantum-mechanical channel, in which the digits ''0'' and ''1'' are represented by quantum states (density matrices) ρ_0 and ρ_1 of independent two-state quantum systems. These can be either pure states or density matrices with equal determinants. Suppose Bob wants to learn the parity bit (exclusive-OR) of the n-bit string and not the specific value of each bit. The parity bit is described by one of two density matrices $\rho_0^{(n)}$ and $\rho_1^{(n)}$ which lie in a 2^n-dimensional Hilbert space $\mathcal{H}_{2^n}$. These parity density matrices $\rho_p^{(n)}$ are the average density matrices, where the average is taken over all strings the source might produce, which have the same parity p. Since the parity bit of the source string x is encoded by $\rho_p^{(n)}$, information about which of the two density matrices was prepared is information about the parity of x.

Let x be any classical string of n such bits, and $\rho_x=\rho_{(1\text{st bit})}\cdots\rho_{(n\text{th bit})}$ be the density matrix made up of the tensor product of the signaling states $\rho_{(i)}$ corresponding to the ith bit of x. Formally, we distinguish between the two density matrices:

$$\rho_0^{(n)}=\frac{1}{2^{n-1}}\sum_{x|p(x)=0}\rho_x \quad \text{and} \quad \rho_1^{(n)}=\frac{1}{2^{n-1}}\sum_{x|p(x)=1}\rho_x, \quad (1)$$

where the sum is over all possible strings with the same parity [each sent with equal probability $(1/2^n)$] and $p(x)$ is the parity function of x. We show a simple way to write the parity density matrices. We find that they are optimally distinguished by a nonfactorizable joint measurement, performed on the composite 2^n-dimensional quantum system, and we calculate the optimal mutual information which can be obtained on the parity bit.

Parity bits are often used in quantum cryptography [7–9], where they play a crucial role in error-correction and privacy amplification [10–12]; for example, the final secret key might be the parity bit of a long string. The question of security of quantum cryptography is yet open, and our results may have several implications for attacking this issue. In particular, we concentrate on the special case where the two signaling states have large overlap, which is important in the analysis of the security of quantum key distribution against powerful multiparticle eavesdropping attacks. We show that the optimal obtainable information decreases exponentially with the length n of the string. This result provides a clue that classical privacy amplification is effective against joint measurements, limiting the ability of an eavesdropper to obtain significant information on the final key.

The first sections deal only with the case where each bit is encoded by a pure state. In Sec. II we find a simple way to write the density matrices of the parity bit for any n when the signaling states are pure; we show that the parity matrices can be put in a block diagonal form and explain the importance of that fact. In Sec. III we investigate the distinguishability of the parity matrices; the optimal measurement which distinguishes them is found to be a standard (von Neumann) measurement in an entangled basis (which is a generalization of the Bell basis of two particles); we calculate exactly the optimal mutual information which is derived on the parity bit by performing that optimal measurement. In Sec. IV we obtain our main result: for two almost fully overlapping states, the optimal mutual information I_M decreases exponentially with the length of the string. While exponentially small, this optimal information is nevertheless considerably greater than the information that would have been obtained by measuring each bit separately and classically combining the results of these measurements, thus, we prove the advantage of such ''joint'' measurements. Going back to the parity matrices obtained in Sec. II we are also able to calculate the maximal deterministic (conclusive) information; this is done in Sec. V, where we also confirm a result previously obtained by Huttner and Peres [13] for two bits. In Sec. VI we repeat the calculation of the optimal mutual information for the more general case where the bits are represented by nonpure states (in $\mathcal{H}_2$ and with equal determinants). In Sec. VII we briefly discuss the implications of our results to the security of quantum cryptography.

II. DENSITY MATRICES FOR PARITY BITS

Let Alice send n bits. The possible values of a single bit $(0$ or $1)$ are represented by

$$\psi_0=\begin{pmatrix}\cos\alpha\\ \sin\alpha\end{pmatrix} \quad\text{and}\quad \psi_1=\begin{pmatrix}\cos\alpha\\ -\sin\alpha\end{pmatrix}, \tag{2}$$

respectively. In terms of density matrices these are

$$\rho_0^{(1)}=\begin{pmatrix}c^2 & sc\\ sc & s^2\end{pmatrix} \quad\text{and}\quad \rho_1^{(1)}=\begin{pmatrix}c^2 & -sc\\ -sc & s^2\end{pmatrix}, \tag{3}$$

where we use a shorter notation $s\equiv\sin\alpha$; $c\equiv\cos\alpha$, for convenience, and the superscript $[\,]^{(1)}$ is explained in the following paragraph.

The parity bit of an n-bit string is the exclusive-OR of all the bits in the string. In other words, the parity is 1 if there are an odd number of 1's and 0 if there are an even number. The parity density matrices of n bits will be denoted as $\rho_0^{(n)}$ and $\rho_1^{(n)}$ in case the parity is ''0'' and ''1,'' respectively. Using these density matrices we define also the *total* density matrix $\rho^{(n)}\equiv\frac{1}{2}(\rho_0^{(n)}+\rho_1^{(n)})$ and the *difference* density matrix $\Delta^{(n)}\equiv\frac{1}{2}(\rho_0^{(n)}-\rho_1^{(n)})$, so that

$$\rho_0^{(n)}=\rho^{(n)}+\Delta^{(n)} \quad\text{and}\quad \rho_1^{(n)}=\rho^{(n)}-\Delta^{(n)}. \tag{4}$$

The one-particle density matrices [Eq. (3)] also describe the parities of one particle, and therefore we can calculate

$$\rho^{(1)}=\frac{1}{2}(\rho_0^{(1)}+\rho_1^{(1)})=\begin{pmatrix}c^2 & 0\\ 0 & s^2\end{pmatrix}, \tag{5}$$

$$\Delta^{(1)}=\frac{1}{2}(\rho_0^{(1)}-\rho_1^{(1)})=\begin{pmatrix}0 & sc\\ sc & 0\end{pmatrix}. \tag{6}$$

The density matrices of the parity bit of two particles are

$$\rho_0^{(2)}=\frac{1}{2}(\rho_0^{(1)}\rho_0^{(1)}+\rho_1^{(1)}\rho_1^{(1)}),$$

$$\rho_1^{(2)}=\frac{1}{2}(\rho_0^{(1)}\rho_1^{(1)}+\rho_1^{(1)}\rho_0^{(1)}), \tag{7}$$

where the multiplication is a tensor product. The total density matrix is

$$\begin{aligned}\rho^{(2)}&=\frac{1}{2}(\rho_0^{(2)}+\rho_1^{(2)})\\ &=\frac{1}{4}[\rho_0^{(1)}(\rho_0^{(1)}+\rho_1^{(1)})+\rho_1^{(1)}(\rho_1^{(1)}+\rho_0^{(1)})]=\rho^{(1)}\rho^{(1)},\end{aligned}$$

which, by using the basis

$$|b_0\rangle\equiv\begin{pmatrix}1\\0\end{pmatrix}_1\begin{pmatrix}1\\0\end{pmatrix}_2=\begin{pmatrix}1\\0\\0\\0\end{pmatrix}, \quad |b_1\rangle\equiv\begin{pmatrix}1\\0\end{pmatrix}_1\begin{pmatrix}0\\1\end{pmatrix}_2=\begin{pmatrix}0\\1\\0\\0\end{pmatrix},$$

$$|b_2\rangle\equiv\begin{pmatrix}0\\1\end{pmatrix}_1\begin{pmatrix}1\\0\end{pmatrix}_2=\begin{pmatrix}0\\0\\1\\0\end{pmatrix} \quad\text{and}\quad |b_3\rangle\equiv\begin{pmatrix}0\\1\end{pmatrix}_1\begin{pmatrix}0\\1\end{pmatrix}_2=\begin{pmatrix}0\\0\\0\\1\end{pmatrix} \tag{8}$$

in $\mathcal{H}_4$, can be written as

$$\rho^{(2)}=\rho^{(1)}\rho^{(1)}=\begin{pmatrix}c^4 & 0 & 0 & 0\\ 0 & c^2s^2 & 0 & 0\\ 0 & 0 & c^2s^2 & 0\\ 0 & 0 & 0 & s^4\end{pmatrix}. \tag{9}$$

The difference density matrix is

$$\begin{aligned}\Delta^{(2)}&=\frac{1}{2}(\rho_0^{(2)}-\rho_1^{(2)})\\ &=\frac{1}{4}[\rho_0^{(1)}(\rho_0^{(1)}-\rho_1^{(1)})+\rho_1^{(1)}(\rho_1^{(1)}-\rho_0^{(1)})]\\ &=\Delta^{(1)}\Delta^{(1)}=\begin{pmatrix}0 & 0 & 0 & c^2s^2\\ 0 & 0 & c^2s^2 & 0\\ 0 & c^2s^2 & 0 & 0\\ c^2s^2 & 0 & 0 & 0\end{pmatrix}.\end{aligned} \tag{10}$$

The density matrices of the parity bit of n particles can be written recursively:

$$\rho_0^{(n)}=\frac{1}{2}(\rho_0^{(1)}\rho_0^{(n-1)}+\rho_1^{(1)}\rho_1^{(n-1)}),$$

$$\rho_1^{(n)}=\frac{1}{2}(\rho_0^{(1)}\rho_1^{(n-1)}+\rho_1^{(1)}\rho_0^{(n-1)}), \tag{11}$$

leading to

$$\rho^{(n)}=\frac{1}{2}(\rho_0^{(n)}+\rho_1^{(n)})=\rho^{(1)}\rho^{(n-1)} \tag{12}$$

and

$$\Delta^{(n)}=\frac{1}{2}(\rho_0^{(n)}-\rho_1^{(n)})=\Delta^{(1)}\Delta^{(n-1)}. \tag{13}$$

Using these expressions recursively we get

$$\rho^{(n)}=(\rho^{(1)})^n, \tag{14}$$

which is diagonal, and

$$\Delta^{(n)}=(\Delta^{(1)})^n, \tag{15}$$

which has nonzero terms only in the secondary diagonal. The density matrices $\rho_0^{(n)}$ and $\rho_1^{(n)}$ are now immediately derived for any n using Eq. (4):

$$\rho_0^{(n)}=(\rho^{(1)})^n+(\Delta^{(1)})^n$$

and

$$\rho_1^{(n)}=(\rho^{(1)})^n-(\Delta^{(1)})^n. \tag{16}$$

As an illustrative example we write ρ_0 and ρ_1 for two particles:

$$\rho_0^{(2)}=\begin{pmatrix} c^4 & 0 & 0 & c^2s^2 \\ 0 & c^2s^2 & c^2s^2 & 0 \\ 0 & c^2s^2 & c^2s^2 & 0 \\ c^2s^2 & 0 & 0 & s^4 \end{pmatrix},$$

$$\rho_1^{(2)}=\begin{pmatrix} c^4 & 0 & 0 & -c^2s^2 \\ 0 & c^2s^2 & -c^2s^2 & 0 \\ 0 & -c^2s^2 & c^2s^2 & 0 \\ -c^2s^2 & 0 & 0 & s^4 \end{pmatrix}. \tag{17}$$

The only nonzero terms in the parity density matrices are the terms in the diagonals for any n, thus the parity density matrices have an X shape in that basis.

The basis vectors can be permuted to yield block-diagonal matrices built of 2×2 blocks. The original basis vectors [see, for example, Eq. (8)], $|b_i\rangle$, are simply 2^n vectors where the ith element of the ith basis vector is 1 and all other elements are 0 (i ranges from 0 to 2^n-1). The new basis vectors are related to the old as follows:

$$|b_i'\rangle=|b_{i/2}\rangle \quad \text{for even } i$$

and

$$|b_i'\rangle=|b_{2^n-(i+1)/2}\rangle \quad \text{for odd } i. \tag{18}$$

The parity density matrices are now, in the new basis (we omit the $'$ from now on as we will never write the matrices in the original basis),

$$\rho_p^{(n)}=\begin{pmatrix} B_p^{[j=1]} & 0 & \dots & 0 \\ 0 & B_p^{[j=2]} & \dots & 0 \\ 0 & 0 & \dots & B_p^{[j=2^{(n-1)}]} \end{pmatrix}, \tag{19}$$

where the subscript p stands for the parity (0 or 1). Each of the 2×2 matrices has the form

$$B_p^{[j]}=\begin{pmatrix} c^{2(n-k)}s^{2k} & \pm c^ns^n \\ \pm c^ns^n & c^{2k}s^{2(n-k)} \end{pmatrix}, \tag{20}$$

with the plus sign for $p=0$ and the minus sign for $p=1$, and $0\leqslant k\leqslant n$, and all these density matrices satisfy $\det B_p^{[j]}=0$. The first block ($j=1$) has $k=0$; there are $\binom{n}{1}$ blocks which have $k=1$ or $k=n-1$; there are $\binom{n}{2}$ j's which have $k=2$ or $k=n-2$, etc. This continues until $k=(n-1)/2$ for odd n. For even n the process continues up to $k=n/2$ with the minor adjustment that there are only $\frac{1}{2}\binom{n}{n/2}$ j's of $k=n/2$. This enumeration groups blocks which are identical or identical after interchange of k and $n-k$ and accounts for all $2^n/2$ blocks. We will see later that blocks identical under interchange of k and $n-k$ will contribute the same mutual information about the parity bit, thus we have grouped them together.

With the density matrices written in such a block-diagonal form of 2×2 blocks the problem of finding the optimal mutual information can be analytically solved. It separates into two parts: (i) Determining in which of $2^n/2$ orthogonal $2d$ subspaces (each corresponding to one of the 2×2 blocks) the system lies; (ii) Performing the optimal measurement within that subspace. The subspaces may be thought of as $2^n/2$ parallel channels, one of which is probabilistically chosen and used to encode the parity by means of a choice between two equiprobable pure states within that subspace (these two states are pure because the B_0 and B_1 matrices each have zero determinant). We shall present in the next section the optimal measurement that yields the optimal mutual information transmissible through such a two-pure-state quantum channel. The channel then corresponds to a classical binary symmetric channel (BSC), i.e., a classical one-bit-in one-bit-out channel whose output differs from its input with some error probability p_j independent of whether the input was 0 or 1. The optimal mutual information in each subchannel is the optimal mutual information of a BSC with error probability p_j and is $I_2(p_j)=1-H(p_j)$, with $H(x)=-x\log_2 x-(1-x)\log_2(1-x)$, the Shannon entropy function. The optimal mutual information I_M for distinguishing $\rho_1^{(n)}$ from $\rho_1^{(n)}$ can thus be expressed as an average over the optimal mutual information of the subchannels:

$$I_M=\sum_{j=1}^{2^n/2} q_j I_2(p_j), \tag{21}$$

where $q_j=\mathrm{Tr}B_0^{[j]}=\mathrm{Tr}B_1^{[j]}$ is the probability of choosing the jth subchannel. The BSC error probability p_j for the jth subchannel depends on the subchannel's 2×2 renormalized density matrices $\hat{B}_p^{[j]}=B_p^{[j]}/q_j$, and is easily calculated once the optimal measurement is found. For each subchannel the q_j and renormalized 2×2 matrices look like

$$q_j=c^{2(n-k)}s^{2k}+c^{2k}s^{2(n-k)} \tag{22}$$

and

$$\hat{B}_p^{[j]}=$$

$$\begin{pmatrix} \frac{c^{2(n-k)}s^{2k}}{c^{2(n-k)}s^{2k}+c^{2k}s^{2(n-k)}} & \frac{\pm c^n s^n}{c^{2(n-k)}s^{2k}+c^{2k}s^{2(n-k)})} \\ \frac{\pm c^n s^n}{c^{2(n-k)}s^{2k}+c^{2k}s^{2(n-k)})} & \frac{c^{2k}s^{2(n-k)}}{c^{2(n-k)}s^{2k}+c^{2k}s^{2(n-k)}} \end{pmatrix}. \qquad 23$$

In our previous example of $n=2$ the matrices are put in a block diagonal form:

$$\rho_0^{(2)}=\begin{pmatrix} c^4 & c^2s^2 & 0 & 0 \\ c^2s^2 & s^4 & 0 & 0 \\ 0 & 0 & c^2s^2 & c^2s^2 \\ 0 & 0 & c^2s^2 & c^2s^2 \end{pmatrix},$$

$$\rho_1^{(2)}=\begin{pmatrix} c^4 & -c^2s^2 & 0 & 0 \\ -c^2s^2 & s^4 & 0 & 0 \\ 0 & 0 & c^2s^2 & -c^2s^2 \\ 0 & 0 & -c^2s^2 & c^2s^2 \end{pmatrix}, \tag{24}$$

so that $q_{j=1}=c^4+s^4$, $q_{j=2}=2c^2s^2$, and

$$\hat{B}_p^{[j=1]}=\begin{pmatrix} \frac{c^4}{c^4+s^4} & \pm\frac{c^2s^2}{c^4+s^4} \\ \pm\frac{c^2s^2}{c^4+s^4} & \frac{s^4}{c^4+s^4} \end{pmatrix},$$

$$\hat{B}_p^{[j=2]}=\begin{pmatrix} 1/2 & \pm 1/2 \\ \pm 1/2 & 1/2 \end{pmatrix}. \tag{25}$$

III. OPTIMAL INFORMATION IN A PARITY BIT

Two pure states or two density matrices in $\mathcal{H}_2$ with equal determinants can always be written (in an appropriate basis) in the simple form

$$\rho_0=\begin{pmatrix} a_1 & a_2 \\ a_2 & a_3 \end{pmatrix}, \quad \rho_1=\begin{pmatrix} a_1 & -a_2 \\ -a_2 & a_3 \end{pmatrix} \tag{26}$$

with a_i real positive numbers such that $\mathrm{Tr}\rho_p=a_1+a_3=1$. For the two pure states of Eq. (2), say, for the polarization states of a photon, it is easy to see (and can be formally proven [5,6]) that a standard measurement in an orthogonal basis symmetric to the two states optimizes the mutual information (and also minimizes the average error probability). The angle between one basis vector and the polarization state is $\pi/4\pm\alpha$. The measurement results in an error with probability

$$P_e=\sin^2\left(\frac{\pi}{4}-\alpha\right)=\frac{1-\cos\left(\frac{\pi}{2}-2\alpha\right)}{2}=\frac{1-\sin(2\alpha)}{2}, \tag{27}$$

and with the same error probability for both inputs, thus leading to a binary symmetric channel (BSC). The optimal information of such a channel is well known and is

$$I_{\mathrm{BSC}}=I_2(P_e). \tag{28}$$

Note that the overlap of the two-states is $\cos(2\alpha)$, thus, for two pure states in any dimension, the optimal information $I_2[1-\sin(2\alpha)/2]$ is a simple function of the overlap. The density matrices of such pure states [Eq. (3)] can be written as $\rho_i=(\mathbb{1}+\sigma\cdot\mathbf{r_i})/2$ with the σ being the Pauli matrices and $\mathbf{r}=(\pm\sin 2\alpha,0,\cos 2\alpha)$ being a three-dimensional vector which describes a spin direction. Using this notation any density matrix is described by a point in a three-dimensional unit ball, called the Bloch sphere. The pure states are points on the surface of that sphere (also called the Poincaré sphere). With the density matrix notation the optimal basis for distinguishing the states is the x basis (note that the angle between the basis vector and the state is doubled in this notation). The measurement of the two projectors

$$A_{\rightarrow}=1/2\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \quad \text{and} \quad A_{\leftarrow}=1/2\begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} \tag{29}$$

yields

$$P_e=\mathrm{Tr}\rho_1 A_{\rightarrow}=\frac{1}{2}-a_2, \tag{30}$$

which recovers the result of Eq. (27) in case of pure states of Eq. (3). However, the treatment of density matrices is more general and this is the optimal measurement also in the case of nonpure states with equal determinants [5,6], when ρ_i of Eq. (3) are replaced by ρ_i^{dm} of Eqs. (69) and (70) of Sec. VI, and this case is also described by a BSC. The only difference between the matrices is that $\det\rho_p=0$ for pure states and $0\leqslant\det\rho_p\leqslant\frac{1}{4}$ for density matrices.

Instead of measuring the density matrices in the x direction we perform the following unitary transformation on the density matrices:

$$U=1/\sqrt{2}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \tag{31}$$

to obtain $\rho'=U\rho U^{\dagger}$ which is then measured in the z basis. Note that the transformation transforms the original z-basis to x-basis (the motivation for this approach will be understood when we discuss the 2×2 blocks of the parity matrices). The new density matrices are

$$\rho_0' = \begin{pmatrix} \frac{1}{2}+a_2 & \frac{a_1-a_3}{2} \\ \frac{a_1-a_3}{2} & \frac{1}{2}-a_2 \end{pmatrix}, \quad \rho_1' = \begin{pmatrix} \frac{1}{2}-a_2 & \frac{a_1-a_3}{2} \\ \frac{a_1-a_3}{2} & \frac{1}{2}+a_2 \end{pmatrix} \tag{32}$$

and their measurement yields the probability $\frac{1}{2}\pm a_2$ to derive the correct (plus) and the wrong (minus) answers (as we obtained before), leading to optimal mutual information of

$$I_2\left(\frac{1}{2}-a_2\right), \tag{33}$$

which depends only on a_2. Note that the same information is obtained in case a_1 and a_3 are interchanged.

The naive way to derive information on a parity bit is to derive the optimal information on each particle separately and calculate the information on the parity bit. We call this *individual* or *single-particle* measurement. It is the best Bob can do in case he has no quantum memory in which to keep the particles (which usually arrive one at a time) or he has no ability to perform more advanced joint measurements. The optimal error-probability for each particle is $r \equiv P_e^{(1)} = (1-\sin 2\alpha)/2$. The probability of deriving the wrong parity bit is equal to the probability of having an odd number of errors on the individual particles

$$P_e^{(n)} = \sum_{j=\mathrm{odd}}^{n} \binom{n}{j} r^j (1-r)^{n-j}.$$

To perform the sum over only odd j we use the formulas

$$(p+q)^n = \sum_{j=0}^{n} \binom{n}{j} p^{n-j} q^j,$$

$$(p-q)^n = \sum_{j=0}^{n} \binom{n}{j} p^{n-j} (-q)^j,$$

to derive

$$\sum_{j=\mathrm{odd}}^{n} \binom{n}{j} p^{n-j} q^j = \frac{(p+q)^n-(p-q)^n}{2}. \tag{34}$$

Assigning $q=r$ and $p=1-r$ we get

$$P_e^{(n)} = \sum_{j=\mathrm{odd}}^{n} \binom{n}{j} r^j (1-r)^{n-j} = \frac{1^n-(1-2r)^n}{2}$$
$$= \frac{1}{2} - \frac{(1-2r)^n}{2}. \tag{35}$$

The mutual information I_S in this single-particle measurement is

$$I_S = I_2(P_e^{(n)}) = I_2\left(\frac{1}{2} - \frac{(\sin 2\alpha)^n}{2}\right) \tag{36}$$

using $r=[1-\sin(2\alpha)]/2$.

A lot of useless side-information is also obtained (e.g., on the individual bits). This fact indicates that Bob might be able to do much better by concentrating on deriving only useful information. The optimal measurement for finding mutual information on the parity bit is not a single-particle measurement, but is instead a measurement on the full 2^n-dimensional Hilbert space of the system. In general, optimizing over all possible measurement is a very difficult task unless the two density matrices in $\mathcal{H}_{2^n}$ are pure states. However, in the preceding section we have shown how to reduce the problem to that of distinguishing the 2×2 blocks of our block-diagonal parity matrices. We now have only to apply the optimal single-particle measurement to the 2×2 $\hat{B}^{[j]}$'s of Eq. (23) and use the result in Eq. (21).

The error probability [Eq. (30)] for distinguishing the $\hat{B}^{[j]}$'s is seen to be

$$p_j = \frac{1}{2} - \frac{c^n s^n}{c^{2(n-k)} s^{2k} + c^{2k} s^{2(n-k)}}, \tag{37}$$

from which the information $I_2(p_j)$ in each channel is obtained.

Plugging the error probability p_j [Eq. (37)] and the probability of choosing the jth subchannel q_j [Eq. (22)] into Eq. (21), the optimal information on the parity bit is now

$$I_M = \sum_{j=1}^{2^n/2} (c^{2(n-k)} s^{2k} + c^{2k} s^{2(n-k)})$$
$$\times I_2\left(\frac{1}{2} - \frac{c^n s^n}{c^{2(n-k)} s^{2k} + c^{2k} s^{2(n-k)}}\right). \tag{38}$$

In the simple case of orthogonal states ($\alpha=\pi/4$) all these density matrices are the same and we get $q_j = (\frac{1}{2})^{n-1}$, $p_j=0$, and $I_M=1$ as expected.

A brief remark is in order at this stage. The transformation to the x basis for each 2×2 matrix, $\hat{B}_p^{(n,k)}$, is actually a transformation from a product basis to a fully entangled basis of the n particles. That basis is a generalization of the Bell basis of [15],

$$\begin{pmatrix}1\\0\end{pmatrix}_1 \begin{pmatrix}1\\0\end{pmatrix}_2 \cdots \begin{pmatrix}1\\0\end{pmatrix}_{n-1} \begin{pmatrix}1\\0\end{pmatrix}_n \pm \begin{pmatrix}0\\1\end{pmatrix}_1 \begin{pmatrix}0\\1\end{pmatrix}_2 \cdots \begin{pmatrix}0\\1\end{pmatrix}_{n-1} \begin{pmatrix}0\\1\end{pmatrix}_n, \tag{39}$$

$$\begin{pmatrix}1\\0\end{pmatrix}_1 \begin{pmatrix}1\\0\end{pmatrix}_2 \cdots \begin{pmatrix}1\\0\end{pmatrix}_{n-1} \begin{pmatrix}0\\1\end{pmatrix}_n \pm \begin{pmatrix}0\\1\end{pmatrix}_1 \begin{pmatrix}0\\1\end{pmatrix}_2 \cdots \begin{pmatrix}0\\1\end{pmatrix}_{n-1} \begin{pmatrix}1\\0\end{pmatrix}_n, \tag{40}$$

etc. The Bell basis for two particles is frequently used and its basis contains the EPR singlet state and three other orthogonal fully entangled states.

For large n, the number of blocks is exponentially large and performing the summation required in Eq. (38) is impractical, since all the 2^{n-1} matrices must be taken into account. However, that problem can be simplified by realizing that all blocks with a given k, as well as all blocks with k and $n-k$ interchanged, contribute the same information to the total. This is easily seen in Eq. (38), where both the weight and the argument of I_2 are symmetric in k and $n-k$. The optimal mutual information for even n is then

$$I_M^{\text{even}}=\sum_{k=0}^{n/2-1}\binom{n}{k}q_kI_2(p_k)+\frac{1}{2}\binom{n}{\frac{n}{2}}q_{n/2}I_2(p_{n/2}),\qquad(41)$$

and for odd n

$$I_M^{\text{odd}}=\sum_{k=0}^{(n-1)/2}\binom{n}{k}q_kI_2(p_k).\qquad(42)$$

As an example we calculate I_M for $n=2$ (of course, the counting argument is not needed in that case). This particular result complements the result in [13] where the deterministic information of such a system is considered (see also Sec. V). In the new basis (32) the density matrices of [Eq. (25)] become

$$\hat{B}_0'^{(n=2,k=0)}=\begin{pmatrix}1/2+\dfrac{c^2s^2}{c^4+s^4} & \dfrac{1}{2}\dfrac{c^4-s^4}{c^4+s^4}\\ \dfrac{1}{2}\dfrac{c^4-s^4}{c^4+s^4} & 1/2-\dfrac{c^2s^2}{c^4+s^4}\end{pmatrix},$$

$$\hat{B}_1'^{(n=2,k=0)}=\begin{pmatrix}1/2-\dfrac{c^2s^2}{c^4+s^4} & \dfrac{1}{2}\dfrac{c^4-s^4}{c^4+s^4}\\ \dfrac{1}{2}\dfrac{c^4-s^4}{c^4+s^4} & 1/2+\dfrac{c^2s^2}{c^4+s^4}\end{pmatrix},$$

and

$$\hat{B}_0'^{(n=2,k=1)}=\begin{pmatrix}1&0\\0&0\end{pmatrix},\quad \hat{B}_1'^{(n=2,k=1)}=\begin{pmatrix}0&0\\0&1\end{pmatrix}.\qquad(43)$$

We use the notation $S=2sc=\sin 2\alpha$; $C=c^2-s^2=\cos 2\alpha$ [hence, $c^4-s^4=C$ and $c^4+s^4=(1+C^2)/2$] to obtain $q_1=2c^2s^2=S/2$, $p_1=0$, $q_0=\frac{1}{2}(1+C^2)$, and $p_0=C^2/(1+C^2)$ (the q_j's were obtained in the preceding section). The mutual information of the parity of two bits is obtained using Eq. (21),

$$\begin{aligned}I_M&=q_0I_2(p_0)+q_1I_2(p_1)\\&=\frac{1}{2}(1+C^2)I_2\left(\frac{C^2}{1+C^2}\right)+\frac{S^2}{2}.\end{aligned}\qquad(44)$$

IV. INFORMATION ON THE PARITY BIT OF ALMOST FULLY OVERLAPPING STATES

The case of almost fully overlapping states is extremely important to the analysis of eavesdropping attacks on any quantum key distribution scheme, as will be discussed in Sec. VII. In this case the angle α is small so $s\equiv\sin\alpha\simeq\alpha$ and $c\equiv\cos\simeq 1-\alpha^2/2$. To observe the advantage of the joint measurement, let us first calculate the optimal information obtained by individual measurements. In that case, Eqs. (35) and (27) yield

$$P_e^{(n)}=\frac{1}{2}-\frac{(2\alpha)^n}{2}.\qquad(45)$$

For small η the logarithmic function is approximated by

$$\log\left(\frac{1}{2}\pm\eta\right)=\frac{\ln\left(\frac{1}{2}\pm\eta\right)}{\ln 2}\approx -1\pm\frac{2}{\ln 2}\eta-\frac{2}{\ln 2}\eta^2,\qquad(46)$$

from which the mutual information

$$\begin{aligned}I_2\left(\frac{1}{2}-\eta\right)&=1-H\left(\frac{1}{2}-\eta\right)\\&=1+\left(\frac{1}{2}+\eta\right)\log\left(\frac{1}{2}+\eta\right)+\left(\frac{1}{2}-\eta\right)\log\left(\frac{1}{2}-\eta\right)\\&\approx\frac{2}{\ln 2}\eta^2\end{aligned}\qquad(47)$$

is obtained. Using this result and assigning $\eta=(2\alpha)^n/2$, the information (to first order) obtained by the optimal single-particle measurement is

$$I_S=\frac{2}{\ln 2}\frac{(2\alpha)^{2n}}{4}=\frac{(2\alpha)^{2n}}{2\ln 2}.\qquad(48)$$

We use the same approximations and Eqs. (37) and (22) to calculate the leading terms in the optimal mutual information (41) and (42). For $k=n/2$ (n even) we get $p_k=0$ (regardless of the small angle) and

$$I_2(p_{n/2})=1.\qquad(49)$$

For $k<n/2$ we get $p_k\approx\frac{1}{2}-s^n/s^{2k}\approx\frac{1}{2}-\alpha^{n-2k}$, which yields [using Eq. (47) with $\eta=\alpha^{n-2k}$]

$$I_2(p_k)\approx\frac{2}{\ln 2}\alpha^{2n-4k}.\qquad(50)$$

The coefficient $q_k=\alpha^{2k}$ for $k<n/2$ and $q_k=2\alpha^{2k}$ for $k=n/2$, so that

$$q_kI_2(p_k)\approx\frac{2}{\ln 2}\alpha^{2(n-k)}\qquad(51)$$

for $k<n/2$, and

$$q_kI_2(p_k)\approx 2\alpha^n\qquad(52)$$

for $k=n/2$. The dominant terms are those with the largest k, that is, k closest to $n/2$. The next terms are smaller by two orders in α. The number of density matrices with these k's are also the largest (up to a factor of 2 in case of even n). Therefore, the terms $k=n/2$ for even n and $k=(n-1)/2$ for odd n are the dominant terms in the final expression. Thus, for almost fully overlapping states, the mutual information is

$$I_M^{\text{even}}\approx\frac{1}{2}\binom{n}{\frac{n}{2}}2\alpha^n=\binom{n}{\frac{n}{2}}\alpha^n$$

for even n, and

$$I_M^{\text{odd}}\approx\binom{n}{\frac{n-1}{2}}\frac{2}{\ln 2}\alpha^{n+1}\qquad(53)$$

for odd n.

These expressions can be further simplified. The number of density matrices of any type is bounded (for large n) using the Stirling formula (see [15] in the chapter on Reed-Solomon codes),

$$\binom{n}{k} < \frac{2^{nH(k/n)}}{\sqrt{2\pi(k/n)(1-k/n)n}}. \tag{54}$$

For k near $n/2$, $\eta \equiv \frac{1}{2} - k/n$ is small, and the standard approximation (47): $H \approx 1 - O(\eta^2) = 1 - O[(\frac{1}{2} - k/n)^2] < 1$ is used to derive $\binom{n}{k} < 2^n / \sqrt{2\pi(k/n)(1-k/n)n}$. Using also $k/n(1-k/n) \approx \frac{1}{4} - \eta^2$, we derive

$$\binom{n}{k} < \frac{2^n}{\sqrt{\frac{\pi}{2}n}}[1 + O(\eta^2)]. \tag{55}$$

Thus the leading term in I_M is

$$I_M < \frac{2^n}{\sqrt{\frac{\pi}{2}n}} \alpha^n = (2\alpha)^n \Big/ \sqrt{\frac{\pi}{2}n} \tag{56}$$

for even n and

$$I_M < \frac{2^n}{\sqrt{\frac{\pi}{2}n}} \frac{2}{\ln 2} \alpha^{n+1} = \frac{2}{\ln 2} \alpha (2\alpha)^n \Big/ \sqrt{\frac{\pi}{2}n}$$

$$< (2\alpha)^n \Big/ \sqrt{\frac{\pi}{2}n} \tag{57}$$

for odd n (using $\alpha < \ln 2/2$). We see that we could keep a better bound for odd n but for simplicity we consider the same bound for both even and odd n's.

We can now compare the optimal information I_M from a joint measurement on all n particles to the optimal information I_S from separate measurements [cf. Eq. (48)]:

$$I_M = O((2\alpha)^n / \sqrt{n}),$$
$$I_S = O((2\alpha)^{2n}). \tag{58}$$

Since α is a small number (corresponding to highly overlapping signal states), the joint measurement is superior to the individual measurement by a factor of $O((2\alpha)^n)$. However, it is only superior by a polynomial factor, since

$$I_M \approx (I_S)^2. \tag{59}$$

V. DETERMINISTIC INFORMATION ON THE PARITY BIT

For a single particle Bob can perform a different kind of individual measurement which is not optimal in terms of average mutual information but is sometimes very useful [9,4]. It yields either a conclusive result about the value of that bit or an inconclusive one, and Bob will know which of the types of information he has succeeded in obtaining. Such a measurement corresponds to a binary erasure channel [4,13,16]. With probability $p_?$ of an inconclusive result, the mutual information is $I_{p_?} = 1 - p_?$. The minimal probability for an inconclusive result is $\cos 2\alpha$ leading to $I_{p_?} = 1 - \cos 2\alpha$ [4]. This result is obtained by performing a generalized measurement (positive operator value measure [4,3,17]) on the system or a standard measurement performed on a larger system which contains the system and an auxiliary particle [4,18]. Note that this results in less mutual information than the optimal measurement for one-particle mutual information. If Bob uses this type of measurement on each particle separately his deterministic single-particle information about the parity bit is $(1 - \cos 2\alpha)^n$.

We now use the block-diagonal density matrices derived in Sec. II to derive also the optimal *deterministic* information on the parity bit. We note that each of the 2×2 blocks in the block-diagonal density matrices is the density matrix of a pure state, so we may replace the optimal measurement in each subchannel with the optimal deterministic measurement and proceed as before. The total optimal deterministic information is easily calculated by replacing $I_2(p_k)$ in Eqs. (41) and (42) by $I(p_{?_k}) = 1 - p_{?_k}$, where $p_{?_k}$ is the probability of an inclusive result in block k. To find the minimal $p_{?_k}$ we write each of the normalized density matrices $\hat{B}_p^{(n,k)}$ as pure states with some angle γ:

$$\begin{pmatrix} \cos\gamma \\ \sin\gamma \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} \cos\gamma \\ -\sin\gamma \end{pmatrix}. \tag{60}$$

Comparing with Eq. (23)

$$p_? = \cos(2\gamma) = \cos^2\gamma - \sin^2\gamma = \frac{c^{2(n-k)}s^{2k} - c^{2k}s^{2(n-k)}}{c^{2(n-k)}s^{2k} + c^{2k}s^{2(n-k)}}, \tag{61}$$

hence

$$I(p_{?_k}) = 1 - \frac{c^{2(n-k)}s^{2k} - c^{2k}s^{2(n-k)}}{c^{2(n-k)}s^{2k} + c^{2k}s^{2(n-k)}}. \tag{62}$$

The total information is

$$I_D^{\text{even}} = \sum_{k=0}^{n/2\ -1} \binom{n}{k} q_k I(p_{?_k}) + \frac{1}{2} \binom{n}{\frac{n}{2}} q_{n/2} I(p_{?_{n/2}}) \tag{63}$$

for even n, and

$$I_D^{\text{odd}} = \sum_{k=0}^{n-1/2} \binom{n}{k} q_k I(p_{?_k}) \tag{64}$$

for odd n.

For $n = 2$ we recover a result previously obtained by Huttner and Peres [14] by performing the optimal POVM on the first pair of density matrices of Eq. (25), and a measurement in the entangled basis (as before) on the second. The probability of an inconclusive result is $\cos^2\gamma - \sin^2\gamma = (c^4 - s^4)/(c^4 + s^4) = 2C/(1 + C^2)$, hence the optimal deterministic information is $1 - 2C/(1 + C^2)$, leading to the total deterministic information

$$I_D = q_1 I_D(p_?) + q_2 I_2(p_2) = \frac{1}{2}(1+C^2)\left(1 - 2\frac{C}{1+C^2}\right) + \frac{S^2}{2} = 1 - C, \tag{65}$$

which is exactly the result obtained by Huttner and Peres (note, however, that they used an angle which is $\pi/4-\alpha$ hence derived $1-S$ for the deterministic information).

For almost overlapping states (small α) the dominant terms are still the same as in the case of optimal information. The term q_k is as before and the information in each port is

$$I(p_{?_k}) = 1 \tag{66}$$

for $k=n/2$ and

$$I(p_{?_k}) = 1 - \frac{c^{2n-4k} - s^{2n-4k}}{c^{2n-4k} + s^{2n-4k}} = 1 - (1-\alpha^{2n-4k})^2 = 2\alpha^{2n-4k} \tag{67}$$

for $k<n/2$. Taking into consideration only the dominant term we get

$$I_D^{\text{even}} \approx \begin{pmatrix} n \\ \frac{n}{2} \end{pmatrix} \alpha^n$$

for even n, which is the same as the optimal information, and

$$I_D^{\text{odd}} \approx \begin{pmatrix} n \\ \frac{n-1}{2} \end{pmatrix} 2\alpha^{n+1} \tag{68}$$

for odd n, which is smaller than the optimal mutual information by a factor of 1/ln2.

VI. PARITY BIT FOR DENSITY MATRICES

The previous discussion assumed that $\rho_p^{(1)}$ are pure states. The generalization to the case of density matrices with equal determinants is straightforward. Let the bit ''0'' and the bit ''1'' be represented by

$$\rho_0^{dm} = \begin{pmatrix} c^2 & sc-r \\ sc-r & s^2 \end{pmatrix} \tag{69}$$

and

$$\rho_1^{dm} = \begin{pmatrix} c^2 & -(sc-r) \\ -(sc-r) & s^2 \end{pmatrix} \tag{70}$$

(with $s=\sin\alpha$, etc., and $r<sc$), which contains the most general density matrices of the desired type. On the Poincaré sphere these density matrices have the same z components as the previously written pure states but smaller x components (hence smaller angle α'). We could choose other ways of representing these density matrices, e.g., with similar x components and smaller z components. Such representations were more appropriate for comparison with pure states (since they yield the same mutual information for a single particle) but less convenient for showing that the previous result is easily generalized.

Clearly

$$\rho^{(1)_{dm}} = \frac{1}{2}(\rho_0^{(1)} + \rho_1^{(1)}) = \begin{pmatrix} c^2 & 0 \\ 0 & s^2 \end{pmatrix}, \tag{71}$$

$$\Delta^{(1)_{dm}} = \frac{1}{2}(\rho_0^{(1)} - \rho_1^{(1)}) = \begin{pmatrix} 0 & sc-r \\ sc-r & 0 \end{pmatrix}. \tag{72}$$

The total density matrix does not change and the difference density matrix has terms $(sc-r)^n$ instead of $(sc)^n$. Reorganizing the basis vectors we again get the block-diagonal matrices where each of the 2×2 matrices has the form

$$B_p^{(n,k)} = \begin{pmatrix} c^{2(n-k)}s^{2k} & \pm(cs-r)^n \\ \pm(cs-r)^n & c^{2k}s^{2(n-k)} \end{pmatrix}. \tag{73}$$

When normalized, these density matrices have the form of Eq. (26) and are optimally distinguished by measuring them in the x direction. Transforming to the x basis as before we get the same

$$q_k = c^{2(n-k)}s^{2k} + c^{2k}s^{2(n-k)} \tag{74}$$

as before, and

$$p_k = \frac{1}{2} - \frac{(cs-r)^n}{c^{2(n-k)}s^{2k} + c^{2k}s^{2(n-k)}}. \tag{75}$$

The total information can now be calculated as before by assigning these p_k and q_k into Eqs. (42) and (41). Thus, the case of mixed states is also analytically solved for any number of bits, and the influence of mixing on the optimal mutual information is through the p_k's [19].

Calculating the optimal information for small α and any r is possible but complicated. Another alternative which is much simpler is to find a bound on the optimal information using pure states with the same angle, α', on the Poincaré sphere, using

$$\tan 2\alpha' = \frac{\sin 2\alpha - 2r}{\cos 2\alpha}, \tag{76}$$

or using an alternative form for the mixed states Eqs. (69) and (70).

VII. IMPLICATIONS

Quantum oblivious transfer [20] and quantum key distribution [10] protocols use parities of publicly announced subsets of the transmitted bits for both error-correction and privacy amplification (PA). When used for error-correction, subset parities are publicly announced in order to identify errors and correct them, and this is a crucial step in real channels since it could leak information to an eavesdropper. When used for PA [11,12] (say, to derive one final bit) a subset parity is agreed to be the final secret bit, and this technique is used to limit the adversary's information to an exponentially small fraction of a bit. PA is effective when particles are not measured together (see discussions in [12] and in [21]), and presumably also if all measurements are completed before the specification of the subsets used in PA is publicly announced. But it is still an open question whether it is effective also when the adversary can use this specification to *choose* her attack. Our result provides the optimal measurement which can be done to find a parity bit and therefore is crucial for such analysis. In particular cases, when almost fully overlapping states are used, we proved two complementary results regarding that optimal measurement: (i) The optimal information is much larger than the one obtained by measuring each bit separately; (ii) The optimal measurement still yields exponentially small information. Thus we proved an *effectiveness result*: classical PA techniques are effective against *any* quantum measurement. The discussion so far treats an imaginary scenario which is very general but is not good as a cryptographic protocol.

Realistic protocols are very complicated (e.g., due to the use of error correction), hence, are more difficult to analyze. However, to emphasize the importance of the ''effectiveness result'' just mentioned let us consider a different scenario which is common in quantum key distribution schemes: Alice and Bob are the legitimate users who try to establish a secret key. They use any binary scheme and Alice sends n particles through a noisy channel to Bob. An adversary, Eve, is trying to learn information on their key. She gets the particles one at a time, interacts with each one of them weakly, and sends it forward to Bob. She must interact weakly with all particles if she wants to induce only small error rate (alternatively, she could attack strongly only a few of the particles but PA is already proven effective against that type of attack [12]). Classical PA is also effective if Eve cannot use the specification of the hash function (i.e., which subset parities Alice and Bob use as their final string) to attack all bits together. However, there is no physical way to prevent her from doing this if she has a quantum memory. Although the specification is announced after the transmission is over, Eve can keep information in the quantum state of a system which has interacted with all the transmitted particles, and use it after all the specification is announced. Security against such joint attacks in error-free channels is shown in [22,23] but the case of real (noisy) channels and devices has never been completely analyzed. Another route to attack the security problem is to attempt a quantum privacy amplification scheme, relying on the recent results on purifying entanglement [24,25]. However, such a scheme would require Alice and Bob to have future technologies such as large quantum memories and sophisticated quantum gates. Traditional quantum cryptography requires only simple one-particle spin or polarization measurements and is therefore far more practical. One typically would like to consider the case where Alice and Bob use existing technologies, while Eve is restricted only by the laws of physics. It is therefore crucial to obtain the security of quantum cryptography based on classical PA techniques.

Our ''effectiveness result'' allows the derivation of strong security results against ''collective'' attacks [26] in which Eve attaches a *separate* probe to each particle via a translucent attack (defined in [16]), keeps the probes in a quantum memory, and uses *all* classical data to choose the optimal measurement of the probes. Eve must attack each transmitted particle weakly since she does not want to induce large error rate. Therefore, for each particle she obtains a probe with two almost overlapping pure states (or density matrices). Hence the ''effectiveness result'' can be used to imply that her information on the final string is exponentially small in the length of the initial string, and security against collective attacks suggests security against the ''joint'' attack [26], the most general attack allowed by quantum mechanics.

ACKNOWLEDGMENTS

Thanks are due to Asher Peres for suggesting the problem and for many helpful discussions. We also thank E. Biham, S. Braunstein, C. Fuchs, and B. Wootters for helpful discussions. We also thank the Institute for Scientific Interchange in Torino, Italy for allowing us to meet and collaborate, and Gilles Brassard and the Universite´ de Montre´al for hosting another productive meeting. Some of these results were derived independently by Dominic Mayers.

[1] A. S. Holevo, Probl. Inf. Transmission **9**, 110 (1973).

[2] E. B. Davies, IEEE Trans. Inf. Theory **IT-24**, 596 (1978).

[3] C. W. Helstrom, *Quantum Detection and Estimation Theory* (Academic Press, New York, 1976).

[4] A. Peres, *Quantum Theory: Concepts and Methods* (Kluwer, Dordrecht, 1993), Chap. 9.

[5] L. B. Levitin, in *Proceedings of the Workshop on Physics and Computation—Phys. Comp. '92* (IEEE Press, New York, 1993), p. 210.

[6] C. A. Fuchs and C. M. Caves, Phys. Rev. Lett. **73**, 3047 (1994).

[7] C. H. Bennett and G. Brassard, in *Proceedings of IEEE International Conference on Computers, Systems and Signal Processing, Bangalore, India* (IEEE, New York, 1984), p. 175.

[8] A. K. Ekert, Phys. Rev. Lett. **67**, 661 (1991); C. H. Bennett, G. Brassard, and N. D. Mermin, *ibid.* **68**, 557 (1992).

[9] C. H. Bennett, Phys. Rev. Lett. **68**, 3121 (1992).

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

[11] C. H. Bennett, G. Brassard, and J-M. Robert, Siam J. Comput. **17**, 210 (1988).

[12] C. H. Bennett, G. Brassard, C. Crepeau, and U. Maurer, IEEE Trans. Inf. Theory **41**, 1915 (1995).

[13] B. Huttner and A. Peres, J. Mod. Opt. **41**, 2397 (1994).

[14] S. L. Braunstein, A. Mann, and M. Revzen, Phys. Rev. Lett. **68**, 3259 (1992).

[15] F. J. MacWilliam and N. J. A. Sloane, *The Theory of Error-Correcting Codes* (North-Holland, Amsterdam, 1977).

[16] A.K. Ekert, B. Huttner, G.M. Palma, and A. Peres, Phys. Rev. A **50**, 1047 (1994).

[17] J. M. Jauch and C. Piron, Helv. Phys. Acta **40**, 559 (1967); E. B. Davies and J. T. Lewis, Commun. Math. Phys. **17**, 239 (1970).

[18] I. D. Ivanovic, Phys. Lett. A **123**, 257 (1987); A. Peres, *ibid.* **128**, 19 (1988).

[19] The case $\alpha=\pi/4$, including the X shape of the parity matrices, was solved independently by D. Mayers (unpublished).

[20] C. H. Bennett, G. Brassard, C. Crepeau, and M. H. Skubiszewska, in *Proceedings of Crypto '91*, Lecture Notes in Computer Science Vol. 576 (Springer, Berlin, 1992), p. 351.

[21] D. Mayers and L. Salvail (unpublished).

[22] A. Yao (unpublished).

[23] D. Mayers, in *Proceedings of Crypto '95*, Lecture Notes in Computer Science Vol. 963 (Springer, Berlin, 1995), p. 124.

[24] C.H. Bennett, G. Brassard, S. Popescu, B Schumacher, J. A. Smolin, and W.K. Wootters, Phys. Rev. Lett. **76**, 722 (1996).

[25] A. Ekert *et al.*, Report No. quant-ph/9604039.

[26] E. Biham and T. Mor, Report No. quant-ph/9605010.