Optimal Detection of Quantum Information

Asher Peres[(1),(2)] and William K. Wootters[(1),(3),(4)]

[(1)]*Santa Fe Institute, 1120 Canyon Road, Santa Fe, New Mexico 87501*

[(2)]*Department of Physics, Technion–Israel Institute of Technology, Haifa, Israel*[(a)]

[(3)]*Center for Nonlinear Studies and Theoretical Division, Los Alamos National Laboratory, Los Alamos, New Mexico 87545*

[(4)]*Department of Physics, Williams College, Williamstown, Massachusetts 01267*[(a)]

(Received 15 February 1990)

Two quantum systems are identically prepared in different locations. An observer's task is to determine their state. A simple example shows that a pair of measurements of the von Neumann type is less effective than a sequence of nonorthogonal probability-operator measures, alternating between the two quantum systems. However, the most efficient set of operations of that type that we were able to design falls short of a single combined measurement, performed on both systems together.

PACS numbers: 03.65.Bz, 89.70.+c

It is well known[1] that composite quantum systems, consisting of noninteracting parts, can possess nonlocal properties. In particular, a composite system can exhibit correlations which cannot be reproduced by any theoretical model that involves only variables belonging to each subsystem separately. A typical example is a pair of spin-$\frac{1}{2}$ particles produced by the decay of a spinless object. Their combined state, $2^{-1/2}(|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle)$, cannot be reduced to a direct product by any transformation of the bases pertaining to each one of the particles.

In this Letter, we consider a different kind of composite system. Its parts never interacted in the past. They may have been prepared in different laboratories. However, they were prepared according to the same set of instructions. Therefore, these subsystems are in the same quantum state—insofar as their internal variables are considered. For example, we may have two noninteracting spin-$\frac{1}{2}$ particles, prepared with the same polarization. We consider in this Letter a particular example in which there are exactly three possible states for the two particles: Both spins may be directed along the z direction, or both may be in the x-z plane, tilted at 120° or $-120°$ from the z axis. (We have chosen this particular setup because the three possible states of each particle satisfy $\langle\psi_1|\psi_2\rangle\langle\psi_2|\psi_3\rangle\langle\psi_3|\psi_1\rangle = -\frac{1}{8}$, and this is the most negative value obtainable for such a triple product.) We refer to the three special directions as the "signal directions."

Suppose now that an outside observer wants to determine which one of these three known preparations was actually implemented. The answer cannot be unambiguous, because the three states are not orthogonal. The observer may nevertheless assign *probabilities* to the various preparations. The problem thus is to design a measurement procedure which minimizes the unavoidable uncertainty of the result. The original purpose of our work was to determine whether more information could be obtained by means of an apparatus interacting with both particles together, than by separate measurements performed on each one of them individually. An example of a measurement of the former type—which we will call a *combined* measurement—is the one represented by the "entangled" operator in Eq. (3) below.

Despite considerable efforts, we have not been able to obtain a complete answer to our question. Nevertheless, the results are intriguing. In particular, we have found a new measurement technique, acting on each particle separately, which yields more information than any separate-particle method hitherto known. Our best strategy of that type is, however, not as efficient as a well-chosen combined measurement. Thus our work suggests that one can indeed obtain more information by measuring the two particles together, but the results are not conclusive. We propose as a challenge to our fellow theorists either to find a better separate-particle strategy, as good as the combined measurement, or to prove that this aim is unattainable.

To give a precise meaning to our problem, we need to quantify the notion of information gained from a measurement. We will use the standard measure of information developed by Shannon.[2] If the probabilities of the various possible states are $P(s)$, the corresponding knowledge is assigned an entropy $H = -\sum_s P(s)\log_2 P(s)$ (measured in bits). Performing a measurement alters the values of the various $P(s)$. The amount of information I, gained from the measurement, is the amount by which the entropy is reduced: $I = H_{\text{initial}} - H_{\text{final}}$. Typically, some outcomes provide more information than others. A measurement scheme is optimal if it maximizes the *average* information gain. This information-theoretic approach has led to interesting insights in the properties of composite quantum systems.[3]

In the particular model which we are considering, there are three possible preparations (or signals). If they are equally probable, then the initial entropy is $\log_2 3 = 1.58496$. After a measurement is performed, with result r, the *a posteriori* probability for preparation i is given by Bayes's theorem as $P(i|r)=P(r|i)P(i)/P(r)$, where $P(r)=\sum_i P(r|i)P(i)$ is the *a priori* probability for the result r. In these formulas, the conditional probabilities $P(r|i)$ are *known* (from quantum mechanics) and the *a priori* probabilities $P(i)$ are *assumed*. The expected final entropy is

$$\langle H_{\text{final}}\rangle = -\sum_r P(r)\left[\sum_i P(i|r)\log_2 P(i|r)\right]. \tag{1}$$

Our problem is to find a measurement procedure which minimizes this expression.

The simplest strategy is to perform separate Stern-Gerlach (SG) measurements on the two particles. The best result is then obtained as follows. The first particle is tested along one of the signal directions. If the result is positive, the second particle is tested along the same direction. If it is negative, it is tested in the perpendicular direction. The resulting average information gain is 1.05228 bit. Note that information has to be carried from the first instrument to the second one, in order to position the latter. This is, however, information of a *classical* nature, such as a paper printout of the first result. This process is essentially different from the transport of quantum information, encoded in *nonorthogonal states*. The difference is that classical information can be amplified at will and can thus be reliably transported over large distances, whereas quantum information cannot be amplified, because single quanta (allowed to be in nonorthogonal states) cannot be "cloned."[4] The question we raise in this Letter is whether one can learn more about the state of the composite system by allowing the transfer of quantum information, and not just of classical information.

Indeed, a much better result can be achieved by treating in a *global* way the quantum information encoded in the two disjoint particles. Let us denote the three preparation states of the pair by $|a\rangle$, $|b\rangle$, and $|c\rangle$ (for example, $|a\rangle=|\uparrow\uparrow\rangle$). Define

$$|A\rangle = \tfrac{1}{9}[54+8(18)^{1/2}]^{1/2}|a\rangle - \tfrac{1}{9}[18-4(18)^{1/2}]^{1/2}(|b\rangle+|c\rangle), \tag{2}$$

and likewise define $|B\rangle$ and $|C\rangle$, by cyclic permutations. The three states $|A\rangle$, $|B\rangle$, and $|C\rangle$ are orthogonal and normalized. They were chosen to be close to the three preparation states. A fourth vector is the singlet state $|S\rangle$ (total spin 0) which is orthogonal to the three preparation states, each of which has total spin 1. Our strategy is to measure, on the composite spin-1 system, a dynamical variable whose eigenstates are $|A\rangle$, $|B\rangle$, and $|C\rangle$. One such variable is

$$|B\rangle\langle B|-|C\rangle\langle C| = (\tfrac{2}{3})^{1/2}J_x - (\tfrac{1}{3})^{1/2}(J_xJ_z+J_zJ_x), \tag{3}$$

where each angular momentum component, $J_i=\frac{1}{2}(\sigma_{1i}+\sigma_{2i})$, can be represented by a Hermitian matrix of order 3. Any variable of this type can be measured by a generalized SG experiment.[5] For this particular measurement, the expected information gain is 1.36907 bit, a substantial improvement over the proceeding result.

The situation described here is just the converse of the one leading to the violation of Bell's inequality.[1] In showing that quantum mechanics does not satisfy that inequality, Bell used as "entangled" (nonfactorizable) state, but ordinary products of spin operators. Here, it is the state which is factorizable, and the operators which are not. However, the result which was just derived is *not* enough to establish what might be regarded as a new kind of quantum nonseparability, namely, that one needs a nonfactorizable measurement to get the most possible information. This is because the SG measurements we considered are not the only measurements one can perform on a spin-$\frac{1}{2}$ particle. They are a special case of von Neumann measurements,[6] which, although the most widely known, are neither exhaustive nor typical of real measurements.

In a von Neumann measurement, the various outcomes are associated with *orthogonal projection* operators P_n satisfying $\sum P_n=\mathbf{1}$ (where $\mathbf{1}$ is the unit operator), and the probability on the nth outcome is given by $\langle\psi|P_n|\psi\rangle$. However, instead of using a set of orthogonal operators, it may be preferable—indeed it is known to be preferable in many cases—to associate the final outcomes with a more general set of *noncommuting positive operators* A_n, satisfying $\sum A_n=\mathbf{1}$. The probability of getting the nth outcome is $\langle\psi|A_n|\psi\rangle$, so that this set of A_n forms a probability-operator measure (POM).[7,8]

To show that such a POM is physically realizable, one relies on Neumark's theorem.[9,10] The latter asserts that one can extend the Hilbert space of states $\mathcal{H}$, in which the POM is defined, in such a way that there exists, in the extended space $\mathcal{K}$, a set of *orthogonal* projection operators satisfying $\sum P_n=\mathbf{1}$, and such that $A_n=\Pi P_n\Pi$, where Π is the projection operator from $\mathcal{K}$ into $\mathcal{H}$. The practical realization of Neumark's theorem involves a method analogous to heterodyne detection in communications engineering: The system (signal) to be observed is combined with another *known* system (signal), sometimes called *ancilla*.[11] Thereafter, a standard von Neumann measurement, corresponding to projection operators P_n, is performed on the combined system. The probability of obtaining the nth outcome is $\langle\psi\Phi|P_n|\Phi\psi\rangle$, where Φ is the ancilla's initial state (which is known). This probability can therefore be written as $\langle\psi|A_n|\psi\rangle$, where $A_n=\langle\Phi|P_n|\Phi\rangle$ is a positive operator. Notice that the number of outcomes may be greater than the dimension of the original Hilbert space $\mathcal{H}$.

The amount of information that can be obtained in this way about a physical system can be larger than if the observer is restricted to von Neumann measurements, without ancilla. There are theorems[12,13] limiting that amount of information. However, these theorems are not directly applicable to our current problem—namely, how much information can be gathered by testing each subsystem separately. This condition restricts the admissible Hamiltonians, as well as the admissible initial states of the ancillas and apparatuses (these states must be factorizable). It therefore restricts the class of operators which are measurable.

In the remaining part of this Letter, we shall consider a number of separate-particle measurement schemes, allowing one to increase the information gain beyond 1.05228 bit (its value for two simple SG measurements) toward successively higher values, but without being able to reach 1.36907 bit, which results from the combined procedure. For the sake of brevity, we shall discuss this matter solely in terms of POM components (namely, the 2-by-2 matrices A_n) without making explicit use of ancillas. An explicit procedure for preparing ancillas is described elsewhere.[10]

The simplest improvement is to replace the first SG measurement by a nonorthogonal POM with[10-12]

$$A_j = \tfrac{2}{3}(1 - |\psi_j\rangle\langle\psi_j|) . \tag{4}$$

Note that $\sum A_j = 1$. For example, if the signal state is $|\psi_1\rangle$, the probabilities of the three outcomes of the POM are $\langle\psi_1|A_1|\psi_1\rangle = 0$ and $\langle\psi_1|A_2|\psi_1\rangle = \langle\psi_1|A_3|\psi_1\rangle = \frac{1}{2}$. The result of performing this POM is thus to rule out definitively one of the three signal states, and to leave equal *a posteriori* probabilities for the two others.[14] The next step is an ordinary SG test performed on the second particle, along a direction perpendicular to the one selected by the first particle's POM (for example, if A_1 yielded "yes" so that one can rule out $|\psi_1\rangle$, the second particle shall be tested along a direction perpendicular to that of signal $|\psi_1\rangle$). The resulting information gain is 1.23038 bit.

This result can further be improved by an *iterative* procedure which involves more than two steps: Instead of executing a conclusive measurement on the first particle, and then on the second one, we shall perform a sequence of "fuzzy" measurements, minimizing the disturbance to the systems under study. This requires a new notion: pure and mixed POM elements. The definition is analogous to that of pure and mixed density matrices. A POM element is *pure* if it is represented by a matrix of rank 1: $A_n = |u\rangle\langle u|$, where $|u\rangle$ may not be normalized. A *mixed* POM element can be represented as $\sum c_n |u_n\rangle\langle u_n|$, where all the c_n are positive. In the particular setup that we are considering, it is convenient to represent POM elements by three parameters:

$$A(w_n, p_n, \alpha_n) = w_n[1 + p_n(\sigma_z \cos\alpha_n + \sigma_x \sin\alpha_n)] , \tag{5}$$

where $0 < w_n \leq 1$ and $0 \leq p_n \leq 1$. The weights w_n and purities p_n must satisfy $\sum w = 1$ and $\sum wpe^{i\alpha} = 0$, so that $\sum A = 1$.

We again introduce a new notion: the *refinement* of a mixed POM element, which is its decomposition

$$A(w,p,\alpha) = \sum_s A(w_s, p_s, \alpha_s) , \tag{6}$$

where $\sum w_s = w$ and $\sum w_s p_s e^{i\alpha_s} = wpe^{i\alpha}$. The term *refinement* originates from the mechanism used to break a POM element into a sum of positive matrices. Consider in the extended Hilbert space $\mathcal{H}$ (representing the system under study together with its ancilla) a unit submatrix P of rank $n > 1$. It can be written as $P = \sum P_s$, where the various P_s have ranks $< n$ and satisfy $P_r P_s = \delta_{rs} P_s$. Define $A_s = \Pi P_s \Pi$, which is a POM element in $\mathcal{H}$. Obviously $\sum A_s = \Pi P \Pi = A$. We do not always have $p_s > p$, but, if this process is repeated until all P_s have rank 1, the result must be either $w_s = 0$ or $p_s = 1$, namely, pure POM elements which can no longer be refined. Note that the method of breaking P into $\sum P_s$ is not unique and is *at the discretion of the observer.*

We shall now explain the iterative separate-particle measurement strategy which yields a result better than 1.23038 bit. We start by performing on one of the particles a low-purity POM, consisting of elements such as $\frac{1}{3}(1-p) + pA_j$, where A_j is given by Eq. (5). According to the result of the test, we choose another low-purity POM for the second particle. Its outcome then instructs us how to refine the first particle's POM, and so on. The process is repeated many times, until we end up with two pure POM elements, with angles α and β. The final information gain depends only on α and β (not on the intermediate results) and it is maximal when α and β lie on different sides of a signal direction, and 60° from it. Reasonably good results are obtained in the range 40° to 75°. The problem is to find a refining method which steers α and β into the favorable configuration.

We have tried numerous strategies and simulated their outcomes by Monte Carlo methods. The best result was obtained as follows. In the first, low-purity step, α and β are brought to different sides of a signal direction. Thereafter, each POM element with angle α is broken into two, with the new angles α' and α'' equally distant from the signal direction which is most distant from β (and vice versa). This tends to keep α and β in the favorable region. We thereby obtained $\langle I\rangle = 1.26205$ bit, with a standard deviation of 0.11748 (there were 10^5 runs, each one involving 100 purity steps).

It is likely that better algorithms can be devised, but we doubt that the result 1.36907 bit, obtainable by means of a combined measurement, can be approached by any variant of our POM refining method. This leaves two possibilities: A radically new separate-particle method yielding 1.36907 bit, or a formal proof that this goal is unattainable.

Work by A.P. at Technion was supported by the Gerard Swope Fund. W.K.W. would like to thank the two groups in the Theoretical Division at Los Alamos that contributed to this research: Complex Systems (T-13) and Theoretical Astrophysics (T-6).

[a]Permanent address.

[1]J. S. Bell, Physics (Long Island City, N.Y.) **1**, 195 (1964).
[2]C. Shannon, Bell Syst. Tech. J. **27**, 379 (1948); **27**, 623 (1948).
[3]S. L. Braunstein and C. M. Caves, Phys. Rev. Lett. **61**, 662 (1988).
[4]W. K. Wootters and W. H. Zurek, Nature (London) **299**, 802 (1982).
[5]A. R. Swift and R. Wright, J. Math. Phys. **21**, 77 (1980).
[6]J. von Neumann, *Mathematical Foundations of Quantum Mechanics* (Princeton Univ. Press, Princeton, 1955). The theoretical feasibility of these measurements is proved by constructing a Hamiltonian which generates the required process.
[7]E. B. Davies and J. T. Lewis, Commun. Math. Phys. **17**, 239 (1970).
[8]P. A. Benioff, J. Math. Phys. **13**, 231 (1972); **13**, 908 (1972); **13**, 1347 (1972).
[9]M. A. Neumark, C. R. Dokl. Acad. Sci. URSS **41**, 359 (1943).
[10]A. Peres, Found. Phys. **20**, 1441 (1990).
[11]C. W. Helstrom, *Quantum Detection and Estimation Theory* (Academic, New York, 1976), pp. 74–83.
[12]A. S. Kholevo, Probl. Inf. Transm. (Engl. Transl.) **9**, 110 (1973).
[13]E. B. Davies, IEEE Trans. Inf. Theory **24**, 596 (1978).
[14]The detailed measurement mechanism, including the ancilla, is described in Ref. 10, p. 1450, and Ref. 11, p. 81. Note that the states called ψ_j in this Letter are *orthogonal* to those called ψ_j in these references.