Storage and Retrieval of Quantum Information

Asher Peres
Department of Physics
Technion—Israel Institute of Technology
32 000 Haifa, Israel

Abstract

Information encoded in non-orthogonal quantum states cannot be duplicated, nor amplified, and in general it is only partly recoverable. The most efficient way of retrieving it is not a direct "quantum measurement" (as defined by von Neumann), but an indirect method similar to heterodyne detection in communications engineering. The mathematical representation of this process requires the introduction of a positive operator valued measure. The optimization of these measures is not yet fully understood. An interesting and potentially important application of quantum information is its use in cryptography.

1: Non-orthogonal information and the limits of objectivity

Information stored in classical form, such as printed text, can be examined objectively without altering it in any detectable way, let alone destroying it. It is impossible to manipulate in this way quantized information encoded in non-orthogonal states, for instance in the polarizations of photons. Therefore quantized information is only partly recoverable [1].

The fundamentally novel feature introduced by quantum theory is that state preparations which are *macroscopically different* can produce quantum states which are *not orthogonal*, and therefore cannot be distinguished unambiguously from each other. As a concrete example, consider a spin-$\frac{1}{2}$ particle whose polarization state is prepared by selecting the upper beam in a Stern-Gerlach experiment. We are given the choice of orienting the magnet along direction $\mathbf{n}_1$, or along direction $\mathbf{n}_2$. The corresponding quantum states of the resulting beams, ψ_1 and ψ_2 respectively, are not orthogonal. Their overlap is $|\langle\psi_1|\psi_2\rangle|^2 = \frac{1}{2}(1 + \mathbf{n}_1 \cdot \mathbf{n}_2)$. This expression is the probability that, following a preparation of state ψ_1, the question "Was the prepared state ψ_2?" will be answered in the affirmative (and vice versa). The answer cannot be predicted with certainty. Once the spin-$\frac{1}{2}$ particle has been severed from the macroscopic apparatus which prepared it, that particle does not carry the full information about its preparation procedure. Some questions become ambiguous, and only *probabilities* can be assigned to their possible answers.

It is therefore impossible to establish the veracity of information supplied about a quantum preparation procedure merely by examining the quantum system that has been so prepared. *The notions of truth and falsehood acquire new meanings in the logic of quantum phenomena.* We may increase the confidence level by testing more than one system but this, in turn, depends on our willingness to believe in the uniformity of the preparations. This issue itself is amenable to a test, but only if other suitable assumptions are made. In general, the residual Shannon entropy (*i.e.*, the uncertainty) left after a quantum test strongly depends on the amount and type of information that was available before the test. This is also true in classical information theory, but the effect is more striking for quantum information which can be supplied in subtler ways.

2: Quantum information gain

How well can we distinguish non-orthogonal states? Consider N different state preparations, represented by known density matrices ρ_i, and let p_i be the known *a priori* probability for preparation i. The testing procedure that we shall use may yield n different outcomes (in general $n \neq N$). The conditional probability $P_{\mu i}$ that preparation i yields result μ is assumed known. Having found a particular result μ, we can compute $Q_{i\mu}$, the likelihood (or *a posteriori* probability) for preparation i. It is given by Bayes's theorem: $Q_{i\mu} = P_{\mu i}\, p_i / q_\mu$, where $q_\mu = \sum_j P_{\mu j}\, p_j$ is the *a priori*

0-8186-3420-0/93 $03.00

probability for occurrence of outcome μ.

Before we found the result μ, we only knew the probabilities p_i. Shannon's entropy, which is a measure of our ignorance, was $-\sum p_i \log p_i$. After we have found the result μ, we can compute the *a posteriori* probabilities $Q_{i\mu}$, and the new Shannon entropy is $H_\mu = -\sum_i Q_{i\mu} \log Q_{i\mu}$. Note that for some outcomes H_μ may be larger than the initial entropy, so that the result of the test is to increase our uncertainty. (Shannon's entropy does not really measure an objective ignorance level, but rather our subjective feeling of ignorance.) On the *average*, however, a quantum test (or "measurement") reduces the Shannon entropy. The average information gain is

$$I_{\rm av} = H_{\rm initial} - \langle H_{\rm final} \rangle$$
$$= -\sum_i p_i \log p_i - \sum_\mu q_\mu H_\mu . \qquad (1)$$

Usually, we are interested in the experimental procedure which maximizes $I_{\rm av}$ for given p_i and ρ_i. However, in some situations where the penalty for errors is large, it is preferable to increase the confidence level of *some* results (e.g., to obtain a *certainty* for them) at the expense of a reduced information gain for other trials. An example will be given below.

3: Positive operator valued measures

The information gain $I_{\rm av}$ defined by Eq. (1) depends on the conditional probabilities $P_{\mu i}$ for obtaining result μ when the system is prepared in state ρ_i. The value of $P_{\mu i}$ is determined by the testing procedure. The most efficient way of obtaining information about the state of a quantum system may not be a direct quantum test (a "measurement" as defined in von Neumann's treatise [2]). It is often preferable to introduce an auxiliary quantum system, called *ancilla* [3], prepared in a *known* state $\rho_{\rm aux}$ (this is similar to heterodyne detection in communications engineering). The combined, uncorrelated state of the original quantum system and the ancilla is

$$(\rho_i \otimes \rho_{\rm aux})_{mr,ns} = (\rho_1)_{mn} (\rho_{\rm aux})_{rs} . \qquad (2)$$

A von Neumann measurement is then performed in the combined Hilbert space, where that measurement is represented by an orthogonal resolution of the identity. Different outcomes μ and ν correspond to *orthogonal* projectors P_μ which satisfy $\mathsf{P}_\mu \mathsf{P}_\nu = \delta_{\mu\nu} \mathsf{P}_\nu$ and $\sum_\mu \mathsf{P}_\mu = \mathbf{1}$. The probability that outcome μ will follow preparation i is

$$P_{\mu i} = {\rm Tr}\left[\mathsf{P}_\mu \, (\rho_i \otimes \rho_{\rm aux})\right]$$
$$= \sum_{mnrs} (\mathsf{P}_\mu)_{mr,ns} \, (\rho_i)_{nm} \, (\rho_{\rm aux})_{sr} . \qquad (3)$$

This can also be written as $P_{\mu i} = {\rm Tr}(\mathsf{A}_\mu \rho_i)$, where $(\mathsf{A}_\mu)_{mn} = \sum_{rs} (\mathsf{P}_\mu)_{mr,ns} (\rho_{\rm aux})_{sr}$ is an operator which acts on the *original* Hilbert space $\mathcal{H}$.

The Hermitian matrices A_μ, which in general do *not* commute, satisfy $\sum_\mu \mathsf{A}_\mu = \mathbf{1}$. The set of A_μ is called a *positive operator valued measure* [4,5] (or POVM, for brevity). The difference between these POVM's and von Neumann's *projection valued measures* is that the number of available preparations and that of available outcomes may be different from each other, and also different from the dimensionality of $\mathcal{H}$. The probability of outcome μ is now given by ${\rm Tr}(\mathsf{A}_\mu \rho)$, instead of von Neumann's ${\rm Tr}(\mathsf{P}_\mu \rho)$.

Conversely, Neumark's theorem [6] asserts that, given any positive operator valued measure, one can extend the Hilbert space of states $\mathcal{H}$, in which the A_μ are defined, so that there exists, in the extended space $\mathcal{K}$, a set of *orthogonal* projectors P_μ satisfying $\sum \mathsf{P}_\mu = \mathbf{1}$, and such that A_μ is the result of projecting P_μ from $\mathcal{K}$ into $\mathcal{H}$.

Moreover, it can be shown explicitly [7] that this extension may always be interpreted as the introduction of an auxiliary, independently prepared quantum system (the ancilla). Therefore a generalized quantum test, defined by an arbitrary POVM, is always equivalent to an ordinary "measurement" performed on a composite system.

4: Optimization

It is often desirable to maximize the average information gain for a given set of p_i and ρ_i. The determination of the optimal strategy is a difficult problem for which no general solution is known. Some partial results are however available. It can be proved [8] that the optimal POVM consists of matrices of rank one, $\mathsf{A}_\mu = \mathsf{u}_\mu \mathsf{u}_\mu^\dagger$, where the vectors u_μ are in general neither normalized nor orthogonal. The required number, n, of different A_μ satisfies the inequality $d \le n \le d^2$, where d is the dimensionality of the subspace of $\mathcal{H}$ spanned by the different preparations ρ_i. The average information gain is bounded by [9,10]

$$I_{\rm av} \le S\Big(\sum_i p_i \rho_i\Big) - \sum_i p_i S(\rho_i), \qquad (4)$$

with equality holding if, and only if, all the density matrices ρ_i commute.

In the above expression, $S(\rho)$ is the von Neumann entropy for state ρ:

$$S(\rho) = -\mathrm{Tr}\,(\rho \log \rho). \tag{5}$$

The latter is closely related to the thermodynamical entropy, and it is in general smaller than the Shannon entropy H (equality between them is reached only if all the ρ_i are pure states and are orthogonal to each other). Therefore, the recoverable information can never exceed the von Neumann entropy.

As a simple example, let us try to distinguish two non-orthogonal states u and v, prepared with equal probabilities. The maximal value of I_{av} can be obtained by measuring the operator $\mathsf{u}\mathsf{u}^\dagger - \mathsf{v}\mathsf{v}^\dagger$, whose eigenvalues are

$$\pm\lambda = \pm\left[1 - |\langle \mathsf{u}|\mathsf{v}\rangle|^2\right]^{1/2}. \tag{6}$$

The probabilities of obtaining these eigenvalues, following a preparation of u or v, are $p_\pm = (1 \pm \lambda)/2$, respectively, and the average information gain is

$$\begin{aligned} I_{\mathrm{av}} &= \log 2 + (p_+ \log p_+ + p_- \log p_-)/2 \\ &= [(1+\lambda)\log(1+\lambda) + (1-\lambda)\log(1-\lambda)]/2. \end{aligned} \tag{7}$$

There are however some situations (such as in quantum cryptography, discussed below) where it is not the average information gain which is of interest, but the certainty that at least *some* of the readings are exact. In that case, the simplest (but not the most efficient) approach is to measure one of the projection operators

$$\mathsf{P}_{\neg\mathsf{u}} = \mathbf{1} - \mathsf{u}\mathsf{u}^\dagger \qquad \text{or} \qquad \mathsf{P}_{\neg\mathsf{v}} = \mathbf{1} - \mathsf{v}\mathsf{v}^\dagger. \tag{8}$$

A positive result for $\mathsf{P}_{\neg\mathsf{u}}$ indicates with certainty that the information carrier was in the v state, and vice versa. A null result is of no use to us if only unambiguous conclusions are acceptable. It only gives the *a posteriori* probabilities

$$Q_{\mathsf{u}0} = \frac{1}{2-\lambda^2} \qquad \text{and} \qquad Q_{\mathsf{v}0} = \frac{1-\lambda^2}{2-\lambda^2}. \tag{9}$$

The probability of getting this inconclusive null result is $1 - \lambda^2/2$. That probability can be reduced somewhat by a more sophisticated measurement process [11,12], which uses an ancilla, prepared in an initial state a. Let b be another state of the ancilla, orthogonal to a. Choose the phases of u and v so that $\langle \mathsf{u}|\mathsf{v}\rangle$ is positive. Let u' and v' be two orthogonal states and let w be any arbitrary state of the information carrier. Then there exists in the product Hilbert space a unitary matrix U such that

$$\mathsf{u}\otimes\mathsf{a} \to \mathsf{U}(\mathsf{u}\otimes\mathsf{a}) = \mu\,\mathsf{u}'\otimes\mathsf{a} + \nu\,\mathsf{w}\otimes\mathsf{b}, \tag{10}$$

and

$$\mathsf{v}\otimes\mathsf{a} \to \mathsf{U}(\mathsf{v}\otimes\mathsf{a}) = \mu\,\mathsf{v}'\otimes\mathsf{a} + \nu\,\mathsf{w}\otimes\mathsf{b}, \tag{11}$$

where

$$\nu^2 = 1 - \mu^2 = \langle \mathsf{u}|\mathsf{v}\rangle. \tag{12}$$

The dynamical evolution in Eqs. (10) and (11) can be generated by a suitable Hamiltonian and therefore it is in principle realizable. After the combined system reaches one of the final states in (10) or (11), we test whether the ancilla is in state a (that is, not in state b). If the answer is positive, the states u' and v' of the information carrier can be distinguished unambiguously. This happens in a fraction μ^2 of cases. The probability for an inconclusive answer then is

$$\nu^2 = \sqrt{1-\lambda^2} \le 1 - \lambda^2/2, \tag{13}$$

and the final Shannon entropy is $\sqrt{1-\lambda^2}\,\log 2$. We thus have

$$I_{\mathrm{av}} = \left(1 - \sqrt{1-\lambda^2}\right)\log 2, \tag{14}$$

which is always smaller than I_{av} in Eq. (7), except in the trivial case $\lambda = 0$ (this is easily seen by differentiating both expressions with respect to λ).

5: Quantum cryptography

An interesting and potentially important application of quantum information is quantum cryptography, which recently entered the experimental era [13]. The objective of cryptography is to allow the transmittal of information in such a way that it cannot be understood by an opponent who might intercept it. An absolutely safe encryption method, whereby code breaking is impossible, involves a random sequence, called a ***key***, which is as long as the message to be transmitted. This key must be used only once, and then discarded. The problem is how to distribute such a key to people who initially share no secret information, by using an insecure communication channel subject to inspection by a hostile eavesdropper. If only classical means are used, this is an impossible task. Quantum information, on the other hand, provide various solutions. It is its elusiveness which makes it ideal for transmitting secrets.

Conceptually, the simplest protocol [14] uses two non-orthogonal states, u and v. One of the parties emits a random sequence of quanta prepared in the

u or v state; the other one randomly tests them by one of the methods discussed above, and then publicly announces the cases in which she was able to definitely identify the quantum state of the particle, without saying of course whether it was u or v. The resulting sequence of u and v, which is known only to the two parties, is the cryptographic key.

It still is necessary to verify that there is no eavesdropper who intercepts some of the information carriers and substitutes other, fake carriers. This would cause a mismatch in the two keys, which could be fatal to the encryption-decryption process. A simple method to ensure that both keys are the same is to use only the first half of each random sequence as the key. That half is combined bitwise, by the Boolean operation XOR (exclusive OR), with the second half of the random sequence, and the resulting bit sequence is publicly announced by each party. If discrepancies between these two sequences are rare, the few mismatching bits are discarded, and it is likely that the remaining parts of the keys are identical. If there are more discrepancies than can be explained by instrumental defects, the presence of a mischievous eavesdropper ought to be suspected. In that case, more efficient methods of key reconciliation and privacy amplification can be used [13].

6: Nonlocal quantum information

I conclude this review by mentioning an *unsolved* problem. Let two quantum systems be identically prepared in different locations. The problem is to find the optimal stategy for determining their common state. It is plausible [15] that a single nonlocal measurement, performed on both systems together, is more efficient than various combinations of POVM's testing each system separately.

However, the problem of finding the best separate-particle strategy is quite difficult and has not yet been solved. It can easily be shown [15] that the most efficient POVM's do *not* consist of matrices of rank one, as in the case of a single system [8]. These are *impure* POVM's of a more general nature (just as density matrices ρ are more general than pure states). These impure POVM's must be gradually refined, by testing alternatively each one of the two systems, and using the result of the last test to choose the parameters of the next one. While the general policy is clear, the most efficient detailed protocol has not been found. This fascinating problem is thus offered as a challenge to quantum theorists.

Acknowledgments

This work was supported by the Gerard Swope Fund, and the Fund for Encouragement of Research at Technion.

References

1. L. B. Levitin, "Information theory for quantum systems," in: *Information Complexity and Control in Quantum Physics*, ed. by A. Blaquière, S. Diner, and G. Lochak, Springer (1987) p. 15.
2. J. von Neumann, *Mathematical Foundations of Quantum Mechanics*, Princeton Univ. Press (1955).
3. C. W. Helstrom, *Quantum Detection and Estimation Theory*, Academic Press, (1976) pp. 74–83.
4. J. M. Jauch and C. Piron, *Helv. Phys. Acta* 40 (1967) 559.
5. E. B. Davies and J. T. Lewis, *Comm. Math. Phys.* 17 (1970) 239.
6. M. A. Neumark, *Izv. Akad. Nauk SSSR, Ser. Mat.* 4 (1940) 53, 277; *C. R. (Doklady) Acad. Sci. URSS (N.S.)* 41 (1943) 359.
7. A. Peres, *Found. Phys.* 20 (1990) 1441.
8. E. B. Davies, *IEEE Trans. Inform. Theory* IT-24 (1978) 596.
9. L. B. Levitin, in *Proc. Fourth All-Union Conf. on Information and Coding Theory*, Tashkent (1969) p. 111 [in Russian].
10. A. S. Holevo, *Probl. Inform. Transmission* 9 (1973) 110, 177 [transl. from the Russian].
11. I. D. Ivanovic, *Phys. Lett.* A 123 (1987) 257.
12. A. Peres, *Phys. Lett.* A 128 (1988) 19.
13. C. H. Bennett, F. Bessette, G. Brassard, L. Salvail, and J. Smolin, *J. Cryptology* 5 (1992) 3.
14. C. H. Bennett, *Phys. Rev. Lett.* 68 (1992) 3121.
15. A. Peres and W. K. Wootters, *Phys. Rev. Lett.* 66 (1991) 1119.