2. From Shannon To Quantum Information Science

2. Mixed States

Rajiah Simon

In this part of the article, we are concerned primarily with mixed states. As was the case with Part 1 of the article, the notion of quantum entanglement plays the centre stage in quantum information processes.

Rajiah Simon is a Professor at the Institute of Mathematical Sciences, Chennai. His primary interests are in classical and quantum optics, geometric phases, group theoretical techniques and quantum information science.

Part 1 Ideas and techniques, *Resonance*, Vol.7, No.2, pp.66-85, 2002.

In practical information processing situations one is forced, very often, to go beyond *pure states* described by state vectors to *mixed states* described by *density operators* (density matrices). As a consequence the unitary Schrödinger evolution gets generalized to completely positive maps, and the von Neumann projection measurement to what are known as positive operator valued measures (fondly called POVM).

Mixed States

Mixed states can arise due to a variety of reasons. For instance, our physical situation on hand may be such that we are not in possession of complete knowledge of the state of the system. Suppose we know that our system has a probability p_k of being in the state $|\chi_k\rangle$, $k = 1, 2, \quad \ell$ and assume we know nothing more. Since the density matrix of a pure state $|\chi\rangle$ is defined simply as the outer product $|\chi\rangle\langle\chi|$, the state in the above situation of partial knowledge is described by the density matrix $\rho = \sum_k p_k |\chi_k\rangle\langle\chi_k|$. Expectation values of dynamical variables are often the quantities of interest in quantum theory. For a pure state $|\Psi\rangle$, the expectation value of a dynamical variable (hermitian operator) $\widehat{\Omega}$ is given by the expression $\langle\Psi|\widehat{\Omega}|\Psi\rangle$. For a mixed state ρ, it is given by $\mathrm{tr}\,(\,\widehat{\Omega}\,\rho\,)$, which takes the form $\sum_k p_k \langle\chi_k|\widehat{\Omega}|\chi_k\rangle$ if the mixed state is realized as a convex sum of pure states as above.

Keywords
Mixed states, entanglement witnesses, partial transpose, quantum computers, von Neumann entropy.

Mixed states often arise in a somewhat different manner. Suppose a bipartite system consisting of subsystems A and B is in a state ρ_{AB}. Let us assume that we are interested in subsystem A alone. This means we are interested in measurement of (local) dynamical variables $\widehat{\Omega}_A$ which act only on the Hilbert space $\mathcal{H}_A$, and do nothing on $\mathcal{H}_B$. The action of such a variable is described by $\widehat{\Omega}_A \otimes \mathrm{Id}_B$ on the extended Hilbert space $\mathcal{H}_S = \mathcal{H}_A \otimes \mathcal{H}_B$. Since the identity or unit operator Id_B has the form $\sum_\alpha |\phi_\alpha\rangle\langle\phi_\alpha|$ (resolution of unity) in any O.N.B. $\{\,|\phi_\alpha\rangle\,\}$ of $\mathcal{H}_B$, the expectation value of the subsystem variable becomes

$$\begin{aligned}\mathrm{tr}(\widehat{\Omega}_A \otimes \mathrm{Id}_B \rho_{AB}) &= \mathrm{tr}_A \mathrm{tr}_B(\widehat{\Omega}_A \otimes \mathrm{Id}_B \rho_{AB})\\ &= \sum_{k\alpha} \langle\psi_k| \otimes \langle\phi_\alpha|\widehat{\Omega}_A \otimes \mathrm{Id}_B \rho_{AB}|\psi_k\rangle \otimes |\phi_\alpha\rangle\\ &= \mathrm{tr}(\widehat{\Omega}_A \rho_A),\end{aligned}$$

where we have written in the last line simply $\mathrm{tr}\,(\,\widehat{\Omega}_A\,\rho_A\,)$, rather than $\mathrm{tr}_A\,(\,\widehat{\Omega}_A\,\rho_A\,)$, since we are left with just the Hilbert space $\mathcal{H}_A$. The operator ρ_A is defined through

$$\rho_A \equiv \mathrm{tr}_B\,\rho_{AB} = \sum_\alpha \langle\phi_\alpha|\rho_{AB}|\phi_\alpha\rangle,$$

which is clearly an operator on the Hilbert space of subsystem A. The trace operation in the last equation is called *partial trace*, and the resulting ρ_A is called the *reduced density matrix* of subsystem A. These notions apply irrespective of whether ρ_{AB} is a pure or mixed state. It is clear that for *every* measurement on subsystem A the state ρ_{AB} behaves exactly as if subsystem A was in the state ρ_A. It is important to appreciate the fact that a pure bipartite state can result in a mixed reduced state ρ_A. Indeed, for a bipartite pure state $\rho_{AB} = |\Psi\rangle\langle\Psi|$, the reduced state ρ_A is pure if and only if ρ_{AB} is a product state, showing that partial trace operation is another entanglement witness for bipartite pure states.

Interaction with the environment can evolve an initially pure state of A into a mixed (reduced) state.

Interaction with the environment can evolve an initially pure state of A into a mixed (reduced) state.

The point being made is that interaction with the environment can evolve an initially pure state of A into a mixed (reduced) state. Let B represent the environment and, to begin with, let the closed total system AB be in a product pure state, so that the initial reduced density matrix ρ_A is a pure state as well. Due to interaction of A with B, the total state will become an entangled (pure) state in course of time, resulting in a mixed reduced state for A (as well as for B). We may note in passing that while the Schrödinger evolution of the (closed) total system remains unitary and hence reversible, irreversible evolution of subsystem A and its decoherence comes about precisely through this route when B is enormously large and its state is thermal (bath).

As an illustration let us apply the partial trace operation to the pure state $\rho_{AB} = |\Psi\rangle\langle\Psi|$, where $|\Psi\rangle$ has the canonical form $|\Psi\rangle = \sum_{j=1}^{r} \sqrt{\lambda_j}\, |\psi'_j\rangle \otimes |\phi'_j\rangle$. We obtain $\rho_A = \sum_{j=1}^{r} \lambda_j\, |\psi'_j\rangle\langle\psi'_j|$ and $\rho_B = \sum_{k=1}^{r} \lambda_k\, |\phi'_k\rangle\langle\phi'_k|$ showing that the λ_j's are the nonzero eigenvalues of ρ_A, *as well as* of ρ_B, and that $|\psi'_j\rangle$'s ($|\phi'_j\rangle$'s) are the corresponding eigenvectors.

There exists a process called *purification*, which is essentially the reverse of partial trace. Given density matrix ρ of a system A, we attach to A an *ancilla* B. The pure state of the composite system, of which the given ρ is the partial trace, is called *a* purification of ρ. While partial trace is unique, purification is not for, referring to the Schmidt form, we see that the ancilla states $\{\, |\phi_j\rangle \,\}$ can be chosen at will. Further, it is clear that purification of ρ will result in a product state if and only if ρ is a pure state.

We recall in passing the three defining conditions on a density matrix ρ:

$$\rho^\dagger = \rho, \quad \rho \geq 0, \quad \mathrm{tr}\, \rho = 1.$$

That is, any and every hermitian positive semidefinite operator of unit trace is a bonafide density operator. These defining properties imply that if ρ_1 and ρ_2 are density operators, then $a\,\rho_1 + (1-a)\,\rho_2$ is a density operator for all $0 < a < 1$. The quantum state space is thus a convex set. Pure states meet the additional condition $\rho^2 = \rho$ or, equivalently, $\operatorname{tr}\rho^2 = 1$. They cannot be realized as nontrivial convex combinations of other states. These are the extremal points of the convex state space. For mixed states $\operatorname{tr}\rho^2 < 1$, and they correspond to non-extremal points of the convex state space. Note that while all interior points are non-extremal, the boundary points can be non-extremal as well. For instance, a solid tetrahedron viewed as a convex set has only four extremal points.

The quantum state space is thus a convex set.

Partial Transpose as an Entanglement Witness

The notion of entanglement is considerably richer in the case of mixed states. Given the density matrix ρ_{AB} of a bipartite mixed state, we can realize it as an ensemble of pure states $\{\,|\Psi\rangle_k, p_k\,\}$. This simply means $\rho_{AB} = \sum_k p_k\,|\Psi_k\rangle\langle\Psi_k|$, $p_k > 0$, $\sum_k p_k = 1$. The non-triviality of ensemble realizations arises from the fact that the $|\Psi\rangle_k$'s need not be orthogonal; indeed, they need not be even linearly independent. Thus the set of ensembles realizing a given mixed state ρ_{AB} is a huge family; this family can be fully characterized, though. A state ρ_{AB} is said to be *separable* if and only if there exists a realization $\{\,|\Psi\rangle_k, p_k\,\}$ in which all the $|\Psi\rangle_k$'s are product states, i.e., if it can be written in the form

$$\rho_{AB} = \sum_k p_k\,\rho_{A\,k} \otimes \rho_{B\,k},$$

where p_k's are positive, and $\rho_{A\,k}$'s ($\rho_{B\,k}$'s) are density matrices of subsystem A (B); without loss of generality, these density matrices of the subsystems can be chosen to correspond to pure states. Stated differently, any convex sum of product states is, by definition, a sep-

The notion of entanglement is considerably richer in the case of mixed states.

arable state. Thus, separable states constitute a convex subset of the convex state space. An inseparable state is said to be *entangled.*

There are correlations between the subsystems even when the composite system is in a state which is separable (but not simply a tensor product). These correlations can, however, be understood in classical terms. An entangled state exhibits genuinely nonclassical correlations. Whereas all classical correlations need to respect Bell's inequalities, the nonclassical ones need not.

All classical correlations need to respect Bell's inequalities, the nonclassical ones need not.

Partial trace is not good enough as entanglement witness in the case of mixed states. A linear map called *partial transpose* has proved to be an immensely useful entanglement witness for bipartite mixed states. This map is conveniently defined in a product basis $\{ |\Psi_{j\alpha}\rangle = |\psi_j\rangle \otimes |\phi_\alpha\rangle \}$ of $\mathcal{H}_A \otimes \mathcal{H}_B$. With ρ_{AB} expressed in this basis as

$$\rho_{AB} = \sum_{j,k,\alpha,\beta} \rho_{j\alpha,k\beta} \, |\Psi_{j\alpha}\rangle\langle\Psi_{k\beta}|,$$

the partial transpose map takes ρ_{AB} to

$$\begin{aligned} \rho'_{AB} &= \sum_{j,k,\alpha,\beta} \rho'_{j\alpha,k\beta} \, |\Psi_{j\alpha}\rangle\langle\Psi_{k\beta}| \\ &= \sum_{j,k,\alpha,\beta} \rho_{j\beta,k\alpha} \, |\Psi_{j\alpha}\rangle\langle\Psi_{k\beta}|. \end{aligned}$$

It turns out that partial transpose of a bipartite density matrix need not be a density matrix, in the sense that ρ_{AB} can acquire negative eigenvalues on partial transposition. Denoting the transpose of any matrix ρ by ρ^{T}, we see that a separable state changes as follows under partial transposition (PT).

$$\mathrm{PT}: \; \rho_{AB} = \sum_k p_k \, \rho_{Ak} \otimes \rho_{Bk} \rightarrow \rho'_{AB} = \sum_k p_k \, \rho_{Ak} \otimes \rho^{\mathrm{T}}_{Bk}.$$

Thus, partial transpose of a separable state is a bonafide density matrix. It follows that if a bipartite density matrix ceases to be a density matrix on partial transposition, the state should have been inseparable to begin with, showing that partial transposition is an entanglement witness.

If a bipartite density matrix ceases to be a density matrix on partial transposition, the state should have been inseparable to begin with, showing that partial transposition is an entanglement witness.

A bipartite system with $\dim \mathcal{H}_A = m$ and $\dim \mathcal{H}_B = n$ is often called an $m \times n$ system. We may use this notation to state an important result due to Peres and Horodecki:

Theorem: Positivity under partial transposition (PPT) is a necessary condition for separability of a bipartite state. This condition is both necessary and sufficient for all 2×2 and 2×3 states.

An Example: Werner States

As an illustration of the use of the partial transpose operation we consider the Werner state which is defined as a mixture of a maximally entangled state and the completely random state; the density matrix of the latter is a multiple of the unit matrix (in any O.N.B). In the particular case of a 2×2 system (a pair of qubits) the maximally random state has the density matrix $\frac{1}{4}\,\mathrm{Id}_{4\times4}$, and let us take $|\Phi_0\rangle$ as the maximally entangled state of interest. Then the Werner state is

$$\rho_W = x\,|\Phi_0\rangle\langle\Phi_0| + (1-x)\,\frac{1}{4}\,\mathrm{Id}_{4\times4}, \quad 0 \le x \le 1.$$

It is possible that, in an attempt to manufacture the state $|\Phi_0\rangle$, we ended up producing this Werner state due to deficiency of our process. In such a case we would like to have a measure which will tell us how close the manufactured state ρ is to the intended pure state $\rho_0 = |\Phi\rangle\langle\Phi|$. The most popular such measure is the *Fidelity* $F \equiv \mathrm{tr}\,(\rho_0\rho) = \langle\Phi|\rho|\Phi\rangle$. Since $\frac{1}{4}\,\mathrm{Id}_{4\times4}$ is an *equal* mixture of the four states $\{\,|\Phi_k\rangle\,\}$, we expect the fidelity of the Werner state to be $F = x+(1-x)/4 = (3x+1)/4$. The above Werner state is often written in terms of the fidelity:

$$\rho_W = \frac{4F-1}{3}|\Phi_0\rangle\langle\Phi_0| + \frac{1-F}{3}\mathrm{Id}_{4\times 4}, \quad 1/4 \leq F \leq 1.$$

In the standard product basis $|00\rangle$, $|01\rangle$, $|10\rangle$, $|11\rangle$, where $|jk\rangle \equiv |j\rangle \otimes |k\rangle$, the Werner state has the matrix representation

$$\rho_W =$$

$$\begin{pmatrix} (1+2F)/6 & 0 & 0 & (4F-1)/6 \\ 0 & (1-F)/3 & 0 & 0 \\ 0 & 0 & (1-F)/3 & 0 \\ (4F-1)/6 & 0 & 0 & (1+2F)/6 \end{pmatrix}$$

On partial transposition this matrix becomes

$$\rho'_W =$$

$$\begin{pmatrix} (1+2F)/6 & 0 & 0 & 0 \\ 0 & (1-F)/3 & (4F-1)/6 & 0 \\ 0 & (4F-1)/6 & (1-F)/3 & 0 \\ 0 & 0 & 0 & (1+2F)/6 \end{pmatrix}$$

It is clear that ρ'_W is positive semidefinite for all $F \leq 1/2$, but picks up a negative eigenvalue if $F > 1/2$. We may thus assert, in view of the above theorem, that the Werner state is entangled if and only if the fidelity $F > 1/2$. This conclusion remains unaltered if we replace $|\Phi_0\rangle$ by any other maximally entanged pure state of a pair of qubits.

Positive and Completely Positive Maps

Let $\mathcal{H}$ be a linear vector space (Hilbert space). Since any linear combination of linear transformations on a vector space is again a linear transformation, the set of all linear transformations (linear operators) on $\mathcal{H}$ is a vector space. If $\mathcal{H}$ is n-dimensional, then this new vector space is clearly n^2-dimensional. Linear transformations on this new vector space will be called linear maps (or simply maps) to distinguish them from linear operators. The phrase super operator is sometimes used in place of maps. Matrix transposition is an important example of linear maps.

Denoting by $\mathcal{T}_A$ the map of matrix transposition in the Hilbert space of system A, we see that $\mathcal{T}_A$ preserves the defining properties of a density matrix, namely hermiticity, positive semidefiniteness, and unit trace, and thus maps quantum states into valid quantum states. Maps with this property are called *positive maps.* Now $\mathcal{T}_A$ can be simply extended to a map on the tensor product space $\mathcal{H}_A \otimes \mathcal{H}_B$ of a bipartite system through the prescription $\mathcal{T}_A \otimes \mathrm{Id}_B$, which means matrix transposition acts only with regard to the first Hilbert space $\mathcal{H}_A$ leaving the second unaffected. This extension is what we have called partial transposition, or *local implementation* of the transposition map. We have seen earlier that matrix transposition does not remain positive under local implementation and, indeed, we used this fact as an entanglement witness.

A linear positive map which remains positive under all local implementations is said to be a *completely positive* map.

A linear positive map which remains positive under all local implementations is said to be a *completely positive* map. Matrix transposition is a distinguished example of a map which is positive but not completely positive. It is analogous to *time reversal* whose implementation on the entire universe is not unphysical, but local implementation is unphysical.

The point to remember is that *a map can be physically realised as a quantum process or evolution if and only if it acts as a trace preserving completely positive map on density operators.* For this reason physicists were interested in earlier days only in completely positive maps. This situation has now changed, for every positive but not completely positive map has use as an entanglement witness.

A map can be physically realised as a quantum process or evolution if and only if it acts as a trace preserving completely positive map on density operators.

Explicit characterization of the family of all completely positive maps turns out to be quite easy. The corresponding task for maps which are positive but not completely positive continues to remain an intractable problem. So also is the problem of finding a finite algorithm for testing if a given mixed state of a bipartite system (of Hilbert space dimensions m, $n \geq 3$) is separable or entangled. It is known, however, that cracking the former will amount to cracking the latter problem too.

This is a convenient point to summarize the shift in paradigm mixed states bring with them. State vectors, unitary Schrödinger evolutions, and von Nuemann projection measurements are the catch phrases for a closed quantum system. But for an open system interacting with the surroundings we are led, at the end of the day, to trace out the surroundings whose evolution we are not interested in. We may mimic the surroundings by attaching a suitable ancilla A to the system B of our interest. In these situations pure states of the total system will appear as mixed states of the subsystem of interest, the unitary evolutions of the total system as completely positive maps on the subsystem, and von Neumann measurements on the total system as POVM on the subsystem.

Quantum 'NOT' Gate is Unphysical

We have noted earlier that pure states of a qubit correspond, in a one-to-one manner, to points on (the surface of) the sphere S^2, and orthogonal states to antipodal points. Since mixed states are convex combinations of pure states, they correspond to points in the interior of the sphere. Indeed, the density matrix of any (pure or mixed) state of a qubit can be expressed in the form

$$\rho(\hat{\mathbf{n}}) = \frac{1}{2}\,(\sigma_0 + \hat{\mathbf{n}}\ \ \vec{\sigma}\,),$$

where $\hat{\mathbf{n}}$ is a three-dimensional real vector of norm ≤ 1.

Thus the state space of a qubit, with mixed states included, becomes the solid sphere. In particular, the centre of the sphere corresponds to the completely random (maximally mixed) state. Recall that unitary $[SU(2)]$ evolutions of the qubit act as $SO(3)$ rotations on this state space (This is a striking manifestation of the classic two-to-one connection between $SU(2)$ and $SO(3)$, familiar from the context of angular momentum in quantum theory).

The state space of a qubit, with mixed states included, becomes the solid sphere.

A classical NOT gate maps 0 to 1 and 1 to 0. It is of interest to ask: what should be called the corresponding quantum gate? We can define this quantum gate to be a linear operator which takes $|0\rangle$ to $|1\rangle$ and $|1\rangle$ to $|0\rangle$, where $|0\rangle$ and $|1\rangle$ may be taken as the eigenstates of σ_3, and represented by the north and south poles. Two features of such a definition could potentially offend the sentiments of the initiated. First, the above definition has nothing unique about it: not only σ_1 and σ_2, but also any linear combination $\cos\theta\,\sigma_1 + \sin\theta\,\sigma_2$, will qualify to be such a linear operator (since $|\psi\rangle$ and $e^{i\theta}\,|\psi\rangle$ are to represent one and the same state). It is not hard to see that all these linear operators are unitary, and correspond to $SO(3)$ rotations by π radians about an axis in the equatorial plane. Secondly, any such rotational implementation of NOT gate is highly basis-dependent. Suffice it to recall that any rotation leaves an orthogonal pair of states, lying on the axis of rotation, invariant (Euler's theorem). One will therefore be tempted to ask: Can we not do better?

It may seem that the following definition does not offend on either score: the gate takes every pure state of the qubit to the unique orthogonal state at the diametrically opposite end of S^2. Let us call this the quantum 'NOT' gate. This gate, if feasible, will act as inversion on the state space of a qubit: 'NOT': $\hat{\mathbf{n}} = (n_1, n_2, n_3) \rightarrow \hat{\mathbf{n}}' = (-n_1, -n_2, -n_3)$. But such an *improper rotation* turns out to be unphysical!

Like the part of energy that can be converted into useful work in thermodynamics, distillable entanglement is the one that can be readily put into information processing use.

To see this, note that the transposition operation $\mathcal{T}$: $(n_1, n_2, n_3) \rightarrow (n_1, -n_2, n_3)$ is an improper rotation and we know that it cannot be locally implemented. On the other hand, proper rotations correspond to unitary evolutions and hence can be locally implemented. That *no* improper rotation on the state space of a qubit can be locally implemented follows from the fact that every improper rotation can be written as a *fixed* improper rotation followed (or preceded) by a proper rotation. In other words, *every improper rotation of our state space is a positive but not completely positive map*, and hence cannot be realized as a physical process.

Bound Entanglement

Entanglement is the primary resource or currency for most quantum information processing tasks. Let us take the entanglement content of one of the maximally entangled states of a pair of qubits ($|\Phi_0\rangle$, for instance), as the unit of this currency (e-bit). Suppose we are given n identical copies of an entangled state ρ of a bipartite system AB, with the subsystem A of each copy in Alice's lab and the subsystem B of each copy in Bob's. Assume n is large. Starting with the n-fold tensor product $\rho \otimes \rho \otimes \quad \otimes \rho$, if Alice and Bob, working in concert but using only local operations and classical communication (LOCC), are able to manufacture k (but not more) copies of $|\Phi_0\rangle$, we call the asymptotic limit of k/n the *distillable entanglement* of the original state ρ. Like the part of energy that can be converted into useful work in thermodynamics, distillable entanglement is the one that can be readily put into information processing use.

It turns out that failure of ρ to remain positive semidefinite under partial transposition (PPT) is a necessary (but not sufficient) condition for ρ to possess distillable entanglement. In other words, entanglement of a state which is PPT is not distillable. Such states are said to have *bound entanglement.* The Peres–Horodecki theorem asserts that 2 × 2 and 2 × 3 systems have no bound entangled states. While examples of states with bound entanglement are known in systems of higher Hilbert space dimensions, complete characterization of the family of all such states remains an active research problem.

Quantum Computers

Quantum computation is an integral and important part of quantum information science. Quantum computers process quantum information encoded in quantum states in much the same way classical computers crunch classical information. The elementary storage gadgets are qubits in place of classical bits. Let $\mathcal{H}_2$ be the Hilbert space of a qubit, so that the 2^N-dimensional tensor product $\mathcal{H}_2^{\otimes N} = \mathcal{H}_2 \otimes \mathcal{H}_2 \otimes \qquad \otimes \mathcal{H}_2$ is the Hilbert space of a system of N qubits. With the initial condition of a given problem coded in the input state vector $|\Psi_I\rangle \in \mathcal{H}_2^{\otimes N}$, the solution gets coded in the output state vector $|\Psi_O\rangle \in \mathcal{H}_2^{\otimes N}$ of the quantum computer.

Quantum computers process quantum information encoded in quantum states in much the same way classical computers crunch classical information.

The task or evolution executed by the quantum computer corresponds to the unitary matrix (one of such unitary matrices) connecting these two states. The program or instruction set or quantum algorithm consists in reducing this unitary matrix into the product of a sequence of primitive unitary matrices or steps, each step acting as identity transformation on all but one or two of the qubits. A transformation that acts nontrivially only on a single qubit is called a 1-qubit gate, though the gate should be viewed as acting on N qubits. An important gate whose action is determined by an ordered pair of qubits is the controlled-NOT or XOR gate. The first member of the pair is called the *control* qubit and the second the *target* qubit. The gate does nothing if the control qubit is in the state $|0\rangle$, but if the control qubit is in the state $|1\rangle$ the gate interchanges the states $|0\rangle$ and $|1\rangle$ of the target qubit, leaving all the other qubits unaffected. That is, in the subspace of these two qubits

Measure of information in the quantum theory is defined parallel to Shannon information in the classical theory.

spanned by the standard basis $|00\rangle$, $|01\rangle$, $|10\rangle$, $|11\rangle$ the XOR or c-NOT gate acts through the unitary (permutation) matrix

$$U_{XOR} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}$$

It turns out that 1-qubit gates and the c-NOT gate together form a *universal set of primitives* for quantum computation in that any unitary matrix can be written as a word in the alphabet consisting of these primitives.

One has to make a measurement on the output or final state at the end of the quantum computation to extract the solution to the problem. And then, as in any quantum measurement process, one will get not a fixed value but a statistical distribution. In a good quantum algorithm, the value of interest occurs with sufficiently large probability so that a small number of runs will be sufficient to extract this value, at the desired level of confidence, as is the case with probabilistic algorithms of classical computation.

Quantum Measure of Information

Measure of information in the quantum theory is defined parallel to Shannon information in the classical theory. The source alphabet $\mathcal{A}$ in the quantum case consists of a set of pure states: $\mathcal{A} = (\rho_1, \rho_2, \quad \rho_N)$ (It is possible to allow these to be mixed states, but we restrict ourselves to the case of pure states). The source now emits every τ_s seconds one of the states from its finite alphabet $\mathcal{A}$ containing N states. Let p_1 be the probability for emitting the state ρ_1, p_2 the probability for emitting ρ_2, and so on. The source as an ensemble is thus described by

$$\rho = \sum_{k=1}^{N} p_k \rho_k.$$

The information of the source, per emission, is defined by the von Neumann entropy

$$S(\rho) = -\operatorname{tr}(\rho \log_2 \rho) \text{ bits.}$$

(Qubit is sometimes used in place of bit as the unit of von Neumann entropy.) The trace is easily evaluated in the eigenbasis of ρ. Thus, if λ_j are the nonzero eigenvalues of ρ, we have

$$S(\rho) = -\sum_j \lambda_j \log_2 \lambda_j \text{ bits} = H(\{\lambda_j\}) \text{ bits,}$$

bringing out the (limited) analogy with the Shannon entropy H: the von Neumann entropy of ρ equals the Shannon entropy of the eigenvalues of ρ. If the states constituting the source alphabet are orthogonal, then the eigenvalues are the same as the probabilities p_k, and in this case the von Neumann entropy coincides with the Shannon entropy of the probabilities, and this matches one's intuition that there is nothing much nonclassical if the states under consideration are mutually orthogonal. On the other hand, if these states are not orthogonal, the von Neumann entropy is strictly less than the Shannon entropy computed from the probabilities. This may be seen from a simple example corresponding to $N = 3$ and $p_1 = p_2 = p_3 = 1/3$, with the following three states of a qubit constituting the alphabet:

The von Neumann entropy of ρ equals the Shannon entropy of the eigenvalues of ρ.

$$\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad \frac{1}{4}\begin{pmatrix} 1 & \sqrt{3} \\ \sqrt{3} & 3 \end{pmatrix} \quad \frac{1}{4}\begin{pmatrix} 1 & -\sqrt{3} \\ -\sqrt{3} & 3 \end{pmatrix}$$

We see that $\rho = \frac{1}{2}\sigma_0$, so that $S = 1$ bit. But the Shannon entropy H calculated from the probabilities turns out to be $\log_2 3$ bits.

In the general case where the source alphabet comprises states from an m-dimensional Hilbert space, the von Neumann entropy attains its maximum possible value when ρ is a multiple of the identity. The maximum value of S is thus $\log_2 m$ bits. But $H(\{p_k\})$ attains its maximum value when all the probabilities coincide, and so equal $1/N$. Thus the maximum value of $H(\{p_k\})$ is $\log_2 N$ bits. And, for a given ρ, the value of N can be made arbitrarily large, and so also the value of $H(\{p_k\})$.

Quantum Version of Shannon's Theorems?

In view of the remarkable success of classical information theory, one of the important objectives in quantum information theory has been to determine and exploit the quantum analogue of every important idea and result in the former. This applies, in particular, to the two fundamental theorems of Shannon.

In view of the remarkable success of classical information theory, one of the important objectives in quantum information theory has been to determine and exploit the quantum analogue of every important idea and result in the former.

A satisfactory quantum version of the source coding theorem dealing with data compression has been achieved through the work of Benjamin Schumacher. The role played by typical sequences in the Shannon theory is now played by typical subspaces. It is beyond the scope of the present exposition to try a self-contained description of this important work. Roughly stated Schumacher's result shows that, for large n, a message consisting of n successive emissions by the source can be compressed to a Hilbert space of dimension $2^{n\,S(\rho)}$. But this space can be viewed as the Hilbert space of $n\,S(\rho)$ qubits. Thus a message of length n is compressed into $n\,S(p)$ qubits, showing that the source gives out, on the average, $S(\rho)$ qubits of information per emission.

As for the channel coding theorem, we have earlier motivated the fact that every physical process corresponds to a trace preserving completely positive map on the state space of density operators. And a quantum channel is a physical process. Thus there is a one-to-one correspondence between quantum channels and trace preserving completely positive maps. For a given (noisy) quantum channel $\mathcal{N}$, one may consider not just one but three distinct capacities: the classical capacity $C(\mathcal{N})$ for sending classical information over the quantum channel;

the quantum capacity $Q(\mathcal{N})$ for sending quantum information over the quantum channel; and the classically assisted quantum capacity $Q_2(\mathcal{N})$ for sending quantum information over the channel, supported by side channel classical communications. It is easy to see that $Q(\mathcal{N}) \leq C(\mathcal{N})$ and $Q(\mathcal{N}) \leq Q_2(\mathcal{N})$, but it is not known if $Q_2(\mathcal{N}) \leq C(\mathcal{N})$.

We do not yet have a general formula for the quantum capacity of a quantum channel.

Quantum version of the channel coding theorem continues to remain a very important open problem. Fundamental advances have been made in this direction, but we do not yet have a general formula for the quantum capacity of a quantum channel.

Measure of Entanglement

Since entanglement is an important resource in most information processing tasks, a quantitative measure for it is desirable. For a bipartite pure state $|\Psi\rangle$ resulting in reduced density matrices ρ_A and ρ_B for the subsystems, this measure is defined as $S(\rho_A)$ e-bits $=$ $S(\rho_B)$ e-bits, i.e., as the von Neumann entropy of either reduced density matrix. We see that the entanglement measure is a local invariant, for it is clear from the above definition that this measure equals the Shannon entropy of the locally invariant parameters λ_k appearing in the Schmidt normal form of the state. Turning to the particular case of a pair of qubits, since $\lambda_1 = \lambda_2 = 1/2$ for each of the four maximally entangled states $|\Phi_k\rangle$ used in the teleportation protocol, we conclude that each one of these states has 1 e-bit of entanglement.

Indeed, the content of this observation is true more generally. Recall that the Schmidt rank r of a pure state of an $m \otimes n$ system cannot exceed $\nu \equiv \min(m, n)$. It follows that the Shannon entropy of the λ-parameters cannot exceed $\log_2 \nu$, the maximum being realized if and only if the Schmidt rank is maximal, $r = \nu$, and all the λ-parameters are equal (and therefore assume the value $1/\nu$). States which saturate this limit are said to be maximally entangled. It may be verified that the four states $|\Phi_k\rangle$ used in the teleportation and super dense coding tasks are indeed maximally entangled, as asserted earlier.

Putting together a large scale quantum computer is one of the principal dreams behind the efforts in QIS.

In the case of mixed states there are several definitions for the measure of entanglement. For pure states all these definitions coincide numerically with the definition considered above in some detail, but are provably different for mixed states. One of them is the *entanglement of formation* (EoF), denoted $E_f(\rho)$. We write the given bipartite density matrix as an ensemble of pure states $\{\,|\Phi_k\rangle,\, p_k\,\}$:

$$\rho = \sum_k p_k\,|\Phi_k\rangle\langle\Phi_k|.$$

It is useful to remind ourselves yet again that the $|\Phi_k\rangle$, need not be orthogonal or linearly independent. Let us denote by $E(|\Phi\rangle)$ the entanglement measure of pure state $|\Phi\rangle$ computed as the von Neumann entropy of the reduced state. The entanglement of formation is defined as

$$E_f(\rho) = \min\left\{\sum_k p_k\,E(|\Phi_k\rangle)\right\}$$

where the minimum is to be evaluated over all possible ensemble realizations of ρ (recall that these realizations form a huge family). It is worth mentioning that Wootters has succeeded in finding a closed form expression for this minimum for an arbitrary mixed state of a pair of qubits. Generalization of this result of Wootters to systems of larger Hilbert space dimensions remains a problem for the future.

Final Remarks

Putting together a large scale quantum computer is one of the principal dreams behind the efforts in QIS. Several candidate systems for quantum computation are being studied intensively. Elementary gates of quantum computation have already been demonstrated. Indeed, in the past weeks a quantum computer has successfully factorized the integer 15. This is a technological milestone! Effective quantum error-correcting codes have been developed to handle the effect of decoherence.

The major hurdle being faced is one of scaling: while the technology works for a few qubits, it does not seem to extrapolate easily to a few hundred qubits. So it may be a while before quantum computers, with large enough power to make the demise of RSA imminent, will be in the market. In the meantime, research in this area is already paying rich dividends: in terms of sophisticated experimental techniques and in terms of the way we understand many fundamental aspects of quantum theory. These developments are bound to reorient the way we teach quantum theory.

It may be a while before quantum computers, with large enough power to make the demise of RSA imminent, will be in the market.

Suggested Reading

[1] Most of the papers dealing with QIS are to be found with the LANL e-Print Archive (quantum physics), as also with its mirror: http:\\ xxx.imsc.ernet.in

[2] Shannon's 1948 work, A Mathematical Theory of Communication, is now available electronically: http:\\cm.bell-labs.com/cm/ms/what/ shannonday/paper.html

[3] Two readable reviews are: C H Bennett and P Shor, *Quantum Information Theory*, and P Shor, *Quantum Information Theory: Results and Open Problems*. Both can be found at http:\\www.research.att.com/~shor/ papers/index.html

[4] John Preskill has a valuable set of extensive lecture notes on Quantum Information and Computation: http:\\www.theory.caltech.edu/people/ preskill/ph219

[5] I have not touched upon the important topic of quantum error correcting codes. This topic has been eloquently presented in: K R Parthasarathy, The Mathematics of Quantum Error Correcting Codes, *Resonance*, Vol.6, No.3, p.34, No.4, p.38, 2001.

[6] Most aspects of QIS are covered in the carefully written book by M A Nielsen and I L Chuang, *Quantum Computation and Quantum Information*, Cambridge Univ. Press, 2000.

Address for Correspondence
Rajiah Simon
The Institute of Mathematical Sciences,
Chennai 600 113, India.
Email: simon@imsc.ernet.in