The Mathematics of Error Correcting Quantum Codes

1. Quantum Probability

K R Parthasarathy

K R Parthasarathy is INSA C V Raman Research Professor at Indian Statistical Institute, Delhi. His interests are quantum probability, mathematical foundations of quantum mechanics and probability theory. He is the author of two classic books in probability theory and one in quantum stochastic calculus. He is a Fellow of all the three academies of sciences in India and is also a Fellow of TWAS (Third World Academy of Sciences). For his contributions to probability theory he was awarded the Bhatnagar Prize in 1976. For his contributions to quantum stochastic analysis he was awarded the TWAS Prize in 1996 and was conferred an honorary doctorate by the Nottingham Trent University, U K in July, 2000.

1. Introduction

The mathematical theory of communication of messages through a quantum information channel is based on the following three basic principles:

(i) Messages can be encoded as states of a quantum system with a finite number of levels. As an example one can think of a single particle spin system with two levels called *spin up* and *spin down* which can be labelled 0 and 1, respectively. A bunch of n such 'independent' systems can be viewed as a single quantum system with 2^n levels, each level labelled by a word of length n from the binary alphabet $\{0, 1\}$.

(ii) Encoded states can be viewed as inputs of a quantum channel and transmitted. However, at the receiving end of the channel the output state can differ from the input state owing to the presence of 'noise' in the channel.

(iii) There is a collection of 'good' states at the input which when transmitted through the channel lead to output states from which the input states can be reconstructed without any error or possibly with a small error.

The good states obeying property (iii) can then be used to encode messages for transmission through the channel. Thus any reasonable theory of error correcting quantum codes should include a proper identification of

the good states and also an algorithm for (almost) error-free reconstruction of the input state from a knowledge of the output state. To formulate such a theory it is essential to understand the notion of a state in quantum theory. In order to make this presentation reasonably self-contained we give here a lightning introduction to quantum probability.

2. Quantum Probability

An elementary quantum system with n levels is described through an n-dimensional complex vector space of all column vectors of the form

$$|u\rangle = \begin{pmatrix} z_1 \\ z_2 \\ \vdots \\ z_n \end{pmatrix}, \; z_j \in C \; \forall \; j. \qquad (2.1)$$

Such a column vector is called a *ket* vector. To any such ket vector one can associate a *bra* vector

$$\langle u| = (\bar{z}_1, \bar{z}_2, \quad , \bar{z}_n)$$

which is a row vector, the bar denoting complex conjugate. If $|u\rangle, |v\rangle$ are two ket vectors then the matrix product

$$\langle u||v\rangle$$

is a scalar called the *scalar product* between $|u\rangle$ and $|v\rangle$ and denoted by $\langle u|v\rangle$. The space C^n of all such ket vectors together with this scalar product is called an n-dimensional *Hilbert space* and denoted by the symbol $\mathcal{H}$. Any ket vector $|u\rangle$ can also be viewed as a function u on the set $\{1, 2, \quad , n\}$ with $u(i) = z_i, i = 1, 2, \quad , n$. If $v(i) = t_i \; \forall \; i$ then $\langle u|v\rangle = \sum_{i=1}^{n} \overline{u(i)} v(i) = \sum_{i=1}^{n} \bar{z}_i t_i$. Sometimes it will be convenient to view the index set $\{1, 2, \quad , n\}$ as an abstract set A of n elements, called an *alphabet*. If T is an $n \times n$ matrix over C then it defines a linear transformation or linear operator on the

Hilbert space by $|u\rangle \longrightarrow T|u\rangle$. Note that $\langle u|T|v\rangle$ is a scalar. The space $M_n(C)$ of all $n \times n$ matrices over C is an algebra equipped with an involution $T \longrightarrow T^\dagger$ where the ij-th element of $T^\dagger$ is the complex conjugate of the ji-th element of T T is said to be *hermitian* if $T = T^\dagger$. One says that T is *nonnegative definite* or $T \geq 0$ in symbols if $\langle u|T|u\rangle \geq 0$ for every $|u\rangle$. If $T = T^2 = T^\dagger$ then T is called a *projection* (onto its range). If T is a projection so is $I - T$, I being the identity matrix. The sum of all the diagonal elements of T is called its *trace* and denoted by Tr T

A nonnegative definite matrix ρ of unit trace is called a *state*. A hermitian matrix T is called an *observable* and for any state ρ, the scalar quantity Tr $\rho T (=$ Tr $T\rho)$ is called the *expectation* of the observable T in the state ρ. Even though T may not be hermitian we still say that Tr ρT is the *expectation* of T in the state ρ. It is important to note that the map

$$T \longrightarrow \text{Tr } \rho T$$

from the space $M_n(C)$ into C has all the features of an averaging procedure:

(i) $\text{Tr}\rho(aT_1 + bT_2) = a\text{Tr}\rho T_1 + b\text{Tr}\rho T_2, a, b \in C, T_1, T_2 \in M_n(C)$;

(ii) If $T \geq 0$ then Tr $\rho T \geq 0$;

(iii) $\text{Tr}\rho I = \text{Tr}\rho = 1$.

It is good to take a pause and compare these three properties with what one is familiar with in classical probability: Consider the algebra $\mathcal{A}_n$ of all complex valued random variables on the probability space $\{1, 2, \quad, n\}$ equipped with the probability distribution $p_1, p_2, \quad, p_n$, p_i being the probability of the elementary outcome i. The map

$$f \longrightarrow E\ f = \sum_{i=1}^{n} f(i)p_i,$$

from $\mathcal{A}_n$ into scalars has the properties

(i') $E\ af_1 + bf_2 = aEf_1 + bEf_2,\ a, b \in C,\ f_1, f_2 \in \mathcal{A}_n,$

(ii') If $f(i) \geq 0\ \forall\ i$ then $E\ f \geq 0$

(iii') $E\ 1 = 1$ where 1 also denotes the random variable identically equal to unity.

The central difference between the algebra $\mathcal{A}_n$ of random variables and the algebra $M_n(C)$ of matrices is that multiplication in $\mathcal{A}_n$ is commutative whereas multiplication in $M_n(C)$ is not. This is the reason why quantum probability is also called noncommutative probability.

The most fundamental theorem concerning hermitian matrices (or observables) is the spectral theorem. According to this theorem every hermitian matrix T has the form

$$T = \sum_{i=1}^{k} \lambda_i E_i, \tag{2.2}$$

where $\lambda_1, \lambda_2, \quad , \lambda_k$ are distinct real scalars and $E_1, E_2, \quad , E_k$ are projections satisfying $\sum_{i=1}^{k} E_i = I,\ E_i E_j = 0$ if $i \neq j$. The set $\{\lambda_1, \lambda_2, \quad , \lambda_k\}$ is called the *spectrum* of T, the elements λ_i are the *eigenvalues* of T and (2.2) is called the *spectral resolution* of T We interpret (2.2) as follows: the observable T assumes values $\lambda_1, \lambda_2, \quad , \lambda_k$ and the 'event' that T assumes the value λ_i is the projection E_i. When the spectral theorem is applied to a state ρ and one also takes into account the fact that every projection E can be expressed as

$$E = \sum_{i=1}^{d} |u_i\rangle\langle u_i|,$$

where $\{|u_i\rangle,\ i = 1, 2, \quad , d\}$ is any orthonormal basis for the subspace $\{|u\rangle \mid E|u\rangle = |u\rangle\}$ of all ket vectors fixed by E, it follows that ρ can be expressed as

$$\rho = \sum_{i=1}^{n} p_i |v_i\rangle\langle v_i|, \tag{2.3}$$

where $(p_1, p_2, \ldots, p_n)$ is a probability distribution on the set $\{1, 2, \ldots, n\}$ and $\{|v_i\rangle,\ i = 1, 2, \ldots, n\}$ is an orthonormal basis for the Hilbert space C^n, i.e., $\langle v_i|v_j\rangle = \delta_{ij}$ for all $i, j = 1, 2, \ldots, n$ (where $\delta_{ij} = 1$ if $i = j$ and is 0 if $i \neq j$). Equations (2.2) and (2.3) lead to the following statistical interpretation. In the state ρ the probability that the observable T assumes the value λ_i is equal to Tr $\rho E_i\ \forall\ i = 1, 2, \ldots k$ and the expectation of T is equal to $\sum_{i=1}^{k} \lambda_i$ Tr $\rho E_i =$ Tr $\rho \sum_{i=1}^{k} \lambda_i E_i =$ Tr ρT which links the classical definition of expectation and the quantum theoretic definition.

If $|u\rangle$ is any unit vector in C^n then $|u\rangle\langle u|$ is a projection whose range is the one dimensional subspace or ray $C|u\rangle$. Such a projection is also a state. According to (2.3) every state ρ can be expressed as a weighted linear combination of states which are the one dimensional projections $|v_i\rangle\langle v_i|$, where the weights p_i constitute a probability distribution. One says that ρ is a *convex combination* or a *mixture* of *pure states* $|v_i\rangle\langle v_i|$. A state of the form $|v\rangle\langle v|$ is called *pure* because if we split $|v\rangle\langle v|$ as $|v\rangle\langle v| = p\rho_1 + (1-p)\rho_2$ where $0 < p < 1$ and ρ_1 and ρ_2 are states then $\rho_1 = \rho_2 = |v\rangle\langle v|$. In other words a pure state cannot be split into a mixture of two distinct states. We say that the set of all states is a convex set whose extreme points are precisely one dimensional projections. When a pure state has the form $|v\rangle\langle v|$ for some unit vector $|v\rangle$ it is customary to call the unit vector $|v\rangle$ (or more precisely the unit ray $\{\lambda|v\rangle,\ |\lambda| = 1\}$) itself as the pure state. What is actually meant is the projection operator $|v\rangle\langle v|$.

The next fundamental notion from quantum probability that we need is the combination of several quantum systems into a single system. Suppose $\mathcal{H}_j = C^{n_j}, j = 1, 2, \ldots k$ are the Hilbert spaces describing quantum systems numbered $1, 2, \ldots, k$, respectively. We wish to describe all of them together as a single system. To

this end we introduce the tensor 'product' of the Hilbert spaces $\mathcal{H}_1, \mathcal{H}_2, \quad \mathcal{H}_k$. If $|u_j\rangle \in \mathcal{H}_j$ is given by

$$|u_j\rangle = \begin{pmatrix} z_{j1} \\ z_{j2} \\ \vdots \\ z_{jn_j} \end{pmatrix} \quad 1 \leq j \leq k, \; z_{jl} \in C \qquad (2.4)$$

define their *tensor product* $|u_1\rangle \otimes |u_2\rangle \otimes \cdot \cdot \otimes |u_k\rangle$ (which is also denoted $|u_1\rangle \, |u_2\rangle \quad |u_k\rangle$) to be the column vector

$$|u_1, u_2, \quad , u_k\rangle = \begin{pmatrix} \vdots \\ z_{\boldsymbol{i}} \\ \vdots \end{pmatrix} \quad \boldsymbol{i} = i_1 i_2 \quad \cdot i_k, \qquad (2.5)$$

where

$$z_{\boldsymbol{i}} = z_{1i_1} z_{2i_2} \quad \cdot z_{ki_k}, \; 1 \leq i_r \leq n_r, r = 1, 2, \quad k \quad (2.6)$$

and the multiindex $\boldsymbol{i}$ runs through in the lexicographic ordering as in a dictionary. For example, when $k = 2, n_1 = 2, n_2 = 3$ the lexicographic ordering for the double index $i_1 i_2$ is $11, 12, 13, 21, 22, 23$ so that

$$|u_1, u_2\rangle = \begin{pmatrix} z_{11} z_{21} \\ z_{11} z_{22} \\ z_{11} z_{23} \\ z_{12} z_{21} \\ z_{12} z_{22} \\ z_{12} z_{23} \end{pmatrix}$$

Similarly, $\langle u_1, u_2, \quad , u_k| = \langle u_1| \langle u_2| \quad \langle u_k| = \langle u_1| \otimes \langle u_2| \otimes \quad \otimes \langle u_k| = (\cdot \quad , \bar{z}_{\boldsymbol{i}}, \quad \cdot)$. The scalar product between two product vectors $|u_1, u_2, \quad , u_k\rangle$ and $|v_1, v_2,$ $v_k\rangle$ is equal to $\prod_{i=1}^{k} \langle u_i | v_i \rangle$. All product vectors of the form (2.5) span the Hilbert space $C^{n_1 n_2 \cdots n_k}$ and we denote it by $\mathcal{H}_1 \otimes \mathcal{H}_2 \otimes \quad \otimes \mathcal{H}_k$. If $\{|u_{i1}\rangle, \; |u_{i2}\rangle, \quad |u_{in_i}\rangle\}$ is an orthonormal basis for $\mathcal{H}_i$ then the collection

$$\{|u_{1j_1}, u_{2j_2}, \quad , u_{kj_k}\rangle, \; 1 \leq j_i \leq n_i, i = 1, 2, \quad k\}$$

is an orthonormal basis for $\mathcal{H}_1 \otimes \mathcal{H}_2 \otimes \quad \otimes \mathcal{H}_k$.

If A_i is a matrix of order $n_i \times n_i$ and each A_i is viewed as a linear transformation or an operator in $\mathcal{H}_i$ then one defines the product linear operator $A_1 \otimes A_2 \otimes \quad \otimes A_k$ in $\mathcal{H}_1 \otimes \mathcal{H}_2 \otimes \quad \otimes \mathcal{H}_k$ by

$$A_1 \otimes A_2 \otimes \cdot \cdot \otimes A_k \, |u_1, u_2, \quad , u_k\rangle = |A_1 u_1, A_2 u_2, \quad , A_k u_k\rangle$$

for all product vectors and extending it linearly to all linear combinations of such product vectors. Such a product operator (or matrix) is well defined and one has

$$(A_1 \otimes A_2 \otimes \cdot \cdot \otimes A_k)(B_1 \otimes B_2 \otimes \cdot \cdot \otimes B_k) = A_1 B_1 \otimes A_2 B_2 \otimes$$

$$\otimes A_k B_k, (A_1 \otimes A_2 \otimes \quad \otimes A_k)^\dagger = A_1^\dagger \otimes A_2^\dagger \otimes \quad \cdot \otimes A_k^\dagger,$$

$$A_1 \otimes A_2 \otimes \quad \otimes A_{i-1} \otimes (\alpha A_i + \beta B_i) \otimes A_{i+1} \otimes \quad \otimes A_k$$

$$= \alpha A_1 \otimes A_2 \otimes \cdot \cdot \otimes A_k + \beta A_1 \otimes A_2 \otimes \cdot \cdot \otimes A_{i-1} \otimes B_i \otimes A_{i+1} \otimes$$

$$\otimes A_k.$$

In particular, if each A_i is hermitian so is their product $A_1 \otimes A_2 \otimes \quad \otimes A_k$. Similarly if each A_i is unitary so is their product. It is a simple exercise to check that

$$\text{Tr } A_1 \otimes \quad \otimes A_k = \Pi_{i=1}^k \text{Tr } A_i.$$

If ρ_i is a state in $\mathcal{H}_i \; \forall i$ then $\rho_1 \otimes \rho_2 \otimes \quad \otimes \rho_k$ is a state in $\mathcal{H}_1 \otimes \mathcal{H}_2 \otimes \quad \otimes \mathcal{H}_k$ with the property that for any product observable $A_1 \otimes \quad \otimes A_k$ its expectation in the product state $\rho_1 \otimes \rho_2 \otimes \quad \otimes \rho_k$ is equal to

$$\begin{aligned} &\text{Tr } (\rho_1 \otimes \rho_2 \otimes \quad \otimes \rho_k)(A_1 \otimes A_2 \otimes \quad \otimes A_k) \\ &= \text{Tr } \rho_1 A_1 \otimes \rho_2 A_2 \otimes \quad \otimes \rho_k A_k \\ &= \Pi_{i=1}^k \text{Tr } \rho_i A_i, \end{aligned}$$

which is the product of the expectation of A_i in the state ρ_i as i varies from 1 to k. Note that a mixture of two distinct product states is not a product state.

We now come down to the special case of the Hilbert space $\mathcal{H} = C^2$ of dimension 2. Write

$$|0\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, |1\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}. \tag{2.7}$$

The ket vectors labelled $|0\rangle$ and $|1\rangle$ constitute an orthonormal basis for C^2 which is viewed as the Hilbert space of a quantum system with two levels denoted 0 and 1. In physical language 0 may denote spin up and 1 spin down for a two level spin system. If $a_1, a_2, \quad , a_k$ is a binary sequence, i.e., each a_i is either 0 or 1 we write

$$|a_1 a_2 \quad \cdot a_k\rangle = |a_1\rangle\ |a_2\rangle \quad |a_k\rangle = |a_1\rangle \otimes |a_2\rangle \otimes \quad \otimes |a_k\rangle. \tag{2.8}$$

As we run through all the words $a_1 a_2 \quad \cdot a_k$ of length k from the binary alphabet $\{0, 1\}$ we see that (2.8) runs through an orthonormal basis for $\mathcal{H} \otimes \cdot \ \cdot \otimes \mathcal{H} = \mathcal{H}^{\otimes^k}$ the tensor product being k-fold. As already mentioned unit vectors can be identified with pure states. Thus we have *encoded* the set of all binary words of length k as a set of pure states in a quantum system which is a product of k elementary systems each of which is described by $h = C^2$. A state in C^2 is called a *one qubit* state, where qubit is an abbreviation for a quantum binary digit. A state in $\mathcal{H}^{\otimes^k}$ is called a k-qubit state. Thus the pure state $|a_1 a_2 \quad a_k\rangle$ is a k-qubit state of the product type. Now consider a ket vector of the form

$$|u\rangle = \sum_{a_1, a_2, \cdots, a_k} \alpha_{a_1 a_2 \cdots a_k} |a_1 a_2 \quad \cdot a_k\rangle \tag{2.9}$$

where

$$\sum_{a_1, a_2, \cdots, a_k} |\alpha_{a_1 a_2 \cdots a_k}|^2 = 1 \tag{2.10}$$

and $a_1, a_2, \quad , a_k$ vary over $\{0, 1\}$. Then $|u\rangle$ defines a pure state which is not a product state. We say that the state $|u\rangle$ is *entangled*. One of the interesting questions of our subject is to define an appropriate measure of this *entanglement*.

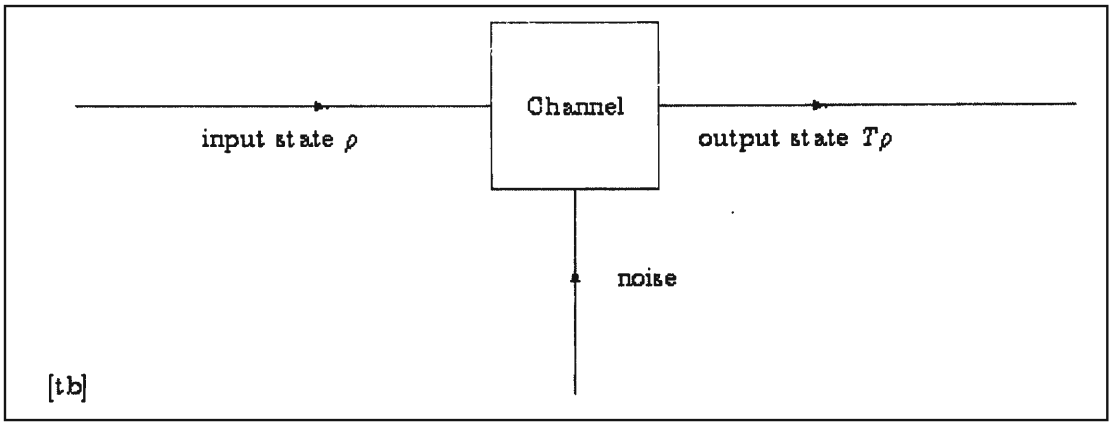


Figure 1.

More generally, one can speak of an *alphabet* A which is a finite set containing, say, N elements. Consider now the Hilbert space C^N with an orthonormal basis $\{|a\rangle, a \in N\}$ of ket vectors in C^N Then $(C^N)^{\otimes k}$ is the k-fold tensor product of copies of C^N which has the orthonormal basis

$$|a_1 a_2 \cdot \cdot a_k\rangle = |a_1\rangle|a_2\rangle \qquad |a_k\rangle = |a_1\rangle \otimes |a_2\rangle \otimes \qquad \otimes |a_k\rangle,$$

$a_1, a_2, \quad \cdot a_k$ varying in the alphabet A. Thus words of length k from the alphabet A are encoded as pure states from an orthonormal basis of $\mathcal{H}^{\otimes k}$

3. Quantum Channels with Noise

A quantum channel can be viewed as a box which, for each input state of a quantum system, produces an output state of another quantum system. See *Figure* 1.

Mathematically speaking, for each input state ρ on a Hilbert space $\mathcal{H}$ one has an output state $T\rho$ on probably some other Hilbert space $\mathcal{H}'$. For simplicity we assume that $\mathcal{H} = \mathcal{H}'$. Thus the channel effects a transformation T on the space of all states of a quantum system. Each time a 'signal' in the form of a state ρ is fed into the channel it is transformed into an output state $T\rho$ but at different times the transformations T may differ! The nature of channel noise is assumed to be such that T

belongs to a well-defined class of transformations. In general, the transformation T may be nonlinear. Since our subject is in a state of infancy (just five years old!) we assume that the transformation T is always 'affine linear', i.e.,

$$T(p\rho_1 + q\rho_2) = pT\rho_1 + qT\rho_2 \tag{3.11}$$

for any two sates ρ_1, ρ_2 on $\mathcal{H}$ and p, q are nonnegative scalars satisfying $p + q = 1$. An example of such a transformation T is given by

$$T\rho = U\rho U^\dagger, \tag{3.12}$$

where U is a unitary operator. This is an example of a reversible transformation in the sense that T has an inverse given by $T^{-1}\rho = U^\dagger \rho U$. Such a T transforms pure states into pure states. Physically speaking, the state ρ undergoes a 'Schrödinger dynamics' for one unit of time with $U = e^{-iH}$, H being a selfadjoint matrix. Suppose there is a bunch of unitary operators $U_1, U_2, \quad , U_k$ and one of them is chosen at random with probabilities $p_1, p_2, \quad , p_k$ respectively and applied to a state ρ. Then one can say that the output state has the structure

$$T\rho = \sum_{j=1}^{k} p_j U_j \rho U_j^\dagger.$$

Such a T need not be reversible. It can transform a pure state into a mixed state. More generally, one can consider transformations of the form

$$T\rho = \sum_{j=1}^{k} L_j \rho L_j^\dagger, \tag{3.13}$$

where

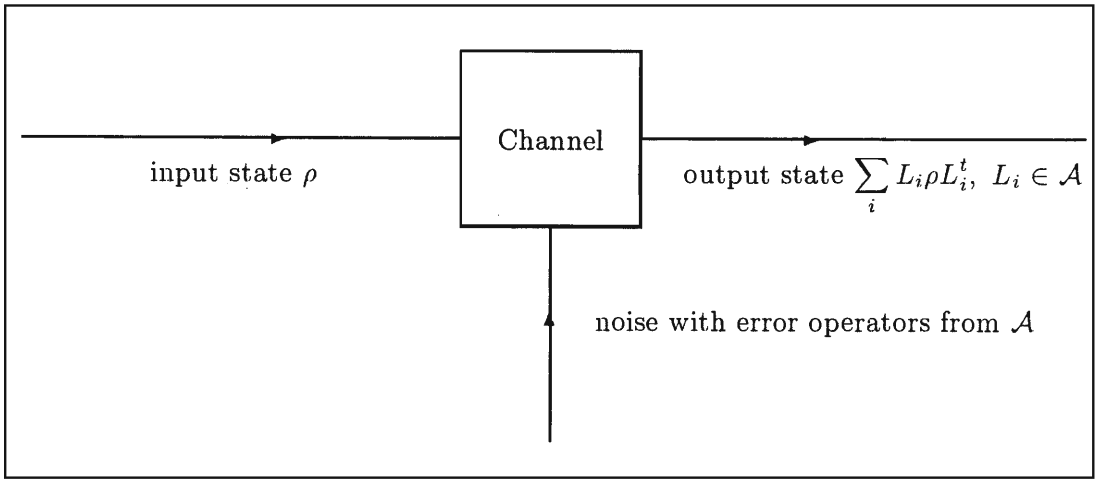
$$\sum_{j=1}^{k} L_j^\dagger L_j = I. \tag{3.14}$$

Since ρ is positive semidefinite each $L_j \rho L_j^\dagger$ is positive semidefinite and so is their sum. Furthermore,

$$\begin{aligned} \text{Tr } T\rho &= \sum_{j=1}^{k} \text{Tr } L_j \rho L_j^\dagger \\ &= \text{Tr } (\sum_{j=1}^{k} L_j^\dagger L_j)\rho \\ &= \text{Tr } \rho \\ &= 1. \end{aligned}$$

Thus $T\rho$ is again a state. Transformations of the form (3.13) with the restriction (3.14) occur extensively in the physical literature and are known as *completely positive maps*. Apparently, one can build (or hope to build) devices which implement transformations of this kind. The matrices L_j in (3.13) are said to *corrupt* the input state ρ and are therefore called *error operators*. The noise in the channel is specified by demarcating a class $\mathcal{A}$ of matrices operating as linear operators in the Hilbert space $\mathcal{H}$ of the quantum system. It is usually assumed that $\mathcal{A}$ is also a linear space, called the space of *error operators* affecting the channel. We now state the *basic hypothesis* concerning the operation of the channel in *Figure 1*. See *Figure* 2.

Figure 2.



Note that the input state ρ can be pure but the output state can be mixed! For example, if $\rho = |u\rangle\langle u|$ the output state $\sum_j L_j|u\rangle\langle u|L_j^\dagger$ is a sum of rank one positive semidefinite operators which need not be a projection.

Suggested Reading

[1] K R Parthasarathy, *An Introduction to Quantum Stochastic Calculus*, Birkhauser Verlag, Basel, 1992.

[2] P A Meyer, Quantum Probability for Probabilists, *Lecture Notes in Mathematics*, Vol. No. 1538, 2nd edition, Springer Verlag, Berlin, 1995.

[3] G S Vijay and Vishal Gupta, Quantum Computing, Part-I, *Resonance*, Vol. 5, No. 9, 69-81, 2000; Part-II, *Resonance*, Vol. 5, No. 10, 66-72, 2000.

[4] A O Pittenger, *An Introduction to Quantum Computing Algorithms* (Progress in Computer Science and Applied Logic, Vol. 19) Birkhauser Verlag, Basel, 1999.

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

Address for Correspondence
K R Parthasarathy
Indian Statistical Institute
7, S J S Sansanwal Marg
New Delhi 110 016, India.
Email: krp@isid.ac.in