An Introduction to Quantum Information Theory

Sibasish Ghosh

The Institute of Mathematical Sciences
C. I. T. Campus, Taramani
Chennai - 600 113
E-mail: sibasish@imsc.res.in

May 27, 2017

What is QIT (also known as QIS)?

- Quantum Information Science (QIS) is the study of storage, transmission, and processing of information using quantum mechanical devices as well as operations allowed by quantum mechanics.

- Thus QIS should reproduce the results of Classical Information Science (CIS) but the former can have much more aspects.

- QIS should have, in general, better efficiency compared to CIS.

What is Classical Information Science (CIS)?

- When we take a look at some huge data (*e.g.*, the whole content of some book), we usually get puzzled towards figuring out which part of this data carries *useful* information and which part does not.
- Which information is *useful*?
- The statement: "Sun rises on the East." contains no useful information!
- The statement: "It will rain tonight here in Chennai." does contain some useful information.
- CIS is all about this useful information in the classical physical setup.

What is CIS? (continued ...)

- Given any data, how to find out the useful information out of it?
- In other words, how to quantify the amount of information necessary for the storage, transmission, and processing of it?
- Contentwise, the following two statements carry same amount of information:
  (i) Two-third of Earth's surface is covered by water.
  (ii) One-third of the entire population of country lies below poverty line.
- Thus information content is all about the probabilities of occurances of events pertaining to the data.

What is CIS? (continued ...)

- If a random variable X takes the value i with probability p_i then the information content corresponding to this i-th value should be a monotonically increasing function of $1/p_i$.
- This function is taken as $\log_2(1/p_i)$ – from the consideration of continuity and additivity of information content.
- Thus the average information content is: $H(X) \equiv -\sum_i p_i \log_2 p_i$.
- $H(X)$ is proportional to the (thermodynamic) entropy of a thermodynamic system with its i-th microstate being associated with probability p_i.
- Although the information theoretic entropy looks (mathematically) similar to the thermodynamic entropy, but perspective wise they are different.

Outline

- Storage of classical information: Shannon's source coding theorem
- Transmission of classical information: Shannon's noisy channel coding theorem
- Brief introduction to QM
- Schumacher's data compression limit: von Neumann entropy
- Entanglement – a fundamental resource for quantum information processing
- Holevo's bound on accessible information
- Different capacities of quantum channels
- Applications of QIT
- Present scenario

Storage of classical information: Shannon's source coding theorem

Storage of classical information

- Two different objects *a* and *b* – appearing with equal probability – can be reliably encoded separately by the two values 0 and 1 of a single bit.

- Four different objects *a*, *b*, *c*, and *d* – appearing with probabilities 3/8, 1/8, 1/8, and 3/8 respectively – can be reliably encoded separately by the bit strings 0, 000, 111, and 1 respectively, with average bit string length being $[1 \times (3/8) + 3 \times (1/8) + 3 \times (1/8) + 1 \times (3/8)] = 3/2$ (selective encoding).

- Encoding – in the last example – by the two bit strings 00, 01, 10, and 11, will give rise to the average bit string length 2 (blind encoding), which is not as good as the last example!

Storage of classical information (continued ...)

- What will be average bit string length for encoding all the values of a sequence of n independent and identically distributed random variable where n is large enough?
- If $\mathrm{Prob}(X = x_i) = p_i$ (for $i = 1, 2, \ldots, L$), then the sequence $x_{i_1} x_{i_2} \ldots x_{i_n}$ occurs with probability $p_{i_1} p_{i_2} \ldots p_{i_n}$.
- In any such sequence of length n, *typically* there will be np_1 no. of x_1, np_2 no. of x_2, ..., np_L no. of x_L – **law of large numbers**.
- How many such typical sequences are there?
- It is: $\frac{n!}{(np_1)! \times (np_2)! \times \ldots (np_L)!} \equiv f_n(p_1, p_2, \ldots p_L)$ (say).
- Using Stirling's approximation $[\log_2 n! \approx n \log_2 n - n$ for large $n]$: $f_n(p_1, p_2, \ldots p_L) \approx 2^{nH(X)}$.
- So $nH(X)$ bits are enough for the storage.

Storage of classical information (continued ...)

- Thus a length of $H(X)$ bit string is sufficient to store *reliably* the data about the different values of the random variable X per single use of it (in the context of using many copies of it).
- It can be shown that this amount is also *necessary for reliable storage*!
- We thus need to get rid of all the *redundancies* in the data about X in order to use minimal amount of space to store the data.
- This provides an alternative interpretation for $H(X)$.
- At the level of reliable storage of values of finite size (n, say) sequence of values of X, one generally requires bit strings of length larger than $nH(X)$.

Transmission of classical information: Shannon's noisy channel coding theorem

Transmission of classical information

- Telephone line, internet networking system, etc. are all examples classical information carrying channels.

- The essential idea behind the working principle of each such channel is:
  (i) the input message (a string of values of some X) is encoded in terms of a string of values of the channel input variable (e.g., a string of n bit values),
  (ii) each such input variable is sent through the channel,
  (iii) at the output of the channel we then have a string (of length same as that of the string of the input variable) of the channel output variable,
  (iv) decode this string to get a string of (possibly modified) values of X.

Transmission of classical information (continued...)

- Higher the correlation between the input string of values of X and the output string of values of X, higher will be the information carrying capacity of the channel.
- (Example 1): $0 \to 0$, $1 \to 1$. There is a perfect correlation between the input and the output, and so the capacity is 1 bit.
- (Example 2): $0 \to 1$, $1 \to 0$. There is a perfect anti-correlation between the input and the output, and so (by simply flipping the output) the capacity is again 1 bit.
- (Example 3): A typing machine for which 'a' is typed as either 'a' or 'b' with equal probability, 'b' is typed as either 'b' or 'c' with equal probability, . . . , 'z' is typed as either 'z' or 'a' with equal probability, while all other characters are being typed properly. By choosing *appropriately* the input text, the machine can work as a noiseless channel.

Transmission of classical information (continued...)

- **Binary Symmetric Channel:** $0 \to 0$ and $1 \to 1$ each separately with probability $(1 - p)$ while $0 \to 1$ and $1 \to 0$ each separately with probability p.
- By looking at the output of the single use of the channel, there is no way of finding out the input correctly.
- Let us use the encoding: $0 \to 000 \equiv 0_L$ and $1 \to 111 \equiv 1_L$. By looking at the three-bit output (for three times usage of the channel), the receiver has to decide upon the input (single) bit.
- If, for example, 001 is the output, it is highly likely that only the 3rd bit has been flipped by the channel action, and consequently, 0 was the corresponding input bit (provided p is small).
- This type of encoding–decoding is called *majority voting*.

Transmission of classical information (continued...)

- Majority voting fails if two or more input bits are flipped by the channel action, the probability of which is $3p^2(1-p)+p^3 = 3p^2 - 2p^3$, which is the error probability p_e in this case.
- Without encoding, the error probability is p.
- So the majority voting will work properly if $p_e < p$, *i.e.*, if $p < 1/2$.
- Above is an example of *repetition code*.
- In any such channel encoding scheme, the basic goal is to introduce *redundancy* in order to reliably decypher the channel input by looking at the output.

Transmission of classical information (continued...)

- Choose the input random variable X to the channel with $\mathrm{Prob}(X = 0) = q$ and $\mathrm{Prob}(X = 1) = 1 - q$.
- Among the 2^n different values of the n-bit input strings, choose only 2^{nR} – forming the *code book* $\mathcal{C}$.
- An arbitrarily chosen *codeword* $x_1 x_2 \ldots x_n$ from $\mathcal{C}$ will get mapped into a bit string $y_1 y_2 \ldots y_n$ under the action of n uses of the channel.
- $y_1 y_2 \ldots y_n$ differes from $x_1 x_2 \ldots x_n$ typically at np places.
- So, typically all the possible outputs corresponding to the input $x_1 x_2 \ldots x_n$ will lie within the *Hamming sphere* of radius np.
- If these Hamming spheres do not overlap (or, overlap very little) with each other, then the input can succesfully be decoded.

Transmission of classical information (continued...)

- How many elements are there within each Hamming sphere: $2^{nH(p)}$, where $H(p) = -[p\mathrm{log}_2 p + (1-p)\mathrm{log}_2(1-p)]$.

- How many typical output sequences are there: $2^{nH(Y)}$.

- For reliable decyphering, we must have: $2^{nR} \times 2^{nH(p)} < 2^{nH(Y)}$, , *i.e.*, $R < H(Y) - H(p)$.

- By maximizing $H(Y)$ – which occurs when $q = 1/2$ – we see that $R < 1 - H(p)$.

- It can be shown that with a rate $R > 1 - H(p)$, reliable transmission of data is not possible.

- $1 - H(p) = \max\{H(X;Y) \equiv H(X) + H(Y) - H(X,Y)|X\}$ is the required classical capacity of the channel.

Transmission of classical information (continued...)

- Any classical channel is represented by an $m \times m$ *doubly stochastic* matrix $M = (\lambda_{ij})_{i,j=1}^{m}$ [*i.e.*, $\sum_{i=1}^{m} \lambda_{ij} = 1$ for $j = 1, 2, \ldots, m$ and $\sum_{j=1}^{m} \lambda_{ij} = 1$ for $i = 1, 2, \ldots, m$].
- If X is the input random variable to the channel (with $\mathrm{Prob}(X = x_i) = p_i$ for $i = 1, 2, \ldots, m$) while Y is the output random variable (with $\mathrm{Prob}(Y = y_i) = q_i$ for $i = 1, 2, \ldots, m$), then $(q_1, q_2, \ldots, q_m)^T = M(p_1, p_2, \ldots, p_m)^T$.
- The capacity of the channel (more correctly, *memoryless channel*) is then given by the maximum of mutual information $H(X; Y)$ over all possible input probability distribution.
- This is so as the mutual information $H(X; Y)$ provides us the amount of left-over ignorance we have about X by knowing Y.

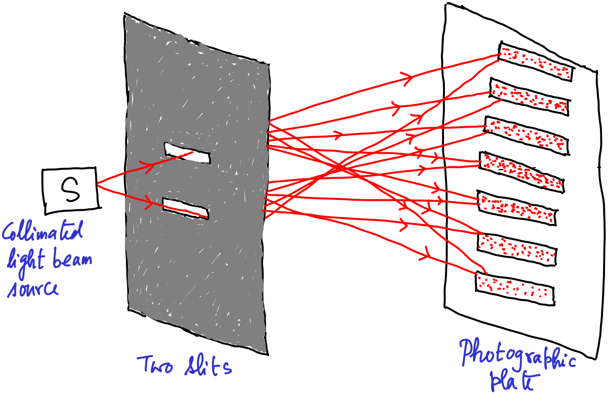
A brief introduction to Quantum Mechanics

Historical background

- By the end of nineteenth century, nature was supposed to be well-understood by most of the scientists.
- In fact, most of the scientists thought that there could not be **any new fundamental law** of nature apart from the existing ones – eventhough there were many natural phenomena (e.g., energy spectrum of black body radiation) without having proper explanation.
- Moreover, by this time, most of the scientists had started believing that the physical universe should be governed by **only deterministic laws**.
- But within next thirty years, both these opinions got completely destroyed.
- And the world observed Quantum Theory taking its complete shape around 1925 with the help of its inventors.

Wave and particle nature of light

- Much of the evidences which supported towards a new way of thinking about the physical universe – in particular, at microspic level – came from studies on some properties of light.
- By Maxwell's electromagnetic theory, the wave nature of light (as an electromagnetic wave) was firmly established:
- We do get to see interference effect of light in double slit experiment irrespective of the intensity of the light source.
- On the other hand, Einstein's photo-electric effect (namely, emmision of electrons from certain metals (in vacuum), acted upon by light beams) establishes particle (*i.e.*, photon – the quanta of light) nature of light.



Double Slit Experiment

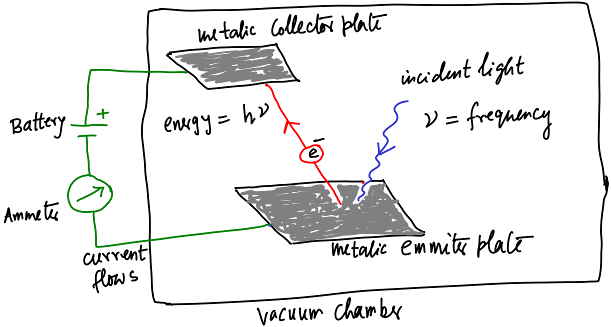


Photo-electric effect

Wave-particle duality

- Quantum Theory helps us to reconcile the wave and particle nature of light:
- In double slit experiment, interference pattern remains intact so far as one does not figure out the slit through which the light ray has passed;
- but the interference pattern gets completely destroyed once we figure out the slit through wich light rays passed.
- According to Feynmann, the light particles (photons) receive a momentum kick as soon as one locates the slit through which light particles have passed, and that leads to destruction of the interference pattern.
- Present-day understanding of this wave-particle dulaity rests on the concept of entanglement between path and polarization degrees of freedom of photons – not on momentum kick!

Uncertainty relation

- Looking at the slit through which light rays are passing amounts to measurement of positions of photons.
- On the other hand, observation of interference patten on the screen corresponds to measurement of photons momenta.
- According to Heisenberg, there should always be constraints on the accuracies of such position as well as momentum measurements: $\Delta x \Delta p \geq \hbar/2$ (Heisenberg's uncertainty relation).
- Note that it does not talk about joint measurement of position and momentum of photon.
- Complementarity and uncertainty relations do not follow from each other, in general.
- Moreover, not every wave property is complementary to every particle property of a quantum system!

Atomic model

- Rutherford came up with the atomic model where electrons are supposed to orbit around atomic nucleus.
- But Maxwell's electromagnetic theory gives rise to radiation of energy by any electrically charged rotating particle.
- And this would lead to falling down of such rotating electrons on the atomic nucleus – causing collapse of he atom itself!
- It was Bohr who then came up with his hypothesis of atomic structure in which each atomic electron can move along some definite **stable** orbit corresponding to 'quantized' angular momentum of such an electron.
- Although such a model could explain energy spectrum of one electron atom (Hydrogen atom), but it fails to do so for other atoms – the stability issue in such cases remained unanswred by Bohr's model.

Matter waves

- For light quanta (photon) having enery $E = h\nu$ (with ν being the light frequency), its momentum is known to be $p = E/c = h\nu/c = h/\lambda$, where λ is the wave length.
- Based on this, Louis de Broglie came up with the radical idea of 'matter waves':
- Microscopic material particles (like electron, neutrons, etc.) should also exhibit wave-like features satisfying the same relation (namely, $p = h/\lambda$).
- It was first confirmed experimentally by Davisson and Germer by scattering of electrons by nickel crystal and thereby looking at the difraction pattern of the scattered electrons.
- Like double slit experiment with photons, interference pattern has been observed in recent past with neutrons.

Schrödinger equation

- **Observation:** When waves are confined within a particular region of space, only particular wavelengths are allowed.
- For example, in any string instrument, only particular notes (*i.e.*, waves of particular frequencies) can be emmited – depending upon the length of the string and the tension applied to it.
- Using the matter wave hypothesis in the context of electron waves in atoms, in which atomic electrons and nucleus are attracted to each other by inverse square law, it was found – analogous to the aforesaid feature – that the electron waves should satisfy a particular equation (namely, the Schrödinger equation) whose solutions exist **only** for specific ('quantized') values of the (electron) energies.
- So, an electron in an atom can not have arbitrarily low energy!

Success of Quantum Theory

- It helped in finding out neclear structure – the structure of the fundamental particles (proton, neutron) involving quarks.
- It helped in understanding phenomena in many-body systems too: chemical bonding, superconductivity, superfluidity, etc.
- Quantum theory became the bread and butter of most of the practicing physicists!
- Still there are foundational issues with Quantum Theory, issues with combining it with Einstein's General Theory of Relativity at two extreme cases (Big Bang and Black hole).
- We still do not have a clear picture to characterise all those natural phenomena which are 'truly' quantum in nature!

Towards axiomatizing Quantum Mechanics

- **Classical systems:** Physical systems which are described by laws of classical physics (namely, Newtonina mechanics and Maxwell's electromagnetism, and their statistical extensions).
- **State space:** States of any classical system S are **minimum** collections (ordered) of measurable quantities – associated to the different degrees of freedom of the system.
- **Examples:** Any classical mechanical system is described by its phase-space variables (x, p), or its generalized form. Any electromagnetic field is completely described by the scalar and vector potentials $(\phi, \vec{A})$. Any thermodynamic system is described by its related thermodynamic variables like pressure, temperature, etc.

Towards axiomatizing Quantum Mechanics · · ·

- **Measurement rule:** Measurement of any observable quantity (like, position, momentum, enery, temperature, pressure, electric field strength, etc.) gives rise to a possible value of the quantity, it supposed to have possessed *apriori*.
- Probabilistic structure of any such measurement result can arise in the case when we consider a statistical ensemble of such systems.
- Moreover, any such measurement can, in principle, be done in a **non-invasive** way. **No change of state** happens!
- **Dynamics:** By Newton's equation of motion, Maxwell's equation, etc. In general, by Hamiltonian evolution.
- The dynamics is fully deterministic.

Axiomatic approach to Quantum Mechanics

- **System space:** Every quantum mechanical system S is associated with a Hilbert space $\mathcal{H}_S$.
- **System states:** Every normalized vector $|\psi\rangle$ of $\mathcal{H}_S$ is a bonafide state of system S.
- **System observables:** Every measurable property of the system is associated with a hermitian operator $\hat{A} : \mathcal{H}_S \mapsto \mathcal{H}_S$.
- **Born rule:** Given that the system S has been prepared in a state $|\psi\rangle$, the measurement of an observable $\hat{A}$ on the system then gives rise to **one** of the eigen values of the observables with associated outcome probability $\langle\psi|\hat{A}|\psi\rangle$, and the system will **immediately collapse** into the corresponding eigenstate of the observable.
- **Dynamics:** By Schrödinger equation: $i\hbar\frac{\partial|\psi\rangle}{\partial t} = \hat{H}|\psi\rangle$, where $\hat{H}$ is the system's Hamiltonian. It is deterministic in nature!

Why do we need to axiomatize Quantum Mechanics?

- Axiomatization of each theory is useful, be it classical mechanics, classical electrodynamics, statistical mechanics, or quantum mechanics – as one always wants to come up with a minimal set of laws describing the the system.
- What is so special about axiomatizing QM?
- It is indeed difficult – and the issue has not yet been settled – to find out a **minimal** set of physical principles which would give rise to Quantum Theory and nothing else!
- On the top of this, we need to address the measurement problem:
- That the issue of collapse in quantum mechanical measurement is not describable by a bonafide evolution.
- Nevertheless, even after axiomatization, measurement problem still remains.

How effective this axiomatization is?

- Quantum systems are really frazile – they interact very often with their surroundings!
- And this interaction leads to noise in the prepared states of the system.
- So, even if we prepare a quantum system in a **pure** state (*i.e.*, in a state vector $|\psi\rangle$), the system will hardly remain in this state after the preparation.
- This leads to the concept of **mixed** states.
- By the same token, **projective measurements** (*i.e.*, measurement in orthonormal bases) are difficult to realize experimentally.
- This leads to the concept of **generalized measurements**.
- Dynamics of an **open** quantum system can not be described by unitary evolution (Schrödinger equation).

What is the way out?

- All these issues can be taken care of **without going beyond** the aforesaid axioms.
- All we need is the concept of **composite quantum system** and a few **mathematical generalizations** of the notions of states, measurements, and dynamics!

Axiom about composite quantum systems

- **Composite systems:** If A and B are two quantum systems with respective Hilbert spaces $\mathcal{H}_A$ and $\mathcal{H}_B$, the Hilbert space associated to the composite system $S = A + B$ is **considered to be** the **tensor product** Hilbert space $\mathcal{H}_S = \mathcal{H}_A \otimes \mathcal{H}_B$.
- Note that we don't care here whether the **subsystems** A and B are interacting or not [unless there is any 'superselection rule'].
- Thus, if $\{|e_\alpha\rangle : \alpha \in \Lambda_A\}$ is an ONB for $\mathcal{H}_A$ and $\{|f_\beta\rangle : \beta \in \Lambda_B\}$ is an ONB for $\mathcal{H}_B$, then the **product basis** $\{|e_\alpha\rangle \otimes |f_\beta\rangle : \alpha \in \Lambda_A, \beta \in \Lambda_B\}$ is an ONB for $\mathcal{H}_S$.
- Thus, a generic pure state of S is of the form: $|\Psi\rangle_S = \sum_{\alpha\in\Lambda_A} \sum_{\beta\in\Lambda_B} a_{\alpha\beta} |e_\alpha\rangle_A \otimes |f_\beta\rangle_B$, where $\sum_{\alpha\in\Lambda_A} \sum_{\beta\in\Lambda_B} |a_{\alpha\beta}|^2 = 1$.

Axiom about composite quantum systems ....

- **Observables of composite quantum systems:** If $\{\hat{A}_\epsilon : \epsilon \in \chi_A\}$ is a basis of hermitian operators $\hat{A}_\epsilon : \mathcal{H}_A \mapsto \mathcal{H}_A$ while $\{\hat{B}_\eta : \eta \in \chi_B\}$ is a basis of hermitian operators $\hat{B}_\eta : \mathcal{H}_B \mapsto \mathcal{H}_B$, then any observable $\hat{S} : \mathcal{H}_S \mapsto \mathcal{H}_S$ for S is of the form: $\hat{S} = \sum_{\epsilon \in \chi_A} \sum_{\eta \in \chi_B} w_{\epsilon\eta} \hat{A}_\epsilon \otimes \hat{B}_\eta$ with real coefficients $w_{\epsilon\eta}$.
- **Action of observables on states:** $\hat{S}|\Psi\rangle = \sum_{\alpha \in \Lambda_A} \sum_{\beta \in \Lambda_B} \sum_{\epsilon \in \eta_A} \sum_{\eta \in \eta_B} a_{\alpha\beta} w_{\epsilon\eta} \hat{A}_\epsilon |e_\alpha\rangle \otimes \hat{B}_\eta |f_\beta\rangle$.
- **Inner product rule:** For any $|\Psi\rangle_S = \sum_{\alpha \in \Lambda_A} \sum_{\beta \in \Lambda_B} a_{\alpha\beta} |e_\alpha\rangle_A \otimes |f_\beta\rangle_B$ and $|\Phi\rangle_S = \sum_{\alpha \in \Lambda_A} \sum_{\beta \in \Lambda_B} b_{\alpha\beta} |e_\alpha\rangle_A \otimes |f_\beta\rangle_B$, we have $\langle\Psi|\Phi\rangle = \sum_{\alpha \in \Lambda_A} \sum_{\beta \in \Lambda_B} a^*_{\alpha\beta} b_{\alpha,\beta}$.

Canonical freedom

- Given the composite system S prepared in state $|\Psi\rangle$, can one **always** assign state vectors to the individual subsystems A and B?
- This boils down to checking whether $|\Psi\rangle$ is a **product state** or an **entangled state**.
- Note that being entangled or not is a physical property of the state **without allowing** the subsystems to 'talk to each other'.
- Nevertheless, the subsystems are allowed to choose their reference frames: they are **free to choose** their respective ONBs.
- Given this freedom, is there a **canonical** choice of **local** ONBs which guarantees **unanimously** about entanglement in $|\Psi\rangle$.

Schmidt decomposition

- **Schmidt decomposition:** One can always an ONB (possibly incomplete) $\{|r_i\rangle_A : i = 1, 2, \ldots, d\}$ for $\mathcal{H}_A$ and an ONB (possibly incomplete) $\{|s_i\rangle_B : i = 1, 2, \ldots, d\}$ for $\mathcal{H}_B$ (with $d = \min\{d_A, d_B\}$) such that $|\Psi\rangle = \sum_{i=1}^{d} \sqrt{\lambda_i}|r_i\rangle_A \otimes |s_i\rangle_B$, with λ_i's being the 'sigular values' of the coefficient matrix $(a_{\alpha\beta})$.
- **Schmidt coefficients:** $\{\lambda_i : i = 1, 2, \ldots d\}$ is a probability distribution (normalization condition).
- $|\Psi\rangle$ is a product state if and only if all but one Schmidt coefficient is non-zero.
- Otherwise, it is entangled.
- The 'thermodynamics' amount of entanglement in $|\Psi\rangle$ is given by the **Shannon entropy** $-\sum_{i=1}^{d} \lambda \log_2 \lambda_i$ of the distribution.

Reduced states

- How to assign a state to each subsystem when the joint state $|\Psi\rangle$ is entangled?
- Why is it necessary?
- When the subsystems are far apart, the individual parties (Alice and Bob) possessing the subsystems separately, can only perform quantum operations on their respective subsystems **only**.
- And, in this case, the individual parties will have access to the measurement statistics (for example) related to their individual subsystems, in case they have the **apriori** information about the states of their respective subsystems.
- In fact, if $\mathcal{A}$ ($\mathcal{B}$) is any operator acting on A (B) then: $\langle\psi|(\mathcal{A} \otimes I_B)|\Psi\rangle = \mathrm{Tr}_A[\rho_A \mathcal{A}]$, where $\rho_A \equiv \mathrm{Tr}_B|\Psi\rangle\langle\Psi|$ is the **reduced state** of A. Same is the case for B.

Density matrices

- Here $\rho_A = \sum_{i=1}^{d} \lambda_i |r_i\rangle\langle r_i|$
- **Mixture of pure states:** $\rho \equiv \sum_\alpha p_\alpha |\psi_\alpha\rangle\langle\psi_\alpha|$, where $\{p_\alpha\}$ is a probability distribution.
- ρ is a **mixed** state.
- **Properties of ρ:** (i) ρ is a linear operator on the system Hilbert space. (ii) ρ is hermitian. (iii) $\mathrm{Tr}\rho = 1$. (iv) $\rho \geq 0$.
- **Generalization of state:** Any state of a quantum system S is described by a density matrix ρ, acting on $\mathcal{H}_S$.

Projective measurement is ideal

- As we are concerned only with measurement probabilities, it is *enough* to consider measurement of any observable $\hat{A}$ on any state ρ of a quantum system A to be 'measurement in the eigenbasis' $\{|\psi_i\rangle : i = 1, 2, \ldots, d_A\}$ of $\hat{A}$. Any such measurement is said to be **projective**.
- But accurate projective measurement is an ideal case!
- Noise is inevitable to happen!
- How to formulate such a measurement with noise, where $|\psi_i\rangle$ becomes noisy?
- In its full generality, $|\psi_i\rangle$ now becomes a **positive** operator E_i such that $\sum_i E_i = I$.
- **POVM:** Given a state ρ of S, a POVM $\{E_i : i = 1, 2, \ldots\}$ on it is a collection of positive operators E_i on $\mathcal{H}_S$ satisfying $\sum_i E_i = I_S$ with i-th outcome probability $\mathrm{Tr}[\rho E_i]$.

Post-measurement states in POVM

- For a projective measurement $\{P_i : i = 1, 2, \ldots n\}$ (with $P_i P_j = \delta_{ij} P_j$ for $i, j = 1, 2, \ldots, n$ where $n \leq d$, the dim. of the system Hilbert space $\mathcal{H}_S$) on a system's state ρ, the post-mesurement state for i-th measurement outcome is: $(P_i \rho P_i)/\mathrm{Tr}[\rho P_i]$.
- For a POVM $\{E_i : i = 1, 2, \ldots m\}$, the post-measuremnt state is **not uniquely** given:
- It depends on the way of **realizing the POVM as a projective measurement** $\{P_i : i = 1, 2, \ldots, m\}$ on an **extended Hilbert space** $\mathcal{H}_{S'}$, which, in principle, can be taken as $\mathcal{H}_S \otimes \mathcal{H}_A$, where A is an 'ancilla' system.
- In fact, if $E_i = M_i^\dagger M_i$ (non-unique), then the post-measurement state for the i-th outcome is: $(M_i \rho M_i^\dagger)/\mathrm{Tr}[\rho E_i]$.

Neumark's dialation

- $M_i U$ will also do the job for any unitary U.
- Neumark's dialation guarantees the existence of such an extension.
- In fact, in the case of **optimal unambiguous discrimination** of two non-orthogonal photon polarized states $|\psi\rangle$, $|\phi\rangle$ (supplied with equal probabilities), the corresponding optimal measurement scheme – represented by a POVM $\{E_1 = x|\phi^\perp\rangle\langle\phi^\perp|, E_2 = x|\psi^\perp\rangle\langle\psi^\perp|, E_3 = I_2 - E_1 - E_2\}$ – has been realized as a projective measurement in four dimension.
- So POVM is not just a theoretical artifact!
- Note that the probability structure remains same in Neumark's dialation: Probability of i-th outcome in POVM is same as that for the i-th outcome in the corresponding projective extension.

Generalized dynamics

- Given the Hamiltonian evolution $\rho_{AB}(o) \mapsto e^{-i\hat{H}t/\hbar}\rho e^{i\hat{H}t/\hbar} \equiv \rho_{AB}(t)$, what is the evolution equation for states of individual subsystems?
- Will it be always a Hamiltonian dynamics? Not always!
- The dynamics, in that case will be, in general, of the form of **master equation**: $i\hbar\frac{\partial\rho}{\partial t} = [\hat{H}_A, \rho] + \mathcal{D}(\rho)$, where $\hat{H}_A$ is a hermitian operator on $\mathcal{H}_A$, and $\mathcal{D}$ is a 'dissipitative' term.
- Such a dynamics **may** sometimes be represented by **Kraus representation**: $\mathrm{Tr}_B\rho_{AB}(0) \equiv \rho_A(0) \mapsto \sum_\alpha A_\alpha(t)\rho_A(0)(A_\alpha(t))^\dagger$ with the trace preservation condition $\sum_\alpha (A_\alpha(t))^\dagger A_\alpha(t) = I_A$.
- Such a representation is possible only under very restricted condition!

Schumacher's data compression limit: von Neumann entropy

Bits vs. Qubits

- In CIS, every information (about state of the system) is stored in terms of string of bits (e.g., 01100010).
- In QIS, every information (about state of the system) is stored in terms of linear superposition of strings of qubits (e.g., $|\Psi\rangle = a|0\rangle \otimes |1\rangle \otimes |1\rangle + b|0\rangle \otimes |1\rangle \otimes |0\rangle + c|1\rangle \otimes |1\rangle \otimes |1\rangle$).
- A bit can only take two values: 0 and 1.
- A qubit (a two-level quantum system) can be in any linear superposition (normalized) of the two orthogonal states $|0\rangle$ and $|1\rangle$.
- Although we mostly use qubits to get a feeling about the quantum mechanical counterpart of Shannon's theory of classical information, it is always better to look the scenario in the general setup.

Schumacher's data compression limit

- If a quantum mechanical system can be prepared in a state $|\psi_i\rangle$ with probability p_i $(i = 1, 2, \ldots, k)$, then system is said to have the density matrix $\rho = \sum_{i=1}^{k} p_i |\psi_i\rangle\langle\psi_i|$ corresponding to the ensemble $\{p_i, |\psi_i\rangle \ : i = 1, 2, \ldots, k\}$.
- In the case of storing the (quantum) information about n identical copies of ρ, it is enough to store the information in terms of states of a system of $nS(\rho)$ qubits, in the large n limit – according to Schumacher's noiseless coding theorem.
- $S(\rho) = -\mathrm{Tr}\ (\rho \log_2 \rho)$ is the von Neumann entropy of ρ.
- $S(\rho) \leq H(p_1, p_2, \ldots, p_k)$ with equality holds if and only if $|\psi_i\rangle$'s are pairwise orthogonal.
- So, in case classical data can be stored in terms of qubits (in a devise like quantum computer), less amount of storage space would be required, in general.

Schumacher's data compression limit (continued ...)

- Given any density matrix ρ of a d dim. quantum system S, find out its spectral decomposition: $\rho = \sum_{i=1}^{d} q_i |\psi_i\rangle\langle\psi_i|$, with $0 \leq q_i \leq 1$, $\sum_{i=1}^{d} q_i = 1$, and $\langle\psi_i|\psi_j\rangle = \delta_{ij}$.
- Thus here we have the correspondense for storing n copies of ρ: $i_1 i_2 \ldots i_n \equiv |\psi_{i_1}\rangle \otimes |\psi_{i_2}\rangle \otimes \ldots \otimes |\psi_{i_n}\rangle$ with $i_1, i_2, \ldots, i_n \in \{1, 2, \ldots, d\}$
- Storing a string $i_1 i_2 \ldots i_n$ requires $nH(X)$ no. of bits where $\mathrm{Prob}(X = i) = q_i$ (for large n).
- So storing the (quantum) information about the string $|\psi_{i_1}\rangle \otimes |\psi_{i_2}\rangle \otimes \ldots \otimes |\psi_{i_n}\rangle$ requires $nH(X) = nS(\rho)$ no. of qubits.
- Thus the quantum information content of ρ is $S(\rho)$.

Entanglement – a fundamental resource for quantum information processing

Separability versus entanglement

- In case one can write (by properly choosing ONBs for $\mathcal{H}_A$ and $\mathcal{H}_B$) $|\Psi\rangle = |\psi\rangle_A \otimes |\phi\rangle_B$, then the system A will have the state vector $|\psi\rangle$ while the sytem B will have the state vector $|\phi\rangle$ – $|\Psi\rangle$ is a *product* state.
- For *most* of the states $|\Psi\rangle$ such a product decomposition is not possible – $|\Psi\rangle$ is an *entangled* state.
- In general, based upon the outcome (x_i with probability p_i, say) of experimenting on a *shared* random variable X, Alice (possessing A) can prepare her system in the density matrix ρ_i while Bob (possessing B) can prepare his system in the density matrix σ_i, so that the state of S becomes: $\rho = \sum_i \rho_i \otimes \sigma_i$ – a *separable* state of S.
- Any density matrix of S which is not separable, is *entangled*.

Separability versus entanglement (continued ...)

- What will be the state of A (B) when the joint state of $S = A + B$ is ρ_{AB}?
- It is the reduced density matrix $\rho_A \equiv \mathrm{Tr}_B \rho_{AB}$ $(\rho_B \equiv \mathrm{Tr}_A \rho_{AB})$.
- This is so because: $\mathrm{Tr}(\rho_A \mathcal{A}) = \mathrm{Tr}((\mathcal{A} \otimes I)\rho_{AB})$ for *any* observable $\mathcal{A}$ on A.
- ρ_A (or, ρ_B) corresponds to a state vector iff *rho*$_{AB}$ is a product state.
- For state vectors $|\Psi\rangle_{AB}$ it is easy to check whether it is entangled or not – look at the Schmidt decomposition: $|\Psi\rangle = \sum_{i=1}^{\min\{d_A, d_B\}} \sqrt{\lambda_i} |e_i\rangle_A \otimes |f_i\rangle_B$ of it and see it it has *more than one* non-zero Schmidt coefficients λ_i's.
- But for mixed states ρ_{AB}: difficult to decide about entanglement/separability.

Separability versus entanglement (continued ...)

- **Horodecki condition:** A bipartite state ρ_{AB} is entangled iff there exists a linear, Hermiticity-preserving, positivity-preserving map $\Phi : \mathcal{B}(\mathcal{H}_A) \to \mathcal{B}(\mathcal{H}_C)$ (with $\mathrm{dim}\mathcal{H}_C = \mathrm{dim}\mathcal{H}_B$) such that $(\Phi \otimes I_{\mathcal{B}(\mathcal{H}_B)})(\rho_{AB})$ has *atleast one* negative eigenvalue.
- Very difficult to scan through all such operators (called *positive but not completely positive* maps), in general.
- But for $\mathrm{dim}\mathcal{H}_A = 2$, the *transpose* map will do the job provided $\mathrm{dim}\mathcal{H}_B$ is either 2 or 3.
- Nevertheless, it is a *mathematical* condition!
- One should look for some *physical* condition to test entanglement in ρ_{AB}.
- Unfortunately no such physical condition is there which would hold good for *all* entangled states ρ_{AB}.

Holevo's bound on accessible information

Accessible information

- Consider a random variable X with $\mathrm{Prob}(X = x_i) = p_i$ for $i = 1, 2, \ldots, N$.
- Use now the encoding: $x_i \to \rho_i$, where ρ_i is a state of a d dim. quantum system A.
- Alice (in possession of A) now send her system to Bob via a noiseless quantum channel without disclosing i.
- Bob's job is to identify i by performing measurement on his system B (he received from Alice).
- How much information Bob can extract about i?
- Given by the max. value (over Y) of the mutual information $H(X; Y)$, where Y is the randomvariable corresponding to the measurement outcome.
- Holevo (in 1970's) showed that this max. value is tightly upper bounded by the quantity: $S(\rho) - \sum_{i=1}^{N} p_i S(\rho_i)$.

Proof of Holevo's bound on accessible information

- See iop-2010-schumacher.pdf (p.30 to p.48) for the proof

Different capacities of quantum channels

Quantum channels

- See iop-2010-channel-capacity.pdf (p.2 to p.19) for the details

Transmission through quantum channel

- Any quantum channel is a physical operation T allowed by quantum mechanics (e.g., unitary evolution, measurement, ignoring a subsystem, etc.) which transforms a quantum system (in state ρ) to another quantum system (in state ρ'): $T(\rho) = \rho'$.
- Encoding the values of a random variable X (with $\mathrm{Prob}(X = x_i) = p_i$ for $i = 1, 2, \ldots, k$) in terms of a quantum ensemble $\rho \equiv \{p_i, |\psi_i\rangle : i = 1, 2, \ldots, k\}$, the rate of transmission of values of X through a noisy quantum channel T – in the event of several independent usage of the channel and disallowing any initial quantum correlation among the input systems to the channel – is given by the Holevo quantity $S(T(\rho)) - \sum_{i=1}^{k} p_i S(T(|\psi_i\rangle\langle\psi_i|))$, as per HSW theorem.
- This is classical information transmission capacity of T. (*)

Transmission of information . . .

- Classical information carrying capacity of a noisy quantum channel T can be, in principle, higher than the Holevo quantity if quantum correlations are allowed in the input states (additivity problem).
- T can carry information about quantum states – quantum capacity of T – not available for classical channels.
- T can, in principle, have better capacity of carrying classical information if the sender and the receiver use some apriori shared quantum correlation – entanglement-assisted classical capacity of T – not available for classical channels.
- T can, in principle, have the capacity of secured transmission of private information (e.g., for sharing private key to be used for secured cryptography)– private capacity of T – not available for classical channels.

Transmission of information . . .

- Finding out analytical expressions of the different capacities of a generic quantum channel is a major challenge in QIS.
- Special cases are there where the values of these capacities are known: quantum teleportation, superdense coding, quantum cryptographic key distribution, etc.
- Additivity question of different capacities of quantum channels remains unresolved – except for a very few cases.
- The situation becomes really difficult if there is some memory effect among the different usage of the channel – almost nothing is known.

Details about different capacities of quantum channels

- See iop-2010-channel-capacity.pdf (p.20 to p.36) for the details

- See also kraus.pdf

Applications of QIT

Processing of information

- In the case of classical information expressed in terms of strings of bits, no further processing of information is necessary – the information is readily available.
- In the case of information (classical or quantum) expressed in terms of states of several qubits, processing of information means performing measurement on the qubits – may be (i) separately and independently on the individual qubits, or (ii) separately but dependently on the individual qubits, or (iii) on the qubits all together.
- In general, measurement of type (iii) gives better information compared to that of type (ii), which, in turn, can give better information compared to measurement of type (i).
- But asymptotically they may not differ.

Usefulness of QIS

- Study of QIS has not only helped in resolving several aspects of physics (e.g., understanding quantum phase transition using many-body quantum correlations, simplifying the task of finding out lower-energy eigen states of many-body quantum Hamiltonian, achieving the so-called Heisenberg limit in precision measurement, understanding CP-violation through quantum correlations, understanding photo-synthesis using quantum coherence, randomness generation, atomic clock precession using entangled states, etc.), it has also given new directions to the study of foundations of quantum mechanics – it is no more a philosophical discourse!

Usefulness of QIS . . .

- In particular, the conflict between epistemic view of quantum states (*i. e.*, quantum states are nothing but a list of all possible outcome probabilities one can have by performing all possible measurements on the state) and ontic view (*i. e.*, quantum states are associated with physical realities, it is not just a list of probabilities) is gradually taking a shape in the context of quantum state discrimination.

- Ontic view turns out to be exponentially better compared to the epistemic view for system dimension ≥ 3 in the context of state discrimination.

Quantum teleportation, superdense coding, etc.

- See madras-univ-talk2011.pdf (p.43 to p.46) for the details

Present scenario

World-wide scenario

- Research is going on in the field of Device-Independent Quantum Information Processing to get security proofs independent of inner mechanism of the devices.
- To find out closed-form formulae of different capacities of useful quantum channels and their orderings.
- Understanding causal structure of quantum correlations and beyond.
- Formulation of thermodynamics laws in the quantum world.
- Applications of QIT in Quantum Many-body Systems
- Etc.

Our activities at IMSc

- One of the major activities in the QIS group here is concentrated towards classification of bipartite as well as multi-partite quantum correlations, their quantification, and their usability in different quantum information processing tasks – both for finite dimensional as well as infinite dimensional quantum systems.
- Another major activity is characterization of quantum channels, finding out their capacities, physical realizations of such channels, etc.
- Study of dynamics of open quantum systems and to use ideas from QIS to characterize these type of dynamics with particular emphasis on photonic and atomic systems.
- Simulation of different quantum phenomena using the ideas from Quantum Walks.

Our activities . . .

- In particular, developing different QIS-based techniques for controlling decoherence of states of open quantum systems.
- Use of QIS techniques in classical wave optics.
- Characterization of non-classicality of states of physical systems.
- Developing Relativistic Quantum Information Processing.
- etc.

Thanks!