УДК 512.6, 517.9, 519.6

BASIC RELATIONS OF QUANTUM INFORMATION THEORY. PT 2: CLASSICAL, QUANTUM AND TOTAL CORRELATIONS IN QUANTUM STATE - MEASURE OF QUANTUM ACCESSIBLE INFORMATION

Ulyanov Sergey[1], Korenkov Vladimir[2], Reshetnikov Andrey[3], Tanaka Takayuki[4], Fukuda Toshio[5]

[1]*Doctor of Science in Physics and Mathematics, professor;*
Dubna State University,
Institute of system analysis and management;
141980, Dubna, Moscow reg., Universitetskaya str., 19;
e-mail: ulyanovsv@mail.ru.

[2]*Doctor of Technical science, professor;*
Deputy Director of the Laboratory;
Joint institute for nuclear researches, Laboratory of Information Technologies;
141980, Moscow reg., Dubna, Joliot-Curie, 6;
e-mail: korenkov@cv.jinr.ru.

[3]*PhD in informatics, associate professor;*
Dubna State University,
Institute of system analysis and management;
141980, Dubna, Moscow reg., Universitetskaya str., 19;
e-mail: agreshetnikov@gmail.com.

[4]*PhD, professor;*
The Graduate School of Information Science and Technology, Hokkaido University;
N14, W9, Sapporo-shi, Hokkaido, Japan;
e- mail: ttanaka@ssc.ssi.ist.hokudai.ac.jp.

[5]*PhD, professor;*
Dept. of Micro System, Dept. of Mechanics- Informatics, Nagoya University;
Furo-cho, Chikusa-ku, Nagoya, Japan;
e- mail: fukuda@mein.nagoya u.ac.jp.

The evolution of a quantum system can be examined from an information theory point of view. The complex vector entering the quantum evolution is considered as an information source both from the classical and the quantum level.

Keywords: quantum computing, quantum information, von Neumann entropy, skew information.

ОСНОВНЫЕ СООТНОШЕНИЯ КВАНТОВОЙ ТЕОРИИ ИНФОРМАЦИИ. Ч. 2: КЛАССИЧЕСКАЯ, КВАНТОВАЯ И ПОЛНАЯ КОРРЕЛЯЦИИ В КВАНТОВОМ СОСТОЯНИИ - КВАНТОВЫЕ МЕРЫ ИЗМЕРЕНИЯ ДОСТУПНОЙ ИНФОРМАЦИИ

Ульянов Сергей Викторович[1], Кореньков Владимир Васильевич[2], Решетников Андрей Геннадьевич[3], Танака Такаюки[4], Фукуда Тошио[5]

[1]*Доктор физико-математических наук, профессор;*
ГБОУ ВО МО «Университет «Дубна»,
Институт системного анализа и управления;
141980, Московская обл., г. Дубна, ул. Университетская, 19;
e-mail: ulyanovsv@mail.ru.

[2]*Доктор технических наук, профессор;*
Объединенный институт ядерных исследований, Лаборатория информационных технологий;

141980, Московская обл., г. Дубна, ул. Жолио-Кюри, 6;
e-mail: korenkov@cv.jinr.ru.

[3]*Доктор информатики (PhD in Informatics), к.т.н., доцент;*
ГБОУ ВО МО «Университет «Дубна»,
Институт системного анализа и управления;
141980, Московская обл., г. Дубна, ул. Университетская, 19;
e-mail: agreshetnikov@gmail.com.

[4]*Доктор наук (PhD in Informatics),*
Высшая школа информатики и технологии,
Университет Хокайдо;
N14, W9, Саппоро-Ши, Хокайдо, Япония;
e-mail: ttanaka@ssc.ssi.ist.hokudai.ac.jp.

[5]*Доктор наук, профессор;*
Факультет микромашин и Факультет механики и информатики
Университет Нагойя;
Furo-cho, Chikusa-ku, Nagoya, Japan;
e- mail: fukuda@mein.nagoya u.ac.jp.

Рассмотрена эволюция квантовой системы с точки зрения квантовой теории информации. Комплексный вектор состояния квантовой системы, описывающий квантовую эволюцию, рассматривается как источник информации, как на классическом, так и на квантовом уровне.

Ключевые слова: квантовые вычисления, квантовая информация, энтропия фон Нёймана, кососимметричная информация

Introduction

In the classical information theory the mutual information measures how much information X and Y has in common. Correlations between two different random variables X and Y are measured by the mutual information, $H(X:Y)=H(X)+H(Y)-H(X,Y)$, where $H(X,Y)=-\sum_{i,j} p_{ij}\log p_{ij}$ is the joint entropy and p_{ij} is the probability of outcomes x_i and y_j both occurring. It may also be defined as a special case of the relative entropy, since it is a measure of how distinguishable a joint probability distribution p_{ij} is from the completely uncorrelated pair of distributions p_i, p_j as

$$H\left(p_{ij}\|p_i p_j\right)=H\left(p_i\right)+H\left(p_j\right)-H\left(p_{ij}\right).$$

In quantum information theory it is common to distinguish between purely classical information, measured in bits, and quantum information, which is measured in qubits. Any bipartite quantum state may be used as a communication channel with some degree of success, and so it is important to determine how to separate the correlations it contains into classical and an entangled part [1-26].

When a measurement is made on a quantum system in which classical information is encoded, the measurement reduced the observer's average Shannon entropy for the encoding ensemble. This reduction, being the mutual information, is always non-negative. For efficient measurements the state is also purified; that is, on average, the observer's von Neumann entropy for the state is also reduced by a non-negative amount. A bound, which is dual to the Holevo bound, one finds that for efficient measurements, the mutual information is bounded by the reduction in the von Neumann entropy. A physical interpretation of this bound can be directly derived from the Schumacher-Westmoreland-Wootters theorem.

The classical mutual information of a quantum state ρ_{AB} can be defined naturally as the maximum classical information that can be obtained by local measurements $M_A \otimes M_B$ on the state ρ_{AB} :

$$I^{(Cl)}(\rho) \equiv \max_{M_A \otimes M_B} I(A:B).$$

Here, the classical information, the entropy functions and the probability distributions of the individual and joint outcomes are defined of performing the local measurement $M_A \otimes M_B$ on ρ. The physical relevance of $I^{(Cl)}(\rho)$ is as following: (1) $I^{(Cl)}(\rho)$ is the maximal classical correlation obtainable from ρ by purely local processing; (2) $I^{(Cl)}(\rho)$ corresponds to the classical definition when ρ is "classical", i.e., diagonal in some local product basis and corresponds to a classical distribution; (3) When ρ is pure, $I^{(Cl)}(\rho)$ is the correlation defined by the Schmidt basis and thus equal to the entanglement of the pure state; and (4) $I^{(Cl)}(\rho)=0$ iff $\rho=\rho_A \otimes \rho_B$.

Any good correlation measure should satisfy certain axiomatic properties:

N	Axiomatic property
I	_Monotonicity_: Correlation is a non-local property and should not increase under local processing
II	_Total proportionality_: A protocol starting from an uncorrelated initial state and using ℓ qubits or 2ℓ classical bits of communication (one-way or two-way) and local operations should not create than 2ℓ bits of correlation
III	_Incremental proportionality_: A small amount of communication should not increase correlation abruptly. One may expect that the transmission of ℓ qubits or 2ℓ classical bits should not increase the correlation of any initial state by more than 2ℓ classical bits
IV	_Continuity in_ ρ: This strengthens total proportionality by allowing all possible initial states, or equivalently by considering the increase in correlation step-wise

For some well-known correlation measures all of these properties are hold. They hold for the classical mutual information $I(A:B)$ when communication is classical as one may expect.

They also hold for the quantum mutual information:

$$I^{(Q)}(\rho) \equiv S(\rho_A)+S(\rho_B)-S(\rho).$$

Remark. For $I^{(Cl)}(\rho)$ the property of incremental proportionality can be violated in some extreme manner for a mixed initial state ρ. A single classical bit, sent from *A* to *B*, can result in an arbitrarily large increase in $I^{(Cl)}(\rho)$. This phenomenon can be viewed as a way of *locking classical correlation in the quantum state* ρ.

In general, the accessible information I_{acc} about an ensemble of states $\mathrm{E}=\{p_i,\eta_i\}$ is the maximum mutual information between i and the outcome of a measurement. The accessible information amount $I_{acc}(\mathrm{E})$ can be maximized by a POVM with rank 1 elements only. Let $M=\left\{\alpha_j |\phi_j\rangle\langle\phi_j|\right\}_j$ stands for a POVM with rank 1 elements where each $|\phi_j\rangle$ is normalized and $\alpha_j>0$. Then $I_{acc}(\mathrm{E})$ can be expressed as:

$$I_{acc}(\mathrm{E})=\max_M\left[-\sum_i p_i \log p_i+\sum_i\sum_j p_i\alpha_j\langle\phi_j|\eta_i|\phi_j\rangle \log\frac{p_i\langle\phi_j|\eta_i|\phi_j\rangle}{\langle\phi_j|\mu|\phi_j\rangle}\right],$$

where $\mu = \sum_i p_i \eta_i$.

Example. Let the initial state ρ is shared between subsystems held by A and B, with respective dimensions $2d$ and d, $\rho = \frac{1}{2d}\sum\sum\left(|k\rangle\langle k|\otimes|t\rangle\langle t|\right)_A \otimes \left(U_t|k\rangle\langle k|U_t^{\dagger}\right)_B$. Here $U_0 = I$ and U_1 changes the computational basis to a conjugate basis: $\forall i,k \left|\langle i|U_1|k\rangle\right| = \frac{1}{\sqrt{d}}$. In this example, B is given a random draw $|k\rangle$ from d states in two possible random bases (depending on $t=0$ or 1), while A has complete knowledge on his state.

Let consider for this case the abovementioned expression of $I_{acc}(\mathrm{E})$. For considered case the ensemble is $\left\{\frac{1}{2d}, U_t|k\rangle\right\}_{k,t}$ with $i=k,t;\ p_{k,t} = \frac{1}{2}d$, $\mu = \frac{I}{d}$ and $\langle\phi_j|\mu|\phi_j\rangle = \frac{1}{d}$.

Putting all these in the expression for $I_{acc}(\mathrm{E})$, we obtain as

$$I^{(Cl)}(\rho) = \max_M\left[\log 2d + \sum_{j,k,t}\frac{\alpha_j}{2d}\left|\langle\phi_j|U_t|k\rangle\right|^2 \log\frac{\left|\langle\phi_j|U_t|k\rangle\right|^2}{2}\right]$$
$$= \max_M\left[\log d + \sum_j \frac{\alpha_j}{d}\left(\underbrace{\frac{1}{2}\sum_{j,k,t}\left|\langle\phi_j|U_t|k\rangle\right|^2 \log\left|\langle\phi_j|U_t|k\rangle\right|^2}_{Entropies\ sum}\right)\right],$$

where $\sum_j \alpha_j = 1$ and $\forall\ j,t \quad \sum_k \left|\langle\phi_j|U_t|k\rangle\right|^2 = 1$ is used to obtain the last line. Since $\sum_j \frac{\alpha_j}{d} = 1$, the second term is convex combination, and can be upper bounded by maximization over just one term:

$$I^{(Cl)}(\rho) \le \log d + \max_{|\phi\rangle}\frac{1}{2}\sum_{k,t}\left|\langle\phi|U_t|k\rangle\right|^2 \log\left|\langle\phi|U_t|k\rangle\right|^2 .$$

Remark. The value $-\sum_{k,t}\left|\langle\phi|U_t|k\rangle\right|^2 \log\left|\langle\phi|U_t|k\rangle\right|^2$ is the sum of entropies of measuring $|\phi\rangle$ in the computational basis and the conjugate basis. Such a sum of entropies is least $\log d$. Lower bounds of these types are called entropic uncertainty inequalities (EUI), which quantify how much a vector $|\phi\rangle$ cannot be simultaneously aligned with states from two conjugated bases. It follows that $I^{(Cl)}(\rho) \le \frac{1}{2}\log d$. Equality can in fact be attained when B measures in the computational basis, so that $I^{(Cl)}(\rho) = \frac{1}{2}\log d$.

The accessible information from m independent draws of an ensemble E of separable states is additive, $I_{acc}\left(\mathrm{E}^{\otimes m}\right) = mI_{acc}(\mathrm{E})$. It follows $I^{(Cl)}\left(\rho^{\otimes m}\right) = mI_{acc}(\rho)$ for this case.

Example. If ρ is a bipartite state on $C^d \otimes C^d$, then: $Tr\left|\rho_{AB} - \rho_A \otimes \rho_B\right| \le (2d)^2\sqrt{2\ln 2 \cdot I^{(Cl)}(\rho)}$.

It means that when $I^{(Cl)}(\rho)$ is small, ρ must be close to an uncorrelated state (in trace distance).

Thus, when there are no restrictions on the allowed measurement strategy, the classical information about the identity of the state in an ensemble $\{p_x, \rho^x\}$, accessible to a measurement is limited by the Holevo bound: $I_{acc} \le S(\rho) - \sum_x p_x S(\rho^x),\ \rho = \sum_x p_x \rho^x$.

For measurements of bipartite ensembles restricted to local operations and classical communications (LOCC) is existed a universal Holevo-like upper bound on the locally accessible information. By "*locally accessible information*" always mean accessible information by LOCC-based measurements. The maximal mutual information $I(X:Y)$ accessible via LOCC between A and B satisfies the following inequality:

$$I_{acc}^{(LOCC)} \le S(\rho_A) + S(\rho_B) - \max_{Z=A,B} \sum_x p_x S(\rho_Z^x),$$

where ρ_A and ρ_B are the reductions of $\rho_{AB} = \sum_x p_x \rho_{AB}^x$, and ρ_Z^x is a reduction of ρ_{AB}^x .

Example: *Interrelations between global and local accessible information*. Consider an ensemble of signal states (not necessarily orthogonal or pure) $\{p_x, \rho_{AB}^x\}$ and pure (not necessarily orthogonal) "detector" states $\{|\phi_{CD}^x\rangle\}$.

Initially, let the signals and the detectors be in a joint state $\rho_{ABCD} = \sum p_x \rho_{AB}^x \otimes |\phi_{CD}^x\rangle\langle\phi_{CD}^x|$ with relative entropy of entanglement $E^{AC:BD}(\rho_{ABCD})$. A measurement restricted to LOCC (between A and B) and obtaining results j with probability q_j, will usually leave the detectors in mixed states $\eta_{CD}^j = \sum_x p_{x|j} |\phi_{CD}^x\rangle\langle\phi_{CD}^x|$, thus giving the accessible information

$$I_{acc}^{(LOCC)} \square H^{(Sh)} - \sum_j q_j H(\{p_{x|j}\}) \le H^{(Sh)} - \sum_j q_j S(\eta_{CD}^j)$$

(equality hold for orthogonal detectors). The general property of relative entropy of entanglement $E(\sigma_{AB}) \ge E(\sigma_A) - E(\sigma_{AB})$, implies that

$$I_{acc}^{(LOCC)} \le H^{(Sh)} + \sum_j q_j E(\eta_{CD}^j) - \sum_j q_j S(\eta_{CD}^j) \equiv H^{(Sh)} + \bar{E}_{out}^{\det} - \bar{S}_{out}^{out} .$$

As $\bar{S}_{out}^{out} = \sum_j q_j S\left(Tr_D \sum p_{x|j} |\phi_{CD}^x\rangle\langle\phi_{CD}^x|\right) \ge \sum_x p_x E(|\phi_{CD}^x\rangle) \equiv \bar{E}_{out}^{\det}$, and $\bar{E}_{out}^{\det} \le E^{AC:BD}(\rho_{ABCD})$ (as LOCC does not increase the relative entropy of entanglement $E^{AC:BD}(\rho_{ABCD})$), we obtain:

$$H^{(Sh)} - I_{acc}^{(LOCC)} \le \Delta E ,$$

where $\Delta E = \bar{E}_{out}^{\det} - E^{AC:BD}(\rho_{ABCD})$.

In the case of orthogonal ensembles, $H^{(Sh)}$ is the global accessible information (I_{acc}^{global}) and for such cases, we have: $I_{acc}^{global} - I_{acc}^{LOCC} \ge \Delta E$. Thus, in general case the difference between globally and locally accessible information for an ensemble of orthogonal (not necessarily pure) states is not less than the amount of the relative entropy of entanglement, which is created in a global measurement to distinguish the ensemble.

Connections between accessible information and quantum operations

A general quantum process cannot be operated accurately. Furthermore, an unknown state of a closed quantum system cannot be operated arbitrarily by unitary quantum operation. A quantum measurement on the state has been to identify X on the measurement result Y. A good measure of how much information has been gained about X form the measurement is the mutual information $H(X:Y)$ between X and the measurement outcome Y. The mutual information $H(X:Y)$ of X and Y measures how much information X and Y have in common. Holevo's theorem states that $H(X:Y)\le\chi$. The quantity χ is an upper bound on the accessible information. But the Holevo bound quantity decreases under quantum operations: $\chi'\left(=S(\varepsilon(\rho))-\sum_x p_x S(\varepsilon(\rho_x))\right)\le\chi\left(=S(\rho)-\sum_x p_x S(\rho_x)\right)$, where ε is a quantum operation.

Suppose we will to perform a quantum process, which can be realized by a quantum operation, ε. After a quantum process, we find the mutual information decreased, in another word, we find that the information we can gain from the quantum system is less than previous. We know the uncertainty of this quantum states increase because of the decreasing of the accessible information. That is, the result after the quantum operation is unreliable. Also, it tell us this quantum process cannot be operated accurately because if this quantum can be operated accurately, the accessible information would not decrease.

Example: *Disentanglement process*. Let us consider disentanglement process: $\varepsilon(\rho_x)\to Tr_2(\rho_{12})\otimes Tr_1(\rho_{12})$, where ρ_{12} is a pure state of two subsystems. An arbitrary state cannot be disentangled by a physical allowable process into a tensor product of its reduced density matrices. Consider ρ_{12} is a pure state, then the Holevo χ-quantity $\chi(\rho_{12})=0$. After the disentangling process, it has that the Holevo χ-quantity $\chi\left[Tr_2(\rho_{12})\otimes Tr_1(\rho_{12})\right]\ge 0$. [We can prove this by using concavity of the entropy: $S\left(\sum_i p_i\rho_i\right)\ge\sum_i p_i S(\rho_i)$]. After the disentangling quantum process, that is

$$\chi\left[Tr_2(\rho_{12})\otimes Tr_1(\rho_{12})\right]=S\left(\sum_i p_i\rho_i\right)-\sum_i p_i S(\rho_i)\ge 0,\quad \chi\left[Tr_2(\rho_{12})\otimes Tr_1(\rho_{12})\right]\ge\chi(\rho_{12}).$$

Since the equality holds iff all the states ρ_i for which $p_i>0$ are identical, we know that a general disentangling quantum states would necessarily increase the Holevo χ-quantity. Thus, it tell us a universal disentangling machine cannot exist.

Total, classical and quantum correlations: Examples of interrelations between of covariance, correlation and entanglement measures

It is important in quantum computation to determine how to separate the correlations it contains into a classical and an entangled part.

Remark. In quantum mechanics, most measurement results are represented by the trace of products of observables in certain quantum states. As example, the superposed vectors as Bell states $|\Phi^\pm\rangle,|\Psi^\pm\rangle$ are maximally entangled in the sense that they are pure on the whole system, and their marginals are maximally mixed:

$$Tr_1\left(|\Psi^\pm\rangle\langle\Psi^\pm|\right)=Tr_2\left(|\Psi^\pm\rangle\langle\Psi^\pm|\right)=Tr_1\left(|\Phi^\pm\rangle\langle\Phi^\pm|\right)=Tr_2\left(|\Phi^\pm\rangle\langle\Phi^\pm|\right)=\frac{1}{2}I.$$

We see readily a string difference between classical states and quantum states. While the marginals of classical pure states are necessarily pure, it is not the case for quantum pure states. This peculiar structure is beautifully described by Schrodinger sense:

”The best possible knowledge of a *whole* does not necessarily include the best possible knowledge of all its *parts*.”

Let us demonstrate this approach with simple examples.

Example. Consider a bipartite separable state of the form: $\rho_{AB} = \sum_i p_i |i\rangle\langle i| \otimes \rho_B^i$, where $\{|i\rangle\}$ are orthogonal states of subsystem A. Clearly the entanglement of this state is zero. The best measurement that A can make to gain information about B's subsystem is a projective measurement onto the states $\{|i\rangle\}$ of subsystem A. Consider the measure of a classical correlation as: $C_B(\rho_{AB}) = \max_{B_i^\dagger B_i} S(\rho_A) - \sum_i p_i S(\rho^i_A)$, where as above $B_i^\dagger B_i$ is a POVM performed on the subsystem B and $\rho^i_A = \dfrac{Tr_B(B_i \rho_{AB} B_i^\dagger)}{Tr_{AB}(B_i \rho_{AB} B_i^\dagger)}$ is the remaining state of A after obtaining the outcome i on B. Clearly $C_B(\rho_{AB}) = C_A(\rho_{AB})$ for all states ρ_{AB} such that $S(\rho_A) = S(\rho_B)$. The measure is a natural generalization of the classical mutual information, which is the difference in uncertainty about the subsystem $B(A)$ before and after a measurement on the correlated subsystem $A(B)$, $H(A:B) = H(B) - H(B|A)$.

Similarly, $C_B(\rho_{AB}), C_A(\rho_{AB})$ are represented the difference in von Neumann entropy before and after the measurement. These measures $C_B(\rho_{AB}), C_A(\rho_{AB})$ are non-increasing under local operations. Note the similarity of the definition to the Holevo bound which measures the capacity of quantum states for classical communication. Therefore the classical correlations are given by:

$$C_A(\rho_{AB}) = S(\rho_B) - \sum_i p_i S(\rho^i_B).$$

For this state, the mutual information is also given by: $I(\rho_{A:B}) = S(\rho_B) - \sum_i p_i S(\rho^i_B)$. This is to be expected since there are no entangled correlations and so the total correlations between A and B should be equal to the classical correlations. We now consider the relations between the classical, total and entangled correlations in some simple cases.

Example: *Maximally entangled pure state*. Let us consider a maximally pure state, $|\phi^+\rangle\langle\phi^+|$, and the family of states that interpolate between it and its completely decohered state $|0\rangle\langle 0| + |1\rangle\langle 1|$. These are states of the form: $\rho_{AB} = p|\phi^+\rangle\langle\phi^+| + (1-p)|\phi^-\rangle\langle\phi^-|$, where $\frac{1}{2} \le p \le 1$.

The mutual information as a function of p is: $I(\rho_{A:B}) = 2 + p\log p + (1-p)\log(1-p)$.
The entanglement is: $E_{RE}(\rho_{AB}) = 1 + p\log p + (1-p)\log(1-p)$.
The classical correlations remain constant at: $C_A(\rho_{AB}) = C_B(\rho_{AB}) = C(\rho_{AB}) = 1$.

This is achieved by a projective measurement onto $\{|0\rangle\langle 0|, |1\rangle\langle 1|\}$, and must be the maximum because C cannot exceed one. The total correlations for this case are just the sum of the entangled and the classical correlations, $I(\rho_{A:B}) = E_{RE}(\rho_{AB}) + C(\rho_{AB})$.

Example: *Werner state*. Consider a Werner state of the form: $\rho_{AB} = p\left|\phi^{+}\right\rangle\left\langle\phi^{+}\right| + \frac{1-p}{4}I$ with $\frac{1}{2} \le p \le 1$.

The mutual information is: $I(\rho_{A:B}) = 2 + f\log f + (1-f)\log\left(\frac{1-f}{3}\right)$, $f = \frac{3p+1}{4}$.
The relative entropy of entanglement is: $E_{RE}(\rho_{AB}) = 1 + f\log f + (1-f)\log(1-f)$.
The classical correlations remain constant at: $C_A(\rho_{AB}) = C_B(\rho_{AB}) \equiv C(\rho_{AB})$.

Any orthogonal projection produces the same value for the classical correlations.

This quantity is called as $C_P(\rho_{AB})$. Clearly, that: $C_P(\rho_{AB}) \le C(\rho_{AB})$.

Example: *Symmetric state*. The state of the form: $\rho_{AB} = p|0\rangle|0\rangle\langle 0|\langle 0| + (1-p)|+\rangle|+\rangle\langle +|$. Same as above, the state is symmetrical with regard to A and B, so: $C_A(\rho_{AB}) = C_B(\rho_{AB}) \equiv C(\rho_{AB})$.

This state provides a simple example where the states on both sides are non-orthogonal. It is not the measurement, which optimizes the classical correlations.

In these two last examples, $C_P(\rho_{AB}) + E_{RE}(\rho_{AB}) < I(\rho_{A:B})$. If the classical correlations are maximized by an orthogonal measurement on one subsystem, the classical and entangle correlations do not account for all the total correlations, and $E(\rho_{AB}) \le C(\rho_{AB})$.

Another possible of classical correlations could be based on the relative entropy (see Appendix 1).

Relative entropy as the measure of classical correlations. Just as measures of total and entangled correlations are both relative entropies, $I(\rho_{A:B}) = S(\rho_{AB}\|\rho_A \otimes \rho_B)$ and $E(\rho_{A:B}) = \min_{\sigma_{AB}\in D} S(\rho_{AB}\|\sigma_{AB})$. Classical correlations could then be given by the relative entropy between the closet separable state, $\sigma^{*}{}_{AB}$, and the product state $\rho_A \otimes \rho_B$, $C_{RE} = S(\sigma^{*}_{AB}\|\rho_A \otimes \rho_B)$.

Example. For a mixture of two Bell states, $C_{RE}(\rho_{AB})$ coincides with $C_{RE}(\rho_{AB}) = 1$. For the separable state $\rho_{AB} = p|0\rangle|0\rangle\langle 0|\langle 0| + (1-p)|+\rangle|+\rangle\langle +|$, $C_{RE}(\rho_{AB}) = I(\rho_{A:B})$, which makes sense, but there is no entanglement.

Example. For Werner state, the relative entropy of classical correlations remains constant at $C_{RE}(\rho_{AB}) = 0.2075$. Therefore for low values of p, $C_{RE}(\rho_{AB}) > E_{RE}(\rho_{AB})$, whereas for high values, $C_{RE}(\rho_{AB}) < E_{RE}(\rho_{AB})$, so that the two types of correlations do not sum to the total.

In general, we have the inequality as $I(\rho_{A:B}) > E_{RE}(\rho_{AB}) + C_{RE}(\rho_{AB})$, so that the two types of correlations do not sum to the total.

Remark. Thus, the measure of the classical correlations of a bipartite state ρ_{AB} is a distance between the nearest state ρ^{*} of ρ_{AB} and $\rho^{*}{}_A \otimes \rho^{*}{}_B$, that is relative entropy $C_1(\rho_{AB}) = S\left(\rho^{*}\|\rho^{*}{}_A \otimes \rho^{*}{}_B\right)$.

This measure of classical correlations, and similar measures based on the distance between ρ^{*} and $\rho_A \otimes \rho_B$ such as $C_2(\rho_{AB}) = S(\rho^{*}\|\rho_A \otimes \rho_B)$ are equal to the von Neumann mutual information, $I(\rho_{AB}) = S(\rho_A) + S(\rho_B) - S(\rho_{AB}) = S(\rho_{AB}\|\rho_A \otimes \rho_B)$ for a set of separable states. A measure of the classical correlations based on the maximum information that could be extracted on a subsystem B of ρ_{AB} by making a measurement on the other subsystem A described above so: $\chi_B = \max_{A_i A_i^\dagger} S(\rho_B) - \sum_i p_i S(\rho^i_B)$, is not proved to be symmetric under interchange of the subsystems A and B and even if it could increase under LOCC this should be expected for classical correlations.

Another measure of classical correlations in quantum state can be defined as the difference between total correlations measured by the von Neumann mutual information and the quantum correlations measured by the relative entropy of entanglement: $\Psi(\rho_{AB}) = S(\rho_{AB} \| \rho_A \otimes \rho_B) - \min_{\rho^* \in D} S(\rho_{AB} \| \rho^*)$, where D is the set of all separable states in the Hilbert space, on which ρ_{AB} is defined. This measure of the classical correlation is superadditive in the sense that $\Psi(\rho \otimes \rho) \geq 2\Psi(\rho)$ due to the fact that the mutual information is additive, whereas the relative entropy of entanglement is subadditive.

For pure states, all the measures of the classical correlations including this measure are equal to the von Neumann entropy of the subsystems A or B: $S(\rho_A) = S(\rho_B)$. According to the definition of the classical correlation, all the correlations contained in separable states are classical.

Covariance and entanglement. In quantum mechanics it is long recognized that there exist correlations between observables, which are much stronger than classical ones. These correlations are usually entanglement, and cannot be accounted for by classical theory.

A. *Covariance*. The notion of classical covariance of two random variables can be naturally extended to quantum mechanics, when the probability is replaced by a quantum state (density matrix), and the random variables by observables. Thus let H be the Hilbert space of a quantum system, let $L(\mathrm{H})$ be the real linear space of all observables on H. Let ρ be a quantum state (mixed state, in general), then for any two observables A and B, their covariance $Cov_\rho(A,B)$ is defined as:

$$Cov_\rho(A,B) = Tr(\rho AB) - Tr(\rho A) \cdot Tr(\rho B).$$

In particular, the variance of A in the state ρ is defined as:

$$Var_\rho A = Cov_\rho(A,A) = Tr(\rho A^2) - (Tr\rho A)^2.$$

Example. When ρ is diagonalized, in spectral decomposition form, $\rho = \sum_j \lambda_j |u_j\rangle\langle u_j|$ and ρ is non-degenerate and thus $\{|u_j\rangle\}$ constitutes an orthonormal base, we have

$$Cov_\rho(A,B) = \sum_{j,k} \lambda_j \langle u_j|A|u_k\rangle\langle u_k|B|u_j\rangle - \sum_{j,k} \lambda_j\lambda_k \langle u_j|A|u_i\rangle\langle u_k|B|u_k\rangle$$

and

$$\begin{aligned} Var_\rho A &= \sum_{j,k} \lambda_j \left|\langle u_j|A|u_k\rangle\right|^2 - \sum_{j,k} \lambda_j\lambda_k \langle u_j|A|u_i\rangle\langle u_k|A|u_k\rangle \\ &= \sum_{j,k} \frac{\lambda_j + \lambda_k}{2} \langle u_j|A|u_k\rangle - \sum_{j,k} \lambda_j\lambda_k \langle u_j|A|u_i\rangle\langle u_k|A|u_k\rangle \end{aligned}.$$

Let $[A,B] = AB - BA$ denote the commutator. Since $Cov_\rho(A,B) - Cov_\rho(B,A) = Tr(\rho[A,B])$, and $Cov_\rho(\cdot,\cdot)$ maybe viewed as an inner product, by Schwarz inequality, we have

$$Var_\rho A \cdot Var_\rho B \ge \frac{1}{4}\left|Tr\left(\rho[A,B]\right)\right|^2.$$

This is the conventional Heisenberg uncertainty relation.

B. *Interrelations between covariance and entanglement*. Physically, covariance is usually used to characterize correlations between two observables in a given quantum state. Alternatively, it can be used to characterize the intrinsic correlations of a quantum state, given the two observables fixed. The situation is usually as follows. Let there be given two quantum systems H_1 and H_2 and two observables a and b for the two quantum systems respectively. Then $A = a \otimes I_2$, $B = I_1 \otimes b$ are two observables for the composite quantum system $\mathrm{H}_1 \otimes \mathrm{H}_2$, and they commute. Now let ρ be any quantum state of the composite quantum system, then $Cov_\rho(A,B)$ maybe used as a measure to quantify the “correlation strength” of the state ρ. The observables A and B serve here as testing observables.

Example. Let $|\psi_1\rangle \in \mathrm{H}_1$ and $|\psi_2\rangle \in \mathrm{H}_2$ be two quantum states, and let the composite quantum state $|\psi\rangle = |\psi_1\rangle \otimes |\psi_2\rangle \in \mathrm{H}_1 \otimes \mathrm{H}_2$ be a product state. For $\rho = |\psi\rangle\langle\psi|$, then we have $Cov_\rho(A,B) = 0$. Let now $\mathrm{H}_1 = \mathrm{C}^2, \mathrm{H}_2 = \mathrm{C}^2$, and the composite quantum system be $\mathrm{H}_1 \otimes \mathrm{H}_2 = \mathrm{C}^2 \otimes \mathrm{C}^2$. Let $A = \sigma_z \otimes I_2$, $B = I_1 \otimes \sigma_z$. Then $\sup_{|\psi\rangle} Cov_{|\psi\rangle\langle\psi|}(A,B) = 1$, $\inf_{|\psi\rangle} Cov_{|\psi\rangle\langle\psi|}(A,B) = -1$. The maximum value is achieved iff: $|\psi\rangle = \frac{1}{\sqrt{2}}\left(e^{i\alpha}|00\rangle + e^{i\beta}|11\rangle\right)$ and the minimum is achieved iff: $|\psi\rangle = \frac{1}{\sqrt{2}}\left(e^{i\alpha}|01\rangle + e^{i\beta}|10\rangle\right)$. Here α and β are any real constants. Thus maximally entangled states as Bell states maximize the magnitude of the covariance $Cov_{|\psi\rangle\langle\psi|}(A,B)$.

The distinctions between classical and quantum correlations are fundamental and subtle, and it is a difficult and thorny problem as how to distinguish them. In this respect, the conventional covariance often gives ambiguous results. Let demonstrate this point by an example.

Example. Let as above $\mathrm{H}_1 = \mathrm{C}^2, \mathrm{H}_2 = \mathrm{C}^2$, and the composite quantum system be $\mathrm{H}_1 \otimes \mathrm{H}_2 = \mathrm{C}^2 \otimes \mathrm{C}^2 = \mathrm{C}^4$. Take a quantum state $\rho = \frac{1}{2}\left(|00\rangle\langle 00| + |11\rangle\langle 11|\right)$ and $A = \sigma_z \otimes I_2$, $B = I_1 \otimes \sigma_z$. Then in the canonical base $\{|00\rangle, |01\rangle, |10\rangle, |1\rangle\}$, we have

$$\rho = \frac{1}{2}\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \text{ and } A = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix},\ B = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}.$$

Now direct calculation leads to $Cov_\rho(A,B) = 1$. On the other hand, if we take $\rho' = \frac{1}{2}\left(|00\rangle + |11\rangle\right)\left(\langle 00| + \langle 11|\right)$, then

$$\rho' = \frac{1}{2}\begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 \end{pmatrix}.$$

We readily compute $Cov_{\rho'}(A,B)=1$.

The state ρ is a mixture of two disentangled (product) states, while ρ' is a Bell state, which is maximally entangled, and the covariance cannot distinguish them. In this sense the conventional covariance has a limited use in characterizing entanglement.

We will investigate entanglement by means of another "covariance" measure, the Wigner-Yanase-Dyson correlation, which is of an informational origin connected with Fisher information and skew information. This correlation measure has some advantages over the conventional covariance in quantifying entanglement (see Appendix 1).

Wigner-Yanase-Dyson (WYD) skew information measure

In the study of measurement theory from an information-theoretic point of view, Wigner and Yanase introduced the quantity:

$$I(\rho,K)=-\frac{1}{2}Tr\left(\left[\sqrt{\rho},K\right]^2\right),$$

which they called skew information (the bracket $[\cdot,\cdot]$ denotes commutator), as the amount of information on the values of observables not commuting with K (which may be a Hamiltonian, a moment, or other conserved quantity). Alternatively, $I(\rho,K)$ may be interpreted as a measure of non-commutativity between ρ and K with asymmetric emphasis on the state ρ and on the conserved observable K.

Properties of the skew information are as following: (i) constant for isolated systems; (ii) decreases when two different ensembles are united that means the information content of the resulting ensemble should be smaller than the average information component of the component ensemble; and (iii) is additive, namely, the information content of two independent pairs is the sum of information of the parts.

The skew information is later on generalized by Dyson to

$$I_\alpha(\rho,K)=-\frac{1}{2}Tr\left(\left[\rho^\alpha,K\right]\left[\rho^{1-\alpha},K\right]\right),\ 0<\alpha<1.$$

The WYD conjecture concerning the convexity of $I_\alpha(\rho,K)$.

The Wigner-Yanase skew information can be rewritten as following:

$$I(\rho,K)=-\frac{1}{2}Tr\left(\left[\sqrt{\rho},K\right]^2\right)=Tr\left(\rho K^2\right)-Tr\left(\sqrt{\rho}K\sqrt{\rho}K\right).$$

In particular case, if $\rho=|\psi\rangle\langle\psi|$ is a pure state, then

$$I(\rho,K)=\Delta_\rho K,\quad \Delta_\rho K=Tr\left(\rho K^2\right)-\left(Tr\left(\rho K\right)\right)^2.$$

Here $\Delta_\rho K$ is the variance of the observable K in the state ρ.

Therefore, for pure states, the Wigner-Yanase skew information reduces to variance. When ρ is a mixed state, we have: $\Delta_\rho K\ge I(\rho,K)$.

A. *Interrelations between quantum Fisher's information and quantum Wigner-Yanase skew information*. In fact, the notion of skew information is very similar to the well-known notion of Fisher information originated from statistical inference. Among concepts describing contents of quantum mechanical density operators, both the Wigner-Yanase skew information and the quantum Fisher information defined via symmetric logarithmic derivatives are natural generalizations of the classical Fisher information.

We will establish a relationship between these fundamental quantities.

Recall that the Fisher information of a parameterized family of probability densities $\{p_\theta : \theta \in \mathrm{R}\}$ on R is defined as

$$I_F(p_\theta) = 4\int_{\mathrm{R}} \left(\frac{\partial \sqrt{p_\theta(x)}}{\partial \theta} \right)^2 dx = \frac{1}{4}\int_{\mathrm{R}} \left(\frac{\partial \log p_\theta(x)}{\partial \theta} \right)^2 p_\theta(x)\, dx$$

is intimately related to the Shannon entropy.

Remark. When we pass from classical theory to quantum mechanics, the integral is replaced by trace, and the probability densities are replaced by density operators.

The Fisher information for a family of quantum states ρ_θ is defined as

$$I_F(\rho_\theta) = 4Tr\left(\frac{\partial \sqrt{\rho_\theta}}{\partial \theta} \right)^2,$$

which may be viewed as a generalization of the classical Fisher information to quantum case.

In particular, if ρ_θ satisfies the Landau-von Neumann equation:

$$i\frac{\partial \rho_\theta}{\partial \theta} = [K, \rho_\theta], \quad \rho_0 = \rho,$$

where $\theta \in \mathrm{R}$ is a (temporal or spatial) parameter, and K may be interpreted as the generator of the temporal shift or the spatial displacement, then $\rho_\theta = e^{-i\theta K} \rho e^{i\theta K}$, and

$$\frac{\partial \sqrt{\rho_\theta}}{\partial \theta} = ie^{-i\theta K}\left[\sqrt{\rho}, K\right] e^{i\theta K},$$

which in turn implies that $I_F(\rho_\theta) = 8I(\rho, K)$.

Therefore, the skew information is essentially a particular kind of quantum Fisher information. In general case, the quantum Fisher information and the Wigner-Yanase information are related by inequalities:

$$I_W(\rho, K) \le I_F(\rho, K) \le 2I_W(\rho, K).$$

Example: *Two-level quantum system*. The quantum state Hilbert space is C^2. A general density operator ρ on C^2 for some $r = (r_1, r_2, r_3) \in \mathrm{R}^3, |r| = \sqrt{r_1^2 + r_2^2 + r_3^2} \le 1$ is of the form: $\rho = \frac{1}{2}\begin{pmatrix} 1+r_3 & r_1 - ir_2 \\ r_1 + ir_2 & 1-r_3 \end{pmatrix}$.

The eigenvalues of ρ are $\lambda_1 = \frac{1-|r|}{2}$ and $\lambda_2 = \frac{1+|r|}{2}$. Let the corresponding eigenvectors be ψ_1 and ψ_2, then $\rho = \lambda_1 |\psi_1\rangle\langle\psi_1| + \lambda_2 |\psi_2\rangle\langle\psi_2|$. Consequently,

$$I_W(\rho, K) = \frac{1}{2}\left(1 - \sqrt{1-|r|^2}\right)\left|\langle\psi_1|K|\psi_2\rangle\right|^2, I_F(\rho, K) = \frac{1}{2}|r|^2 \left|\langle\psi_1|K|\psi_2\rangle\right|^2$$

and

$$I_F(\rho,K)=\left(1-\sqrt{1-|r|^2}\right)I_W(\rho,K).$$

Thus $I_F(\rho,K)$ may vary continuously from $I_W(\rho,K)$ to $2I_W(\rho,K)$. Moreover, in this case, if $\langle\psi_1|K|\psi_2\rangle\neq 0$ and K does not commute with ρ, then $I_F(\rho,K)=I_W(\rho,K)$ iff $|r|=1$, that is, ρ is a pure state.

B. *The Wigner-Yanase correlation*. Motivated by Fisher information a skew information, the Wigner-Yanase correlation can be introduced as following (in analogue with the measure of correlation $\Delta_\rho K$):

$$Corr_\rho(A,B)=Tr(\rho AB)-Tr\left(\sqrt{\rho}A\sqrt{\rho}B\right).$$

In particular,

$$I(\rho,K)\equiv Corr_\rho(A,A)=Tr\left(\rho A^2\right)-Tr\left(\sqrt{\rho}A\right)^2=-\frac{1}{2}Tr\left[\sqrt{\rho},A\right]^2$$

is exactly the skew information introduced by Wigner and Yanase.

Since $Corr_\rho(A,B)-Corr_\rho(B,A)=Tr[A,B]$ and $Corr_\rho(\cdot,\cdot)$ maybe viewed as an inner product, by the Schwarz inequality, we also have:

$$I(\rho,A)\cdot I(\rho,B)\geq\frac{1}{4}\left|Tr[A,B]\right|^2.$$

This inequality is more strong than the Heisenberg uncertainty relation since: $Var_\rho A\geq I(\rho,A)$. Thus, just like the Fisher information, the Wigner-Yanase correlation can also to improve the conventional Heisenberg uncertainty relations (see Appendix 2).

C. *The Wigner-Yanase correlation and covariance*. The conventional covariance and the Wigner-Yanase correlation are two different correlation measures: Wigner-Yanase correlation has more information character and there are some intrinsic relations between them.

The inequality $Cov_\rho(A,A)\geq Corr_\rho(A,A)$ is true, but $\left|Cov_\rho(A,A)\right| \lesseqgtr \left|Corr_\rho(A,A)\right|$ and it is not true in general: the magnitude of the conventional covariance can be less, or large, than the Wigner-Yanase correlation.

Example. Let

$$\rho=\begin{pmatrix}\lambda_1 & 0\\ 0 & \lambda_2\end{pmatrix},\quad A=\begin{pmatrix}a_{11} & a\\ \bar{a} & a_{22}\end{pmatrix},\quad B=\begin{pmatrix}b_{11} & b\\ \bar{b} & b_{22}\end{pmatrix}.$$

Here $\lambda_1+\lambda_2=1,\ \lambda_1\geq 0,\ \lambda_2\geq 0$ and $a_{11},a_{22},b_{11},b_{22}$ are all real numbers, while a and b may be complex. Then

$Cov_\rho(A,B)=\lambda_1 a\bar{b}+\lambda_2\bar{a}b+\lambda_1\lambda_2(a_{11}-a_{22})(b_{11}-b_{22})$
$Corr_\rho(A,B)=\lambda_1 a\bar{b}+\lambda_2\bar{a}b-\sqrt{\lambda_1\lambda_2}\left(a\bar{b}+\bar{a}b\right)$

In particular, if a and b are real, then $\left|Cov_\rho(A,A)\right|\geq\left|Corr_\rho(A,A)\right|$ iff

$$\lambda_1\lambda_2(a_{11}-a_{22})(b_{11}-b_{22})+2\sqrt{\lambda_1\lambda_2}ab\geq 0.$$

However, if ρ is a pure state, these two correlation measures coincide: $Cov_\rho\left(A,B\right)=Corr_\rho\left(A,B\right)$ for any observables A and B.

D. *The Wigner-Yanase correlation and entanglement.* Let $\rho,\ \rho'$ and A,B be the same as above:

$$\rho=\frac{1}{2}\begin{pmatrix}1&0&0&0\\0&0&0&0\\0&0&0&0\\0&0&0&1\end{pmatrix},\ \rho'=\frac{1}{2}\begin{pmatrix}1&0&0&1\\0&0&0&0\\0&0&0&0\\1&0&0&1\end{pmatrix}.$$

Then we may readily compute $Corr_\rho\left(A,B\right)=0,\ Corr_{\rho'}\left(A,B\right)=1$. Thus the Wigner-Yanase correlation indeed distinguishes between the mixture of disentangled states (classical correlation) and the maximally entangled Bell states (quantum correlation).

Example: $H_1=C^2, H_2=C^2$ and the composite quantum system $H_1\otimes H_2=C^2\otimes C^2=C^4$. Take a quantum state

$$\rho=\frac{1}{2}\left|00\right\rangle\left\langle 00\right|+\frac{1}{2}\left|\Phi^+\right\rangle\left\langle\Phi^+\right|,\left|\Phi^+\right\rangle=\frac{1}{\sqrt{2}}\left(\left|00\right\rangle+\left|11\right\rangle\right).$$

Let us as above $A=\sigma_z\otimes I_2,\quad B=I_1\otimes\sigma_z$. Then in the canonical base $\left\{\left|00\right\rangle,\left|01\right\rangle,\left|10\right\rangle,\left|11\right\rangle\right\}$, we have

$$\rho=\frac{1}{4}\begin{pmatrix}3&0&0&0\\0&0&0&0\\0&0&0&0\\0&0&0&1\end{pmatrix}.$$

The calculation leads to $Tr\left(\rho A\right)=Tr\left(\rho B\right)=\frac{1}{2},\quad Tr\left(\rho AB\right)=1$, thus $Cov_\rho\left(A,B\right)=\frac{3}{4}=0.75$.

On the other hand, ρ can be diagonalized as

$$\rho=U\begin{pmatrix}\frac{\sqrt{2}+1}{2\sqrt{2}}&0&0&0\\0&0&0&0\\0&0&0&0\\0&0&0&\frac{\sqrt{2}-1}{2\sqrt{2}}\end{pmatrix}U^{-1}$$

with unitary matrix

$$U=\begin{pmatrix} \frac{1}{\sqrt{4-2\sqrt{2}}} & 0 & 0 & \frac{1}{\sqrt{4+2\sqrt{2}}} \\ 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ 0 & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 0 \\ \frac{\sqrt{2}-1}{\sqrt{4-2\sqrt{2}}} & 0 & 0 & -\frac{\sqrt{2}+1}{\sqrt{4+2\sqrt{2}}} \end{pmatrix}.$$

Thus,

$$\sqrt{\rho}=\begin{pmatrix} \left(\frac{\sqrt{2}+1}{2\sqrt{2}}\right)^{\frac{3}{2}}+\left(\frac{\sqrt{2}-1}{2\sqrt{2}}\right)^{\frac{3}{2}} & 0 & 0 & \frac{\left(\sqrt{2}+1\right)^{\frac{1}{2}}}{\left(2\sqrt{2}\right)^{\frac{3}{2}}}-\frac{\left(\sqrt{2}-1\right)^{\frac{1}{2}}}{\left(2\sqrt{2}\right)^{\frac{3}{2}}} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ \frac{\left(\sqrt{2}+1\right)^{\frac{1}{2}}}{\left(2\sqrt{2}\right)^{\frac{3}{2}}}-\frac{\left(\sqrt{2}-1\right)^{\frac{1}{2}}}{\left(2\sqrt{2}\right)^{\frac{3}{2}}} & 0 & 0 & \frac{\left(\sqrt{2}+1\right)^{\frac{1}{2}}}{\left(2\sqrt{2}\right)^{\frac{3}{2}}}+\frac{\left(\sqrt{2}-1\right)^{\frac{1}{2}}}{\left(2\sqrt{2}\right)^{\frac{3}{2}}} \end{pmatrix}.$$

Simple calculation leads to: $Corr_{\rho}\left(A,B\right)=\frac{1}{2}-\frac{\sqrt{2}}{4}\approx 0.15$.

This is in sharp contrast with the covariance: $Cov_{\rho}\left(A,B\right)=0.75$. Indeed, the state ρ is half mixture of a disentangled state $\left|00\right\rangle\left\langle 00\right|$ and an entangled state: $\left|\Phi^{+}\right\rangle\left\langle\Phi^{+}\right|$.

Thus the entanglement in ρ should be less than 0.5 if we take $\left|\Phi^{+}\right\rangle$ (a bell state) as a state with a unit of entanglement (indeed, $Cov_{\rho=\left|\Phi^{+}\right\rangle\left\langle\Phi^{+}\right|}\left(A,B\right)=Corr_{\rho=\left|\Phi^{+}\right\rangle\left\langle\Phi^{+}\right|}\left(A,B\right)=1$) and assume reasonable that classical mixing will reduce entanglement (mixing the disentangled state $\left|00\right\rangle\left\langle 00\right|$ with the entangled state $\left|\Phi^{+}\right\rangle\left\langle\Phi^{+}\right|$ will corrupt the entanglement).

Efficient measurements and bounds on the accessible mutual information

In general case, when a measurement is made on a quantum system in which classical information is encoded, the measurement reduced the observer's average Shannon entropy for the encoding ensemble. This reduction, being the mutual information, is always non-negative. For efficient measurements the state is also purified; that is, on average, the observer's von Neumann entropy for the case of the system is also reduced by a non-negative amount.

By rewritten a bound derived by Hall, which is dual to the Holevo bound, one finds that for efficient measurements, the mutual information is bounded by the reduction in the von Neumann entropy.

A. *Hall*'s *dual Holevo bound* is as following:

$$\Delta I_{in} \equiv H\left(I:J\right) \le S\left(\rho\right) - \sum_j Q_j S\left(\frac{\sqrt{\rho}\,\mathrm{U}_j\sqrt{\rho}}{Q_j}\right); \quad Q_j = Tr\left[\mathrm{U}_j\rho\right]; \mathrm{U}_j = A_j^{\dagger}A_j$$

where Q_j is the probability that outcome j will result; each of operators A_j corresponds to the measurement outcome, and the outcomes are therefore labeled by j.

B. *Schumacher-Westmoreland-Wootters (SWW) bound*. The Holevo bound and Hall's bound, may both be derived from the more general SWW-bound

$$\Delta I_{in} \le S\left(\rho\right) - \sum_i p_i S\left(\rho_i\right) - \sum_j Q_j \underbrace{\left[S\left(\rho_i'\right) - \sum_i p_{i|j} S\left(\rho_{ij}'\right)\right]}_{\chi-quantity},$$

where all quantities are as defined above, and the quantity ρ_{ij}' is introduced, which is the *final* state that the receiver would have had, if he knew that the *initial* state was ρ_i.

Thus, $\rho_{ij}' = \dfrac{A_j\rho_i A_j^{\dagger}}{Q\left(j|i\right)}$, where $Q\left(j|i\right)$ is naturally the probability density for the measurement outcomes, given that the initial state is ρ_i. Because of the final term on the right-hand side of this inequality, this bound is, in general, stronger than the Holevo bound.

Remark. The expression in the square brackets $\left[S\left(\rho_i'\right) - \sum_i p_{i|j} S\left(\rho_{ij}'\right)\right]$ is the Holevo χ quantity for the ensemble ε_j, which results from measurement outcome j. Thus their bound may be written as: $\Delta I_{in} \le \chi\left[\varepsilon\right] - \sum_j Q_j \chi\left[\varepsilon_j\right]$. Now, $\chi\left[\varepsilon_j\right]$ is the Holevo bound on the information that the receiver could extract when making a subsequent measurement after obtaining result j.

While ΔI_{in} quantities the information which the observer obtains about the initial preparation, there exists another quantity which can be said to characterize the average amount of information which receiver obtains about the final state which he is left with after the measurement. Denote this by ΔI_{fin}, expression for which is as:

$$\Delta I_{fin} = S\left(\rho\right) - \sum_j Q_j S\left(\rho_j'\right).$$

This is the average difference between the receiver's initial von Neumann entropy of the quantum system, and his final von Neumann entropy. A more fundamental difference between ΔI_{in} and ΔI_{fin} is that the former is the average change in the observer Shannon entropy regarding the ensemble, where as the latter is the average change in the observers von Neumann entropy regarding the overall state of the quantum system.

That is, for any ensemble ε, $\Delta I_{in} = \left\langle \Delta H\left(\varepsilon\right)\right\rangle$ and $\Delta I_{fin} = \left\langle \Delta S\left(\varepsilon\right)\right\rangle$, and the result is $\Delta I_{in} \le \Delta I_{fin}$.

One can interpret this as saying that the observer cannot learn more about the classical information encoded in a quantum system than he learns about the state of the quantum system. This provides a physical interpretation for Hall's bound.

Further, this bound can only be saturated when all operators $\mathrm{U}_j = A_j^{\dagger}A_j$ commute.

Remark. One consequence of the relation $\Delta I_{in} \le \Delta I_{fin}$ is that, if we choose an ensemble, which has the maximal accessible information for a fixed ρ, we can only obtain all this information if all the final states are pure. Measurements, which leave the final state impure, leave some information in the system. That is, if the final state is mixed, in general it depends on the initial ensemble, and as the result subsequent measurements can obtain further information about the initial preparation, whereas this is not possible if the final state is pure.

For a given ρ not all ensembles have an accessible information equal to $S(\rho)$. In fact, this is possible if the encoding satisfies special conditions; in general, incomplete measurements will not even extract the accessible information from an ensemble. Consider the final states, ρ'_j, which result from the measurement. Each of these consists of an ensemble, ε_j over the states $\rho_{i|j}$, and $\rho'_j = \sum_i p(i|j)\rho_{i|j}$.

Since these ensembles consist of states indexed by $\boldsymbol{i}$, they can, in general, be measured to obtain further information about the initial preparation.

Since the accessible information is the maximal amount of information that can be obtained about i by making measurements, we have the inequality:

$$\Delta I_{in}(\varepsilon, \mathrm{U}) \le \Delta I_{acc}(\varepsilon) - \sum_j Q_j \Delta I_{acc}(\varepsilon_j).$$

Thus, the amount of extracted information by measurement U, $\Delta I_{in}(\varepsilon, \mathrm{U})$, can only be equal to $\Delta I_{acc}(\varepsilon)$ if the amount $\Delta I_{acc}(\varepsilon_j)$ are zero for all j. If ρ'_j is pure, then $\Delta I_{acc}(\varepsilon_j)$ is zero. If ρ'_j is not pure, then the accessible information of ε_j is only zero if, for any given j, the $\rho_{i|j}$ are the same for all i.

The SWW bound shows that, if the initial ensemble ε is chosen so that its accessible information is maximal [i.e., equal to $\chi[\varepsilon]$], then the information obtained by an incomplete measurement will be reduced by the maximal amount of information which could be accessible from the final ensemble ε_j, and not merely the actual information available in these ensembles, which would imply the bound given in:

$$\Delta I_{in}(\varepsilon, \mathrm{U}) \le \Delta I_{acc}(\varepsilon) - \sum_j Q_j \Delta I_{acc}(\varepsilon_j).$$

In general, there is a gap between the information lacking in an incomplete measurement, and that which can be recovered by subsequent measurements: no matter what incomplete measurement is performed on it, the information which is not retrieved by the measurement can always be extracted by subsequent measurements.

Remark. For inefficient quantum measurements, however, the inequality $\Delta I_{in} \le \Delta I_{fin}$ does not hold. The reason for this is that for inefficient measurements ΔI_{fin} can be negative (whereas ΔI_{in} is always non-negative). An example of such a situation is one in which the initial state ρ is not maximally mixed, and the observer performs a von Neumann measurement in a basis unbiased with respect to eigenbasis of ρ. If the observer has no knowledge of the outcome, then his final state is maximally mixed. Further, if one mixes this measurement with one whose measurement operators commute with ρ, it is not hard to obtain a measurement in which both ΔI_{in} and ΔI_{fin} are positive, but which violates the inequality $\Delta I_{in} \le \Delta I_{fin}$.

These quantities are useful when considering quantum state preparation and, more general quantum feedback control.

Information (communication) capacity of quantum computing

Any computation (both classical and quantum) is formally identical to a communication in time. By considering quantum computation as a communication process, we relate its efficiency to its classical communication capacity. At time $t = 0$, the programmer sets the computer to accomplish any one of several possible tasks. Each of these tasks can be regarded as embodying a different message. Another programmer can obtain this message by looking at the output of the computer when the computation is finished at time $t = T$. Computation based on quantum principles allows for more efficient algorithms for solving certain problems than algorithms based on pure classical principles. The classical capacity of a quantum communication channel is connected with the efficiency of quantum computing using *entropic* arguments.

This formalism allows us to derive lower bounds on the computational complexity of quantum control and search algorithms in the most general context.

Communication model of quantum computing

In this model two programmers (the sender and the receiver) and two registers (the memory ($\underline{M}$) register and the computational ($\underline{C}$) register) are applied. The sender prepares the memory register in a certain quantum state $|i\rangle_{\underline{M}}$, which encodes the problem to be solved. For example, in the case of search, this register will store the state of the list to be searched. The number N of possible states $|i\rangle_{\underline{M}}$ will be limited by the target list that the given computer could search. The receiver then prepares the computational register in some initial state $\rho_{\underline{C}}^{0}$. Both the sender and the receiver feed the registers (prepare by them) to the quantum computer. The quantum computer implements the following general transformation on the registers:

$$\left(|i\rangle\langle i|\right)_{\underline{M}} \otimes \rho_{\underline{C}}^{0} \to \left(|i\rangle\langle i|\right)_{\underline{M}} \otimes U_i \rho_{\underline{C}}^{0} U_i^{\dagger},$$

where $\rho_{\underline{C}}(i) = U_i \rho_{\underline{C}}^{0} U_i^{\dagger}$ is the resulting state of the computational register that contains the answer to the computational basis and is measured by the receiver according to measurement basis (as example, $(|0\rangle, |1\rangle)$ or $\left(|0'\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle), |1'\rangle \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle)\right)$). The quantum computation should also work for any mixture $\sum_{i=1}^{N} p_i \left(|i\rangle\langle i|\right)_{\underline{M}}$, where p_i are probabilities. For the sender to use the above computation as a communication protocol, he has to prepare any one of the state $|i\rangle_{\underline{M}}$ with an *a priori* probability p_i . The entire input ensemble is thus: $\sum_{i=1}^{N} p_i \left(|i\rangle\langle i|\right)_{\underline{M}} \otimes \rho_{\underline{C}}^{0}$. Because of the quantum computation, this becomes

$$\sum_{i=1}^{N} p_i \left(|i\rangle\langle i|\right)_{\underline{M}} \otimes \rho_{\underline{C}}^{0} \to \sum_{i=1}^{N} p_i \left(|i\rangle\langle i|\right)_{\underline{M}} \otimes \rho_{\underline{C}}(i).$$

Whereas before the quantum computation, the two registers were completely uncorrelated (the amount of mutual information is zero), at the end, the amount of the mutual information becomes: $I_{\underline{MC}} := S(\rho_{\underline{M}}) + S(\rho_{\underline{C}}) - S(\rho_{\underline{MC}}) = S(\rho_{\underline{C}}) - \sum p_i S(\rho_{\underline{C}}(i))$, where $\rho_{\underline{M}}$ and $\rho_{\underline{C}}$ are the reduced density operators for the two registers, $\rho_{\underline{MC}}$ is the density operator of the entire $(M + C)-$system, and $S(\rho) = -Tr(\rho \log \rho)$ is the von Neumann entropy.

Remark. The sender conveys the maximum information when all the message states have equal *a priori* probability (which also maximizes the channel capacity). In that case the mutual information (channel capacity) at the end of the computation is $\log N$.

Thus, the communication capacity $I_{\underline{MC}}$, defined above, gives an index of efficiency of a quantum computation: *A necessary target of a quantum computation is to achieve the maximum possible communication capacity consistent with given initial states of the quantum computing*.

Remark. If one breaks down the general unitary transformation U_i of a quantum algorithm into a number of successive unitary blocks, then the maximum capacity may be achieved only after the number of applications of the blocks. In each of the smaller unitary blocks the mutual information between $\underline{M}$ and $\underline{C}$ registers (i.e., the communication capacity) increases by a certain amount. When its total value reaches the maximum possible value consistent with a given initial state of the quantum computing, the computation is regarded as being complete.

Application of information formalism to any general quantum search algorithm

Any general quantum algorithm has to have a certain number of queries into the memory register. This is necessitated by the fact that the transformation on the computational register has to depend on the problem at hand, encoded in the state $|i\rangle_{\underline{M}}$. These queries are considered to be implemented by a block box into which the states of both the memory and the computational registers are fed. The number of such queries (needed in a certain quantum algorithm) gives the black box complexity of that algorithm and is a lower bound on the complexity of the whole algorithm.

If the memory register was prepared initially in the superposition $\sum_{i=1}^{N}|i\rangle_{\underline{M}}$, then, in a search algorithm, $O\left(\sqrt{N}\right)$ queries would needed to completely entangle it with the computational register (*Ambainis*, 2000). This gives a lower bound on the number of queries in a search algorithm. We can calculate the change in mutual information between the memory and the computational registers in one query step. The number of queries needed to increase the mutual information to *log N* (the perfect communication between the sender and the receiver), is then a lower bound on the complexity of the algorithm.

Any search algorithm (whether quantum or classical) will have to find a match for the state $|i\rangle_{\underline{M}}$ of the $\underline{M}$ register among the state $|j\rangle_{\underline{C}}$ of the $\underline{C}$ register and associate a marker to the state that matches (here we suppose that $|j\rangle_{\underline{C}}$ is a complete orthonormal basis for the $\underline{C}$ register). The most general way of doing such a query in the quantum case is the black box unitary transformation:

$$U_{\underline{B}}|i\rangle_{\underline{M}}|j\rangle_{\underline{C}} = (-1)^{\delta_{ij}}|i\rangle_{\underline{M}}|j\rangle_{\underline{C}}.$$

Any other unitary transformation performing a query matching the states of the $\underline{M}$ and the $\underline{C}$ registers could be constructed from the above type of queries.

We can to put a bound on the change of the mutual information in one such black box step. Let the memory states $|i\rangle_{\underline{M}}$ be available to the sender with equal *a priori* probability so that the communication capacity is a maximum. The initial ensemble of the sender then is $\frac{1}{N}\sum_{i=1}^{N}\left(|i\rangle\langle i|\right)_{\underline{M}}$. Let the receiver prepare the register $\underline{C}$ in an initial pure state $|\psi_0\rangle$. In fact, the power of quantum computing stems from the ability of the receiver to prepare pure superposition of form $\frac{1}{\sqrt{N}}\sum_{j=1}^{N}|j\rangle_{\underline{C}}$. This is an equal weight superposition of all $|j\rangle_{\underline{C}}$.

This can be done by performing a Hadamard transformation to each qubit of the $\underline{C}$ register. In general, there will be many black box steps on the initial ensemble before a perfect correlation is set up between the $\underline{M}$ and the $\underline{C}$ registers. Let, after the $k-th$ black box step, the state of the system be: $\rho^k = \frac{1}{N}\sum_{i=1}^{N}\left(|i\rangle\langle i|\right)_{\underline{M}} \otimes \left[\,\left|\psi^k(i)\right\rangle\left\langle\psi^k(i)\right|\,\right]_{\underline{C}}$, where $\left|\psi^k(i)\right\rangle_{\underline{C}} = \sum_j \alpha_{ij}^k \,|\, j\rangle_{\underline{C}}$.

The $(k+1)-th$ black box step changes this state to

$$\rho^{k+1} = \frac{1}{N}\sum_{i=1}^{N}\left(|i\rangle\langle i|\right)_{\underline{M}} \otimes \left[\,\left|\psi^{k+1}(i)\right\rangle\left\langle\psi^{k+1}(i)\right|\,\right]_{\underline{C}} \text{ with} \left|\psi^{k+1}(i)\right\rangle_{\underline{C}} = \sum_j \alpha_{ij}^k (-1)^{\delta_{ij}} \,|\, j\rangle_{\underline{C}}.$$

Thus, *only the difference of mutual information* between the $\underline{M}$ and the $\underline{C}$ registers for states are evaluated.

Remark. This difference of mutual information can be defined as follows. The amount of information lost may be quantified by the difference in mutual information between the respective states. Mutual information is a measure of correlation between the memory $\underline{M}$ and the $\underline{C}$ registers, giving the amount of information about the $\underline{C}$ register, which may be obtained from a measurement on the $\underline{M}$ register. The quantum mutual information between the $\underline{M}$ and the $\underline{C}$ registers is defined as above. The mutual information loss for the step k is $\Delta I_k = S\left(\rho_{\underline{C}}{}^k\right)$. Then the difference of mutual information between the step k and the step $k+1$ can be shown to be the difference: $\left|\Delta I_{k+1} - \Delta I_k\right| = \left|S\left(\rho_{\underline{C}}{}^{k+1}\right) - S\left(\rho_{\underline{C}}{}^k\right)\right|$.

Remark. To understand how the entanglement (quantum correlation) between the $\underline{M}$ and the $\underline{C}$ registers varies as we vary the density matrix for the combined system, we need to introduce some distance measures on density matrices. We will make use of three closely related distance measures: the *trace distance* $Tr(\rho,\sigma)$, the *fidelity* $F(\rho,\sigma)$, and the *Bures distance* $D(\rho,\sigma)$. These distances between density matrices ρ and σ are defined to be as follows:

$T(\rho,\sigma)$	$=$	$\lvert\rho-\sigma\rvert$
$F(\rho,\sigma)$	$=$	$\sqrt{\sqrt{\rho}\sigma\sqrt{\rho}}$
$D_1(\rho,\sigma)$	$=$	$2\sqrt{1-F(\rho,\sigma)}$
$D_2(\rho,\sigma)$	$=$	$\sqrt{1-F^2(\rho,\sigma)}$

where we define $|A| \equiv \sqrt{A^\dagger A}$ to be the positive square root of $A^\dagger A$.

The trace distance $Tr(\rho,\sigma)$ is a metric on the space of density matrices and it is nonincreasing under quantum operations: $T\left(L(\rho),L(\sigma)\right) \le T(\rho,\sigma)$ for all density matrices. The Bures distance $D(\rho,\sigma)$ is also to be a metric on the space of density matrices and agrees with the trace distance for pure states. Bures metric does not increase under general complete positive maps (which is what the query represents when we trace out the $\underline{M}$ register). For our purpose, it is especially important to note that this is true for the case where L is a partial trace operation, as the partial trace is a trace – preserving quantum operation. The fidelity $F(\rho,\sigma)$ is not a metric and for pure states $|\psi\rangle$ and $|\phi\rangle$ reduces to the overlap between the states,

$$F(\psi,\phi) = \left|\langle\psi|\sigma\rangle\right|.$$

It was proved a useful continuity relation relating trace distance and entropy.

Fannes' inequality states that for any density matrices ρ and σ such that

$$Tr(\rho,\sigma) \le \frac{1}{e}, \qquad |S(\rho)-S(\sigma)| \le T(\rho,\sigma)\log d + h\left(\left(T(\rho,\sigma)\right)\right),$$

where d is the dimensional of the Hilbert space, and $h(x) = -x\log x$.

Recall that $T(\rho,\sigma) \le D(\rho,\sigma)$ the quantity $|\Delta I_{k+1} - \Delta I_k| = \left|S\left(\rho_{\underline{C}}{}^{k+1}\right) - S\left(\rho_{\underline{C}}{}^{k}\right)\right|$ is bounded from Fannes' inequality by

$$\left|S\left(\rho_{\underline{C}}{}^{k+1}\right) - S\left(\rho_{\underline{C}}{}^{k}\right)\right| \le D_2\left(\rho_{\underline{C}}^{k}, \rho_{\underline{C}}^{k+1}\right)\log N - D_2\left(\rho_{\underline{C}}^{k}, \rho_{\underline{C}}^{k+1}\right)\log D_2\left(\rho_{\underline{C}}^{k}, \rho_{\underline{C}}^{k+1}\right).$$

It can be shown that $F^2\left(\rho_{\underline{C}}^{0}, \rho_{\underline{C}}^{1}\right) \ge \frac{N-2}{N}$ from which it follows that the change in the first step:

$$\left|S\left(\rho_{\underline{C}}{}^{1}\right) - S\left(\rho_{\underline{C}}{}^{0}\right)\right| \le \frac{3}{\sqrt{N}}\log N.$$

The change $\left|S\left(\rho_{\underline{C}}{}^{k+1}\right) - S\left(\rho_{\underline{C}}{}^{k}\right)\right|$ in the subsequent steps has to be less than or equal to the change in the first step. This is because Bures metric does not increase under general complete positive maps as above is mentioned. Any other operation performed only on the $\underline{C}$ register in between two queries can only reduce the mutual information between the $\underline{M}$ and the $\underline{C}$ registers. This means that at least $O\left(\sqrt{N}\right)$ steps are needed to produce full correlations (maximum mutual information of value $\log N$ as a measure of a maximum entanglement) between two registers. This gives the black box lower bound on the complexity of any quantum algorithms.

Intelligent coherent states with minimum uncertainty and maximal information

The minimum-uncertainty coherent states (as example, for the harmonic-oscillator potential) can be defined as those states that minimize the uncertainty relation of Heisenberg (leading to the equality in the uncertainty relations), subject to the added constraint that the ground state is a member of the set. They are considered to be as close as possible to the classical states. Beyond the harmonic-oscillator system, coherent states have also been developed for quantum (Schrodinger) systems with general potentials and for general Lie symmetries. These states are called (general) minimum-uncertainty coherent states and (general) displacement-operator coherent states. There is also a different generalization of the coherent states of the harmonic-oscillator system. This is the concept of “squeezed” states. (Squeezing is a reduction of quadrature fluctuations below the level associated with the vacuum.)

A. The *even* and *odd* coherent states for one-mode harmonic oscillator were introduced in 1970s. These states, which have been called Schrodinger cat states, were studied in detail. These states are representatives of non-classical states. Schrodinger cat states have properties similar to those of the squeezed states, i.e. the squeezed vacuum state and the even coherent state contain Fock states with an even number of photons.

Definition: Intelligent states are quantum states, which satisfy the equality in the uncertainty relation for non-commuting observables.

In quantum mechanics two non-commuting observables cannot be simultaneously measured with arbitrary precision. This fact, often called the Heisenberg uncertainty principle, is a fundamental restriction that is related neither to imperfection of the existing real-life measuring devices nor to the experimental errors of observation. It is rather the intrinsic property of the quantum states itself.

The uncertainty principle provides (paradoxically enough) the only way to avoid many interpretation problems. The uncertainty principle specified for given pairs of observables finds its mathematical manifestation as the uncertainty relations. The first rigorous derivation of the uncertainty relation from the basic non-commuting observables (i.e., for the position and moment, $[\hat{x}, \hat{p}] = i\hbar$) is due to Kennard (1927). This derivation (repeated in most textbooks on quantum mechanics ever since) leads to the inequality: $\Delta\hat{x}\Delta\hat{p} \geq \frac{1}{2}\hbar$. In fact, it can be considered as a simple consequence of the properties of the Fourier transform that connects the wave functions of the system in the position and momentum representation (more general form of uncertainty inequality with Wigner-Yanase-Dyson skew information in Appendix 2 are described).

B. It is possible to present quantum uncertainty relations (UR) in terms of entropy or information ("entropic UR" –EUR). The usual "standard UR" (for standard deviations)

$$(\Delta_\phi A)^2 (\Delta_\phi B)^2 \geq \frac{1}{4}\left|\langle [A,B]_- \rangle_\phi\right|^2 + \frac{1}{4}\left|\langle \{A,B\}_+ \rangle_\phi - 2\langle A\rangle_\phi \langle B\rangle_\phi\right|^2$$

(note that the second term in this inequality represents the covariance, or correlation,

$$\mathrm{cov}_\phi(A,B) := \frac{1}{2}\langle\phi|AB + BA|\phi\rangle - \langle\phi|A|\phi\rangle\langle\phi|B|\phi\rangle$$

between the observables A and B in the state $|\varphi\rangle$) presented by an inequality of the entropic form

$$S^{(A)} + S^{(B)} \geq S_{AB}$$

or in information form

$$I_\phi(A) + I_\phi(B) \leq I_\phi(A,B)$$

as more adequate expressions for the "uncertainty principle".

It is known that given two non-commuting observables, we can derive an uncertainty relation for them and the class of states that satisfy the equality sign in the inequality are called intelligent states (see, Definition).

Example. If we have any continuous parameter λ and any hermitian observable $A(\lambda)$ which is the generator of the parametric evolution, then UR give us $\langle\Delta A(\lambda)\rangle\Delta\lambda \geq \frac{\hbar}{4}$, where $\langle\Delta A(\lambda)\rangle = \frac{1}{(\lambda_2 - \lambda_1)}\int_{\lambda_1}^{\lambda_2}\Delta A(x)\,dx$ is the parameter average of the observable uncertainty and $\Delta\lambda = \frac{\pi}{s_0}(\lambda_2 - \lambda_1)$ is the scaled displacement in the space of the conjugate variable of A. This generalized UR would hold for position-momentum, phase-number or any combinations. For the case when initial and final states are orthogonal we know that all states of the form

$$|\psi(\lambda)\rangle = \frac{1}{\sqrt{2}}\left(e^{-\frac{i}{\hbar}a_i\lambda}|\psi_i\rangle + e^{-\frac{i}{\hbar}a_j\lambda}|\psi_j\rangle\right), i \neq j$$

are the only intelligent states which satisfy the equality $\langle\Delta A(\lambda)\rangle\Delta\lambda = \frac{\hbar}{4}$.

However, these states do not satisfy the equality when the initial and final states are non-orthogonal. In this case, if the generator of the parametric evolution A can be split into two parts $A_0 + A_1$ such that A_0 has a complex basis of normalised eigenvectors $\{|\psi_i\rangle\}_{i\in I}$ which degenerate spectrum $\{a_0\}$, with I a set of quantum numbers and A_1 has matrix elements $(A_1)_{ii} = 0 = (A_1)_{jj}$, and $(A_1)_{ij} = (A_1)_{ji} = a_1$, then all states of the form

$$|\psi(\lambda)\rangle = e^{-\frac{i}{\hbar}a_0\lambda}\left[\cos\left(a_1\frac{\lambda}{\hbar}\right)|\psi_i\rangle - i\sin\left(a_1\frac{\lambda}{\hbar}\right)|\psi_j\rangle\right], i \neq j$$

are intelligent states for non-orthogonal initial and final states.

C. It has been shown that the "Everett (entropic) UR" implies the famous Heisenberg UR as $\Delta q\Delta p \geq \frac{\hbar}{2}$. We shall compare various characterisations of "maximal information" and point out their connection with "minimum uncertainty". In the following we restrict ourselves mainly to "simple" observables (defined on the smallest non-trivial Boolean algebra $\Sigma = \{0, a, \neg a, 1\}$): we are interested in information with respect to single effect

$$E : I_\phi(E) = E_\phi \ln(E_\phi) + E'_\phi \ln(E'_\phi), E' = I - E.$$

Non-commutativity or incompatibility of (unsharp) properties E and F will, in general, exclude the possibility of measuring or preparing both of them simultaneously. In particular, if $E = E^Q(X), F = F^P(Y)$ are position and momentum spectral projections associated with bounded measurable sets X, Y, then $E^Q(X) \wedge E^P(Y) = 0$ holds or, equivalently

$$\langle\phi|E^Q(X)|\phi\rangle = 1 \Rightarrow \langle\phi|E^P(Y)|\phi\rangle < 1,$$

$$\langle\phi|E^P(Y)|\phi\rangle = 1 \Rightarrow \langle\phi|E^Q(X)|\phi\rangle < 1.$$

Thus "certain" position and momentum determinations exclude each other, and the question arises as to what "degree of uncertainty" they can be "known" simultaneously. One may take any reasonable characterisation of maximal joint knowledge, or joint information. In this case above mentioned statement can be put into the following equivalent form

$$\left.\begin{matrix}\langle\phi|E^Q(X)|\phi\rangle = 1 \Rightarrow \langle\phi|E^P(Y)|\phi\rangle < 1 \\ \langle\phi|E^P(Y)|\phi\rangle = 1 \Rightarrow \langle\phi|E^Q(X)|\phi\rangle < 1\end{matrix}\right\} \Rightarrow \begin{cases} E_\phi + F_\phi < 2 \\ E_\phi \cdot F_\phi < 1 \end{cases}.$$

The "state of maximal information" can be defined through three values.

The first expression $E_\varphi + F_\varphi$ can be maximised and an explicit construction procedure for the corresponding "state of maximal information" has been given below. Here we shall study the question of maxima for this quantity as well as for $E_\varphi \cdot F_\varphi$ and for $I_\varphi(E) + I_\varphi(F)$ for an arbitrary pair of effects, E and F. In particular, we shall show that each quantity can be maximal only if there exist states which lead to minimal uncertainty product in UR.

Furthermore, in the case of projections the maxima of $I_\varphi(E) + I_\varphi(F)$ (if they exist) coincide with those of one of the quantities $E_\varphi^\nu + F_\varphi^\eta$ and $E_\varphi^\nu \cdot F_\varphi^\eta$ $\left(E^\nu \in \{E, E'\}, F^\eta \in \{F, F'\}\right)$.

For maximal $E_\varphi + F_\varphi$ the variation of $\langle\varphi|E|\varphi\rangle + \langle\varphi|F|\varphi\rangle - \lambda\langle\varphi|\varphi\rangle$ must vanish which implies the following equations: $(E + F)|\varphi\rangle = (E_\varphi + F_\varphi)|\varphi\rangle$. Multiplying with E or with F and taking the expectations yields

$$(\Delta_\phi E)^2 = (\Delta_\phi F)^2 = -(\langle\phi|EF|\phi\rangle - E_\phi \cdot F_\phi) = -\mathrm{cov}_\phi(E, F),$$

which leads to a minimal UR: $(\Delta_\varphi E)^2 \cdot (\Delta_\varphi F)^2 = [\mathrm{cov}_\varphi(E,F)]^2$.

Similarly, maximising the product $E_\varphi \cdot F_\varphi$ gives $(F_\varphi E + E_\varphi F)|\varphi\rangle = 2E_\varphi \cdot F_\varphi|\varphi\rangle$ and

$$(\Delta_\phi E)^2 F_\phi{}^2 = (\Delta_\phi F)^2 E_\phi{}^2 = -E_\phi \cdot F_\phi \,\mathrm{cov}_\phi(E,F)$$

which leads again to a minimal UR, $E_\varphi \neq 0 \neq F_\varphi$.

Finally, maximal information sum $I_\varphi(E)+I_\varphi(F)$ will be realised in states satisfying

$$(\ln E_\phi - \ln E'_\phi)(E - E'_\phi)|\phi\rangle + (\ln F_\phi - \ln F'_\phi)(F - F'_\phi)|\phi\rangle = 0.$$

Generally this equation contains all stationary points, e.g. the minimum $E_\varphi = E'_\varphi = F_\varphi = F'_\varphi = \frac{1}{2}$, or the joint eigenstates. Since we are looking for states of maximal information with respect to positive outcomes for E, F we shall assume $E_\varphi > \frac{1}{2}$ and $F_\varphi > \frac{1}{2}$. Then this equality implies:

$$(\alpha E + F)|\phi\rangle = (\alpha F_\phi + F_\phi)|\phi\rangle, \alpha = \frac{\ln\left(\frac{E_\phi}{E'_\phi}\right)}{\ln\left(\frac{F_\phi}{F'_\phi}\right)} \geq 0$$

and $\alpha(\Delta_\varphi E)^2 = \frac{1}{\alpha}(\Delta_\varphi F)^2 = -\mathrm{cov}(E,F)$ which again gives rise to the minimal uncertainty product in UR.

We have thus shown that all three notions of maximal information are consistent in so far as they imply minimal uncertainty product.

Example. Let E, F denote position and momentum spectral projections, respectively: $E = E^Q(X), F = F^P(Y)$. The sum of probabilities $E_\varphi + F_\varphi$ has been shown to be maximal in the state

$$\varphi = \varphi_{\min} \text{ with } |\phi_{\min}\rangle = \left(\frac{1+a_0}{2a_0{}^2}\right)^{1/2} E|g_0\rangle + \left(\frac{1-a_0}{2(1-a_0{}^2)}\right)^{1/2} E'|g_0\rangle$$

provided that X, Y are bounded measurable sets. Here a_0^2 is the maximal eigenvalue of the compact operator (FEF) and g_0 is the corresponding eigenvector satisfying

$$FEF|g_0\rangle = a_0^2|g_0\rangle, F|g_0\rangle = |g_0\rangle, \||g_0\rangle\|_2{}^2 = 1.$$

It is clear from above description that $\varphi_{\min}$ must be an eigenstate of $(E+F)$. This can also be seen directly in the following way. Introduce

$$|f_0\rangle = a_0^{-1}E|g_0\rangle, \|f_0\|_2{}^2 = a_0^{-2}\langle g_0|FEF|g_0\rangle = 1, E|f_0\rangle = |f_0\rangle.$$

Then we have

$$EFE|f_0\rangle = a_0^2|f_0\rangle, |g_0\rangle = a_0^{-1}F|f_0\rangle$$

and $\varphi_{\min}$ can be written in the symmetric form

$$|\phi_{\min}\rangle = \frac{1}{\sqrt{2(1+a_0)}}\left[|f_0\rangle + |g_0\rangle\right].$$

We conclude that $\varphi_{\min}$ maximises all the three quantities $(E_\varphi \cdot F_\varphi)$, $(E_\varphi + F_\varphi)$ and $(I_\varphi(E) + I_\varphi(F))$, and it minimises the uncertainty product $\Delta_\varphi E \cdot \Delta_\varphi F$.

Thus maximal information (minimal entropy) and minimal uncertainty can be achieved on intelligent coherent states and will again coincide.

Conclusions

We discuss the role of entropy changing in quantum evolution as information data flow processing and how the classical and quantum information amount changes in the dynamics of some quantum control algorithms. We introduce the following qualitative axiomatic description of dynamic evolution of information flow in QA's:

(1) The information amount of successful result increases while the quantum algorithm is in execution;

(2) The quantity of information becomes the fitness function for recognition of successful results and introduces a measure of accuracy for them: in this case the Principle of Minimum of Classical / Quantum Entropy corresponds to recognition of success results on intelligent output states of quantum algorithm computation;

(3) If the classical entropy of the output vector is small, the degree of order for this output state is great and the output of measurement process on intelligent states of quantum algorithm's gives us the necessary information to solve with success the initial problem.

These information axioms mean that the quantum algorithms should automatically guarantee convergence of information amount to a precise value. This is a necessary condition in order to get robust and stable results for fault-tolerant computation in quantum control.

Appendix 1: Relative entropies and divergence functions

Divergence functions can be used to define metric tensors on the space of invertible states of a quantum system. These divergence functions play a central role in quantum information theory.

As an example, let consider the quantum relative entropy, also known as von Neumann relative entropy:

$$S_{vN}(\rho|\mathrm{c}) = Tr\rho(\log\rho - \log\mathrm{c}) \qquad \text{(A1.1)}$$

It is the quantum generalization of the Kullback-Leibler divergence function used in classical information geometry and, in the asympotic, memoryless setting, it yields fundamental limits on the performance of information-processing tasks. Another important family of relative entropies is the *q-Renyi relative entropies* (*q*-RRE)

$$S_{RRE}(\rho|\mathrm{c}) = \frac{1}{q-1}\log Tr\left(\rho^q \mathrm{c}^{1-q}\right) \qquad \text{(A1.2)}$$

where $q \in (0,1) \cup (1,\infty)$. These divergence functions are able to describe the cut-off rates in quantum binary state discrimination. Two other examples, which are relevant for the definition of metric tensors, can be given. The potential functions[5] of the Bures metric tensor

$$S_B\left(\rho|\mathrm{c}\right)=4\left[1-Tr\left(\rho\mathrm{c}\right)^{\square}\right] \tag{A1.3}$$

and the potential function of the Wigner-Yanase metric tensor:

$$S_{WY}\left(\rho|\mathrm{c}\right)=4\left[1-Tr\left(\rho^{½}\mathrm{c}^{½}\right)\right] \tag{A1.4}$$

Several efforts were done in order to find a common mathematical framework to unify this plethora of different divergence functions. A first (partial) result was achieved by the *q-quantum Renyi divergence* (*q*-QRD)

$$S_{QRD}\left(\rho|\mathrm{c}\right)=\frac{1}{q-1}\log Tr\left(\mathrm{c}^{\frac{1-q}{2q}}\rho\mathrm{c}^{\frac{1-q}{2q}}\right) \tag{A1.5}$$

where again $q\in(0,1)\cup(1,\infty)$. However, it has two important limitations: the data processing inequality (DPI)

$$S_{QRD}\left(\Phi(\rho)|\Phi(\mathrm{c})\right)\le S_{QRD}\left(\rho|\mathrm{c}\right) \tag{A1.6}$$

where Φ is a completely positive trace preserving map (CPTP) acting on a pair of semidefinite Hermitian operators ρ and c, is not satisfied for $q\in(0,½)$ and it does not contain the *q*-RRE family.

Recently, a new family of two-point functions which includes all the previous examples was defined. It is the so-called *q-z-Renyi Relative Entropy* (*q-z*-RRE)

$$S_{q,z}\left(\rho|\mathrm{c}\right)=\frac{1}{q-1}\log Tr\left(\mathrm{c}^{\frac{q}{2z}}\rho^{\frac{1-q}{z}}\mathrm{c}^{\frac{q}{2z}}\right) \tag{A1.7}$$

that can be recast as:

$$S_{q,z}\left(\rho|\mathrm{c}\right)=\frac{1}{q-1}\log Tr\left(\mathrm{c}^{\frac{q}{z}}\rho^{\frac{1-q}{z}}\right)^{z} \tag{A1.8}$$

Remark: In general, the product of two Hermitian matrices is not a Hermitian matrix. However, the product matrix $\mathrm{c}^{\frac{q}{z}}\rho^{\frac{1-q}{z}}$ has real, non-negative eigenvalues, even though it is not in general a hermitian matrix. It means that the trace functional

$$f_{q,z}\left(\rho|\mathrm{c}\right)=Tr\left(\mathrm{c}^{\frac{q}{z}}\rho^{\frac{1-q}{z}}\right)^{z} \tag{A1.9}$$

is well defined as the sum of the the *z*-th power of the eigenvalues of the product matrix and it can be developed in Taylor series.

In particular limits of the parameters q and *z* it is possible to recover the *q*-RRE family

$$S_{q,1}\left(\rho|\mathrm{c}\right):=\lim_{z\to1}S_{q,z}\left(\rho|\mathrm{c}\right)\equiv S_{RRE}\left(\rho|\mathrm{c}\right)=\frac{1}{q-1}\log Tr\left(\rho^{q}\mathrm{c}^{1-q}\right) \tag{A1.10}$$

the *q*-QRD family

$$S_{q,q}(\rho|c) := \lim_{z\to q} S_{q,z}(\rho|c) \equiv S_{QRD}(\rho|c) = \frac{1}{q-1}\log Tr\left(c^{\frac{1-q}{2q}}\rho c^{\frac{1-q}{2q}}\right) \quad (A1.11)$$

and the von Neumann relative entropy:

$$S_{1,1}(\rho|c) := \lim_{z=q\to 1} S_{q,z}(\rho|c) \equiv S_{vN}(\rho|c) = Tr\rho(\log\rho - \log c) \quad (A1.12)$$

The data processing inequality for the *q*-*z*-RRE was studied and it is not established yet in full generality. To prove it, one has to show that the trace functional (A1.9) is jointly concave when $q \le 1$, or jointly convex when $q \ge 1$. The results of these analysis are well summarized and it results that the DPI holds only for certain range of the parameters as sketched in Fig. A1.1.

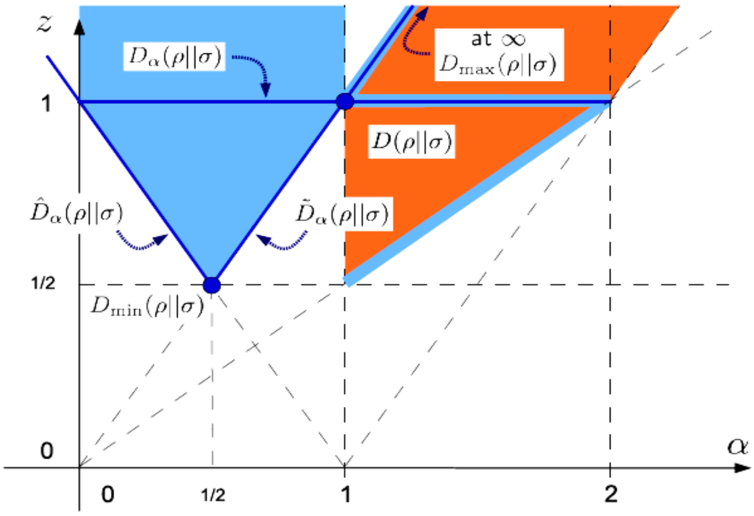

Figure A1.1. A schematic overview of the various relative entropies unified by the q-z-relative entropy is shown. [The blue region indicates the range of the parameters in which the DPI was proven, while the orange region indicates where it is just conjectured. To keep contact with the notation , the divergence functions S are indicated with the letter D, the parameter q is indicated with $\alpha, D(\rho\|\sigma)$ *is the von Neumann relative entropy,* $D_{\min}$ *is the logarithm of the fidelity and* $D_{\max} := \inf\{\gamma : \rho \le 2^\gamma \sigma\}$ *].*

Since we are interested in computing the metric tensors starting from this two-parameter family of two-point functions, it is convenient to consider the following regularization of the logarithm, the so-called *q*-logarithm:

$$\log_q \rho = \frac{1}{1-q}\left(\rho^{1-q} - 1\right) \text{ with } \lim_{q\to 1}\log_q \rho = \log\rho \quad (A1.13)$$

Moreover, inspired by Petz, we will consider a rescaling by a factor $1/q$. In this way, the resulting family of functions will be symmetric under the exchange of $q \to (1-q)$. Let us denote the resulting two-point function with the same symbol of the *q*-*z*-RRE, that is:

$$S_{q,z}(\rho|c) = \frac{1}{q(q-1)}\left[1 - Tr\left(c^{\frac{q}{z}}\rho^{\frac{1-q}{z}}\right)^z\right] \quad (A1.14)$$

Since the analysis of the DPI involves only the trace functional, we are ensured that the DPI holds for the same range of parameters of the *q*-*z*-RRE. Moreover, in the limit $z \to 1$, it is possible to recover the expression for the Tsallis relative entropy as

$$S_{q,1}\left(\rho|\mathrm{c}\right):=\lim_{z\to1}S_{q,z}\left(\rho|\mathrm{c}\right)\equiv S_{Ts}\left(\rho|\mathrm{c}\right)=\frac{1}{q(1-q)}\left[1-Tr\left(\rho^{q}\mathrm{c}^{1-q}\right)^{z}\right] \tag{A1.15}$$

in the limit $z=q\to1$, we recover the von Neumann relative entropy (A1.12) as

$$S_{1,1}\left(\rho|\mathrm{c}\right):=\lim_{z=q\to1}S_{q,z}\left(\rho|\mathrm{c}\right)\equiv S_{vN}\left(\rho|\mathrm{c}\right)=Tr\rho\left(\log\rho-\log\mathrm{c}\right) \tag{A1.16}$$

in the limit $z=q=\square$, we recover the divergence function of the Bures metric tensor

$$S_{\frac{1}{2},\frac{1}{2}}\left(\rho|\mathrm{c}\right):=\lim_{z=q\to\frac{1}{2}}S_{q,z}\left(\rho|\mathrm{c}\right)\equiv S_{B}\left(\rho|\mathrm{c}\right)=4\left[1-Tr\left(\rho\mathrm{c}\right)^{\square}\right] \tag{A1.17}$$

and finally, in the limit $z=1,q\to\square$, we recover the divergence function of the Wigner-Yanase metric tensor:

$$S_{\frac{1}{2},1}\left(\rho|\mathrm{c}\right):=\lim_{z=1,q\to\frac{1}{2}}S_{q,z}\left(\rho|\mathrm{c}\right)\equiv S_{WY}\left(\rho|\mathrm{c}\right)=4\left[1-Tr\left(\rho^{\square}\mathrm{c}^{\square}\right)\right] \tag{A1.18}$$

All these special cases belong to the range of parameters for which $S_{q,z}$ is actually a quantum divergence function satisfying the DPI. Consequently, the family of associated quantum metric tensors satisfies the monotonicity property.

Appendix 2: Entropic-like uncertainty relations

The uncertainty principle is an essential feature of quantum mechanics, characterizing the experimental measurement incompatibility of non-commuting quantum mechanical observables in the preparation of quantum states. Heisenberg first introduced variance-based uncertainty. Later, Robertson proposed the well-known formula of the uncertainty relation, $Var(\rho,R)Var(\rho,S)\geq\frac{1}{4}\left|Tr\rho\left[R,S\right]\right|^{2}$, for arbitrary observables *R* and *S*, where $\left[R,S\right]=RS-SR$ and $Var(\rho,R)$ is the standard deviation of *R*. Schrödinger gave a further improved uncertainty relation:

$$Var(\rho,R)Var(\rho,S)\geq\frac{1}{4}\left|\left\langle\left[R,S\right]\right\rangle\right|^{2}+\frac{1}{2}\left|\left\langle\left\{R,S\right\}\right\rangle-\left\langle R\right\rangle\left\langle S\right\rangle\right|^{2}$$

where $\left\langle R\right\rangle=Tr\left(\rho R\right)$, and $\left\{R,S\right\}=RS+SR$ is anti-commutator. Since then many kinds of uncertainty relations have been presented. In addition to the uncertainty of the standard deviation, entropy can be used to quantify uncertainties. The first entropic uncertainty relation was given by Deutsch and was then improved by Maassen and Uffink:

$$H\left(R\right)+H\left(S\right)\geq\log_{2}\frac{1}{c}$$

where $R=\left\{\left|u_{j}\right\rangle\right\}$ and $S=\left\{\left|\upsilon_{k}\right\rangle\right\}$ are two orthonormal bases on d-dimensional Hilbert space *H*, and $H\left(R\right)=-\sum_{j}p_{j}\log p_{j}$ ($H\left(S\right)=-\sum_{k}q_{k}\log q_{k}$) is the Shannon entropy of the probability distribution $p_{j}=\left\langle u_{j}\right|\rho\left|u_{j}\right\rangle$ ($q_{k}=\left\langle\upsilon_{k}\right|\rho\left|\upsilon_{k}\right\rangle$) for state ρ of *H*. The number c is the largest overlap among all $c_{jk}=\left|\left\langle u_{j}|\upsilon_{k}\right\rangle\right|^{2}$ between the projective measurements *R* and *S*. Berta et al. bridged the gap between cryptographic scenarios and the uncertainty principle and derived this landmark uncertainty relation for measurements R and S in the presence of quantum memory B: $H(R|B)+H(S|B)\ge\log_2\frac{1}{c}+H(A|B)$ where $H(R|B)=H(\rho_{RB})-H(\rho_B)$ is the conditional entropy with $\rho_{RB}=\sum_j(|u_j\rangle\langle u_j|\otimes I)\rho_{AB}(|u_j\rangle\langle u_j|\otimes I)$ (similarly for $H(S|B)$), and d is the dimension of the subsystem A. The term $H(A|B)=H(\rho_{AB})-H(\rho_B)$ appearing on the right-hand side is related to the entanglement between the measured particle A and the quantum memory B. The bound of Berta et al. has been further improved. Moreover, there are also some uncertainty relations given by the generalized entropies, such as the Rényi entropy and the Tsallis entropy, and even more general entropies such as the (h,Φ) entropies. These uncertainty relations not only manifest the physical implications of the quantum world but also play roles in entanglement detection, quantum spin squeezing and quantum metrology.

An uncertainty relation based on Wigner–Yanase skew information $I(\rho,H)$ has been obtained with quantum memory, where $I(\rho,H)=\frac{1}{2}Tr\left[\left(i\left[\sqrt{\rho},H\right]^2\right)\right]=Tr(\rho H^2)-Tr\left(\sqrt{\rho}H\sqrt{\rho}H\right)$ quantifies the degree of non-commutativity between a quantum state ρ and an observable H, which is reduced to the variance $Var(\rho,H)$ when ρ is a pure state. In fact, the Wigner–Yanase skew information $I(\rho,H)$ is generalized to Wigner–Yanase–Dyson skew information:

$$I_\alpha(\rho,H)=\frac{1}{2}Tr\left[\left(i\left[\rho^\alpha,H\right]\right)\left(i\left[\rho^{1-\alpha},H\right]\right)\right]=Tr(\rho H^2)-Tr\left(\rho^\alpha H\rho^{1-\alpha}H\right)\quad \alpha\in[0,1]. \quad \text{(A2.1)}$$

Here theWigner–Yanase–Dyson skew information $I_\alpha(\rho,H)$ reduces to theWigner–Yanase skew information $I(\rho,H)$, when $\alpha=$. The Wigner–Yanase–Dyson skew information $I_\alpha(\rho,H)$ reduces to the standard deviation $Var(\rho,H)$ when ρ is a pure state. The convexity of $I_\alpha(\rho,H)$ with respect to ρ has been proven by Lieb. Kenjiro introduced another quantity:

$$I_\alpha(\rho,H)=\frac{1}{2}Tr\left[\left(\left\{\rho^\alpha,H_0\right\}\right)\left(\left\{\rho^{1-\alpha},H_0\right\}\right)\right]=Tr(\rho H_0^2)+Tr\left(\rho^\alpha H_0\rho^{1-\alpha}H_0\right)\quad \alpha\in[0,1] \quad \text{(A2.2)}$$

where $H_0=H-Tr(\rho H)I$ with I being the identity operator. For a quantum state r and observables R, S and $0\le\alpha\le1$, the following inequality holds:

$$\mathrm{U}_\alpha(\rho,R)\,\mathrm{U}_\alpha(\rho,S)\ge\alpha(1-\alpha)\left|Tr\rho[R,S]\right|^2$$

where $\mathrm{U}_\alpha(\rho,R)=\sqrt{I_\alpha(\rho,R)I_\alpha(\rho,R)}$ can be regarded as a kind of measure for quantum uncertainty. For a pure state, a standard deviation-based relation is recovered from Eq. (A2.3).

Let $\phi_k=|\phi_k\rangle\langle\phi_k|$ and $\psi_k=|\psi_k\rangle\langle\psi_k|$ be the rank 1 spectral projectors of two non-degenerate observables R and S with the eigenvectors $|\phi_k\rangle$ and $|\psi_k\rangle$, respectively. We can define $\mathrm{U}N_\alpha(\rho,\phi)=\sum_k\mathrm{U}_\alpha(\rho,\phi_k)=\sum_k\sqrt{I_\alpha(\rho,\phi_k)I_\alpha(\rho,\phi_k)}\mathrm{U}_\alpha(\rho,\psi)$ as the uncertainty of ρ associated to the projective measurement $\{\phi_k\}$, and $\mathrm{U}_\alpha(\rho,\psi)$ to $\{\psi_k\}$.

Let ρ_{AB} be a bipartite state on $H_A\otimes H_B$, where H_A and H_B denote the Hilbert space of subsystems A and B, respectively. Let V be any orthogonal basis space on H_A and $|\phi_k\rangle$ be an orthogonal basis of H_A. We define a quantum correlation of ρ_{AB} as

$$\tilde{D}_\alpha(\rho_{AB}) = \min_V \sum_k \left[I_\alpha(\rho_{AB}, \phi_k \otimes I_B) - I_\alpha(\rho_A, \phi_k) \right] \tag{A2.4}$$

where the minimum is taken over all the orthogonal bases on H_A, $\rho_A = Tr_B \rho_{AB}$. For any bipartite state ρ_{AB} and any observable X_A on H_A, we have $I_\alpha(\rho_{AB}, X_A \otimes I_B) \geq I_\alpha(\rho_A, X_A)$. Therefore, $\tilde{D}_\alpha(\rho_{AB}) \geq 0$.

Furthermore, $\tilde{D}_\alpha(\rho_{AB}) = 0$ when ρ_{AB} is a classical quantum correlated state.

$\tilde{D}_\alpha(\rho_{AB})$ has a measurement on subsystem *A*, which gives an explicit physical meaning: it is the minimal difference of incompatibility of the projective measurement on the bipartite state ρ_{AB} and on the local reduced state ρ_A. $\tilde{D}_\alpha(\rho_{AB})$ quantifies the quantum correlations between the subsystems *A* and *B*. We have the following.

Theorem A2.1. Let ρ_{AB} be a bipartite quantum state on $H_A \otimes H_B$ and $\{\phi_k\}$ and $\{\psi_k\}$ be two sets of rank 1 projective measurements on H_A. Then

$$\mathrm{UN}_\alpha(\rho_{AB}, \phi \otimes I)\, \mathrm{UN}_\alpha(\rho_{AB}, \psi \otimes I) \geq \sum_k L^2_{\alpha,\rho_A}(\phi_k, \psi_k) + \tilde{D}^2_\alpha(\rho_{AB}) \tag{A2.5}$$

where $L_{\alpha,\rho_A}(\phi_k, \psi_k) = \alpha(1-\alpha) \dfrac{\left| Tr \rho_A [\phi_k, \psi_k] \right|^2}{\sqrt{I_\alpha(\rho_A, \phi_k) I_\alpha(\rho_A, \psi_k)}}$.

Theorem A2.1 gives a *product form* of the uncertainty relation. Comparing the results (Eq. (A2.3)) without quantum memory with those (Eq. (A2.5)) with quantum memory, one finds that if the observables *A* and *B* satisfy $[A, B] = 0$, the bound is trivial in Eq.(A2.3), while in Eq. (A2.5), even if the projective measurements ϕ_k and ψ_k satisfy $[\phi_k, \psi_k] = 0$, that is, $L_{\alpha,\rho_A}(\phi_k, \psi_k) = 0$, $\tilde{D}_\alpha(\rho_{AB})$ may still not be trivial because of correlations between the system and the quantum memory.

Corresponding to the product form of the uncertainty relation, we can also derive the *sum form* of the uncertainty relation:

Theorem A2.2. Let ρ_{AB} be a quantum state on $H_A \otimes H_B$ and $\{\phi_k\}$ and $\{\psi_k\}$ be two sets of rank 1 projective measurements on H_A. Then

$$\mathrm{UN}_\alpha(\rho_{AB}, \phi \otimes I) + \mathrm{UN}_\alpha(\rho_{AB}, \psi \otimes I) \geq 2\sum_k L_{\alpha,\rho_A}(\phi_k, \psi_k) + 2\tilde{D}^2_\alpha(\rho_{AB}) \tag{A2.6}$$

From Theorems A2.1 and A2.2, we obtain uncertainty relations in the form of the product and sum of skew information, which are different from the uncertainty, which only deals with the single partite state. However, we treat the bipartite case with quantum memory *B*. It is shown that the lower bound contains two terms: one is the quantum correlation $\tilde{D}_\alpha(\rho_{AB})$, and the other is $\sum_k L_{\alpha,\rho_A}(\phi_k, \psi_k)$ which characterizes the degree of compatibility of the two measurements, just as for the meaning of $\log_2 \frac{1}{c}$ in the entropy uncertainty relation.

For the Shannon entropy, Rényi entropy, Tsallis entropy, (h, Φ) entropies and Wigner-Yanase skew information, theWigner-Yanase-Dyson skew information characterizes a special kind of information of a system or measurement outcomes, which needs to satisfy certain restrictions for given measurements and correlations between the system and the memory. Different α parameter values give rise to different kinds of information.

The uncertainty relations both in product and summation forms in terms of the Wigner-Yanase-Dyson skew information with quantum memory have investigated. It has been shown that the lower bounds contain two terms: one is the quantum correlation $\tilde{D}_{\alpha}\left(\rho_{AB}\right)$, and the other is $\sum_{k} L_{\alpha,\rho_A}\left(\phi_k,\psi_k\right)$, which characterizes the degree of compatibility of the two measurements.

References

1. Килин С.Я. Квантовая информация // УФН. - 1999. - Т. 169. - № 5.
2. Холево А.С. Введение в квантовую теорию информации. - М.: 2002.
3. Keyl M. Fundamentals of quantum information theory // Physical Reports. - 2002. - V. 369. - № 5.
4. Cerf N.J., Adami C. Negative entropy and information in quantum mechanics. // Physical Review Letters. - 1997. - V. 79. - № 26.
5. Horodecki M., Oppenheim J., Winter A. Partial quantum information. // Nature. - 2005. - V. 436. - № 7051.
6. Henderson L., Vedral V. Classical, quantum and total correlation. // J. Phys. A: Math. Gen. - 2001. - Vol. 34. - № 35.
7. Sanctuary B.C. Correlations in entangled states. // arXiv: quant-ph/0508238v1. 2005.
8. Hamieh S., Kobes R., Zaraket H. Positive-operator-valued measure optimization of classical corelations. // Phys. Review. - 2004. - Vol. A70. - № 5.
9. Hamieh S., Qi J., Siminovitch D. et all. Extraction of classical corelations from a bipartite quantum system. // Phys. Review. - 2003. - A67. - № 1.
10. Groisman B., Popescu S., Winter A. Quantum, classical, and total amount of correlations in quantum state. // Phys. Review. - 2005. -Vol. A72. - № 3.
11. DiVincenzo D.P., Horodecki M., Leung D.W. et all. Locking classical correlation in quantum states. // Physical Review Letters. - 2004. – Vol. 92. - No 6.
12. Horodecki M., Horodecki P., Horodecki R., et all. Local versus non-local information in quantum-information theory: Formalism and phenomena. // Phys. Review. - 2005. - Vol. A71. - № 6.
13. Maruyama K., Brukner C., Vedral V. Thermodynamical cost of accessing quantum information. // J. Phys. A: Math. Gen. - 2005. - Vol. 38. - № 32.
14. Christandl M., Winter A. Uncertainty, monogamy, and locking of quantum correlations. // IEEE Transactions on Information Theory. - 2005. - Vol. 51. - № 9.
15. Adesso G., Bromley T.R., Cianciaruso V. Measures and applications of quantum correlations // J. Phys. A: Math. Theor. - 2016. - Vol. 49 - pp. 473001 (82pp)
16. Levitin L.B. Conditional entropy and information in quantum systems. // Chaos, Solitons and Fractals. 1999. - Vol. 10. - № 10.
17. Ohya M. and Petz D. Quantum Entropy and Its Use. - Springer-Verlag, Berlin Heidelberg. - 1993.
18. Gibilisco P. and Isola T. A characterization of Wigner-Yanase skew information among statistically monotone metrics // Infin. Dimens. Anal. Quantum. Probab. Relat. Top. - 2001. - Vol. 4. - No 4. - pp 553 - 557.
19. Gibilisco P. and Isola T. Wigner-Yanase information on quantum state space: the geometric approach // Journal of Mathematical Physics - 2003. - Vol. 44. - pp. 3752-3762.
20. Ciaglia F.M, Di Cosmo F., Laudato M., Marmo G., et all, A pedagogical intrinsic approach to relative entropies as potential functions of quantum metrics: the q-z family // arXiv:1711.09769v1 [quant-ph] 27 Nov 2017.

21. Wigner E.P., Yanase M.M. Information contents of distributions // Proc. Natl. Acad. Sci. USA. - 1963. - Vol. 49. - pp. 910-918.

22. Luo S. Heisenberg uncertainty relation for mixed states // Phys. Rev. A - 2005. - Vol.72. - pp.042110.

23. Luo S. Notes on superadditivity of Wigner–Yanase–Dyson information // J. Stat. Phys. - 2007. - Vol.128. - pp. 1177-1188.

24. Luo S., Fu S., Oh C.H. Quantifying correlations via the Wigner-Yanase skew information. // Phys. Rev. - 2012. - Vol. 85. - pp. 2117.

25. Li J., Fei Sh-M. Uncertainty relation based on Wigner-Yanase-Dyson skew information with quantum memory // Entropy. - 2018. - Vol. 20. - No 132. - pp. 1-9.

26. Zozor S., Bosyk G. M. and Portesi M. General entropy-like uncertainty relations in finite dimensions // arXiv:1311.5602v3 [quant-ph] 7 Oct 2014.

27. Muller-Hermes A. and Franca D. Sandwiched Renyi convergence for quantum evolutions // arXiv:1607.00041v3 [quant-ph] 26 Feb 2018.