BASIC RELATIONS OF QUANTUM INFORMATION THEORY. PT 1: MAIN DEFINITIONS AND PROPOERTIES OF QUANTUM INFORMATION

Ulyanov Sergey[1], Korenkov Vladimir[2], Reshetnikov Andrey[3], Tanaka Takayuki[4], Rizzotto Giovanni[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;*
ST Microelectronics;
20041 Agrate Brianza, Italy, Via C. Olivetti, 2;
e-mail: gianguido.rizzotto@st.com

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. 1: ОСНОВНЫЕ ОПРЕДЕЛЕНИЯ И СВОЙСТВА КВАНТОВОЙ ИНФОРМАЦИИ

Ульянов Сергей Викторович[1], Кореньков Владимир Васильевич[2], Решетников Андрей Геннадьевич[3], Танака Такаюки[4], Риззотто Джиовани[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 Института системного анализа и управления;*
Директор Лаборатории информационных технологий (ЛИТ, ОИЯИ),
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;
ST Microelectronics;
Италия, 20041 Agrate Brianza, Via C. Olivetti, 2;
e-mail: gianguido.rizzotto@st.com.

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, косо-симметричная информация

Introduction: Basic concept of quantum entropy and information theory - Interrelations with measures of entanglement

Information-theoretic analysis of quantum evolution is based on the results of classical and quantum information theory. Let us consider the main results of classical/quantum information theory and its role in information analysis of quantum evolution of successful solution searching.

The most important tools and results of classical and quantum information theory Table 1 shows the most important results from classical/quantum information theory for information analysis and design of quantum evolution.

Table 1: Summary of classical and quantum information theory

Information theory	
Classical	**Quantum**
Shannon entropy: $$H(X) = -\sum_x p_x \log p_x$$	*Von Neumann entropy*: $$S(\rho) = -Tr(\rho \log \rho)$$
Distinguishability and accessible information	
Letters always distinguishable: $$N = |X|$$	*Holevo-Levitin bound*: $$H(X:Y) \le S(\rho) - \sum_x p_x S(\rho_x), \rho = \sum_x p_x \rho_x$$
Information-theoretic relations	
Fano inequality: $$H(p_e) + p_e \log(|X|-1) \ge H(X|Y)$$	*Quantum Fano inequality*: $$H(F(\rho,E)) + (1-F(\rho,E))\log(d^2-1) \ge S(\rho,E)$$
Mutual information: $$H(X:Y) = H(Y) - H(Y|X)$$	*Coherent information*: $$I(\rho,E) = S(E(\rho)) - S(\rho,E)$$
Data processing inequality: $$X \to Y \to Z$$ $$H(X) \ge H(X:Y) \ge H(X:Z)$$	*Quantum data processing inequality*: $$\rho \to E_1(\rho) \to (E_2 \circ E_1)(\rho)$$ $$S(\rho) \ge I(\rho, E_1) \ge I(\rho, E_2 \circ E_1)$$
Noiseless channel coding	
Shannon's theorem: $n_{bits} = H(X)$	*Schumacher's theorem*: $n_{qubits} = S\left(\sum_x p_x \rho_x\right)$
Capacity of noisy channels for classical information	

Shannon's noisy coding theorem:	Holevo-Schumacher-Westmoreland theorem:
$$C(\mathrm{N}) = \max_{p(x)} H(X:Y)$$	$$C^{(1)}(\mathrm{E}) = \max_{\{p_j, \rho_j\}} \left[S(\rho') - \sum_x p_x S(\rho'_x) \right],$$ $$\rho'_x = \mathrm{E}(\rho_x), \rho' = \sum_x p_x \rho'_x$$

The main problems that are investigated with these tools in information theory (that is important in an information analysis of quantum evolution) are the following:

- Classical, quantum and total correlations: Interrelations with entanglement measures;
- Accessible information and information-theoretical models of quantum measurement;
- Extraction of information with efficient measurements and bounds on the mutual information.

Design of quantum evolution and successful solutions is based on the same results.

Let us consider briefly these results [1-17].

Information-theoretical models of quantum system evolution and irreversible measurements

Mathematically output solutions of quantum evolution are described as quantum system entities and represented by Hilbert spaces vectors $|\psi\rangle$ usually normalized $\langle\psi|\psi\rangle = 1$. Quantum systems evolve unitarily, that is, if a system is initially in a state $|\psi_1\rangle$, it becomes later state $|\psi_2\rangle$ after unitary operations: $U|\psi_1\rangle = |\psi_2\rangle$. Unitary operations are reversible, since $UU^\dagger = 1$ and previous states can be reconstructed by: $U|\psi_1\rangle = U^\dagger|\psi_2\rangle$. Because, $\langle\psi_2|\psi_2\rangle = \langle\psi_1|UU^\dagger|\psi_1\rangle = \langle\psi_1|I|\psi_1\rangle = 1$, the operation of normalization, which soon will be interpreted as probability, is conserved. The measurement of the properties of these objects is described by a collection of operators $\{M_m\}$. The quantum object will found to be in the m-th state with probability $p(m) = \langle\psi|M_m^\dagger M_m|\psi\rangle$, and after this measurement it is going to be in definite, possibly different from the starting one: $|\psi'\rangle = \dfrac{M_m}{\sqrt{\langle\psi|M_m^\dagger M_m|\psi\rangle}}|\psi\rangle$.

In this case m should be viewed as the measurement outcome, hence information extracted from physical system. This way information can be assigned to each state of the quantum object.

Example. In each energy state of an atom one can map four numbers or in each polarization state of a photon one can map two numbers say 0 and 1. It should be stressed here, that the measurement operators should satisfy the completeness relation: $\sum_m M_m^\dagger M_m = 1$, which results $\sum_m p(m) = \sum_m \langle\psi|M_m^\dagger M_m|\psi\rangle = 1$ as instructed by probability theory. However this implies that $\langle\psi|M_m^\dagger M_m|\psi\rangle \neq 1$ and looking at equation $|\psi'\rangle = \dfrac{M_m}{\sqrt{\langle\psi|M_m^\dagger M_m|\psi\rangle}}|\psi\rangle$ one understands that measurement is an *irreversible operation*.

Measurement process, trace and entropy Structurally, quantum mechanics has two parts: one part concerned with quantum states, the other with quantum dynamics. To find out what is going on inside a quantum system, one must perform a quantum measurement. And a partial trace operation can give the correct description of observable quantities for subsystems of a composite system: a partial trace operation consists to a projective measurement on a subsystem. The quantum operations formalism is a general tool for describing the evolution process of quantum systems, such process include unitary evolution, quantum measurement, and even more general processes.

Partial trace. The partial trace operator is the operation, which gives the description of observable quantities for subsystems of a composite system. The partial trace is defined by

$$Tr_B\left(|a\rangle\langle a|\otimes|b\rangle\langle b|\right) \equiv |a\rangle\langle a|Tr\left(|b\rangle\langle b|\right).$$

The trace is cyclic and linear.

Measurement. It is familiar that quantum mechanics describe projective measurements, "Positive-Operator-Valued Measure" (POVM), and general measurements. A projective measurement is described by a Hermitian operator, M, on the state space of the system being observed $M = \sum_m mP_m$, where P_m is the projector onto the eigenspace of M with eigenvalue m. The possible outcomes of measurement correspond to the eigenvalues, m, of the observable. Upon measuring the state $|\psi\rangle$, the probability of getting result m is given by $p(m) = \langle\psi|P_m|\psi\rangle$.

Example: The standard (projective) quantum measurement. We start by reminding the modalities of the standard quantum measurement (of one qubit). Let us consider a qubit in the superposition state: $|\psi\rangle = a|0\rangle + b|1\rangle$, where $|0\rangle$ and $|1\rangle$ form an orthogonal basis, called the computational basis, and a and b, called probability amplitudes, are complex numbers such that the probabilities sum up to one: $|a|^2 + |b|^2 = 1$. The standard quantum measurement of the qubit $|\psi\rangle$ gives either $|0\rangle$ with probability $|a|^2$, or $|1\rangle$ with probability $|b|^2$. This is achieved by the use of projector operators.

A projector operator P is defined by: $\mathrm{P}^2 = \mathrm{P}$, $\mathrm{P}^\dagger = \mathrm{P}$, $\mathrm{P}^\dagger = \mathrm{P}^{T^*}$.

Thus a projector P is idempotent and Hermitian. Let us consider a general superposed quantum state: $|\psi\rangle = \sum_{i=1}^n c_i|\psi_i\rangle$ in the Hilbert space C^n, with $\sum_{i=1}^n |c_i|^2 = 1$. The probability $Pr(i)$ of finding the state $|\psi\rangle$ in one of the basis states $|\psi_i\rangle$ is, after a measurement: $Pr(i) = |\mathrm{P}_i|\psi\rangle|^2$. After the measurement, the state $|\psi\rangle$ has "collapsed" to the state $|\psi\rangle' = \dfrac{\mathrm{P}_i|\psi\rangle}{\sqrt{Pr(i)}}$.

The n projectors P_i $(i = 1, 2, \ldots, n)$ are orthogonal: $\mathrm{P}_i\mathrm{P}_j = \delta_{ij}\mathrm{P}_i$, $\sum_{i=1}^n \mathrm{P}_i = 1$. For the case $i = 0,1$ we have the two 2D-projectors: $\mathrm{P}_0 = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$, $\mathrm{P}_1 = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$.

For which it holds: $\mathrm{P}_0\mathrm{P}_1 = \mathrm{P}_1\mathrm{P}_0 = 0$, $\mathrm{P}_0^2 = \mathrm{P}_0$, $\mathrm{P}_1^2 = \mathrm{P}_1$, $\mathrm{P}_0 + \mathrm{P}_1 = 1$. The actions of P_0 and P_1 on the basis states are, respectively:

$$\mathrm{P}_0|0\rangle = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} = |0\rangle, \mathrm{P}_0|1\rangle = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} = 0$$

$$\mathrm{P}_1|0\rangle = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} = 0, \mathrm{P}_1|1\rangle = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix} = |1\rangle$$

From which it follows that their action on the superposed state is, respectively:

$$\mathrm{P}_0|\psi\rangle = a|0\rangle, \quad \mathrm{P}_1|\psi\rangle = b|1\rangle.$$

The probability of finding the qubit state in the state $|0\rangle$ is, for example:

4

$$Pr(0) = \left| \mathrm{P}_0 \left| \psi \right\rangle \right|^2 = \left| a \left| 0 \right\rangle \right|^2 = \left| a \right|^2.$$

After the measurement, the qubit has "collapsed" to the state:

$$\left| \psi \right\rangle' = \frac{\mathrm{P}_0 \left| \psi \right\rangle}{\sqrt{Pr(0)}} = \frac{a \left| 0 \right\rangle}{\sqrt{\left| a \right|^2}} = \left| 0 \right\rangle.$$

Then, a lot of quantum information that encoded in qubit state is made hidden by the standard quantum measurement. As a projector is not a unitary transformation, the standard quantum measurement is not a reversible operation. This means that the hidden quantum information will never be recovered (i.e., we will not be able to get back the superposed state).

We can see that a project measurement on a subsystem is the same as a partial trace. To see it clearly, let us consider an example.

Exampe: *GHZ state*. In this case the GHZ state is as

$$\left| \psi_{GHZ} \right\rangle = \frac{1}{\sqrt{2}} \left(\left| 0_A 0_B 0_C \right\rangle + \left| 1_A 1_B 1_C \right\rangle \right)$$

of a three-qubit system ABC. Then trace one qubit (suppose of a system A) out the three- qubit system. It has

$$Tr_A \left(\left| \psi_{GHZ} \right\rangle \left\langle \left| \psi_{GHZ} \right\rangle \right| \right) = \frac{1}{2} \rho_{00}^{BC} + \frac{1}{2} \rho_{11}^{BC},$$

where $\rho_{00}^{BC} = \left| 0_B 0_C \right\rangle \left\langle 0_B 0_C \right|$, $\rho_{11}^{BC} = \left| 1_B 1_C \right\rangle \left\langle 1_B 1_C \right|$. After a partial trace on system A the system of BC has a probability $p = \frac{1}{2}$ in state ρ_{00}^{BC} or in the state ρ_{11}^{BC}. Obviously, the quantum measurement postulate tells us that is we perform a project measurement in basis $\left\{ \left| 0 \right\rangle, \left| 1 \right\rangle \right\}$ on the system A, the measurement result is in state $\left| 0 \right\rangle$ (consists to ρ_{00}^{BC}) or in state $\left| 1 \right\rangle$ (consists to ρ_{11}^{BC}) with a probability $p = \frac{1}{2}$. Consider the case the three qubit in the state $\left| w \right\rangle$: $\left| w \right\rangle = \frac{1}{\sqrt{3}} \left(\left| 1_A 0_B 0_C \right\rangle + \left| 0_A 1_B 0_C \right\rangle + \left| 0_A 0_B 1_C \right\rangle \right)$.

Let us suppose we trace out the qubit of the system A. Then we have $Tr_A \left(\left| w \right\rangle \left\langle w \right| \right) = \frac{1}{3} \rho_{00}^{BC} + \frac{2}{3} \left| \Psi_{BC}^+ \right\rangle \left\langle \Psi_{BC}^+ \right|$, where $\left| \Psi_{BC}^+ \right\rangle = \frac{1}{\sqrt{2}} \left(\left| 0_B 1_C \right\rangle + \left| 1_B 0_C \right\rangle \right)$ is a *Bell* state.

Clearly, after a project measurement on a system A in basis $\left\{ \left| 0 \right\rangle, \left| 1 \right\rangle \right\}$, the system A has a probability $p = \frac{1}{3}$ in the state $\left| 1 \right\rangle$ (consists to ρ_{00}^{BC}) and a probability $p = \frac{2}{3}$ in the state $\left| 0 \right\rangle$ (consists to $\left| \Psi_{BC}^+ \right\rangle$).

Remark: Physical interpretation of quantum measurement processes. Complementarity principle tells us the microscopic world has the behavior, wave-particle duality. One cannot draw pictures of individual quantum processes. To gain information of a quantum system, one has to perform a quantum measurement. A unitary operation on a quantum system will keep the wave behavior of the system. But, a non–unitary operation will destroy the wave behavior of the system. In the quantum case, the completeness relation requires that trace of ρ equal to one, $Tr(\rho) = 1$. We can see it consists to the quantum measurement:

$$1 = \sum_m p(m) = \sum_m \left\langle \psi \left| M_m^\dagger M_m \right| \psi \right\rangle.$$

A trace process can be treated as a notion that we can find out the particle in the whole space to a certain. After a trace process on a quantum system, the wave behavior disappeared completely. If the quantum

system is trace-preserving ($Tr(\rho) = 1$), which means that the trace of this quantum system is unit, the probability of find out the particle is 1.

Example: Quantum dualism and quantum measurement. Consider the case of quantum measurement. When we perform a projective measurement on a quantum system, the result of the measurement is in a state with the probability $p(m) = \langle \psi | M_m^\dagger M_m | \psi \rangle$. Since the state is in an orthogonal state after a projective measurement, according to the definition $|\psi\rangle = \sum_i a_i |i\rangle$ of quantum system, the behavior of this system is *particle*. The wave behavior disappeared. But, POVM maybe shows some *wave* behavior. Suppose a particle is a two-dimension $\{|0\rangle, |1\rangle\}$ quantum system. Consider a POVM containing three elements:

$$E_1 = \frac{\sqrt{2}}{1+\sqrt{2}} |1\rangle\langle 1|; \quad E_2 = \frac{\sqrt{2}}{1+\sqrt{2}} \cdot \frac{1}{2} \Big[\big(|0\rangle - |1\rangle\big)\big(\langle 0| - \langle 1|\big) \Big]; \quad E_3 = I - E_1 - E_2.$$

We can see that E_1, E_2 and E_3 are not orthogonal to each other. Then after a POVM, the quantum system will keep some wave behavior and appear some particle behavior.

Example: Information erasure and quantum measurement. We will assume to have a large number of bits but they will be erased individually, one by one. Landauer's principle argues that since information erasure is a logical function, which does not have a single-valued, inverse, it must be associated with physical irreversibility and require heat dissipation. Suppose we have a quantum system, which is in an unknown state ρ. An information erasure process is that we prepare this system in a standard state, $E : \rho \to |0\rangle\langle 0|$, where ρ is an arbitrary state and $|0\rangle\langle 0|$ is a standard state. This information erasure process is non-unitary generally. If the system is in a known state, this information erasure process can be realized by a unitary operation. But in the case we have a large number of bits to erase, the information erasure operation must be irreversible. In this case information erasure will be considered as the quantum state erasure of a quantum system. Information erasure operation can be realized by a quantum projective measurement and a unitary operation: $(i)\, M : \rho \to |j\rangle\langle j|; \quad (ii)\, U : |j\rangle\langle j| \to |0\rangle\langle 0|$.

First, we perform a projective measurement M. Then the quantum system will be in a known state (suppose in the state $|j\rangle\langle j|$) after the quantum measurement. We can perform a unitary operation U on the state $|j\rangle\langle j|$ to prepare the system in state $|0\rangle\langle 0|$. Suppose we want to erase the state information of quantum system A, which is in a known state ρ.

We can perform a projective measurement described by projectors P_j. We can gain the measurement result (suppose in state $|j\rangle$). Then we can perform a unitary operation U to prepare the system in a state $|0\rangle$. This is an information erasure process.

An information erasure operation will destroy the wave behavior completely. After an information erasure operation, the system is in state $|0\rangle\langle 0|$. There is no wave behavior in this quantum system. Since the information erasure operation is physical irreversible, the wave behavior of the quantum system cannot be recurred.

Quantum operations and von Neumann entropy

The entropy of a system measures the amount of uncertainty about the system before we learn its value. It is a measure of the "amount of chaos" or of the lack of information about a system. If one has complete information, i.e., if one is concerned with a pure state, $S(\rho) = 0$. A unitary operation does not change the entropy of a system because a unitary transformation does not change the eigenvalues of ρ.

In general, a non-unitary transformation would change the eigenvalues of ρ. So a non-unitary operation would change the von Neumann entropy of a quantum system.

Suppose P_j is a complete set of orthogonal projectors and ρ is a density operator. Then the entropy of the state $\rho' = \sum_i P_i \rho P_i^\dagger$ of the system after measurement is at least as great as the original entropy, $S(\rho') - S(\rho) \geq 0$, with equality iff $\rho = \rho'$. So, the uncertainty of the system increases after under the projective measurement if we never learn the result of the measurement. The entropy of the system would increase under a projective measurement. How does the entropy behave depends on the type of measurement which we perform. Projective measurement increases entropy of a quantum system. But, generalized measurements can decrease entropy of a quantum system.

Example. Consider a trace operation. The information of quantum system disappeared completely after a trace operation. So the uncertainty of the quantum system would increase. In another view, a trace process can be realized by a projective measurement if we do not know the measurement result. So a trace operation would increase the entropy of the system. Information erasure process will induce the entropy of the environment increase. Suppose a quantum system is in a state ρ. The von Neumann entropy of the system is: $S(\rho) \geq 0$, where $S(\rho) = 0$ iff ρ is a pure state. After the information erasure operation, the system is in the known state $|0\rangle\langle 0|$. The entropy of the system is zero. The entropy of the system is non-increasing. Since entropy does not decrease, the entropy of the environment must increase. To realize the information erasure process, there must be interaction between system and environment. An information erasure process has to exchange momentum-energy with environment randomly. In other view, an information erasure process can be realized by a projective measurement and a unitary operation. A unitary operation does not change the entropy of the system. The change of entropy is derived by the measurement. A measurement will induce exchange of momentum-energy between system and environment. Here, the projective measurement will decrease the entropy of a system because we have to know the result of the measurement if we want to realize the information erasure operation.

Remark. The probabilistic features of quantum states are different from that of classical thermodynamics. The difference is demonstrated by taking two totally independent facts A and B. Then the probability of both of them occurring would be: $P_{AB}^{Cl} = P(A) + P(B)$. In contrast, in quantum mechanics the wave function should be added, and by calculating the inner product probabilities are obtained. More analytically,

$$P_{AB}^Q = \left(\langle \psi_A | + \langle \psi_B | \right)\left(| \psi_A \rangle + | \psi_B \rangle \right) = \left(\langle \psi_A | \psi_A \rangle + \langle \psi_B | \psi_B \rangle \right) + 2 \operatorname{Re} \langle \psi_A^* | \psi_B \rangle$$

$$= P(A) + P(B) + \underbrace{2 \operatorname{Re} \langle \psi_A^* | \psi_B \rangle}_{Interference} \overset{in\ general}{\neq} P_{AB}^{Cl}.$$

For this reason a quantum version of thermodynamics for description of quantum probability is needed. In the sequel the name "*thermodynamical* probabilities" is used to distinguish the statistical mixture of several quantum states, from the quantum probabilities occurring by observation of a quantum state.

The unitary evolution of the system can be generalized by quantum operations $\mathrm{E}(\rho) = \sum_i E_i^\dagger \rho E_i$, where $\sum_i E_i^\dagger E_i = I$. Quantum operations are sometimes referred as superoperators. Quantum thermodynamics just can be viewed as a transition from classical to quantum thermodynamics.

Example. Suppose an ensemble of n particles is given in equilibrium. For this ensemble says that for the i-th particle, there is a probability: $p_i = \dfrac{1}{\sum_j e^{-\beta E_j}} e^{-\beta E_i}$, to be in this state, where $\beta = (k_B T)^{-1}$ with k_B Boltzmann's constant and T the temperature. If the particles, where described by quantum mechanics, then if the i-th would have an eigenenergy E_i, given by the solution of the Schrodinger equation $H | \psi_i \rangle = E_i | \psi_i \rangle$, where H is the Hamiltonian of the system.

With the help of equation $p_i = \dfrac{1}{\sum\limits_j e^{-\beta E_j}} e^{-\beta E_i}$ the density matrix $\rho \overset{\Delta}{=} \sum\limits_i p_i \left| \psi_i \right\rangle \left\langle \psi_i \right|$ can be written as:

$$\rho = \sum_i p_i \left| \psi_i \right\rangle \left\langle \psi_i \right| = \frac{1}{\sum\limits_j e^{-\beta E_j}} \sum_i \left| \psi_i \right\rangle e^{-\beta E_i} \left\langle \psi_i \right| = \frac{1}{Tr\left(e^{-\beta H}\right)} e^{-\beta H} \leftrightarrow p_i = \frac{1}{\sum\limits_j e^{-\beta E_j}} e^{-\beta E_i},$$

which is a generalization of classical thermodynamic probability $p_i = \dfrac{1}{\sum\limits_j e^{-\beta E_j}} e^{-\beta E_i}$.

Von Neumann entropy and relevant measures of quantum information theory

Willing to describe quantum information and quantum disorder, one could use quantum version of entropy, and in order to justify its mathematical definition, recall how Shannon entropy was given $H(X) = -\sum\limits_i p_i \log p_i$, and assume that a density matrix is diagonalized by $\rho = \sum\limits_x \lambda_x \left| x \right\rangle \left\langle x \right|$. Then naturally quantum entropy of ρ is defined as: $S(\rho) \overset{\Delta}{=} -\sum\limits_x \lambda_x \log \lambda_x$.

Translating this into the mathematical formalism, the von Neumann entropy is defined as

$$S(\rho) \overset{\Delta}{=} -Tr\left(\rho \log \rho\right).$$

The last formula is often used for providing theoretical results and equation

$$S(\rho) \overset{\Delta}{=} -\sum_x \lambda_x \log \lambda_x$$

is used for calculations.

Example. Let us consider the von Neumann entropy of the density matrix

$$\rho = p\left(\left|0\right\rangle\left\langle 0\right|\right) + (1-p)\frac{1}{2}\left(\left|0\right\rangle + \left|1\right\rangle\right)\left(\left\langle 0\right| + \left\langle 1\right|\right)$$

for the quantum state $\left|\psi\right\rangle = \sqrt{p}\left|0\right\rangle + \dfrac{\sqrt{1-p}}{\sqrt{2}}\left(\left|0\right\rangle + \left|1\right\rangle\right)$ that is found

$$S(\rho) = S\left(\begin{bmatrix} p + \dfrac{1-p}{2} & \dfrac{1-p}{2} \\ \dfrac{1-p}{2} & \dfrac{1-p}{2} \end{bmatrix}\right) = -\lambda_1 \log \lambda_1 - \lambda_2 \log \lambda_2,$$

where $\lambda_1 = \dfrac{1}{2}\left[1 + \sqrt{1 + 2p^2 - 2p}\right]$ and $\lambda_2 = \dfrac{1}{2}\left[1 - \sqrt{1 + 2p^2 - 2p}\right]$ are the eigenvalues of the corresponding matrix.

In general case, $S(\rho) \neq H(p, 1-p)$ even if the same probabilities where assigned for both of them.

This shows that quantum probabilities are not expelled by quantum thermodynamics. The equality could only hold if the probabilities written in Shannon's entropy are the eigenvalues of density matrix.

Main information measures

Following the same path as for classical information theory, in the quantum case it is straightforward to define the *joint entropy*, the *relative entropy* of ρ and σ, the *entropy of A conditional on knowing B*, and the *common* or *mutual information of A and B*. Each case is correspondingly:

$$(i)\ S(A,B) \overset{\Delta}{=} Tr(\rho^{AB} \log \rho^{AB}); \quad (ii)\ S(\rho\|\sigma) \overset{\Delta}{=} Tr(\rho \log \rho) - Tr(\rho \log \sigma);$$

$$(iii)\ S(A|B) \overset{\Delta}{=} S(A,B) - S(B); \quad (iv)\ S(A:B) \overset{\Delta}{=} S(A) + S(A) - S(A,B) = S(A) - S(A|B)$$

One can see that there are a lot of similarities between Shannon's and von Neumann's entropies.

Example: *Shannon and von Neumann measure of entropy*. As measure of classical information we shall use *Shannon entropy* H_{Sh}. Consider a complex vector of modulus 1 in the Hilbert space $Hil_{Q_1} \otimes ... \otimes Hil_{Q_n}$, where Hil_{Q_k} has dimension 2 for every k, written as a complex linear combination of basis vectors: $|\psi\rangle = \sum\limits_{i_1,i_2,...,i_n \in \{0,1\}} \alpha_{i_1 i_2 ... i_n} |i_1\rangle \otimes |i_2\rangle \otimes ... \otimes |i_n\rangle$. Then, the Shannon Entropy of $|\psi\rangle$ with respect to the basis $\{|i_1\rangle \otimes ... \otimes |i_n\rangle\}_{i_1,...,i_n \in \{0,1\}}$ is so defined: $H_{Sh}(|\psi\rangle) = - \sum\limits_{i_1,i_2,...,i_n \in \{0,1\}} \|\alpha_{i_1 i_2 ... i_n}\|^2 \log\left(\|\alpha_{i_1 i_2 ... i_n}\|^2\right)$, where $\|\alpha_{i_1 i_2 ... i_n}\|^2$ is interpreted as the probability of measuring vector $|i_1 i_2 ... i_n\rangle$.

The *Von Neumann entropy* is used to measure the information stored in quantum correlation. Let $\rho = |\psi\rangle\langle\psi|$ be the density matrix associated to state $|\psi\rangle$ and $T \subseteq \{1,...,n\}$. Then we define: $\rho_T = Tr_{\{1...n\}-T}\rho$, where $Tr_{\{1...n\}-T}(...)$ is the *partial trace* operator (see below).

The *Von Neumann entropy* of qubit j in $|\psi\rangle$ is defined as: $S_{|\psi\rangle}(T) = -Tr(\rho_T \log \rho_T)$

The following definitions are useful:

1	$S^{VN}_{|\psi\rangle}(h|k;l)$	=	$S^{VN}_{|\psi\rangle}(\{h\}) - I_{|\psi\rangle}(h;k) - I_{|\psi\rangle}(h;l) + I_{|\psi\rangle}(h;k;l)$
2	$I_{|\psi\rangle}(h;k;l)$	=	$S^{VN}_{|\psi\rangle}(\{h\}) + S^{VN}_{|\psi\rangle}(\{k\}) + S^{VN}_{|\psi\rangle}(\{l\}) - S^{VN}_{|\psi\rangle}(\{h,k\})$ $-S^{VN}_{|\psi\rangle}(\{h,l\}) - S^{VN}_{|\psi\rangle}(\{k,l\}) - S^{VN}_{|\psi\rangle}(\{h,k,l\})$
3	$S^{VN}_{|\psi\rangle}(h|k)$	=	$S^{VN}_{|\psi\rangle}(\{h\}) - I_{|\psi\rangle}(h;k)$
4	$S^{VN}_{|\psi\rangle}(h|k;l)$	=	$S^{VN}_{|\psi\rangle}(\{h\}) - I_{|\psi\rangle}(h;k) - I_{|\psi\rangle}(h;l) + I_{|\psi\rangle}(h;k;l)$

for $h,k,l \in \{1,..,n\}$. These quantities can be represented in a Wenn diagram of Figure 1.

Remark. Measures of entropy are different from most physical quantities. In quantum mechanics one has to distinguish between *observables* and *states*. Observables (like position, momentum, etc.) are mathematically described by self-adjoint operators in Hilbert space. States (which generally are mixed) are characterised by a density matrix $\rho \geq 0$, i.e. an Hermitian operator, with trace $Tr(\rho) = 1$. The expectation value of an observable A in the state ρ is $\langle A \rangle = Tr(\rho A)$. Now entropy is not an observable that means that there does not exist an operator with the property that its expectation value in some state would be its entropy. It is rather a *function of state*.
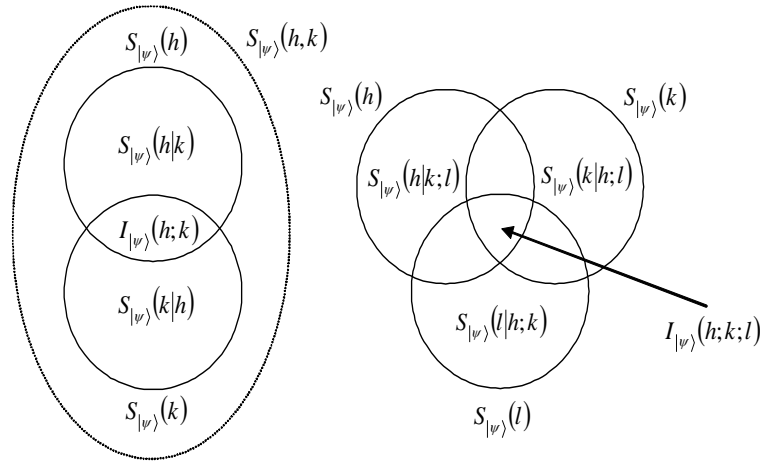
Figure 1. Wenn diagram for entropy and mutual information

Due to the Jaynes relation between the information-theoretical and physical entropy via Boltzmann's constant, k_B, one can ascribe to any quantum object a certain value of its physical entropy $S_{ph}^{Cl} = k_B H_{Sh}$.

The classical limit S_{ph}^{Cl} of the expression for the entropy can be justified mathematically on coherent states. The best thing one can do is to measure the probability of finding a particle in a state with minimum uncertainty centred on the classical values, i.e. in a coherent state.

Remark. In general case $S_{ph}^{Cl} \geq S$. Let, with the usual observation $z := \dfrac{(q + ip)}{\sqrt{2\hbar}}$, $|z\rangle = W(z)|0\rangle$ be a coherent state with expectation values of position, or momentum, q or p, respectively. In configuration space, $|0\rangle$ is explicitly given by the wave function $\dfrac{1}{\sqrt[4]{\pi\hbar}} e^{-\frac{x^2}{2\hbar}}$ and $W(z)$ is the unitary operator $W(z) = e^{\frac{1}{\hbar}(pQ - qP)}$ with Q, P operators of position, or momentum, respectively.

We now define the classical density distribution corresponding to the density matrix ρ by

$$\rho(z) := \langle z|\rho|z\rangle.$$

For every function $f(z)$ there exists at most one density matrix ρ that $\rho(z) = f(z)$ and $Tr\rho = \int \dfrac{dz}{\pi} \langle z|\rho|z\rangle = 1$. The relation $S_{ph}^{Cl} \geq S$ is true because, for $s(x) := \{-x \ln x (x > 0); 0 (x = 0)\}$ due to concavity, $s(\langle z|\rho|z\rangle) \geq \langle z|s(\rho)|z\rangle$.

Hence: $S_{ph}^{Cl} = \int \dfrac{dz}{\pi} s(\rho(z)) \geq \int \dfrac{dz}{\pi} \langle z|s(\rho)|z\rangle = Tr(s(\rho)) = S(\rho)$.

More generally, for any convex (concave) function f, $Tr(f(\rho)) \leq (\geq) \int \dfrac{dz}{\pi} f(\rho(z))$. By continuity of $\rho(z)$, $S_{ph}^{Cl} = S$ would imply $s(\langle z|\rho|z\rangle) = \langle z|s(\rho)|z\rangle$ for all z, i.e. regarding the strict concavity of $S(\cdot)$, every $|z\rangle$ must be an eigenvector of ρ, which is impossible.

Hence, $S_{ph}^{Cl} > S$.

Remark. The classical entropy is not invariant under every unitary transformation, i.e. we do have not $S_{ph}^{Cl}(U^*\rho U) = S_{ph}^{Cl}(\rho)$ for every U but for a restricted class only. For instance if $U = W(z_0)$, then

$$S_{ph}^{Cl}(W(-z_0)\rho W(z_0)) = \int \frac{dz}{\pi} S(\langle 0|W(-z)W(z_0)W(z_0)W(z_0)|0\rangle) = \int \frac{dz}{\pi} S(\rho(z+z_0)) = S_{ph}^{Cl}(\rho).$$

This argument also works for all unitary U such that $UW(z) = W(z')$ times a phase factor provided that $dz = dz'$ (canonical transformation).

If $|\psi\rangle$ is a pure state (unit vector), $\rho = |\psi\rangle\langle\psi|$. Then $\rho(z) = |\langle\psi|z\rangle|^2$, and

$$S_{ph}^{cl}(\rho) = -2\int \frac{dz}{\pi} |\langle\psi|z\rangle|^2 \ln|\langle\psi|z\rangle|^2.$$

Inserting for $|\psi\rangle = |z_0\rangle$, we obtain: $S_{ph}^{cl}(\rho) = \int \frac{dz}{\pi} e^{-|z|^2} \cdot |z|^2 = 1$.

On the other hand, there exist pure states with arbitrary high classical entropy: it suffices to show that for every $\varepsilon > 0$ one can find unit vectors $|\psi\rangle$ such that $\langle\psi|z\rangle < \varepsilon$ for all z.

For them a well-known inequality tells us that: $S_{ph}^{cl}(\rho) \geq -\ln\varepsilon^2$.

We conjecture that the states with *minimal classical entropy* are exactly given by the *density matrices* $|z\rangle\langle z|$, and consequently, $S_{ph}^{cl}(\rho) \geq 1$.

In order that $S_{ph}^{cl}(\rho)$ be small, $Sup|\rho(z)|$ must be close to one, otherwise the inequality mentioned before gives a value too large for classical entropy. Now if $Sup|\rho(z)|$ is exactly equal to 1, then, by continuity, there is some z_0 with $\rho(z_0) = 1$, i.e. $\langle z_0|\rho|z_0\rangle = 1$. Since $\|\rho\| \leq 1$ this implies $\|\rho\| = 1$ and $\rho(z_0) = |z_0\rangle$; on the other hand, $Tr(\rho) = 1$, hence $\rho = |z_0\rangle\langle z_0|$ because all other eigenvalues of ρ must be 0.

Thus in general case $S_{ph}^{Cl} \geq S_{ph}^{vN}$. These entropies are used below in information analysis of the quantum evolution.

The *quantum conditional entropy* $S(X|Y)$ is defined by a natural generalization of the classical case as

$$S^{vN}(X|Y) = -Tr_{XY}\left[\rho_{XY} \lg \rho_{X|Y}\right], \text{ where } \rho_{X|Y} = \lim_{n\to\infty}\left[\rho_{XY}^{\frac{1}{n}}(I_X \otimes \rho_Y)^{-\frac{1}{n}}\right]^n,$$ which should be an analogue to the

conditional probability $p(x|y)$. The quantity I_X is the unit matrix in the Hilbert space for X and $\rho_Y = Tr[\rho_{XY}]$ denotes a "marginal" density matrix – an analogue to the "marginal" probability $p_y = \sum_x p(x,y)$. The above definition of $S(X|Y)$ leads to an analogue of the classical Shannon theory holds also for quantum entropy: $S^{vN}(X|Y) = S^{vN}(X,Y) - S^{vN}(Y)$.

The complicated expression of $\rho_{X|Y}$ and the fact that a limit has to be used are due to the fact that the joint density matrix (ρ_{XY}) and the marginal matrices $I_X \otimes \rho_Y$ do not commute in general. If they do commute, the whole expression gets much simpler, as discussed below. In spite of the apparent similarity between quantum $S^{vN}(X|Y)$ and classical $S^{Sh}(X|Y)$ the fact that in the quantum case we deal with (density) matrices, rather than with numbers, as in the classical case, brings quite a different situation for quantum information theory and potential far exceeding the classical one.

While $p(x|y)$ is a probability distribution in x (i.e., $0 \leq p(x|y) \leq 1$), its quantum analogue $\rho(X|Y)$ is not a density matrix. It is Hermitian and positive but its eigenvalues can be larger than 1 and, consequently, the conditional entropy can be *negative*.

This helps to explain the well-known fact that quantum entropy is non-monotonic, and it can be the case that $S^{vN}(X,Y) < S^{vN}(Y)$, i.e. the quantum entropy of the entire system can be smaller than entropy of one of its

subparts (what is not possible in the classical case). This happens, for example, in the case of quantum entanglement as shown in the following example.

Example. Consider the Bell state $|\psi\rangle = \frac{1}{\sqrt{2}}\big(|00\rangle + |11\rangle\big)$ of the Hilbert space $H_{AB} = H_A \otimes H_B$, where $H_A = H_B = H_2$. The density matrices $\rho_{AB} = |\psi\rangle\langle\psi|$, ρ_A, $\rho_{A|B}$ are as follows:

$$\rho_A = \begin{pmatrix} \dfrac{1}{2} & 0 \\ 0 & \dfrac{1}{2} \end{pmatrix}, \quad \rho_{A|B} = \begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 \end{pmatrix}, \text{ and } \rho_{AB} = \begin{pmatrix} \dfrac{1}{2} & 0 & 0 & \dfrac{1}{2} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ \dfrac{1}{2} & 0 & 0 & \dfrac{1}{2} \end{pmatrix}.$$

Hence: $S^{vN}(A) = S^{vN}(B) = 1$. The density matrix is $\rho_{A|B} = \rho_{AB}(I_A \otimes \rho_B)^{-1}$ (because in this case the joint and marginal matrices ρ_{AB} and $I_A \otimes \rho_B$ commute).

Hence, $S(AB) = S(B) + S(A|B) = 1 - 1 = 0$ because $S(A|B) = -1$.

Unlike in classical (Shannon) information theory, quantum (von Neumann) conditional entropies *can be negative* when considering quantum entangled systems, a fact related to quantum non-separability, so that *entanglement* might be viewed as *super-correlation*.

Quantum mutual information $I(X:Y)$, as an analogue of the classical mutual information $I(X:Y)$ is defined on the base of "mutual" density matrix $\rho_{X:Y} = \lim_{n\to\infty}\left[(\rho_X \otimes \rho_Y)^{\frac{1}{n}}\rho_{XY}^{-\frac{1}{n}}\right]^n$, and the definition implies the standard relation: $I(X:Y) = S(X) - S(X|Y) = S(X) + S(Y) - S(XY)$.

This definition is reduced to the classical one for the case where ρ_{XY} is a diagonal matrix.

However, not all basic equalities and inequalities of the classical information theory transfer from the classical to the quantum case. For example, in the classical case: $I(x:y) \le \min\big[S(x), S(y)\big]$, but in the quantum case the best upper pound possible is $S(X:Y) \le 2\min\big[S(X), S(Y)\big]$.

Additional basic properties of quantum von Neumann entropy are summarized below:

Invariance: $S^{vN}(\rho) = S^{vN}(U^*\rho U)$ for any unitary matrix *U*;
2.1. Concavity: if $\rho = \lambda_1\rho_1 + \lambda_2\rho_2$, $\lambda_1 + \lambda_2 = 1$, $\lambda_1 \ge \lambda_2 \ge 0$, then $S^{vN}(\rho) \ge \lambda_1 S^{vN}(\rho_1) + \lambda_2 S^{vN}(\rho_2)$;   2.2. Strong concavity: **a.** $S\left(\sum_i p_i \rho_A^i \otimes \rho_B^i\right) \ge \sum_i p_i S\left(\rho_A^i\right) + \sum_i S\left(p_i \rho_B^i\right)$;   **b.** $S\left(\sum_i p_i \rho_A^i \otimes \rho_B^i\right) \ge \sum_i p_i S\left(\rho_B^i\right) + \sum_i S\left(p_i \rho_A^i\right)$;   **c.** $S\left(\sum_i p_i \rho_A^i \otimes \rho_B^i\right) \ge \max\left\{\sum_i p_i S\left(\rho_A^i\right) + \sum_i S\left(p_i \rho_B^i\right), \sum_i p_i S\left(\rho_B^i\right) + \sum_i S\left(p_i \rho_A^i\right)\right\}$
3.1. Additivity: $S(\rho_1 \otimes \rho_2) = S(\rho_1) + S(\rho_2)$;   **3.2. Subadditivity**: if ρ is a density matrix of a bipartite system $A \otimes B$,

$\rho_1 = Tr_B\rho, \rho_2 = Tr_A\rho$, then $S^{vN}(\rho) \leq S^{vN}(\rho_1) + S^{vN}(\rho_2) \leq S^{vN}(\rho_1 \otimes \rho_2)$;
Lower bound: $S^{vN}(\rho) \geq -\log\lambda_1$, where λ_1 is the largest eigenvalue of ρ.
Triangle inequality: $\left| S_A - S_B \right| \leq 2S_C \leq S_A + S_B, \quad S_C = \frac{1}{2}(S_A + S_B)$

Relative entropy

The classical Kullback-Leibler information admits several extensions to the quantum case. The most successful appeared the most straightforward definition due to *Umegaki* as:

$$S(\rho\|\sigma) \overset{\Delta}{=} Tr(\rho\log\rho) - Tr(\rho\log\sigma).$$

This entropy is an important tool in the quantum theory of noise channels. Its basic property is monotonicity (*Uhlmann inequality*). We have for any quantum operation Λ:

$$S(\Lambda\rho\|\Lambda\sigma) \leq S(\rho\|\sigma).$$

This inequality can be viewed the counterpart of the Boltzmann H-theorem in statistical physics.

As such one can prove the following result:

$$S(\rho\|\sigma) > 0, S(\rho\|\sigma) = 0 \Leftrightarrow \rho = \sigma,$$

and is known as *Klein*'s inequality. This inequality provides evidence of why von Neumann relative entropy is close to the notion of metric. Relative entropy satisfies also the following inequality:

$$S(\rho\|\sigma) \geq \frac{1}{2}\|\rho - \sigma\|,$$

which is generalization of the similar one for Kullback-Leibler relative entropy.

Main properties of relative entropy are as following:

Monotonicity of relative entropy under completely positive, trace preserving maps Φ $$S(\Phi(\rho)\|\Phi(\sigma)) \leq S(\rho\|\sigma)$$
Monotonicity of relative entropy under partial trace $$S(\rho_A\|\sigma_B) \leq S(\rho_{AC}\|\sigma_{BC})$$
Strong subadditivity of von Neumann entropy I and II, where I and II are equivalent (I) $S(\rho_A) + S(\rho_B) \leq S(\rho_{AC}) + S(\rho_{BC})$; (II) $S(\rho_{ABC}) + S(\rho_B) \leq S(\rho_{AB}) + S(\rho_{BC})$
Joint convexity of relative entropy $$S\left(\sum_i p_i\rho^i \left\| p_i\sigma^i\right.\right) \leq \sum_i p_i S(\rho^i\|\sigma^i)$$
Generalized (equivalent) inequality, concavity of conditional entropy

$$S\left(\sum_i p_i \rho_{AB}^i\right) - S\left(\sum_i p_i \rho_B^i\right) \geq \sum_i p_i \left[S\left(\rho_{AB}^i\right) - S\left(\rho_B^i\right) \right]$$

The entropy exchange $S_{ex}(\rho, \mathrm{E})$ for a quantum operation

$$S(\mathrm{E}(\rho)) - S(\rho) + S_{ex}(\rho, \mathrm{E}) \geq 0$$

Since these five inequalities are equivalent, we can obtain any one of them from one of the other four inequalities.

Remark. The *Kullback-Leibler* divergence offers an information-theoretic basis for measuring the divergence between two given distributions. Its quantum analog fails, however, to ply a corresponding role for comparing two density matrices, if the reference states are pure states. The non-additive quantum information theory inspired by nonextensive statistical mechanics is free from such a difficulty and the associated quantity, termed the quantum q-divergence, can in fact be a good information-theoretic measureof the degree of state purification. The corresponding relation between the ordinary divergence and the q-divergence is violated for the pure states, in general.

Example: *Quantum Kullback-Leibler divergence*. In classical information theory, a comparison of two distributions is customarily discussed by employing the Kullback-Leibler divergence. Its quantum-mechanical counterpart is the quantum divergence of a density matrix ρ with respect to a reference density matrix σ, which is given by:

$$K\left[\rho \| \sigma\right] = Tr\left[\rho(\ln \rho - \ln \sigma)\right],$$

where the equality holds iff $\rho = \sigma$. However, since $\ln \sigma$ is a singular quantity if the reference state σ is a pure state, this quantity turns out to be inadequate for a measuring degree of purification. The quantum q-divergence is obtained by replacing the derivative

$$K\left[\rho \| \sigma\right] = \frac{d}{dx} Tr\left(\rho^x \sigma^{1-x}\right)\Big|_{x \to 1_{-0}},$$

with the Jackson q-derivative: $K_q\left[\rho \| \sigma\right] = D_q Tr\left(\rho^x \sigma^{1-x}\right)\Big|_{x \to 1_{-0}}$, where $D_q = \dfrac{f(qx) - f(x)}{x(q-1)}$

denotes the Jackson differential operator. Than $K_q\left[\rho \| \sigma\right]$ can be defined as

$$K_q\left[\rho \| \sigma\right] = \frac{1}{1-q} Tr\left[\rho^q\left(\rho^{1-q} - \sigma^{1-q}\right)\right] = Tr\left[\rho^q\left(\ln_q \rho - \ln_q \sigma\right)\right].$$

The crucial difference between $K\left[\rho \| \sigma\right]$ and $K_q\left[\rho \| \sigma\right]$ is that, in marked contrast with $\ln \sigma$, $\ln_q \sigma$ is a well-defined quantity for a pure reference state $\sigma = |\psi\rangle\langle\psi|$. In fact, in this case, $\sigma^{1-q} = \sigma (0 < \sigma < 1)$, whereas $\ln \sigma = (\sigma - I)\zeta(1)$, which is divergent, where I and $\zeta(1)$ are the unit matrix, and the Riemann ζ function, respectively.

Accordingly, equation $K_q\left[\rho \| \sigma\right] = Tr\left[\rho^q\left(\ln_q \rho - \ln_q \sigma\right)\right]$ is seen to be:

$$K_q\left[\rho \| \sigma = |\psi\rangle\langle\psi|\right] = \frac{1}{1-q}\left(1 - \langle\psi|\rho^q|\psi\rangle\right).$$

The additive limit $q \to 1$ cannot be taken in this equation anymore.

Particular case. Let ρ is also a pure state, $\rho = |\phi\rangle\langle\phi|$. Then

$$K_q\left[|\phi\rangle\langle\phi|\,\big\|\,|\psi\rangle\langle\psi|\right] = \frac{d_{FS}^2}{1-q},$$

where $d_{FS}^2 = 1 - |\langle\phi|\psi\rangle|^2$ is the *Fubini-Study metric* in the projective Hilbert space, which may give the geometric interpretation to quantum uncertainty and correlation.

Example: *Purification of Werner state*. The Werner state is a state of a bipartite spin-$\frac{1}{2}$ system, i.e., two qubit, given as follows:

$$\rho_W = F|\Psi^-\rangle\langle\Psi^-| + \frac{1-F}{3}\left(|\Psi^+\rangle\langle\Psi^+| + |\Phi^+\rangle\langle\Phi^+| + |\Phi^-\rangle\langle\Phi^-|\right),$$

where $|\Psi^\pm\rangle = \frac{1}{\sqrt{2}}(|01\rangle \pm |10\rangle), |\Phi^\pm\rangle = \frac{1}{\sqrt{2}}(|00\rangle \pm |11\rangle)$ are the *Bell* states, and F is the fidelity with respect to the reference state $\sigma = |\Psi^-\rangle\langle\Psi^-|$.

Its allowed range is $\frac{1}{4} \le F \le 1$, and ρ_W is known to be separable iff $F \le \frac{1}{2}$.

The quantum q-divergence of ρ_W with respect to the reference state $\sigma = |\Psi^-\rangle\langle\Psi^-|$ is calculated to be

$$K_q\left[\rho_W\,\big\|\,\sigma = |\Psi^-\rangle\langle\Psi^-|\right] = \frac{1}{1-q}\left(1 - F^q\right) \ge 0,$$

where the zero value is realized when $F = 1$ or $q \to 0_+$. However, as already stressed, the limit $q \to 1_{-0}$ is singular and does not commute with the limit $F \to 1_{-0}$.

Remark. The relative entropy is used as a distance; this distance is not a metric because it is not symmetric under interchange of the states. Corresponding to Shannon's entropy, $S(\rho) < \log d, S(\rho) = \log d \Leftrightarrow \rho = \frac{1}{d}$. In addition to this from definition of von Neumann entropy it follows that: $S(\rho) > 0, S(\rho) = 0 \Leftrightarrow \rho\ is\ pure$.

One can also prove that supposing some ρ_i states, with probabilities p_i have support on orthogonal subspaces, then:

$$S\left(\sum_i p_i\rho_i\right) = H(p_i) + \sum_i p_i S(\rho_i).$$

Directly from this relation the joint entropy theorem can be proved, where supposing that p_i are probabilities, $|i\rangle$ are orthogonal states for a system A, and ρ_i is an set of density matrices of another system B, then:

$$S\left(\sum_i p_i|i\rangle\langle i| \otimes \rho_i\right) = H(p_i) + \sum_i p_i S(\rho_i).$$

Example: *Splitting of information in a particular quantum state into classical and quantum part*. Consider performing a general measurement on the state, $A_i A_i^\dagger$, such that $\rho_B^i = \dfrac{A_i \rho_B A_i^\dagger}{Tr\left(A_i \rho_B A_i^\dagger\right)}$. The final state of subsystem B is then $\sum_i A_i \rho_B A_i^\dagger = \sum_i p_i \rho_B^i$. The entropy of the residual states is $\sum_i p_i S\left(\rho_B^i\right)$. The classical information obtained by measuring outcomes i with probabilities p_i is $H(p)$. If the states ρ_B^i have support o orthogonal subspaces, then the entropy of the final state is the sum of the residual entropy and the classical information, i.e.,

$$S\left(\sum_i p_i \rho_B^i\right) = \underbrace{H(p)}_{Classical} + \underbrace{\sum_i p_i S\left(\rho_B^i\right)}_{Quantum}.$$

It has been shown that the state $\rho_B = \sum_i p_i \rho_B^i$ can be reconstructed with arbitrary high fidelity from the classical measurement outcomes and the residual states iff the residual states ρ_B^i are on orthogonal subspaces. We see then that the information in a quantum state may be split into a *quantum* and a *classical* part.

The entropy of tensor product $\rho \otimes \sigma$ is found to be $S(\rho \otimes \sigma) = S(\rho) + S(\sigma)$. Another result can be derived by *Schmidt decomposition* (see, in details Vol. 83).

If a composite system AB is in a pure state, it has a subsystems A and $\boldsymbol{B}$ with density matrices of equal eigenvalues, and $S(A) = S(B)$.

A great discrepancy between classical and quantum information theory is quantum *entanglement* and its measures. The tools described in above can help reveal a great discrepancy between classical and quantum information theory: *entanglement*.

Entanglement

In quantum mechanics two states are named entangled if they cannot be written as a tensor product of other states. For the demonstration of the above mentioned discrepancy, let for example a composite system AB be in an entangled pure state $|AB\rangle$, then, because of entangled, in the Schmidt decomposition it should be written as the sum of more than one terms: $|AB\rangle = \sum_{i \in I} \lambda_i |i_A\rangle |i_B\rangle$, with $|I| > 1$, where $|i_A\rangle$ and $|i_B\rangle$ are orthonormal bases. The corresponding density matrix is obviously

$$\rho^{AB} = |AB\rangle\langle AB| = \sum_{i,j \in I} \lambda_i \lambda_j |i_A\rangle |i_B\rangle \langle j_A| \langle j_B|.$$

As usually the density matrix of the subsystem B can be found by tracing out system A,

$$\rho^B = Tr_A\left(\rho^{AB}\right) = \sum_{i,j,k \in I} \lambda_i \lambda_j \langle k_A | i_A\rangle |i_B\rangle \langle j_A | k_A\rangle \langle j_B| = \sum_{i \in I} \lambda_i^2 |i_B\rangle \langle i_B|.$$

Because of assumption $|I| > 1$ and the fact that $|i_B\rangle$ are orthonormal bases and it is impossible to collect them together in a tensor product, subsystem B is not pure. Thus, according to the property: $S(\rho) > 0$, $S(B) > 0$, AB is pure and $S(\rho) = 0 \Leftrightarrow \rho$ *is pure*, thus $S(A,B) = 0$ and obviously by the definition

(iii) $S(A|B) \overset{\Delta}{=} S(A,B) - S(B)$, $S(A|B) < 0$,

i.e., *negative*.

The last steps can be repeated backward and conclusion, which can be drawn, is that a pure composite system AB is entangled iff $S(A|B) < 0$.

Remark. In classical information theory conditional entropy only be $H(X|Y) > 0$ and that is obviously the reason why entangled states did not exit at all. This is an exclusive feature of quantum information theory. Entangled states are named after the fact $S(A|B) < 0 \Leftrightarrow S(A,B) < S(B)$, which means that the ignorance about a system B can be in quantum mechanics more than the ignorance than of both A and B. This proposes some correlation between these two systems.

Example. Imagine a simple pair of quantum particles, with two possible states each $|0\rangle$ and $|1\rangle$. Then a possible formulation of entanglement can be a state $|\psi\rangle = \frac{1}{\sqrt{2}}|0\rangle \otimes |0\rangle + \frac{1}{\sqrt{2}}|1\rangle \otimes |1\rangle$.

After a measurement M_m of the first particle for example, according to

$$|\psi'\rangle = \frac{M_m}{\sqrt{\langle\psi|M_m^\dagger M_m|\psi\rangle}}|\psi\rangle, \quad M_m|\psi\rangle = 1\left(|m\rangle \otimes |m\rangle\right) + 0\left(1 - |m\rangle\right) \otimes \left(1 - |m\rangle\right) = |m\rangle \otimes |m\rangle,$$

hence they collapse to state $|m\rangle$.

This example sheds light into the quantum property, where ignorance of both particles is greater than the ignorance of one of them, since perfect knowledge about the second.

Basic properties of von Neumann entropy and Wigner-Yanase-Dyson skew entropy

The basic properties of von Neumann entropy (which can be compared to the properties of Shannon entropy) shown in Table 2.

Table 2. Main properties of von Neumann entropy

	Property of von Neumann entropy
	Symmetry: $S(A,B) = S(B,A);\quad S(A:B) = S(B:A)$
	Unitary operations preserve entropy: $S(U\rho U^\dagger) = S(\rho)$
	Subadditivity: $S(A,B) < S(A) + S(B)$
	Triangle (or *Araki-Lieb*) *inequality*: $S(A,B) \geq |S(A) - S(B)|$
	Strict concavity of the entropy: Suppose that are probabilities $p_i \geq 0$ and the corresponding density matrices ρ_i, then $$S\left(\sum_i p_i \rho_i\right) < \sum_i p_i S(\rho_i) \text{ and } S\left(\sum_i p_i \rho_i\right) < \sum_i p_i S(\rho_i) \Leftrightarrow \rho_i \text{ for which } p_i > 0 \text{ are all identical}$$
	Upper bound of a mixture of states: Suppose $\rho = \sum_i p_i \rho_i$ where $p_i \geq 0$ are probabilities and the corresponding density matrices ρ_i, then $$S(\rho) < \sum_i p_i S(\rho_i) + H(p_i) \quad \text{and} \quad S(\rho) = \sum_i p_i S(\rho_i) + H(p_i) \Leftrightarrow \rho_i \text{ have support}$$

	on orthogonal subspaces
	Strong subadditivity: $S(A,B,C)+S(B)<S(A,B)+S(B,C)$, or equivalently $$S(A)+S(B)<S(A,C)+S(B,C)$$
	<u>*Conditioning reduced entropy*</u>: $S(A|B,C)<S(A|B)$
	<u>*Discarding quantum system never increase mutual information*</u>: Suppose ABC is a composite quantum system, then $S(A:B)<S(A:B,C)$
0	<u>*Trace preserving quantum operations never increase mutual information*</u>: Suppose AB is a composite quantum system and E is a trace preserving quantum operation on system B. Let $S(A:B)$ denote the mutual information between systems A and B before E applied to system B, and $S(A':B')$ the mutual information after E is applied to system B. Then $$S(A':B')\le S(A:B)$$
1	Relative entropy is jointly convex in its arguments: Let $0\le\lambda\le1$, then $$S\big(\lambda A_1+(1-\lambda)A_2\|\lambda B_1+(1-\lambda)B_2\big)\ge\lambda S(A_1\|B_1)+(1-\lambda)S(A_2\|B_2)$$
2	The relative entropy is monotonic: $S\big(\rho^A\|\sigma^B\big)\le S\big(\rho^{AB}\|\sigma^{AB}\big)$

Partial trace, entropy inequalities and skew-entropy

The partial trace operation described above and the strong subadditivity property of entropy in quantum information theory can be explained also in linear algebra terms. Let A be a density matrix and K any self-adjoint operator. For $0<t<1$ let

$$S_t(A,K)=\frac{1}{2}Tr\big[A^t,K\big]\big[A^{1-t},K\big],$$

where $[X,Y]=XY-YX$ stands for commutator.

This quantity is a measure of non-commutativity of A and K, and is called skew-entropy of *Wigner-Yanase-Dyson* (see below). This too is a concave function of A, a consequence of a more general theorem.

Theorem 1 (Lieb): The function $f(A,B)=Tr\big(X^*A^tXB^{1-t}\big)$ of positive matrices A,B is jointly concave for each matrix X and for $0\le t\le1$

Using the familiar identification of $\mathrm{L}(\mathrm{H})$ with $\mathrm{H}\otimes\mathrm{H}^*$, this statement can be reformulated as: the function $g(A,B)=A^t\otimes B^{1-t}$ is jointly concave for $0\le t\le1$.

Another quantity of interest is the relative entropy:

$$S(A|B)=TrA(\log A-\log B)$$

associated with a pair of density matrices A,B. If f is any convex function on the real line, then for Hermitian matrices A,B

$$Tr\big[f(A)-f(B)\big]\ge Tr\big[(A-B)f'(B)\big].$$

This inequality (called *Klein*'s *inequality*) applied to the function $f(t) = t \log t$ on the positive half-line shows that for positive matrices A, B,

$$TrA(\log A - LogB) \geq Tr\left[(A-B)\right].$$

If A, B are density matrices, then $Tr\left[(A-B)\right] = 0$, and hence $S(A|B) \geq 0$. It is easy corollary of Lieb's concavity theorem A9.1 that $S(A|B) \geq 0$ is jointly convex in A, B.

Remark. Choose $X = I$ in $f(A,B) = Tr\left(X^* A^t X B^{1-t}\right)$ and differentiate this function at $t = 0$.

Finally, we come to the properties of entropy von Neumann $S(A)$ and $S(A|B)$ related to *partial trace*. Let A_1, A_2 be density matrices. It is easy to see that $S(A)$ is additive over tensor product: $S(A_1, A_2) = S(A_1) + S(A_2)$. Now let A be a density matrix on $H_1 \otimes H_2$ and let A_1, A_2 be its partial traces. The subadditivity of $S(A)$ says that: $S(A) \leq S(A_1) + S(A_2)$.

This can be proved as follows. From $S(A|B) \geq 0$ we have

$$0 \leq S(A|A_1 \otimes A_2) = TrA(\log A - log(A_1 \otimes A_2)) = TrA(\log A - log(A_1 \otimes I) - log(I \otimes A_2))$$
$$= Tr(A \log A - A_1 \log A_1 - A_2 \log A_2)$$
$$= -S(A) + S(A_1) + S(A_2)$$

It is obvious that $S(UAU^* | UBU^*) = S(A|B)$ for every unitary matrix U.

Hence, from the representations

$$\Phi_1(A) = \frac{1}{m}\sum_{k=0}^{m-1} W^{*k} A W^k, \quad \Phi_2(A) = \frac{1}{m}\sum_{k=0}^{m-1} X^{*k} A X^k,$$

$W = U \oplus U \oplus \ldots \oplus U \ (n \ copies)$, $X = V \oplus V \oplus \ldots \oplus V \ (n \ copies)$,

if $A = \left[A_{ij}\right] 1 \leq 1, j \leq n$, is partial trace we see that

$$S\left(\Phi_2 \circ \Phi_1(A) | \Phi_2 \circ \Phi_1(B)\right) \leq S\left(\Phi_1(A) | \Phi_1(B)\right) \leq S(A|B).$$

Thus we have $S(A_1|B_1) \leq S(A|B)$.

More general result can be proved using same techniques: if Φ is any completely positive, trace-preserved (CTP) map, then $S\left(\Phi(A) | \Phi(B)\right) \leq S(A|B)$.

Now consider a tensor product $H_1 \otimes H_2 \otimes H_3$ of three Hilbert spaces. For simplicity of notation let us use the notation A_{123} for an operator on this Hilbert space, and drop the index j when we take a partial trace $Tr_{H_{ij}}$. Thus $Tr_{H_3} A_{123} = A_{12}, Tr_{H_3} A_{12} = A_2$, etc. Using the diminish property

$$S\left(\Phi(A) | \Phi(B)\right) \leq S(A|B)$$

with respect to the partial trace $Tr_{H_3} A_{123} = A_{12}$ we see that

$$S\left(A_{12}\middle|A_1\otimes A_2\right)\le S\left(A_{123}\middle|A_1\otimes A_{23}\right).$$

We have seen providing $S\left(A\right)\le S\left(A_1\right)+S\left(A_2\right)$ that

$$S\left(T\middle|T_1\otimes T_2\right)=-S\left(T\right)+S\left(T_1\right)+S\left(T_2\right).$$

So the above inequality can be written as

$$-S\left(A_{12}\right)+S\left(A_1\right)+S\left(A_2\right)\le -S\left(A_{123}\right)+S\left(A_1\right)+S\left(A_{23}\right)$$

and on rearranging terms, we are received strong subadditivity of entropy is the following statement.

Theorem 2 (Lieb-Ruskai): Let A_{123} be a density matrix on $\mathrm{H}_1\otimes\mathrm{H}_2\otimes\mathrm{H}_3$. Then

$$S\left(A_{123}\right)+S\left(A_2\right)\le S\left(A_{12}\right)+S\left(A_{23}\right).$$

The partial trace operation described above and the strong subadditivity property of entropy in quantum information theory can be explained also in linear algebra terms.

Information-theoretic models of quantum measurements. There are two types of evolution a quantum system can be undergone: *unitary*; *measurement*.

The first one is needed to preserve probability during evolution.

The information-theoretic interpretation of entropy is in this case as following: Entropy is the amount of knowledge one has about a system.

One is relieved by seeing that knowledge can decrease or increase by measurement, as seen by Results 1 and 2 (see below), and that is what measurement was meant to be in the first place.

Result 3 instructs that if only unitary evolutions were present in quantum theory, then no knowledge on any physical system could exist.

Thus information theory can explain why a second type of evolution is needed as following.

Measurement	Entropy law of evolution
1. Projective measurement	*Projective measurement can increase entropy*: This is derived using strict concavity. Let P be a projector and $Q\overset{\Delta}{=}I-P$ the complementary projector, then exist unitary matrices U_1, U_2 and a probability p such that for all ρ, $$P\rho P+Q\rho Q=pU_1\rho U_1^{\dagger}+\left(1-p\right)U_2\rho U_2^{\dagger},$$ thus $$\begin{aligned}S\left(P\rho P+Q\rho Q\right)&=S\left(pU_1\rho U_1^{\dagger}+\left(1-p\right)U_2\rho U_2^{\dagger}\right)\\&\ge pS\left(U_1\rho U_1^{\dagger}\right)+\left(1-p\right)S\left(U_2\rho U_2^{\dagger}\right)\\&=pS\left(\rho\right)+\left(1-p\right)S\left(\rho\right)=S\left(\rho\right)\end{aligned}$$ Because of strict concavity the equality holds iff $P\rho P+Q\rho Q=\rho$
	General measurement can decrease entropy: One can convenient by considering a qubit in state $\rho=\frac{1}{2}\left|0\right\rangle\left\langle 0\right|+\frac{1}{2}\left|1\right\rangle\left\langle 1\right|$, which is not pure thus $S\left(\rho\right)>0$, which is measured using the measurement matrices $M_1=\left|0\right\rangle\left\langle 0\right|$

2. General measurement	and $M_2 = \lvert 0 \rangle \langle 1 \rvert$. If the result of the measurement is unknown then the state of the system afterwards is $\rho' = M_1 \rho M_1^\dagger + M_2 \rho M_2^\dagger = \lvert 0 \rangle \langle 0 \rvert$, which is pure, hence $S(\rho') = 0 < S(\rho)$
3. Unitary evolution	*Unitary evolution preserves entropy*: This is already seen as von Neumann's entropy property

Accessible information

Some very important results derived by information theory, which will be useful for the development, concern the amount of accessible information and how data can be processed.

Accessible classical information: Fano's inequality. Let $\tilde{X} = f(Y)$ is some function, which is used as the best guess for X and let $p_e \overset{\Delta}{=} p(X \neq \tilde{X})$ be the probability that this guess is incorrect. Then an "error" random variable can be defined as following:

$$E = \begin{cases} 1, & X \neq \tilde{X} \\ 0, & X = \tilde{X} \end{cases},$$

thus $H(E) = H(p_e)$. Of major importance, in classical information theory, is the amount of information that can be extracted from a random variable X based on the knowledge of another random variable Y.

That should be given as an upper bound for $H(X|Y)$ and is known as *Fano's inequality*:

$$H(p_e) + p_e \log(|X| - 1) \geq H(X|Y).$$

Accessible quantum information: quantum Fano's inequality and the Holevo bound

There exists analogous relation to above classical Fano's inequality, in quantum information theory, named *quantum Fano's inequality*:

$$S(\rho, \mathrm{E}) \leq H(F(\rho, \mathrm{E})) + (1 - F(\rho, \mathrm{E})) \log(d^2 - 1),$$

where $F(\rho, \mathrm{E})$ is the entanglement fidelity of a quantum operation defined as $F(\rho, \mathrm{E}) = \sum_i \lvert Tr(\rho E_i) \rvert^2$, E_i are the operation elements of E. Quantum fidelity quantifies how much entanglement between subsystems sent a quantum channel E is preserved (see measure definition of quantum fidelity below). In the above equation, the entropy exchange of the operator E upon ρ was introduced as $S(\rho, \mathrm{E}) = S(R', Q')$, purified by R. The prime notation is used to indicate the states after the application of E. The entropy exchange does not depend upon the way in which the initial state of subsystem Q is purified by R. This is because any two purifications of Q into RQ are related by a unitary operation on the system R, and because of von Neumann entropy property.

Remark. Quantum Fano's inequality is proven by taking an orthonormal basis $|i\rangle$ for the system RQ, chosen so that the first state in the set $|1\rangle = |RQ\rangle$. Forming the quantities $p_i = \langle i | \rho^{R'Q'} | i \rangle$, then it follows that $S(R',Q') \leq H(p_1, \ldots, p_{d^2})$, and with some algebra

$$H(p_1, \ldots, p_{d^2}) = H(p_1) + (1 - p_1) H\left(\frac{p_2}{1 - p_1}, \ldots, \frac{p_{d^2}}{1 - p_1}\right) \leq \log(d^2 - 1)$$

and since $p_1 = F(\rho, E)$.

Holevo bound

Another result giving an upper bound of accessible quantum information is the Holevo bound as:

$$H(X:Y) \leq S(\rho) - \sum_x p_x S(\rho_x), \rho = \sum_x p_x \rho_x.$$

The right side of this inequality is useful in quantum information theory, and hence it is given a special title: *Holevo χ-quantity*. While the von Neumann entropy is intuitively associated with information content of the pure state ensemble, for ensembles of mixed states it is not a good candidate. A natural generalization is here the Holevo χ-quantity (called also H-information) given by:

$$\chi = I_H = S\left(\sum_x p_x \rho_x\right) - \sum_x p_x S(\rho_x).$$

The second term in this formula reflects the fact that the information content of the ensemble is lower if the components are impure. When ρ_x are pure then I_H reduces to the von Neumann entropy. The H-information has physical meaning in terms of the so-called classical capacity of quantum channel. Namely, it was shown that if we wish to communicate classical bits via quantum states, then I_H maximized over the *a priori* probabilities associated with the states gives the classical capacity of such channel. In contrast with the von Neumann entropy, the H-information does not have interpretation in terms of quantum bits so far.

It is a lower bound for the optimal compression rate of ensemble – the quantity that may be regarded as the ensemble information content.

Holevo-Schumacher-Westmoreland (HSW) theorem

HSW-theorem tell us the asymptotic rate at which classical information can be transmitted over a quantum channel E per channel use is given by the maximum output Holevo quantity χ across all possible signaling ensembles: $\quad C^{(1)}(E) = \max_{\{p_j, \rho_j\}} \left[S(\rho') - \sum_x p_x S(\rho'_x) \right], \rho'_x = E(\rho_x), \rho' = \sum_x p_x \rho'_x ,\quad$ or $C = \max_{\{p_i, \rho_i\}} \chi\left(\{p_i, \tilde{\rho}_i = E(\rho_i)\}\right)$. Using the unique nature of the average output state of an optimal signaling ensemble for a special class of qudit unital channels the HSW channel capacity is

$$C = \log_2(d) - \min_\rho S(E(\rho)),$$

where d is the dimension of the qudit. Thus, the connection between the minimum von Neumann entropy at the channel output and the transmission rate for classical information over quantum channels extends beyond the qubit domain.

An alternative, but equivalent, description of HSW channel capacity can be made using relative entropy:

$$C_1 = \max_{\{all \ possible \ \{p_k, \varphi_k\}\}} \sum p_k S\big(E(\varphi_k)\|E(\varphi)\big),$$

where the φ_k are the quantum input to the channel and $\varphi = \sum_k p_k \varphi_k$.

Example: *Relative entropy in the Bloch sphere representation*. In the Bloch sphere representation the key formula is relative entropy. The respective Bloch sphere representation for two density matrix ρ and ψ have $\rho = \frac{1}{2}\big(I + \vec{W}\cdot\vec{\sigma}\big)$, $\psi = \frac{1}{2}\big(I + \vec{V}\cdot\vec{\sigma}\big)$. We can define $\cos(\theta)$ as: $\cos(\theta) = \frac{\vec{W}\cdot\vec{V}}{r\,q}$ where $r = \sqrt{\vec{W}\cdot\vec{W}}$ and $q = \sqrt{\vec{V}\cdot\vec{V}}$, and θ is the angle between $\vec{W}$ and $\vec{V}$. The following formula for the relative entropy

$$S(\rho\|\psi) = \frac{1}{2}\log_2\big(1-r^2\big) + \frac{r}{2}\log_2\left(\frac{1+r}{1-r}\right) - \frac{1}{2}\log_2\big(1-q^2\big) - \frac{\vec{W}\cdot\vec{V}}{2q}\log_2\left(\frac{1+q}{1-q}\right)$$

$$= \frac{1}{2}\log_2\big(1-r^2\big) + \frac{r}{2}\log_2\left(\frac{1+r}{1-r}\right) - \frac{1}{2}\log_2\big(1-q^2\big) - \frac{r\cos(\theta)}{2}\log_2\left(\frac{1+q}{1-q}\right)$$

Ordinarily, $S(\rho\|\psi) \neq S(\psi\|\rho)$. However, when $r = q$, we can see from above formula that $S(\rho\|\psi) = S(\psi\|\rho)$. A few special case of $S(\rho\|\psi)$ are worth examining. Consider the case when ψ in $S(\rho\|\psi)$ is the maximally mixed state: $\psi = \frac{1}{2}I$. In this case, $q = 0$, and $S(\rho\|\psi)$ becomes the radially symmetric function

$$S(\rho\|\psi) = S\left(\rho\left\|\frac{1}{2}I\right.\right) = \frac{1}{2}\log_2\big(1-r^2\big) + \frac{r}{2}\log_2\left(\frac{1+r}{1-r}\right)$$

$$= \frac{1+r}{2}\log_2\left(\frac{1+r}{2}\right) + \frac{1+r}{2} + \frac{1-r}{2}\log_2\left(\frac{1-r}{2}\right) + \frac{1-r}{2} = 1 - S(\rho)$$

Thus, $S\left(\rho\left\|\frac{1}{2}I\right.\right) = 1 - S(\rho)$, where $S(\rho)$ is the von Neumann entropy of ρ, the first density matrix in the relative entropy function.

E. *Fannes type inequalities*. We will present basic inequalities relating two "information-like" quantities (von Neumann entropy and Holevo H-information) with the "fidelity-like" quantities. All of them are variations of the Fannes inequality that has been recognized as an important tool in quantum information theory and information analysis of QA's evolution.

Fannes inequality: Let $\|\rho - \rho'\| \leq \frac{1}{3}$. Then the following inequality holds

$$\big|S(\rho) - S(\rho')\big| \leq \log\dim H\|\rho - \rho'\| - H\big(\|\rho - \rho'\|\big),$$

where $H(x) = -x\log x$.

Remark. The physical interpretation of the inequality is the following. Assume that ρ and ρ' are the states of n particles so that they act on the Hilbert space $H = H_1^{\otimes n}$, where H_1 is the single particle space. Then the inequality says that if the states are closely to each other, then their entropies per particle are also close to each other.

Fannes inequality with fidelity is as follows: $\left|S(\rho)-S(\rho')\right| \le 2\log\dim H\sqrt{1-F(\rho,\rho')}+1,$

and *Fannes inequality for ensembles* one can derive the following version:

$$\left|S(\mathrm{E})-S(\mathrm{E}')\right| \le 4\log\dim H\sqrt{1-F(\rho,\rho')}+1.$$

Other useful inequality, obtained by applying concavity of square root to the above inequality with fidelity, is the following one: $\sum_i p_i\left|S(\rho_i)-S(\rho_i')\right| \le 2\log\dim H\sqrt{1-F(\mathrm{E},\mathrm{E}')}+1.$

Finally, from these inequalities we obtain Fannes inequality for Holevo H-information:

$$\left|I_H(\mathrm{E})-I_H(\mathrm{E}')\right| \le 6\log\dim H\sqrt{1-F(\mathrm{E},\mathrm{E}')}+2.$$

Recent development in quantum information theory has motivated extensive study of entanglement. Furthermore, an exciting subject of characterizing other types of correlations has emerged. For example, quantum correlation, classical one, or quantum and classical correlation have been studied.

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. V. 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. V. A70, № 5.

9. Hamieh S., Qi J., Siminovitch D. et all. Extraction of classical corelations from a bipartite quantum system. // Phys. Review. 2003. V. A67, № 1.

10. Groisman B., Popescu S., Winter A. Quantum, classical, and total amount of correlations in quantum state. // Phys. Review. 2005. V. A72, № 3.

11. DiVincenzo D.P., Horodecki M., Leung D.W. et all. Locking classical correlation in quantum states. // Physical Review Letters. 2004. V. 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. V. A71, № 6.

13. Maruyama K., Brukner C., Vedral V. Thermodynamical cost of accessing quantum information. // J. Phys. A: Math. Gen. 2005. V. 38, № 32.

14. Christandl M., Winter A. Uncertainty, monogamy, and locking of quantum correlations. // IEEE Transactions on Information Theory. 2005. V. 51, № 9.

15. Levitin L.B. Conditional entropy and information in quantum systems. // Chaos, Solitons and Fractals. 1999. V. 10, № 10.

16. 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