INEQUALITIES FOR NONNEGATIVE NUMBERS
AND INFORMATION PROPERTIES
OF QUDIT TOMOGRAMS

Margarita A. Man’ko* and Vladimir I. Man’ko

P. N. Lebedev Physical Institute, Russian Academy of Sciences
Leninskii Prospect 53, Moscow 119991, Russia

*Corresponding author e-mail: mmanko@sci.lebedev.ru

Abstract

We discuss some inequalities for \( N \) nonnegative numbers. We use these inequalities to obtain known inequalities for probability distributions and new entropic and information inequalities for quantum tomograms of qudit states. The inequalities characterize the degree of quantum correlations in addition to noncontextuality and quantum discord. We use the subadditivity and strong subadditivity conditions for qudit tomographic-probability distributions depending on the unitary-group parameters in order to derive new inequalities for Shannon, Rényi, and Tsallis entropies of spin states.

Keywords: uncertainty relations, entropy and information, qudits, spin tomography, nonnegative numbers, Rényi entropic inequalities.

1 Introduction

There exist quantum phenomena related to the presence of quantum correlations. The quantum correlations are responsible for entanglement [1], violation of the Bell inequalities [2-4], noncontextuality (see, for example, [5,6]), and discord [7]. In some cases, the existence of quantum correlations can be expressed in terms of the tomographic-probability distributions (spin tomograms) [8,11] and their specific properties. These properties were discussed for Shannon entropy [12] and \( q \)-entropy [13,14] and information associated with the tomographic-probability distributions in [15,29].

The idea of our approach is to consider three different but closely connected objects. The probability distributions are determined by a set of nonnegative numbers. In view of this fact, our first object is the set of nonnegative numbers not related to any applications. It is a purely mathematical object with specific properties. These properties can be studied considering some functions on the set of nonnegative numbers. The functions can satisfy some inequalities that are generic inequalities characterizing both the set of nonnegative numbers and the functions. The second object is the standard probability distributions that are identified with the set of nonnegative numbers with an additional interpretation that these numbers are the probability of some observable measurements, and the observables themselves are associated with other numbers, which code the results of the measurements. The third object to be considered is the nonnegative functions defined, for example, on unitary matrices. For each unitary matrix (or a point on the sphere), one has the discussed set of nonnegative numbers. Also the nonnegative functions can be associated with the probability distributions considered as the probability distributions depending
on extra parameters like the unitary matrices or point on the sphere. This means that the parameter-dependent probability distributions are the nonnegative functions, and for each parameter there exists another set of numbers that code the outcome of experiments where some observables are measured.

We try to consider entropic and information inequalities analyzing what system properties are connected with only mathematical properties of the sets of nonnegative numbers and what properties are associated with an extra information contained in the probability distributions of measurable variables and the results of experiments for the cases where the dependence on some parameters like unitary matrices or coordinates of a point on the sphere play a role.

The aim of this paper is to connect the entropic and information inequalities (uncertainty relations) with some general properties of a set of \( N \) positive numbers and properties of unitary matrices.

This paper is organized as follows.

In Sec. 2, we discuss the properties of nonnegative numbers and some inequalities for these numbers and consider the interpretation of the nonnegative numbers in terms of the probability distributions. In Sec. 3, we apply the obtained results to the tomographic-probability distributions of quantum systems.

In Sec. 4, we study a qudit system and consider the Shannon and \( q \)-entropies in Sec. 5. In Sec. 6, we review known entropic inequalities and obtain new information inequalities in Sec. 7. In Sec. 8, we study the probability properties, in view of the vector and matrix properties. In Sec. 9, we discuss the influence of permutations of nonnegative numbers on the properties of entropies and consider the relation between the strong subadditivity condition and matrices in Sec. 10. Our conclusions are presented in Sec 11.

## 2 Nonnegative Numbers

We consider a set of \( N \) nonnegative numbers \( P_1, P_2, \ldots, P_N \). Let these numbers satisfy the additional normalization condition \( \sum_{k=1}^{N} P_k = 1 \). There are different functions \( f(P_1, P_2, \ldots, P_N) \), which have the index permutation symmetry, i.e., \( f(P_1, P_2, \ldots, P_N) = f(\hat{P}_1, \hat{P}_2, \ldots, \hat{P}_N) \), where \( \hat{P}_j \) means the result of a permutation-operator action on the \( j \)th nonnegative number. The Shannon [12] and Rényi [13] entropies have such a symmetry in the case where the nonnegative numbers \( P_j \) are associated with the probability distributions describing the results of measurements in different nondeterministic processes. Independently of the interpretation in terms of the probability distributions, it is worth pointing out that the set of positive numbers can be characterized by some inequalities for functions that can be considered in the applications as entropies, information, etc.

The permutation of the nonnegative numbers can be visualized if the numbers are organized in a vector \( \vec{P} \). Then permutation is described by a stochastic matrix acting on this vector. Thus, there exist \( N! \) different permutation \( N \times N \)-matrices. The permutation matrices are unitary real matrices with matrix elements equal either to zero or to unity. We discuss also the action of the permutations onto an nonnegative Hermitian \( N \times N \)-matrix considered as a complex \( N^2 \)-vector with components constructed of rows of the matrix [30]. Then the permutation \( N^2 \times N^2 \)-matrix acting on the vector is the direct product of two unitary permutation \( N \times N \)-matrices. The matrix realizes a specific positive map of the Hermitian matrix. The map does not change the matrix eigenvalues but yields the permutation of the matrix eigenvectors.
Joint Probability Properties

In this section, we discuss the probability distributions on the example of a bipartite system consisting of two subsystems.

It is known that, if one has the joint probability distribution \( w(m_1, m_2) \geq 0 \) of two discrete random variables \( m_1 \) and \( m_2 \) related to a system containing two subsystems 1 and 2, there exist two marginals

\[
P_1(m_1) = \sum_{m_2} w(m_1, m_2), \quad P_2(m_2) = \sum_{m_1} w(m_1, m_2),
\]

which are associated with Shannon entropies

\[
H(k) = -\sum_{m_k} P_k(m_k) \ln P_k(m_k) \geq 0, \quad k = 1, 2.
\]

The entropy of the system \( H(1, 2) \) reads

\[
H(1, 2) = -\sum_{m_1, m_2} w(m_1, m_2) \ln w(m_1, m_2) \geq 0.
\]

There exists the inequality called the subadditivity condition

\[
H(1) + H(2) \geq H(1, 2),
\]

and the Shannon mutual information is defined as the difference

\[
I = H(1) + H(2) - H(1, 2) \geq 0.
\]

There exist other probability distributions determined by the initial distribution \( w(m_1, m_2) \). For example, two conditional probability distributions \( P_1(m_1 \mid m_2) \) and \( P_2(m_2 \mid m_1) \) are defined as

\[
P_1(m_1 \mid m_2) = \frac{P_1(m_1, m_2)}{P_2(m_2)}, \quad P_2(m_2 \mid m_1) = \frac{P_1(m_1, m_2)}{P_1(m_1)}.
\]

The meaning of the conditional probability distribution follows from the obvious statement, which is the essence of the Bayesian formula, namely, the joint probability \( P(m_1, m_2) \) to obtain the values of two random variables \( m_1 \) and \( m_2 \) (measurable simultaneously) is equal to the product of the probability \( P_2(m_2) \) to obtain the variable \( m_2 \) and the probability \( P_1(m_1 \mid m_2) \) to obtain the value \( m_1 \) of the first random variable under the condition that the value of variable \( m_2 \) is known.

The conditional probability distributions determine the Shannon entropies

\[
H_1(1 \mid m_2) = -\sum_{m_1} P_1(m_1 \mid m_2) \ln P_1(m_1 \mid m_2),
\]

\[
H_2(2 \mid m_1) = -\sum_{m_2} P_2(m_2 \mid m_1) \ln P_2(m_2 \mid m_1).
\]

One can calculate average entropies

\[
H(1 \mid 2) = \overline{H}_1 = \sum_{m_2} P_2(m_2) H_1(1 \mid m_2),
\]

\[
H(2 \mid 1) = \overline{H}_2 = \sum_{m_1} P_1(m_1) H_2(2 \mid m_1).
\]
One can check that the following equalities are valid:

\[ H_1 = H(1, 2) - H(2) = H(1 \mid 2), \quad H_2 = H(1, 2) - H(1) = H(2 \mid 1). \]  

(9)

Also the mutual information can be expressed as the difference

\[ I = -H_1 + H(1) = -H_2 + H(2). \]  

(10)

The nonnegativity of the mutual information means that

\[ H(1) \geq H_1, \quad H(2) \geq H_2. \]  

(11)

3.1 Example of the Classical Coins

We illustrate the discussed notions on the example of two classical coins.

Let the first and second coins have the outcomes of the experiment (up and down) labeled by \( \pm 1 \). This means that one has the probability distribution \( w(m_1, m_2) \) determined by four nonnegative numbers

\[ w(+1, +1) = a, \quad w(+1, -1) = b, \quad w(-1, +1) = c, \quad w(-1, -1) = d. \]

The normalization of the probability distribution \( w(m_1, m_2) \) means that \( a + b + c + d = 1 \). One can consider the probability distribution \( w(m_1, m_2) \) as a probability column vector \( \vec{w} \) with four components. The marginals \( P_1(m_1) \) and \( P_2(m_2) \) are the probability distributions

\[ P_1(+1) = a + b, \quad P_1(-1) = c + d, \quad P_2(+1) = a + c, \quad P_2(-1) = b + d. \]

These marginals can be considered either as the probability column vectors \( \vec{P}_1 \) and \( \vec{P}_2 \),

\[ \vec{P}_1 = \begin{pmatrix} a + b \\ c + d \end{pmatrix}, \quad \vec{P}_2 = \begin{pmatrix} a + c \\ b + d \end{pmatrix}, \]

or as the probability column 4-vectors obtained as the qubit portrait \([31][33]\) of the initial vector \( \vec{w} \), using stochastic matrices

\[ M_1 = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad M_2 = \begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}. \]

Thus, one has \( \vec{P}_1 = M_1 \vec{w} \) and \( \vec{P}_2 = M_2 \vec{w} \), where the column vectors \( \vec{P}_1 \) and \( \vec{P}_2 \) read

\[ \vec{P}_1 = \begin{pmatrix} \vec{p}_1 \\ \vec{0} \end{pmatrix}, \quad \vec{P}_2 = \begin{pmatrix} \vec{p}_2 \\ \vec{0} \end{pmatrix}, \quad \vec{0} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}. \]
3.2 The Probability Vector as a Rectangular Matrix

Now we formulate the general matrix rule for constructing the marginals for the probability \( N \)-vector \( \vec{P} \) given as a column with \( N = mn \) components, i.e.,
\[
\vec{P} = (P_{11}, P_{12}, \ldots P_{1n}, P_{21}, P_{22}, \ldots P_{2n}, \ldots, P_{m1}, P_{m2}, \ldots, P_{mn}).
\]

First we represent \( \vec{P} \) in the form of a rectangular matrix
\[
P_{kj} = \begin{pmatrix}
P_{11} & P_{12} & \cdots & P_{1n} \\
P_{21} & P_{22} & \cdots & P_{2n} \\
\vdots & \vdots & \ddots & \vdots \\
P_{m1} & P_{m2} & \cdots & P_{mn}
\end{pmatrix}.
\]

(12)

Then the two marginals are represented by the probability \( m \)-vector \( \vec{P}_1 \) and \( n \)-vector \( \vec{P}_2 \). The components of the \( m \)-vector \( \vec{P}_1 \) are obtained by the sum of the matrix elements in the \( k \)th rows,
\[
(\vec{P}_1)_k = \sum_{k=1}^{n} P_{ks}, \quad k = 1, 2, \ldots, m,
\]

(13)

and the components of the \( n \)-vector \( \vec{P}_2 \) are obtained by the sum of the matrix elements in the \( j \)th columns,
\[
(\vec{P}_2)_j = \sum_{l=1}^{m} P_{lj}, \quad j = 1, 2, \ldots, n.
\]

(14)

The subadditivity condition (4) can be formulated as an inequality for the matrix \( P_{kj} \) (12) following the statement: Given a rectangular matrix with nonnegative matrix elements \( P_{kj} \) such that \( \sum_{k} \sum_{j} P_{kj} = 1 \), one has
\[
\sum_{j=1}^{n} \left( \sum_{l=1}^{m} P_{lj} \right) \ln \left( \sum_{l'=1}^{m} P_{lj} \right) + \sum_{k=1}^{n} \left( \sum_{s=1}^{m} P_{ks} \right) \ln \left( \sum_{s'=1}^{m} P_{ks'} \right) \leq \sum_{j=1}^{n} \sum_{k=1}^{m} P_{kj} \ln P_{kj}.
\]

(15)

Entropy (3) reads
\[
H(1, 2) = -a \ln a - b \ln b - c \ln c - d \ln d \equiv -\vec{w} \ln \vec{w},
\]

and the entropy associated with the marginals are
\[
H(1) = -(a + b) \ln(a + b) - (c + d) \ln(c + d) = -\vec{P}_1 \ln \vec{P}_1 = -(M_1 \vec{w}) \ln(M_1 \vec{w}),
\]
\[
H(2) = -(a + c) \ln(a + c) - (b + d) \ln(b + d) = -\vec{P}_2 \ln \vec{P}_2 = -(M_2 \vec{w}) \ln(M_2 \vec{w}).
\]

(16)

We use here the notation \( \ln \vec{A} = \vec{A} \ln \), which means the \( N \)-vector with components \( (\ln \vec{A})_k \equiv \ln A_k \). Also we use for any function \( f(x) \) the notation for a vector \( \vec{A}_f \equiv f(\vec{A}) \) with components \( (\vec{A}_f)_k \equiv f(A_k) \), \( k = 1, 2, \ldots, N \). Precisely in our case, for the Shannon entropy the function \( f(x) = \ln x \). The scalar product of real vectors \( \vec{A}_f \vec{B}_\varphi \) is defined as
\[
f(\vec{A}) \varphi(\vec{B}) \equiv \vec{A}_f \vec{B}_\varphi = \sum_{k=1}^{n} f(A_k) \varphi(B_k).
\]
3.3 Conditional Probability Distribution for Two Coins

The mutual information is given by the expression

\[
I = \vec{w} \ln \vec{w} - (M_1 \vec{w}) \ln(M_1 \vec{w}) - (M_2 \vec{w}) \ln(M_2 \vec{w}) \\
= a \ln a + b \ln b + c \ln c + d \ln d - (a + b) \ln(a + b) - (c + d) \ln(c + d) \\
- (a + c) \ln(a + c) - (b + d) \ln(b + d).
\]  

(17)

We make a general statement that follows from inequality (15).

Given a probability vector \( \vec{P} \), i.e., \( N \) nonnegative numbers \( P_\alpha \), \( \sum_{\alpha=1}^{N} P_\alpha = 1 \), we distribute these numbers in any way in a rectangular matrix \( P_{kj} \) with the number of matrix elements larger than \( N \) and put the number zero for matrix elements in empty positions. Then we have inequality (15).

The conditional probability distributions \( P_1(m_1 | m_2) \) read

\[
P_1(+1 | +1) = \frac{a}{a + c}, \quad P_1(-1 | +1) = \frac{c}{a + c}, \\
P_1(+1 | -1) = \frac{b}{b + d}, \quad P_1(-1 | -1) = \frac{d}{b + d},
\]

(18)

and the two entropies are

\[
H(+1 | +1) = \left( -a \frac{\ln a}{a + c} - c \frac{\ln c}{a + c} \right), \\
H(+1 | -1) = \left( -b \frac{\ln b}{b + d} - d \frac{\ln d}{b + d} \right).
\]

(19)

We define the entropy \( \overline{H}(1) = H(1 | -1) \) as

\[
\overline{H}(1) = H(1 | +1)P_2(+1) + H(1 | -1)P_2(-1) \\
= \left( -a \frac{\ln a}{a + c} - c \frac{\ln c}{a + c} \right)(a + c) \\
+ \left( -b \frac{\ln b}{b + d} - d \frac{\ln d}{b + d} \right)(b + d) \\
= -a \ln a - c \ln c - b \ln b - d \ln d \\
= H(1, 2) - H(2).
\]

(20)

Analogously,

\[
\overline{H}(2) = H(2 | 1) = H(1, 2) - H(1),
\]

(21)

and we obtain the following rule.

We consider a probability vector \( \vec{P} \) with \( N \) components \( P_1, P_2, \ldots P_N \) and construct the portrait of this probability vector, which is a new probability vector \( \vec{\Pi} \) given by the action of the fiducial stochastic matrix \( M_f \), such that

\[
\Pi_1 = (P_1 + P_2 + \cdots + P_j), \Pi_2 = (P_{j+1} + P_{j+2} + \cdots + P_{j_2}), \ldots, \Pi_s = (P_{j_{s-1}+1} + P_{j_{s-1}+2} + \cdots + P_N),
\]

(265)
and all the other vector components are zeros. Such a map provides the Shannon entropy of the portrait probability distributions
\[ H_{\Pi} = - \sum_{k=1}^{s_{N}} \Pi_k \ln \Pi_k. \]

Also there are \( s_{N} \) conditional probability distributions created by the map \( M \), namely,
\[ P(1 | 1) = P_1/\Pi_1, \quad P(2 | 1) = P_2/\Pi_1, \quad \ldots, \quad P(j_1 | 1) = P_{j_1}/\Pi_1, \]
\[ P(1 | 2) = P_{j_1+1}/\Pi_2, \quad P(2 | 2) = P_{j_1+2}/\Pi_2, \quad \ldots, \quad P(j_2 | 2) = P_{j_2}/\Pi_2, \]
\[ \ldots \]
\[ P(1 | s) = P_{j_s+1}/\Pi_s, \quad P(2 | s) = P_{j_s+2}/\Pi_s, \quad \ldots, \quad P(j_s | s) = P_N/\Pi_s, \]
where \( j_1 + j_2 + \cdots + j_s = N \).

The entropies defined as analogs of the entropies associated with conditional probability distributions read
\[ H(k) = - \sum_{l} P(l | k) \ln P(l | k). \]

Then the average entropy \( \overline{H} \) is expressed as
\[ \overline{H} = H - H_{\Pi} = - \overline{P} \ln \overline{P} + \overline{\Pi} \ln \overline{\Pi}. \]

The conditional probability distribution means the probability distribution to have the outcome of the event if it is known that the event belongs to the \( k \)th group given by the \( k \)th row of the portrait matrix \( M_f \) of the stochastic matrix.

### 4 Spin Tomograms (Qubit and Qudit Tomograms)

Given an \( N \)-dimensional space of states of spin system. One can interpret this space either as the state space for one particle with spin \( j = (N - 1)/2 \) (qudit) or, in the case of the product representation of the number \( N = n_1 n_2 \cdots n_M \), as the space of multipartite spin system (multipartite qudit system) with \( j_1 = (n_1 - 1)/2, j_2 = (n_2 - 1)/2, \ldots, j_M = (n_M - 1)/2 \).

The \( N \times N \) density matrix \( \rho \) of the quantum state can be represented by the unitary tomogram of the spin state \[ 10 \]. In the case of the spin state with \( j = (N - 1)/2 \), the tomogram is defined by the relation
\[ w(m, u) = \langle m | u^\dagger \rho u | m \rangle, \] (23)
where \( \rho \) is the density matrix, \( u \) is the \( N \times N \) unitary matrix, and semi-integers \( m = -j, -j + 1, \ldots, j \) are values of the spin projection on the \( z \) axis. Tomogram (23) is the nonnegative probability-distribution function of the random spin-projection variable satisfying the normalization condition \( \sum_{m=-j}^{j} w(m, u) = 1 \) and the equality \( \int w(m, u) du = 1 \), where \( du \) is the Haar measure on the unitary group with the normalization \( \int du = 1 \). An important property of tomogram \( w(m, u) \) is that its connection with the density matrix \( \rho \) reads \( \rho \leftrightarrow w(m, u) \). This means that the quantum state is given if the tomogram is known [8,9].
Following standard definitions of the probability theory, one can introduce Shannon \cite{12} tomographic entropy \cite{34} and Rényi \cite{13} tomographic entropy \cite{35}.

The Shannon tomographic entropy is the function on the unitary group
\[
H_u = - \sum_{m=-j}^{j} w(m,u) \ln w(m,u). \tag{24}
\]

The Rényi tomographic entropy is also the function on the unitary group and it depends on an extra parameter \(q\)
\[
R_{u}^{(q)} = \frac{1}{1-q} \ln \left(\sum_{m=-j}^{j} (w(m,u))^q\right). \tag{25}
\]

The Tsallis tomographic entropy is determined as
\[
T_{u}(q) = \frac{1}{1-q} \left(\sum_{m=-j}^{j} (w(m,u))^q - 1\right). \tag{26}
\]

For two spin tomograms \(w_1(m,u)\) and \(w_2(m,u)\), we define the relative tomographic \(q\)-entropy
\[
H_{q}(w_1(u)|w_2(u)) = - \sum_{m=-j}^{j} w_1(m,u) \ln_q \frac{w_2(m,u)}{w_1(m,u)}, \tag{27}
\]
with
\[
\ln_q x = \frac{x^{1-q} - 1}{1-q}, \quad x > 0, \quad q > 0, \quad \ln_{q \to 1} x = \ln x.
\]

The relative tomographic \(q\)-entropy is a nonnegative function for any admissible deformation parameter \(q\). For \(q \to 1\), \(R_{u} \to H_{u}\) and the relative tomographic \(q\)-entropy becomes the relative entropy associated to the two tomographic-probability distributions
\[
H(w_1(u)|w_2(u)) = - \sum_{m=-j}^{j} w_1(m,u) \ln \frac{w_2(m,u)}{w_1(m,u)}. \tag{28}
\]

As was shown in \cite{35}, the minimum over the unitary group of the Rényi tomographic entropy \(25\) is equal to the quantum Rényi tomographic entropy
\[
\min R_{u} = \frac{1}{1-q} \ln \text{Tr} \rho^q. \tag{29}
\]

The minimum over the unitary group of the Shannon tomographic entropy \(24\) is equal to the von Neumann entropy \cite{34,35}, i.e.,
\[
\min H_{u} = - \text{Tr} \rho \ln \rho. \tag{30}
\]

One has for \(\min R_{u}\) the corresponding quantum Tsallis entropy
\[
\frac{1}{1-q} (\text{Tr} \rho^q - 1) = \frac{1}{1-q} \left\{\exp[\min R_{u}(1-q)] - 1\right\}. \tag{31}
\]
6 Known Inequalities for Bipartite and Tripartite Systems

The tomographic entropies satisfy some known inequalities found in [35].

For example, if the spin system is bipartite, i.e., one has spins $j_1$ and $j_2$, the basis in the tensor-product space reads $|m_1 m_2⟩ = |m_1⟩ |m_2⟩$. In this case, the tomogram is the joint-probability distribution of two random spin projections $m_1 = -j_1, -j_1 + 1, \ldots, j_1$ and $m_2 = -j_2, -j_2 + 1, \ldots, j_2$ depending on the $(2j_1 + 1)(2j_2 + 1)$ states. The tomogram reads

$$w(m_1, m_2, u) = ⟨m_1 m_2 | u^† ρ(1, 2) u | m_1 m_2⟩,$$

(32)

where $ρ(1, 2)$ is the density matrix of the bipartite-system state with matrix elements

$$ρ(1, 2)_{m_1 m_2, m'_1 m'_2} = ⟨m_1 m_2 | ρ(1, 2) | m'_1 m'_2⟩.$$

(33)

For this tomogram, one can introduce the Shannon entropy $H_{12}(u)$ as

$$H_{12}(u) = - \sum_{m_1 = -j_1}^{j_1} \sum_{m_2 = -j_2}^{j_2} w(m_1, m_2, u) \ln w(m_1, m_2, u).$$

(34)

The Shannon entropy $H_{12}(u)$ satisfies the subadditivity condition for all elements of the unitary group

$$H_{12}(u) ≤ H_1(u) + H_2(u),$$

(35)

where $H_1(u)$ and $H_2(u)$ are Shannon entropies associated with subsystem tomograms

$$w_1(m_1, u) = \sum_{m_2 = -j_2}^{j_2} w(m_1, m_2, u), \quad w_2(m_2, u) = \sum_{m_1 = -j_1}^{j_1} w(m_1, m_2, u)$$

(36)

as follows:

$$H_k(u) = - \sum_{m_k = -j_k}^{j_k} w_k(m_k, u) \ln w_k(m_k, u), \quad k = 1, 2.$$

(37)

From this inequality, in view of the relation between the von Neumann and tomographic entropies, follows the known inequality [35], namely, the subadditivity condition for corresponding von Neumann entropy for the bipartite system

$$S_{12} ≤ S_1 + S_2,$$

(38)

where

$$S_k = -\text{Tr} \rho_k \ln ρ_k, \quad k = 1, 2 \quad ρ_1 = -\text{Tr} \rho(1, 2), \quad ρ_2 = -\text{Tr} \rho(1, 2).$$

(39)

For tripartite spin system with spins $j_1$, $j_2$, and $j_3$ and the density matrix $ρ(1, 2, 3)$, the spin tomogram reads

$$w(m_1, m_2, m_3, u) = ⟨m_1 m_2 m_3 | u^† ρ(1, 2, 3) u | m_1 m_2 m_3⟩.$$

(40)

One associates the Shannon entropy $H_{123}(u)$ with this tomogram. This entropy satisfies the inequality, which is the strong subadditivity condition on the unitary group. It reads [35]

$$H_{123}(u) + H_2(u) ≤ H_{12}(u) + H_{23}(u),$$

(41)

268
where

$$H_{123}(u) = - \sum_{m_1=-j_1}^{j_1} \sum_{m_2=-j_2}^{j_2} \sum_{m_3=-j_3}^{j_3} w(m_1, m_2, m_3, u) \ln w(m_1, m_2, m_3, u)$$  \hspace{1cm} (42)

and entropies $H_{12}(u)$, $H_{23}(u)$, and $H_2(u)$ are defined by means of projected tomograms

$$w_{12}(m_1, m_2, u) = \sum_{m_3=-j_3}^{j_3} w(m_1, m_2, m_3, u), \quad w_{23}(m_2, m_3, u) = \sum_{m_1=-j_1}^{j_1} w(m_1, m_2, m_3, u),$$

$$w_2(m_2, u) = \sum_{m_1=-j_1}^{j_1} w_{12}(m_1, m_2, u).$$

Our new inequality (41) is compatible with the known strong subadditivity condition for the von Neumann entropy presented in [36,37]

$$S_{123} + S_2 \leq S_{12} + S_{23},$$  \hspace{1cm} (43)

where $S_{123} = -\text{Tr} \rho_{123} \ln \rho_{123}$, and other entropies are von Neumann entropies for reduced density matrices $\rho(1,2) = \text{Tr}_3 \rho(1,2,3)$ and $\rho(2,3) = \text{Tr}_1 \rho(1,2,3)$.

Inequalities (35) and (41) are new inequalities for composite quantum finite-dimensional systems obtained in [35].

### 7 Quantum Correlations and New Local-Transform Dependent Information Inequalities

In view of (35), the Shannon tomographic information is defined as

$$I(u) = H_1(u) + H_2(u) - H_{12}(u),$$

and in view of (38), the quantum information is defined as

$$I_q = S_1 + S_2 - S_{12}.$$  \hspace{1cm} (45)

If we consider equality (44) for the unitary matrix $u = u_{10} \otimes u_{20}$, corresponding to local unitary transforms $u_{10}$ and $u_{20}$ for which $H_1(u_{10}) = S_1$ and $H_2(u_{20}) = S_2$, i.e., the unitary matrices $u_{10}$ and $u_{20}$ are acting in the first and second qudit Hilbert spaces and are providing the minima of entropies $H_1(u_{10}) = S_1$ and $H_2(u_{20}) = S_2$, we obtain the following equality:

$$I(u_{10} \otimes u_{20}) = S_1 + S_2 - H_{12}(u_{10} \otimes u_{20}).$$  \hspace{1cm} (46)

Since $S_{12}$ is the minimum of $H_{12}(u)$, we have the inequality $S_{12} \leq H_{12}(u_{10} \otimes u_{20})$, which provides a new inequality for entropies

$$S_1 + S_2 \geq H_{12}(u_{10} \otimes u_{20}) \geq S_{12}.$$  \hspace{1cm} (47)

Also we obtain a new inequality for informations $I_q \geq I(u_{10} \otimes u_{20})$.

For the two-qudit product state with the density matrix $\rho(1,2) = \rho_1(1) \otimes \rho_2(2)$, we have the equality $I_q = I(u_{10} \otimes u_{20})$. Thus, the difference in information

$$D = (S_1 + S_2 - S_{12}) - I(u_{10} \otimes u_{20}) \geq 0$$  \hspace{1cm} (48)
is a characteristic of correlations of the qudit subsystems of the bipartite two-qudit systems. It is an additional characteristic of correlations in the qudit system, which, in its spirit, is analogous to discord.

Recently [35, 40], we pointed out that tomograms \( w(m, u) \) and \( w(m_1, m_2, u) \) can be interpreted as conditional probability distributions, i.e.,

\[
w(m, u) \equiv w(m \mid u), \quad w(m_1, m_2, u) \equiv w(m_1, m_2 \mid u).
\]

Also for \( u = u_1 \otimes u_2 \), the tomogram \( w(m_1, m_2, u_1, u_2) \equiv w(m_1, m_2 \mid u_1, u_2) \).

Thus, all the inequalities discussed can be considered as inequalities for the entropies and information corresponding to the tomographic conditional probability distributions.

The properties of Tsallis entropies associated with a joint probability distribution were discussed in [41]. We apply these results to the tomogram \( w(m_1, m_2, u) \) of two qudit systems \( A \) and \( B \). The tomographic \( q \)-entropy reads

\[
T_q(A, B, u) = \frac{1}{1-q} \left( \sum_{m_1=-j_1}^{j_1} \sum_{m_2=-j_2}^{j_2} w(m_1, m_2, u)^q - 1 \right).
\]

In the limit \( q \to 1 \), this entropy becomes the Shannon tomographic entropy.

We have the equalities

\[
T_q(A, B, u) = T_q(A \mid B, u) + T_q(B, u),
\]

where \( T_q(A \mid B, u) \) is the conditional tomographic \( q \)-entropy defined as

\[
T_q(A \mid B, u) = \sum_{m_2=-j_2}^{j_2} w(m_2, u)^q \ T_q(A \mid m_2, u),
\]

\[
T_q(A \mid m_2, u) = \sum_{m_1=-j_1}^{j_1} w(m_1 \mid m_2, u) \ln_q \frac{1}{w(m_1 \mid m_2, u)}.
\]

In the last formula [52], the conditional tomographic-probability distribution \( w(m_1 \mid m_2, u) \) is defined by the Bayesian formula

\[
w(m_1 \mid m_2, u) = \frac{w(m_1, m_2, u)}{\sum_{m_1=-j_1}^{j_1} w(m_1, m_2, u)}.
\]

The function \( \ln_q(x) \) reads \( \ln_q(x) = (x^{1-q} - 1)(1-q)^{-1} \), and in the limit \( q \to 1 \), \( \ln_q(x) = \ln x \). Also the above relations [49]–[52] become in this limit the relations for Shannon tomographic entropies.

Using the known inequalities (see, for example, [41]), we obtain the inequalities for tomographic entropies

\[
T_q(A, u) \leq T_q(A, B, u), \quad T_q(A \mid B, u) \leq T_q(A, u).
\]

A new aspect of these inequalities is that one can consider the minima of the Tsallis entropy for particular unitary tomograms \( u = u_{10} \otimes u_{20} \) for which

\[
T_q(A, u_{10} \otimes u_{20}) = \frac{1}{1-q} (\text{Tr} \rho^q(A) - 1), \quad T_q(A, B, u_{10} \otimes u_{20}) \geq \frac{1}{1-q} (\text{Tr} \rho^q(A, B) - 1).
\]

In the limit \( q \to 1 \), \( T_q(A, u_{10} \otimes u_{20}) \to S(A) \).

Then we have the inequalities for the von Neumann \( S \) and Shannon \( H \) entropies as well as the conditional Shannon tomographic entropy \( H(A \mid B, u) \) as follows:

\[
S(A) \leq H(A, B, u_{10} \otimes u_{20}), \quad H(A \mid B, u_{10} \otimes u_{20}) \leq S(A).
\]
8 Vectors and Matrices with Nonnegative Numbers

The discussed properties of entropies and their inequalities can be related to the properties of vectors and matrices. Suppose that one has a rectangular matrix $P_{jk}, j = 1, 2, \ldots, n$ with nonnegative matrix elements such that $\sum_{j,k} P_{jk} = 1$. This means that one can interpret the vector $\vec{P}$ constructed as a column with rows taken as subsequent pieces of this vector. Among the matrix elements one can have zeros.

An analog of the subadditivity inequality reads

$$-\sum_{j=1}^{m} \left( \sum_{k=1}^{n} P_{jk} \right) \left( \ln \sum_{k'=1}^{n} P_{jk'} \right) - \sum_{k=1}^{n} \left( \sum_{j=1}^{m} P_{jk} \right) \left( \ln \sum_{j'=1}^{m} P_{jj'} \right) \geq -\sum_{j=1}^{m} \sum_{k=1}^{n} P_{jk} \ln P_{jk}. \quad (55)$$

One can illustrate this inequality for the probability 4-vector $\vec{P} = (a, b, c, d)$. We construct the $2 \times 2$ matrix of the form

$$P_{jk} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}. \quad (56)$$

Then inequality (55) reads

$$-(a + b) \ln(a + b) - (c + d) \ln(c + d) - (a + c) \ln(a + c) - (b + d) \ln(b + d) \geq -a \ln a - b \ln b - c \ln c - d \ln d. \quad (57)$$

One can use all the permutations of numbers $a, b, c,$ and $d$ to obtain other inequalities, but the right-hand side of (57) is invariant under the permutation.

Now we suppose that the vector $\vec{P}$ describes a joint probability distribution for two coins (subsystems $A$ and $B$)

$$a = w(\text{++}), \quad b = w(\text{+-}), \quad c = w(\text{-+}), \quad d = w(\text{--}). \quad (58)$$

Then the Shannon entropy

$$H(A, B) = -w(\text{++}) \ln w(\text{++}) - w(\text{+-}) \ln w(\text{+-}) - w(\text{-+}) \ln w(\text{-+}) - w(\text{--}) \ln w(\text{--}) \quad (59)$$

is smaller than the sum of entropies

$$H(A) = -(w(\text{++}) + w(\text{-+})) \ln (w(\text{++}) + w(\text{-+})) - (w(\text{+-}) + w(\text{-+})) \ln (w(\text{+-}) + w(\text{-+})) \quad (60)$$

and

$$H(B) = -(w(\text{++}) + w(\text{-+})) \ln (w(\text{++}) + w(\text{-+})) - (w(\text{-+}) + w(\text{-+})) \ln (w(\text{-+}) + w(\text{-+})) \quad (61)$$

i.e.,

$$H(A, B) \leq H(A) + H(B). \quad (62)$$

But if we take

$$a = w(\text{++}), \quad b = w(\text{--}), \quad c = w(\text{+-}), \quad d = w(\text{-+}) \quad (63)$$

the general inequality (57) provides the inequality for the functions

$$H_1 = -(w(\text{++}) + w(\text{-+})) \ln (w(\text{++}) + w(\text{-+})) - (w(\text{-+}) + w(\text{-+})) \ln (w(\text{-+}) + w(\text{-+})) \quad (64)$$

271
and

\[ H_2 = -(w(++) + w(+-)) \ln(w(++) + w(+-)) - (w(--)+w(+-)) \ln(w(--)+w(+-)), \tag{65} \]

which reads

\[ H_1 + H_2 \geq H(A, B). \tag{66} \]

### 9 Permutations of Factorized Joint Probability Distributions

Inequality (66) is different from inequality (62). We can see this difference for the 4-vector, which is the tensor product of two probability vectors

\[ \vec{P} = \left(\begin{array}{c} x \\ y \end{array}\right) \otimes \left(\begin{array}{c} \alpha \\ \beta \end{array}\right). \tag{67} \]

We rewrite (58) as

\[ a = x\alpha, \quad b = x\beta, \quad c = y\alpha, \quad d = y\beta. \tag{68} \]

In this case, instead of \( H(A, B) = H(A) + H(B) \), which follows from (62) for the joint probability vector, we obtain the inequality, which in terms of number \( \alpha, \beta, x, \) and \( y \), reads

\[ -(x\alpha + y\beta) \ln(x\alpha + y\beta) - (x\beta + y\alpha) \ln(x\beta + y\alpha) - x \ln x - y \ln y \]

\[ \geq -x \ln x - y \ln y - \alpha \ln \alpha - \beta \ln \beta. \tag{69} \]

Since \( H(A) = H_1 \), the sum \( H_1 + H_2 \geq H(A) + H(B) \) means that

\[ H_2 \geq H(B). \tag{70} \]

The meaning of this new inequality for bipartite systems without correlations in their subsystems needs to be clarified.

### 10 Matrices with Three Indices and the Strong Subadditivity Condition

Another extension of the strong subadditivity condition can be formulated in terms of the matrix \( P_{jkm} \) with three indices. We suppose that, for \( j = 1, 2, \ldots n_1 \), \( k = 1, 2, \ldots n_2 \), and \( m = 1, 2, \ldots n_3 \), all numbers \( P_{jkm} \) are nonnegative and satisfy the condition \( \sum_{j=1}^{n_1} \sum_{k=1}^{n_2} \sum_{m=1}^{n_3} P_{jkm} = 1 \). This means that initially we have the probability \( N \)-vector \( \vec{P} (N \leq n_1 \cdot n_2 \cdot n_3) \) and apply the labels \( jkm \) to all components of the vector. Then an analog of the strong subadditivity condition reads

\[ -\sum_{j=1}^{n_1} \sum_{k=1}^{n_2} \sum_{m=1}^{n_3} P_{jkm} \ln P_{jkm} - \sum_{k=1}^{n_2} \left( \sum_{j=1}^{n_1} \sum_{m=1}^{n_3} P_{jkm} \right) \ln \left( \sum_{j=1}^{n_1} \sum_{m=1}^{n_3} P_{jkm} \right) \]

\[ \leq -\sum_{j=1}^{n_1} \sum_{k=1}^{n_2} \left( \sum_{m'=1}^{n_3} P_{jkm'} \right) \ln \left( \sum_{m=1}^{n_3} P_{jkm} \right) - \sum_{k=1}^{n_2} \sum_{m=1}^{n_3} \left( \sum_{j=1}^{n_1} P_{jkm} \right) \ln \left( \sum_{j=1}^{n_1} P_{jkm} \right), \tag{71} \]

272
In the case where $P_{jkm} = w(j, k, m)$ is a joint probability distribution for three random variables (three subsystems $A$, $B$, and $C$ of a composite system), inequality (71) is the strong subadditivity condition.

In the particular case of a system without correlations, this inequality becomes the equality. Nevertheless, this inequality takes place for any set $P_{jkm}$ of nonnegative numbers, which can also label the components of arbitrary probability vectors $\vec{P}$, even for a single system or for a system consisting of several subsystems.

Any tomogram for $N$ qudit states $w(m_1, m_2, \ldots, m_N, u)$ with the density matrix $\rho$ can be considered as the probability $n$-vector

$$\vec{w}(u) = |uu_0|^2 \vec{\rho},$$

where $\vec{\rho} = (\rho_1, \ldots, \rho_n)$ is the column vector with eigenvalues of the density matrix $\rho_k$, $k = 1, 2, \ldots, n$, and columns of the unitary matrix $u_0$ are the eigenvectors of $\rho$. If one labels the vector components of the vector $\vec{w}$ writing it as a matrix $P_{jk}$ (if necessary, adding the corresponding number of zero components to the vector $\vec{w}$), one obtains the inequality for the unitary matrix.

Analogously, one can label the vector components of the vector $\vec{w}(u)$ as $P_{jkm}$. In this case, the discussed inequalities are entropic inequalities for tomograms and also the inequalities for the unitary matrices.

11 Conclusions

We point out our main results presented here.

We formulated some inequalities for sets of nonnegative numbers and matrices with nonnegative matrix elements.

For qudits, we studied relations between Shannon and $q$-entropies known in the conventional probability theory. We applied these relations to the tomographic-probability distributions determining the qudit states. Taking the minima of the entropies with respect to the local unitary transforms, we obtained the inequalities containing the von Neumann entropies and their $q$-generalizations. The new inequalities, such as (48), (53), and (54), can be used to characterize the degree of quantum correlations.

The obtained new entropic and information inequalities for qudit systems can be considered as some analogs of the quantum discord properties, which provide an extra clarification of the properties of quantum correlations. For continuous variables, the quantum evolution equations for optical tomograms of quantum systems were obtained in [42] and studied in [43]. The dynamical maps describing the evolution of hybrid classical–quantum systems were studied in [40]. The evolution of tomograms yields the evolution of entropies.

It is worth pointing out that some entropic inequalities for optical tomograms of photon states were checked experimentally [44], in addition to the photon-quadrature uncertainty relations checked in [45].

We will study the evolution of continuous variables in the future work.

Acknowledgments

This work was partially supported by the Russian Foundation for Basic Research under Project No. 11-02-00456_a.
References

[1] E. Shrödinger, Naturwissenschaften, 23, 807 (1935).
[2] J. S. Bell, Physica, 1, 195 (1965).
[3] J. S. Bell, Speakable and Unspeakable in Quantum Mechanics, Cambridge University Press (1987).
[4] J. F. Clauser, M. A. Horne, A. Shimony, and R. A. Holt, Phys. Rev. Lett., 23, 880 (1969).
[5] A. A. Klyachko, H. Ali Can, S. Binicioglu, and A. S. Shumovsky, Phys. Rev. Lett., 101, 020403 (2008).
[6] R. Lapkiewicz, P. Li, C. Schaeff, et al., Nature, 460, 490 (2011).
[7] H. Olliver and W. Zurek, Phys. Rev. Lett., 88, 017901 (2002).
[8] V. I. Man’ko and O. V. Man’ko, J. Exp. Theor. Phys., 85, 430 (1997).
[9] V. V. Dodonov and V. I. Man’ko, Phys. Lett. A, 239, 335 (1997).
[10] V. I. Man’ko, G. Marmo, E. C. G. Sudarshan, and F. Zaccaria, Phys. Lett. A, 327, 353 (2004).
[11] A. Ibort, V. I. Man’ko, G. Marmo, et al., Phys. Scr., 79, 065013 (2009).
[12] C. E. Shannon, Bell Syst. Tech. J., 27, 379 (1948).
[13] A. Rényi, Probability Theory, North-Holland, Amsterdam (1970).
[14] C. Tsallis, “Nonextensive statistical mechanics and thermodynamics: historical background and present status,” in: S. Abe and Y. Okamoto (Eds.), Nonextensive Statistical Mechanics and Its Applications, Lecture Notes in Physics, Springer, Berlin (2001), Vol. 560, p. 3.
[15] M. A. Man’ko, V. I. Man’ko, S. De Nicola, and R. Fedele, Acta Phys. Hung. B, 26/1-2, 71 (2006).
[16] S. De Nicola, R. Fedele, M. A. Man’ko, and V. I. Man’ko, Eur. J. Phys. B, 52, 191 (2006).
[17] M. A. Man’ko, J. Russ. Laser Res., 27, 405 (2006).
[18] S. De Nicola, R. Fedele, M. A. Man’ko, and V. I. Man’ko, J. Phys. Conf. Ser., 70, 012007 (2007).
[19] S. De Nicola, R. Fedele, M. A. Man’ko, and V. I. Man’ko, Theor. Math. Phys., 152, 1081 (2007) [tr. Teor. Mat. Fiz., 152, 241 (2007)].
[20] M. A. Man’ko, “Tomographic entropy and new entropic uncertainty relations,” in: G. Adenier, A. Yu. Khrennikov, P. Lahti, V. I. Man’ko, and T. Nieuwenhuizen (Eds.), Quantum Theory: Reconsideration of Foundations-4, AIP Conference Proceedings, New York (2007), Vol. 962, p. 132.
[21] V. N. Chernega and V. I. Man’ko, J. Russ. Laser Res., 29, 505 (2008).
[22] M. A. Man’ko and V. I. Man’ko, Theor. Math. Phys., 160, 995 (2009) [tr. Teor. Mat. Fiz., 160, 143 (2009)].
[23] M. A. Man’ko, J. Russ. Laser Res., 30, 514 (2009).
[24] M. A. Man’ko, Phys. Scr., 82, 038109 (2010).
[25] M. A. Man’ko, Phys. Scr., T143, 014017 (2011).
[26] M. A. Man’ko, Theor. Math. Phys., 168, 985 (2011) [tr. Teor. Mat. Fiz., 168, 151 (2011)].
[27] M. A. Man’ko and V. I. Man’ko, J. Russ. Laser Res., 32, 439 (2011).
[28] M. A. Man’ko and V. I. Man’ko, Found. Phys., 41, 330 (2011), [DOI 10.1007/s10701-009-9403-9].
[29] M. A. Man’ko and V. I. Man’ko, Phys. Scr., T147, 014020 (2012).
[30] V. I. Man’ko, G. Marmo, E. C. G. Sudarshan, and F. Zaccaria, “Entanglement structure of adjoint representation of unitary group and tomography of quantum states,” J. Russ. Laser Res., 24, 507 (2003).
[31] V. N. Chernega and V. I. Man’ko, J. Russ. Laser Res., 28, 103 (2007).
[32] C. Lupo, V. I. Man’ko, and G Marmo, J. Phys. A: Math. Theor., 40, 13091 (2007).
[33] S. N. Filippov and V. I. Man’ko, J. Russ. Laser Res., 30, 55 (2009).
[34] O. V. Man’ko and V. I. Man’ko, J. Russ. Laser Res., 25, 11 (2004).
[35] M. A. Man’ko, V. I. Man’ko, and R. Vilela Mendes, J. Russ. Laser Res., 27, 507 (2006), [arXiv: quant-ph/0602129].
[36] E. H. Lieb and M. B. Ruskai, J. Math. Phys., 14, 1938 (1973).
[37] M. B. Ruskai, “Lieb’s simple proof of concavity (A, B) → TR A^p K^+ B^{1−p} K and remarks on related inequalities,” arXiv: quant-ph/0404126 v4.
[38] M.A. Man’ko and V.I. Man’ko, “Tomographic entropic inequalities in the probability representation of quantum mechanics,” in: R. Bijker (Ed.), Beauty in Physics: Theory and Experiment, Conference Proceedings, New York (2012), Vol. 1488, p. 110 [arXiv:1208.5695 quant-ph].
[39] M. A. Man’ko, Phys. Scr., T153, 014045 (2013).
[40] M. A. Man’ko and V. I. Man’ko, “Entropy of conditional tomographic probability distributions for classical and quantum systems,” in: Proceedings of the Sixth International Workshop “Spacetime – Matter – Quantum Mechanics from the Planck Scale to Emergent Phenomena (DICE 2012), Castello Pasquini/Castiglioncello, Tuscany, Italy (September 17–21, 2012), J. Phys. Conf. Ser. (2013, in press).
[41] A. E. Rastegin, “Tests for quantum contextuality in terms of q-entropies,” arXiv:1210.6742 quant-ph.
[42] Ya. A. Korennoy and V. I. Man’ko, J. Russ. Laser Res., 32, 74 (2011).
[43] G. G. Amosov, Ya. A. Korennoy, and V. I. Man’ko, Phys. Rev. A, 85, 052119 (2012).
[44] M. Bellini, A. S. Coelho, S. N. Filippov, et al., Phys. Rev. A, 85, 052129 (2012).

[45] V. I. Man’ko, G. Marmo, A. Porzio, et al., Phys. Scr., 83, 045001 (2011).