MUTUAL INFORMATION OF BIPARTITE STATES
AND QUANTUM DISCORD
IN TERMS OF COHERENCE INFORMATION

FEDOR HERBUT *
Serbian Academy of Sciences and Arts, Knez Mihajlova 35,
11000 Belgrade, Serbia and Montenegro

April 1, 2022

In relation of observable and quantum state, the entity $I_C$ from previous work quantifies simultaneously coherence, incompatibility and quantumness. In this article its application to quantum correlations in bipartite states is studied. It is shown that Zurek’s quantum discord can always be expressed as excess coherence information (global minus local). Strong and weak zero-discord cases are distinguished and investigated in terms of necessary and sufficient and sufficient conditions respectively. A unique string of relevant subsystem observables, each a function of the next, for “interrogating” the global state about the state of the opposite subsystem is derived with detailed entropy and information gain discussion. The apparent disappearance of discord in measurement is investigated, and it is shown that it is actually shifted from between subsystems 1 and 2 to between subsystems 1 and $(2 + 3)$, where 3 is the measuring instrument. Finally, it is shown that the global coherence information $I_C(A_2, \rho_{12})$ is shifted into the global coherence information $I_C(A_2, \rho^f_{123})$ in the final state $\rho^f_{123}$ of the measurement interaction.

1 Introduction

The investigation in this article is directed at the quantum correlations contained in a general, i.e., pure or mixed, bipartite state. By “contained” is meant the von Neumann mutual information of the state. There are numerous other important approaches in the literature that are not limited to the mutual information [1]. These will not be touched upon in this work.

No need to expand on the importance of this problem for quantum information theory, quantum communications, and quantum computers.

We will distinguish the two subsystems by 1 and 2. The former will be called "the distant" subsystem, and the latter "the nearby" one. We will distinguish "local" properties of the nearby subsystem (or of the distant one), and "global" ones of the bipartite state.

The approach of this article is based on the concept of coherence information. Coherence of an observable $A$ with respect to a quantum state $\rho$ and the incompatibility of the two have been simultaneously quantified by the concept of coherence

*e-mail: fedor@infosky.net
information $I_C(A, \rho)$ [2]. It is defined in three equivalent ways:

$$I_C(A, \rho) = S(\sum P_l \rho P_l) - S(\rho), \quad (1a)$$

where $A = \sum a_l P_l$ is the spectral form of the Hermitian operator $A$ in terms of distinct eigenvalues $a_l$, and $S(\ldots)$ is the von Neumann entropy of a state. Further,

$$I_C(A, \rho) = S(\rho|| \sum P_l \rho P_l), \quad (1b)$$

where $S(\rho||\sigma)$ is the relative entropy, a known function of two states, and finally,

$$S(\rho) = S(A, \rho) + \sum p_l S(P_l \rho P_l / p_l) - I_C(A, \rho), \quad (1c)$$

where $S(A, \rho) = H(p_l) \equiv - \sum p_l \log p_l$ quantifies the uncertainty of $A$ in $\rho$ in terms of the Shannon entropy $H(p_l)$ of the probability distribution $\forall l: p_l \equiv \text{tr}(P_l \rho)$.

The coherence information $I_C(A, \rho)$ quantifies also the quantumness in the relation between observable and state: The relation is quasi-classical if and only if $A$ and $\rho$ are compatible $[A, \rho] = 0$; in this and only in this case $I_C(A, \rho) = 0$.

It will turn out that the coherence-information approach of this paper is closely connected with the Zurek concept of quantum discord. (It will be called shortly "discord"). Zurek introduced an approach in which the bipartite state $\rho_{12}$ is investigated by "interrogating" it with a complete nearby subsystem observable $A^c_2$ [3], [4], [5], [6]. The associated discord $\delta_{A^c_2}(\rho_{12})$ appeared as the natural quantification of quantumness of the correlations. It is not entanglement. Also separable states, which, by definition of entanglement, do not contain it, are stated to have positive discord, showing quantumness in "interrogation" by a concrete subsystem observable. Nevertheless, discord addresses, just like entanglement, though in a different way, the same basic problem of quantum correlations: What is there typically quantum mechanical in them?

In [5] Zurek takes a thermodynamical approach to the study of the physical meaning of least discord $\delta \equiv \inf_{A^c_2} \delta_{A^c_2}(\rho_{12})$. He does this using the idea of a quantum demon extracting locally work from $\rho_{12}$. He finds that $\delta$ equals the (nonnegative) excess of work that a quantum demon can extract in comparison with a classical one. He also discusses how his approach relates to a similar thermodynamical approach of Oppenheim and the Horodecki family [7].

In a recent review article the Horodecki family, Oppenheim et al. gave a detailed presentation on "local versus non-local information" [8]. They discuss the connection between their approach and results with those of Zurek and his discord. Indirectly, the results of this article are connected also with this work.

Uhlmann gives an elementary presentation of an analogous approach to quantum correlations studies independently from both Zurek and the Horodecki school of thought [9].

It will be shown that discord is actually coherence-information excess (global minus local). This will make it possible to throw new light on the zero-discord problem. The "interrogating" complete observable $A^c_2$ will be generalized to include also incomplete observables $A_2$. Then a string of relevant observables, each a function of the next, will be derived that will eliminate, what will be called, redundant noise, eliminate the garbled part of the information gain (on the state of the distant subsystem), and,
finally, eliminate all quantumness - all this at the cost of diminishing the information gain.

The state \( \rho_L \equiv \sum_l P_l \rho P_l \) appearing in definitions (1a) and (1b) is the so-called Lüders mixture of \( \rho \) with respect to \( A \) [10], [11] (relation (14.16) on p. 225 there). It is the non-selective (or entire-ensemble) version. (Some authors call it "dephasing operation, e.g., [7].) The admixed Lüders state is \( P_l \rho P_l / p_l \), where \( p_l \equiv \text{tr}(\rho P_l) \) is the corresponding probability. It appears in the selective (or definite-result) version (utilized in (1c) e.g.).

To avoid unnecessary repetitions in the exposition, the following will be understood throughout the article: the physical term will be given priority, like "observable" instead of "Hermitian operator", "state" instead of "statistical operator", "mixture" instead of "decomposition of a statistical operator" (into a finite or infinite convex combination of statistical operators), "compatibility" instead of "commutation", etc.

Observables will be restricted to discrete ones, and as a rule, given in spectral form like \( A = \sum_i a_i P_i \) with all eigenvalues \( a_i \) distinct. This will always be tacitly accompanied by the completeness relation (decomposition of the identity) \( \sum_i P_i = 1 \). The sum "\( \sum_i \)" is finite or infinite as the case may be. If the sum is necessarily restricted to be finite for some claim to be valid, then it will be written, e.g., like \( \sum_{i=1}^m \) and it will be understood that \( m \) is an integer. If the spectrum may be finite or infinite, we will write \( \{ a_i : \forall l \} \), etc.; if it is necessarily finite, we will write \( \{ a_i : l = 1, 2, \ldots, m \} \). Complete observables \( A' = \sum_i a_i |l\rangle \langle l| \) are written with the suffix "c".

If the given state \( \rho \) has an infinite-dimensional null space, then also observables \( A \) that have a continuous part in their spectrum can be considered for coherence-information studies under the restriction that the subspace spanned by the eigenspaces of \( A \) contains the range of \( \rho \) [12].

Functions of an observable amount to coarsenings in the spectrum of the latter. We will prefer the term "coarsening" because it has the simple opposite "refinement". Both are order relations like "smaller or equal" and "larger or equal respectively".

When an observable \( A = \sum_i a_i P_i \) and a state \( \rho \) are given, we will speak of "detectable" eigenvalues \( a_i \) or index values \( l \) meaning those for which \( p_l \equiv \text{tr}(\rho P_l) > 0 \). The spectrum \( \{ a_i : \forall l \} \), and the set of the index values \( \{ l : \forall l \} \) will always be understood to be connected by a fixed one-to-one map, enabling us to talk of "corresponding" eigenvalue etc.

Mixtures like \( \rho = \sum_k w_k \rho_k \) finite or infinite, will be understood in a formal, not operational sense, i.e., they express the fact that one can write \( \rho \) in that way. The statistical weights will be called only "weights"; they can be positive or zero; in the latter case \( \rho_k \) need not be defined, nevertheless by definition \( w_k \rho_k = 0 \) (and analogously for other entities than \( \rho_k \)). The states \( \rho_k \) will be referred to as "admixed states".

A mixture \( \rho = \sum_k w_k \rho_k \) is orthogonal if \( k \neq k' \Rightarrow \rho_k \rho_{k'} = 0 \). An example is the Lüders mixture \( \rho_L = \sum_l p_l (P_l \rho P_l / p_l) \). Then the mixing property of entropy is valid: \( S(\rho_L) = H(p_l) + \sum_l p_l S(P_l \rho P_l / p_l) \) (see p. 242 in [13]).

Both for mixtures and for observables the subsystem will be exhibited in the index, e.g., \( \sum_k w_k \rho_k P_{l_2} \sum_i a_i P_{l_2} \). Both in mixtures and in observables we will deal with coarsenings, and binary relations "linked" and "chained". To distinguish the two cases, we will occasionally use the terms "m-coarsening", "m-linked" and "m-chained" for mixtures, and "o-coarsening", "o-linked", and "o-chained" for observables.

Mentioning subsystems, we will often omit "nearby", and only say "subsystem".
2 Role of Coherence Information in Bipartite Quantum Correlations

We take a bipartite state $\rho_{12}$ with its reductions $\rho_s \equiv \text{tr}_s \rho_{12}$, $s, s' = 1, 2, s \neq s'$, and a subsystem observable $A_2 = \sum_i a_i P^i_2$. Two coherence informations $I_C(A_2, \rho_2)$ and $I_C(A_2, \rho_{12})$ appear. Also two Lüders mixtures $\rho^l_2 \equiv \sum_i P^i_2 \rho_2 P^i_2$ and $\rho^l_{12} \equiv \sum_i P^i_2 \rho_{12} P^i_2$ enter the scene. (Here $P^i_2$ is short for $(1 \otimes P^i_2)$.) We utilize the notation:

$$\forall l : \quad p_l \equiv \text{tr}(\rho_{12} P^l_2),$$

(2a)

$$\forall l, p_l > 0 : \quad \rho_{12}^l \equiv \frac{P^l_2 \rho_{12} P^l_2}{p_l},$$

(2b)

$$\forall l, p_l > 0 : \quad \rho^l_s \equiv \text{tr}_s(\rho_{12}^l), \quad s, s' = 1, 2, s \neq s'.$$

(2c)

Next, we’ll need the entropy additivity accompanying any mixture $\rho = \sum_k w_k \rho_k$:

$$S(\rho) = J + \sum_k (w_k S(\rho_k)),$$

(3a)

$$J = \sum_k (w_k S(\rho_k||\rho)),$$

(3b)

(If proof is wanted for the known relation (3b), see proof of Lemma 4 in [14].) If the mixture is orthogonal, then $J$ takes the special form of the Shannon entropy $H(w_k)$ due to the mixing property. (See proposition 7 below for more on $J$.)

Now we consider a relevant decomposition of the mutual information $I_{12} \equiv I(\rho_{12}) \equiv S_1 + S_2 - S_{12}$, where $S_s \equiv S(\rho_s)$, $s = 1, 2, 12$.

**Theorem 1**: A) The mutual information $I_{12}$ of any bipartite state $\rho_{12}$, when viewed in relation to any given discrete second-subsystem observable $A_2$, can be decomposed as follows:

$$I_{12} = J_{A_2} + \left(I_C(A_2, \rho_{12}) - I_C(A_2, \rho_2)\right) + \sum_l p_l I(\rho_{12}^l),$$

(4a)

where

$$J_{A_2} \equiv \sum_l p_l S(\rho_{12}^l||\rho_1),$$

(4b)

and

$$\rho_1 = \sum_l p_l \rho_{12}^l$$

(4c)

is the distant mixture induced by $A_2$.

B) Each of the three terms on the RHS of (4a) is always nonnegative.

**Proof**: A) We utilize the entropy decompositions (3a) for $\rho_1$ and (1c) for $\rho_s$, $s = 2, 12$:

$$I_{12} \equiv S_1 + S_2 - S_{12} \equiv \left[J_{A_2} + \sum_l p_l S(\rho_{12}^l)\right] + \left[H(p_l) + \sum_l p_l S(\rho_{12}^l) - I_C(A_2, \rho_2)\right] -$$

4
\[
\left[ H(p) + \sum_i p_i S(\rho^i_{12}) - I_C(A_2, \rho_{12}) \right] = \text{RHS}(4a).
\]

This completes the proof of part A).

B) The first and the third terms on the RHS of (4) are obviously non-negative. To prove that also the second term is non-negative we need two auxiliary claims.

Corollary 1: Decomposition (4a) in application to the Lüders mixture \(\rho^L_{12} \equiv \sum_l P^l_{12} P^l_{22} \) gives:

\[
I(\rho^L_{12}) = J_{A_2} + \sum_i p_i I(\rho^i_{12}).
\] (5)

Proof: Straightforward evaluation gives \(I_C(A_2, \rho^L_{12}) = I_C(A_2, \rho^L_{22})\) (or see proposition 1 below).

Lemma 1: The inequality \(I(\rho^L_{12}) \leq I(\rho_{12})\) is always valid.

Proof: As it is well known, the mutual information can be written in the form of relative entropy \(I_{12} = S(\rho_{12}||\rho_1 \otimes \rho_2)\). By this same formula also \(I(\rho^L_{12}) = S\left(\sum_l P^l_{12} P^l_{22}||\rho_1 \otimes (\sum_l P^l_{22} P^l_{22})\right)\). It was proved by Lindblad [15] for the finite-dimensional case (Theorem on p. 149 there) that \(S(\Phi\rho||\Phi\sigma) \leq S(\rho||\sigma)\) for every two states \(\rho\) and \(\sigma\) and every completely positive trace-preserving map \(\Phi\). The inequality was extended to the infinite-dimensional case by Uhlmann [16]. (It is unjustly called a theorem of Uhlmann instead of one of Lindblad and Uhlmann.)

Since \(\Phi \equiv \sum_l P^l_{22}...P^l_{22}\) is such a map, the lemma is proved.

End of proof of part B) of Theorem 1: Comparing (4a) and (5) and taking into account Lemma 1, one obtains

\[
I(\rho_{12}) - I(\rho^L_{12}) = I_C(A_2, \rho_{12}) - I_C(A_2, \rho_2) \geq 0
\]
in the general case.

In the classical discrete case one has a relation analogous to (3a) and (3b), and one analogous to (1c), but the latter with \(I_C\) missing. Then a relation analogous to (4a) is obtained (analogously as in the proof of theorem 1), but without the excess of coherence information (the second term) on the RHS. Following Zurek [3], this term quantifies the quantumness in the mutual information and is called the quantum discord with respect to a complete or incomplete second-subsystem discrete observable \(A_2\), and it is denoted by \(\delta_{A_2}(\rho_{12})\).

The following physical interpretation of (4a) suggests itself. The observable \(A_2\) is a probe (or an “interrogation”, cf [5]) into the quantum correlations in \(\rho_{12}\) making subsystem 2 the nearby one (the instrument measuring \(A_2\) interacts directly with it), and subsystem 1 the distant one (no interaction with the measuring apparatus). Applying (3a) and (3b) to the mixture (4c), one obtains

\[
S(\rho_1) = \sum_i \left(p_i S(\rho^i_{11}||\rho_1)\right) + \sum_i \left(p_i S(\rho^i_{11})\right).
\] (6)

In view of (6), the first term on the RHS of (4a) is obviously the information gain about the distant subsystem acquired by the probe (cf [17], [3], [4]). The detectable eigenvalues \(a^i\) of \(A_2\) distinguish and enumerate the admixed states \(\rho^i_{11}\), and the acquired information is the gain in the distant mixture (4c).
The third term on the RHS of (4a) is the amount of quantum correlations in \( \rho_{12} \) that is *inaccessible* by the probe used. (As easily seen, it is zero if \( A_2 \) is complete.) We shall call it residual correlations. Both the first and the third term are entropy terms, i.e., as easily seen, they are concave with respect to mixtures. But since the mutual information appears with a minus sign in the subsystem entropy decomposition \( S(\rho_{12}) = S(\rho_1) - I(\rho_{12}) + S(\rho_2) \), the mentioned terms are actually convex as information quantities should be.

Discord \( \delta_{A_2}(\rho_{12}) \), being, in general, excess coherence information, i.e., a difference of two information quantities:

\[
\delta_{A_2}(\rho_{12}) = [I_C(A_2, \rho_{12}) - I_C(A_2, \rho_2)],
\]

is neither convex nor concave (because coherence information is convex, cf proposition 5 in [2]). This fact gives some insight into Lieb’s result that mutual information is neither convex nor concave in the general case [18]. Some authors call \( I_{12} \) "mutual entropy". Having its behavior under mixing in view, it is neither information nor entropy. (See [19] - subsection III.c there - for a different point of view.)

Discord is a necessary accompaniment of the described probing into \( \rho_{12} \) by \( A_2 \). It is due to quantumness of the correlations.

Assuming that the observable \( A_2 = \sum a_l P_l^2 \) is incomplete, one may wonder how the terms in (4a) behave when \( A_2 \) is refined (down to a complete observable or just to a more complete one). By refinement is meant another observable

\[
A'_2 = \sum_{l,q} a_{l,q} P_{2}^{l,q}
\]

(8a)

(the range of \( q \) depends on the value of \( l \); for simplicity, this is omitted in notation). It is by definition such that it further decomposes the eigenprojectors of \( A \), i.e.,

\[
\forall l : \quad P_{2}^{l} = \sum_{q} P_{2}^{l,q}.
\]

This is refinement in an absolute sense, i.e., it does not depend on any state \( \rho_2 \). We need the generalization of this notion to state-dependent refinement.

Let besides \( A'_2 \) (cf (8a)) also \( A_2 \) and \( \rho_2 \) be given. Let \( l' \) enumerate the detectable and \( l'' \) the undetectable eigenvalues of \( A_2 \) in \( \rho_2 \). Then

\[
A_2 = \sum_{l'} a_{l'} P_{2}^{l'} + \sum_{l''} a_{l''} P_{2}^{l''}.
\]

(8b)

If

\[
\forall l' : \quad P_{2}^{l'} = \sum_{q} P_{2}^{l',q},
\]

(8c)

then we say that \( A'_2 \) is a (state-dependent) refinement of \( A_2 \) with respect to \( \rho_2 \), and we write \( A'_2 \geq \rho_2 \) \( A_2 \). (The symbol \( \geq \) is to remind us that we are dealing with a reflexive and transitive binary relation - like "larger or equal" - that is state dependent.)

**Theorem 2:** In refinement of \( A_2 \) by \( A'_2 \) with respect to \( \rho_2 \) (cf (8a)-(8c)), the reduction of a given arbitrary bipartite state \( \rho_{12} \), the information gain and the discord remain equal or become larger, and the residual correlations remain the same or become smaller. To be explicit quantitatively, one can write (4a) and (4b) with respect to \( A'_2 \).
as a two-step expression (as if the probing took place first with $A_2$, and then it was continued to $A'_2$):

$$I_{12} = \left\{ \sum_i \left( p_i S(\rho_i^1 || \rho_1) \right) + \sum_{l,q} \left( p_{l,q} S(\rho_{l,q}^1 || \rho_1) \right) \right\} + \left\{ \delta_{A_2}(\rho_{12}) + \sum_l p_l \delta_{A'_2}(\rho_{l12}) \right\} + \left\{ \sum_{l,q} \left( p_{l,q} I(\rho_{l,q}^{12}) \right) \right\},$$  \hspace{1cm} (9)

where the expressions in the large brackets are the information gain, the discord and the residual correlations respectively (and $p_{l,q} \equiv \text{tr}(P_{l,q}^2 \rho_{12})$).

**Proof** is given in Appendix A.

Information gain is the basic purpose of the probe, hence, one wants it to be as large as possible. This is the reason why most studies are restricted to complete observables $A_c^2$. Then, whenever $p_l > 0$, the state $|l\rangle_2 \rho_2 |l\rangle_2 (l_2)$ is pure, $\rho_{l12} = \rho^1_l \otimes |l\rangle_2 \langle l_2|$ is uncorrelated, and $I(\rho_{l12}) = 0$. Then (4a) is simplified to become

$$I(\rho_{12}) = J_{A_c^2} + \left( I_C(A_c^2, \rho_{12}) - I_C(A_c^2, \rho_2) \right).$$  \hspace{1cm} (10)

It was argued in [20] that taking the infimum of the discords in (10) (cf (7))

$$\hat{\delta}(\rho_{12}) \equiv \text{inf}_{A_c^2} \delta_{A_c^2}(\rho_{12})$$  \hspace{1cm} (11)

one may obtain an observable-independent quantum measure of quantumness in $I_{12}$. Vedral et al. [17] take into account also generalized observables, and then, taking the supremum of the $J_{A_c^2} = \sum_l p_l S(\rho_l^1 || \rho_1)$ expressions, they define the classical part of $I_{12}$.

3 On Zero Discord

As it is obvious from (7), a discord $\delta_{A_c^2}(\rho_{12})$ can be zero either if both coherence informations are zero, then we call it strong zero, or if both coherence informations are positive but equal. We call this case weak zero.

A detailed analysis including open problems (at least for the author) on unachieved results is now presented.

3.1 Strong zero discord with an incomplete or complete observable

**Proposition 1:** Each of the following two relations is a necessary and sufficient condition for an observable $A_2 = \sum a_l P_{l2}^j$ to have a strong zero discord in a given bipartite state $\rho_{12}$:

$$[A_2, \rho_{12}] = 0,$$  \hspace{1cm} (12)

$$\rho_{12} = \sum_l P_{l2}^j \rho_{l2} P_{l2}^j.$$  \hspace{1cm} (13)

**Proof:** Upon partial trace over the first subsystem, the commutation (12) becomes $[A_2, \rho_2] = 0$. Hence the sufficiency and the necessity of this condition is obvious.
Relation (12) is equivalent to
\[ \forall l : \quad [P^l_2, \rho_{12}] = 0. \quad (14) \]
The identity \( \rho_{12} = \sum_k w_k \rho^k_{12} \) idempotency and commutation then give (13). Conversely, (13) implies (14).

\[ \square \]

**Remark 1:** Relation (12) implies the local necessary condition \([A_2, \rho_2] = 0\) for strong zero discord. A local sufficient condition is not possible in a nontrivial way. Namely, if such a condition were given in terms of strong zero discord. A local sufficient condition is not possible in a nontrivial way. Then, relation (14) would imply, as easily seen,
\[ \exists l : \quad \forall l : \quad (1 \otimes P^l_2) |\Psi\rangle_{12} = \delta_{l,l} |\Psi\rangle_{12}, \]
and further
\[ \forall l : \quad P^l_2 \rho_2 = \delta_{l,l} \rho_2. \]
This gives zero discord, but it also gives zero information gain \( J = 0 \) because it does not decompose \( \rho_1 \) at all, and thus it is a trivial probe.

One wants to know what kind of state \( \rho_{12} \) has a strong zero discord.

**Definition 1:** If a bipartite state \( \rho_{12} \) is a nontrivial mixture of admixed states \( \rho^k_{12} \)
\[ \rho_{12} = \sum_k w_k \rho^k_{12} \quad (15a) \]
(all weights \( w_k \) being positive) so that
\[ k \neq k' \implies \rho^k_{12} \rho^{k'}_{12} = 0, \quad (15b) \]
where \( \forall k : \quad \rho^k_2 \equiv \text{tr}_1 \rho^k_{12}, \) then \( \rho_{12} \) is said to be mono-orthogonal (in the second subsystem).

**Proposition 2:** A bipartite state \( \rho_{12} \) has a strong zero discord if and only if it is mono-orthogonal (in the second subsystem).

**Proof:** Sufficiency. Let us assume that a state \( \rho_{12} \) for which (15a) and (15b) are valid is given. Let us, further, for each \( k \) value denote by \( Q^k_2 \) the range-projector of \( \rho^k_2 \). Finally, let us define \( A_2 \equiv \sum_k a_k Q^k_2 \) with arbitrary but distinct eigenvalues \( a_k \). Then one has \( \forall k : \quad \rho^k_{12} = Q^k_2 \rho_{12} Q^k_2 \) (This is a known but not well known general relation. For proof cf relation (12a) in [21].) Hence (14) (changing what has to be changed) holds true.

Necessity. If \( \rho_{12} \) has a strong zero discord with respect to an observable \( A_2 = \sum_l a_l P^l_2 \), then, according to the necessary condition (13), one can write \( \rho_{12} = \sum_l p_l \rho^l_{12} \), where the prim on the sum denotes that all \( (p_l = 0) \)-terms are omitted, and \( \forall l, p_l > 0 : \quad \rho^l_{12} = P^l_2 \rho_{12} P^l_2 / p_l \). This is of the form (15a). Further, \( \forall l, p_l > 0 : \quad \rho^l_2 \equiv \text{tr}_1 \rho^l_{12} = P^l_2 \rho_{12} P^l_2 / p_l \), and requirement (15b) (with \( l \) instead of \( k \)) is obviously satisfied.

**Remark 2:** Let it be locally known that \( \rho_{12} \) is mono-orthogonal. This means that besides \( \rho_2 \) also an orthogonal projector decomposition \( \sum_k Q^k_2 = Q_2 \) of the range projector \( Q_2 \) of \( \rho_2 \) is given and it is known that it is associated with mono-orthogonality, i.e., \( Q^k_2 \) is the range projector of \( \rho^k_2 \equiv \text{tr}_1 \rho^k_{12} \), where \( \rho^k_{12} \) are the admixed mono-orthogonal states in (15a). Then, as easily seen, a local sufficient condition for strong
zero discord is that each eigenprojector $P_l^2$ of $A_2$ be a sum of $Q_k^2$ projectors. This implies the necessary condition $[A_2, \rho_2] = 0$ (because the $Q_k^2$ projectors commute with $\rho_2$). Nevertheless, it is not a necessary and sufficient condition, because it may require too much. A necessary and sufficient local condition cannot be given in view of lack of knowledge of the admixed mono-orthogonal states $\rho_{12}^k$ (cf remark 1).

3.2 Strong zero discord with a complete observable

The necessary and sufficient condition (12) is unchanged, but, since now $A_2 = \sum a_l |l\rangle_2 \langle l|_2$, (13) and (14) take the respective forms:

$$\rho_{12} = \sum_l |l\rangle_2 \langle l|_2 \rho_{12} |l\rangle_2 \langle l|_2,$$

and

$$\forall l : \quad [|l\rangle_2 \langle l|_2, \rho_{12}] = 0.$$  

Condition (16) was highlighted in [4] (in a less elaborate context, without distinguishing strong and weak zero discord).

**Proposition 3:** A bipartite state $\rho_{12}$ has a strong zero discord with respect to a complete observable $A_2 = \sum a_l |l\rangle_2 \langle l|_2$ *if and only if* it is a mixture of the form

$$\rho_{12} = \sum_l p_l \rho_1^l \otimes |l\rangle_2 \langle l|_2.$$  

**Proof:** Sufficiency. If (18) is valid, then so is (16).

Necessity. Since $\forall l : \quad [|l\rangle_2 \langle l|_2, \rho_{12}] = ([|l\rangle_2 \langle l|_2, \rho_{12}, \rho_{12}^2] \langle l|_2 \langle l|_2 = p_l \rho_1^l \otimes |l\rangle_2 \langle l|_2$ (cf (2b) and (2c) with $|l\rangle_2 \langle l|_2$ instead of $P_l$). Thus, (16) implies (18).

**Proposition 4:** A bipartite state $\rho_{12}$ has a strong zero discord with respect to some complete observable $A_2$ *if and only if* the state is mono-orthogonal (cf (15a) and (15b)), and

$$\forall k : \quad \rho_{12}^k = \rho_1^k \otimes \rho_2^k,$$

i. e., if it is simultaneously also separable.

**Proof:** Sufficiency. Let (15a) and (15b) be given, and let $\rho_{12}$ be simultaneously also separable as stated. Substituting each $\rho_2^k$ by its spectral form in terms of eigen-ray-projectors, one obtains $\rho_{12}$ as a mixture of the form (18) (changing what has to be changed).

Necessity. The form (18) is mono-orthogonal and simultaneously separable.

**Proposition 5:** Let $\rho_{12}$ be a mixture of the form

$$\rho_{12} = \sum_k w_k \rho_1^k \otimes \rho_2^k$$

with the validity of (15b) (cf definition 1 and proposition 4). Then a *local sufficient condition* for $A_2 = \sum a_l P_l^2$ to give a strong zero discord for $\rho_{12}$ is:

$$\forall k : \quad [A_2, \rho_{12}^k] = 0.$$

**Proof:** It is obvious in (19) that, on account of (20), $A_2$ commutes with $\rho_{12}$ (cf proposition 1).
3.3 Weak zero discord

We begin by two general results, which play an auxiliary role in this subsection.

**Lemma 2:** Let $\rho$ be a state and $A = \sum l a_l P_l$ an observable. Let, further, $\sum_n P_n = 1$ be an (orthogonal projector) decomposition of the identity such that

$$\forall n: \quad [P_n, \rho] = [P_n, A] = 0. \tag{21}$$

Then the following statistical decomposition of the coherence information ensues:

$$I_C(A, \rho) = \sum_n w_n I_C(A, P_n \rho/w_n), \tag{22}$$

where $\forall n: \quad w_n \equiv \text{tr}(\rho P_n)$.

**Proof:** On account of (21), one has the mixture $\rho = \sum_n w_n (P_n \rho/w_n)$, and, $[P, P_n] = 0$. Hence,

$$I_C(A, \rho) \equiv S\left(\sum_l P_l \rho P_l\right) - S(\rho) =$$

$$S\left(\sum_n w_n \sum_l P_l (P_n \rho/w_n) P_l\right) - S\left(\sum_n w_n (P_n \rho/w_n)\right) =$$

$$H(w_n) + \sum_n w_n S\left(\sum_l P_l (P_n \rho/w_n) P_l\right) -$$

$$[H(w_n) + \sum_n w_n S(P_n \rho/w_n)] = \sum_n w_n I_C(A, P_n \rho/w_n).$$

The symbol $H(w_n)$ denotes the Shannon entropy $-\text{tr}(w_n \log w_n)$, and the mixing property of entropy has been made use of. $\square$

**Lemma 3:** Let $\rho_{12}$ be a bipartite state and $A_2 = \sum l a_l P^2_l$ a subsystem observable. Besides, let $\sum_n P^2_n = 1$ be a subsystem (orthogonal projector) decomposition of the identity such that

$$\forall n: \quad [P^2_n, \rho_{12}] = 0 \quad \text{and} \quad [P^2_n, A_2] = 0. \tag{23}$$

Then the following statistical decomposition of the discord is valid:

$$\delta_{A_2}(\rho_{12}) = \sum_n w_n \delta_{A_2}(P^2_n \rho_{12}/w_n), \tag{24}$$

where the mixture $\rho_{12} = \sum_n [w_n (P^2_n \rho_{12}/w_n)]$ is due to (23).

**Proof:** Taking the first-subsystem partial trace in the first commutation relation in (23), one obtains $[P^2_n, \rho_2] = 0$. Hence, according to (7) and lemma 2,

$$\delta_{A_2}(\rho_{12}) = I_C(A_2, \rho_{12}) - I_C(A_2, \rho_2) = \sum_n w_n \delta_{A_2}(P^2_n \rho_{12}/w_n).$$

$\square$

**Proposition 6:** A sufficient condition for a weak zero discord of $A_2$ in $\rho_{12}$ is the mixture (19) of the latter with (15b) valid, further,

$$\forall k: \quad [A_2, Q^k_2], \tag{25a}$$

where $Q^k_2$ is the range projector of $\rho_2^k$, and finally, for at least one detectable value $\bar{k}$ of $k$ one has

$$[A_2, \rho_2^\bar{k}] \neq 0. \tag{25b}$$
Proof: Since \( \forall k : \rho_k^{12} = Q_k^{12} \rho_0^{12} Q_k^{12} \), and \( \rho_1^{12} \otimes \rho_2^{12} = Q_1^{12} (\rho_1^{12} \otimes \rho_2^{12}) Q_1^{12} \), the assumptions of lemma 3 are satisfied with the decomposition \( \sum_k Q_k^{12} = 1 \). (The null-space projector of \( \rho_2 \), if it is nonzero, is joined to the \( Q_k^{12} \).) Hence, one can write
\[
\delta_{A_2}(\rho_{12}) = \sum_k w_k \delta_{A_2}(\rho_k^{12} \otimes \rho_k^{12}) = 0,
\]
because uncorrelated states have zero mutual information, and this is an upper bound for the (nonnegative) discord (cf (7) and (4a)).

On the other hand, also the assumptions of lemma 2 are satisfied. Thus
\[
I_C(A_2, \rho_2) = \sum_k w_k I_C(A_2, \rho_k^{12}) \geq w_k I_C(A_2, \rho_k^{12}) > 0.
\]
In view of (7), the zero discord must be weak.

Remark 3: One would like to know if the condition in Proposition 6 is also necessary, or if some other at least partially local necessary and sufficient condition is valid.

As it is well known, in quantum mechanics, unlike in the classical discrete case, the von Neumann mutual information \( I_{12} \) can exceed the subsystem entropies, actually \( I_{12} \leq 2 \min \{S(\rho_1), S(\rho_2)\} \). Any correlated pure bipartite state is a good example, because, as it is also well known, there \( I_{12} = 2S(\rho_1) = 2S(\rho_2) \).

Substituting (7) in (4a) and utilizing (3a), (4a) implies for any complete subsystem observable \( A_2^c \)
\[
\delta_{A_2^c} = \sum_i p_i S(\rho_i) + (S_{12} - S_1).
\]
(26)
If \( I_{12} \) exceeds \( S_1 \), then (26) gives rise to a lower bound
\[
\delta_{A_2^c} \geq (I_{12} - S_1) > 0.
\]
(27)
Thus, for such typically quantum states \( \rho_{12} \) no choice of \( A_2^c \) can give zero discord.

Cerf and Adami introduced quantum conditional entropies \( S(1|2) \) [22]. One has \( S(1|2) = S_1 - I_{12} \). If (27) is valid, then \( S(1|2) < 0 \). It is what Adami and Cerf call "supercorrelations" [19].

The opposite-sign entity \(-S(1|2) \equiv E(1 \rightarrow 2)\) is called "directed entanglement" by Devetak and Staples [23]. Its properties are discussed and it is applied to quantum communication. The same entity was called "coherent quantum information" (not to be confused with "coherence information" of the present study) by Schumacher and Nielsen [24] with analogous discussion and application.

Remark 4: One would like to know if there can be zero discord between the case of mono-orthogonal and the case of states for which (27) is valid. In other words, one wonders if for some separable but not mono-orthogonal states and for some nonseparable but states for which \( I_{12} \leq S_1 \), one can find a complete subsystem observable \( A_2^c \) giving zero discord.

Remark 5: It is desirable to learn if in the definition of the least discord \( \delta \equiv \inf(A_2^c) \delta_{A_2^c}(\rho_{12}) \) one can replace "\( \inf \)" by "\( \min \)" or not. In other words, it might be that there exist states \( \rho_{12} \) for which \( \delta \) is "irrational" in the sense that it can be reached by no \( A_2^c \), but it can be arbitrarily well approximated by some \( \delta_{A_2^c} \). One wants to see such states if they exist, or to see a proof that they do not exist. This is, of course, important also for \( \delta = 0 \).
The investigation in this section reveals that there is a number of open problems about the zero discord (contrary to a false impression one might mistakenly get, e. g., from [4]).

4 String of Relevant Coarsenings

4.1 Elaborate subsystem entropy decomposition

When a bipartite state \( \rho_{12} \) is given and a subsystem observable \( A_2 \) is selected, then the subsystem entropy decomposition

\[
S_{12} = S_1 - I_{12} + S_2
\]  

(28a)
can be viewed in the more elaborate way

\[
S_{12} = \left\{ \sum_i p_i S(\rho_1^i) + J_{A_2}(\rho_1) \right\} = \left\{ J_{A_2}(\rho_1) + \delta_{A_2}(\rho_{12}) + \sum_i p_i I(\rho_{12}^i) \right\} + \left\{ H(p_1) - I_C(\rho_{12}) + \sum_i p_i S(\rho_{2}^i) \right\}
\]  

(28b)

(cf (2a)-(2c), (3a) and (3b), (4a), (7), and (1c)). Naturally, \( J_{A_2}(\rho_1) = J_{A_2}(\rho_{12}) \). It is understood that each expression in large brackets in (28b) equals the corresponding term on the RHS of (28a).

The elaborate subsystem entropy decomposition (28b) can be interpreted physically as follows. The subsystem observable \( A_2 \) is chosen to "interrogate" the uncertainty in the distant subsystem 1; the measure of the latter is \( S_1 \). On account of this, \( S_1 \) is broken up into \( \sum_i p_i S(\rho_1^i) \), the part of \( S_1 \) that is inaccessible to our "interrogation" (or the residual part), and \( J_{A_2}(\rho_1) \), the information gain. The mutual information \( I_{12} \), which quantifies the total quantum correlations in \( \rho_{12} \), is decomposed into the mentioned information gain \( J_{A_2}(\rho_1) \), the discord \( \delta_{A_2}(\rho_{12}) \), and \( \sum_i p_i I(\rho_{12}^i) \), which is the part that is not made use of in the chosen "interrogation" (residual correlations). The appearance of the information gain in \( I_{12} \) shows that the quantum correlations in \( \rho_{12} \) act as an information channel, transferring the information gain from subsystem 1 to subsystem 2. The discord appears because, unless \( A_2 \) is compatible with \( \rho_{12} \), there is a part of the correlations that is unsuitable for the mentioned transfer of the information gain, which is a quasi-classical notion. This is why it is said that the discord quantifies the quantumness of the correlations (regarding \( A_2 \)). Finally, the uncertainty in \( \rho_2 \), i. e., \( S_2 \) is broken up into \( H(p_1) \equiv -\sum_i p_i \log p_1 = S(\rho_2), \) the entropy (or amount of uncertainty) of \( A_2 \) in the state of the second subsystem; into the coherence or incompatibility information \( I_C(\rho_{12}) \), which is again a necessary accompaniment of our "interrogation" due to the quantumness of \( \rho_2 \); and into \( \sum_i p_i S(\rho_{2}^i) \), which is the amount of uncertainty in \( \rho_2 \) inaccessible to \( A_2 \) (residual uncertainty).

It should be noted that (28b) does not describe a process; it only gives a relevant quantitative view of \( \rho_{12} \). In other words, what the quantum correlations in \( \rho_{12} \) do, among other things, is to transfer the information gain \( J_{A_2}(\rho_1) \) from subsystem 1 to subsystem 2. Now it is natural to ask how we can extract it from subsystem 2. Evidently, the thing to do is to measure \( A_2 \) on the nearby subsystem 2, i. e., locally (see section VI). But then one extracts the amount of information \( H(p_1) \), and not \( J_{A_2}(\rho_1) \). This motivates the rest of investigation in this section.
4.2 Information gain $J$

It is the aim of this subsection to understand how the uncertainty $H(p_i) = S(A_2, \rho_2)$ and the information gain $J_{A_2}(\rho_1)$ relate to each other. We begin by a precise understanding of $J_{A_2}(\rho_1)$.

**Proposition 7:** If $\rho = \sum_{l=1}^{m} p_l \rho^l$ is an arbitrary mixture of a finite number of admixed states, then (3a) and (3b) are valid. Besides,

$$0 \leq J(\rho) \leq H(p_i),$$

(29)

and $J(\rho) = 0$ if and only if $\forall l, p_l > 0 : \rho^l = \rho$ (total overlap), and $J(\rho) = H(p_i)$ if and only if $\forall (l \neq l'), p_l > 0 < p_{l'} : \rho^l \rho^{l'} = 0$ (pairwise orthogonality).

**Proof:** The first inequality in (29) is obvious from (3b). The second one is proved in the review article of Wehrl [13] (relation (2.3) there).

One has $J = 0$ if and only if in (3b) (changing what has to be changed) $p_l > 0 \Rightarrow S(\rho^l|\rho) = 0$. It is well known that relative entropy is zero if and only if the two states in it coincide.

It is standard knowledge that the so-called mixing property holds true: if the admixed states $\rho^l$ are pairwise orthogonal, then $J = H(p_i)$. The converse statement, that $J = H(p_i)$ implies orthogonality of the $\rho^l$, is not proved anywhere known to the author of this study. Therefore, its somewhat lengthy proof, through auxiliary lemmas, is given in Appendix B. 

The quantity $H(p_i)$ is called the mixing entropy of the mixture at issue. But it is only the upper possible extreme value of the information gain $J$. It is obvious from proposition 7 that the excess $(H(p_i) - J)$ (or how much is missing in the information gain) quantifies the overlap of the admixed states. It is zero if and only if there is no overlap (the admixed states are orthogonal). It is maximal, i.e., equal to $H(p_i)$, in case of total overlap, when one is dealing only with an apparent mixture.

**Remark 6:** It is desirable to have the extension of proposition 7 to the case of infinitely many admixed states.

To clarify what is apparent and what is genuine in a mixture, we consider two trivial lemmata.

**Lemma 4:** Let us take a mixture

$$\rho = \sum_s p_s \rho^s, \quad S(\rho) = \sum_s p_s S(\rho^s) + J,$$

(30a)

and a refinement of it

$$\forall s, p_s > 0 : \rho^s = \sum_{k_s} w_{k_s} \rho^{k_s}, \quad \rho = \sum_s \sum_{k_s} p_s w_{k_s} \rho^{k_s}.$$  

(30b)

Then the residual entropy is non-increasing, whereas the information gain and the mixing entropy are non-decreasing. More precisely (in obvious notation):

$$S(\rho) = \sum_s \sum_{k_s} p_s w_{k_s} S(\rho^{k_s}) + \{ \sum_s (p_s J^s) + J \},$$

(31a)

$$H(p_s w_{k_s}) = H(p_s) + \sum_s p_s H(w_{k_s}).$$

(31b)
Proof is straightforward.

Lemma 5: If the refinement in a mixture is done through mere repetition, i.e., if \( \forall s, p_s > 0 : k_s \neq k'_s \Rightarrow \rho^{k_s} = \rho^{k'_s} \), then the residual entropy and the information gain remain the same.

Proof is obvious from (31a) if one takes into account that \( \forall s, p_s > 0 : J_s = 0 \).
\[ \square \]

It is now seen that the information gain is insensitive to apparent mixing (or repetition of the admixed states); it depends only on the genuine mixing, i.e., on the distinct admixed states. Contrariwise, the mixing entropy is insensitive to the distinction between genuine and apparent mixing, i.e., it increases whenever at least one of the refined probability distributions is nontrivial. Therefore, in spite of the fact that \( (H(p_l) - J) \) does quantify the overlap in the given mixture, which may contain repetition of admixed states, it can be diminished on the basis of (31b).

4.3 Essential noise and garbled information

Definition 2: If a given mixture \( \rho = \sum_l p_l \rho_l \) is rewritten without repetition of the admixed states with the use of a new index \( s \), the expression \( (H(p_l) - J) \) quantifies the essential overlap in the mixture, i.e., the one due to the genuine mixing of the distinct admixed states. The original quantity of overlap is the sum of the quantity of essential overlap and of that of redundant overlap:
\[
(H(p_l) - J) = (H(p_s) - J) + (H(p_l) - H(p_s)).
\]

One can see in (31b) that \( (H(p_l) - H(p_s)) \) is the increase in the mixing entropy due to repetition of admixed states.

Returning to the elaborate subsystem entropy decomposition (28b), we see that at best we can extract the information gain \( H(p_l) \) from subsystem 2 by measuring the subsystem observable \( A_2 \) (which is simultaneously the measurement of \( (1 \otimes A_2) \) in the bipartite state \( \rho_{12} \)). The difference \( (H(p_l) - J_{A_2}) \), corresponding to the overlap in the distant mixture \( \rho_1 = \sum_l p_l \rho'_l \), appears now as noise. In accordance with definition 2, this noise is the sum of an essential term and a redundant term. One cannot eliminate the former (without changing drastically \( A_2 \), i.e., without taking another subsystem observable that is not a function of \( A_2 \)) because the essential term is due to the overlap of the distinct admixed states in \( \rho_1 \), but one can dispose of the redundant noise by sheer coarsening.

Theorem 3: There exists one and only one coarsening \( B^{ess}_{ess} \) of \( A_2 \) in which the redundant noise is and the essential noise is not eliminated, and the induced distant mixture \( \rho_1 = \sum_s p_s \rho'_s \) is equal to the one obtained due to \( A_2 \) but rewritten with positive weights and without repetitions in the admixed states. To obtain the subsystem observable \( B^{ess}_{ess} \), one defines the following equivalence relation in the detectable spectrum of \( A_2 \): \( l \sim l' \) if \( p'_l = p'_l' \) (cf (2c)). Further, enumerating by \( s \) the obtained equivalence classes \( \{C_s : \forall s\} \), one defines
\[
B^{ess}_{ess} = \sum_s b_s P^s_2,
\]
where \( \{b_s : \forall s\} \) is an arbitrary set of distinct nonzero real numbers, and
\[
\forall s : P^d_s = \sum_{l \in C_s} P^d_{2 l}.
\]
Proof: Since $\forall s: \quad p_s \equiv \text{tr}(\rho_{12} P^s_2) = \left( \sum_{i \in C} p_i \right) > 0$, and

$$\rho^s_1 \equiv p^{-1}_s \text{tr}_2(\rho_{12} P^s_2) = \sum_{i \in C} (p_i/p_s) p^{-1}_i \text{tr}_2(\rho_{12} P^i_2) = \sum_{i \in C} (p_i/p_s) \rho^i_1 = \rho^\bar{l}_1,$$

where $\bar{l}$ is any index value from the class $C_s$. Thus, $B_{2}^{\text{ess}}$ does induce the desired mixture for $\rho_1$. It is evidently the unique coarsening of $A_2$ doing this because every coarsening has to break up the detectable spectrum of $A_2$ into classes, and the desired purpose cannot be achieved in any other way. \qed

In general, the information gain $J$ is garbled because in the measurement of $A_2$ it appears necessarily with (inseparable) essential noise $(H(p_s) - J)$. (For a precise definition of "garbled information gain" see the last but one term in (38) below.)

Needles to say that the expounded procedure of eliminating redundant noise is analogous in the classical discrete case of probability distributions.

### 4.4 Orthogonal distant mixture, pure information gain and twin observables

As it is obvious from proposition 7 and (28b), if the distant mixture $\rho_1 = \sum_{i=1}^{n} p_i \rho^i_1$ is orthogonal, and only in this case, the essential noise is zero. Then, one has pure information: $J_{A_2} = H(p_1) = S(A_2, \rho_2)$. In this case there is no redundant noise either. It may happen that orthogonality is achieved only after disposing of the redundant noise. Therefore, we concentrate on $B^{\text{ess}}_2 = \sum b_s P^s_2$ and the corresponding distant mixture $\rho_2 = \sum_s p_s \rho^s_2$, but to make the results more general, the suffix "ess" is omitted.

Let $Q^s_1$ be the range projector of $\rho^s_1$. Orthogonality of the above mixture amounts to

$$Q^\dagger_1 Q^s_1 = \delta_{s,s^\prime} Q^s_1,$$

and one has

$$\sum_s Q^s_1 = Q_1,$$

where $Q_1$ is the range projector of the distant state $\rho_1$. In this case, we prove the following result.

**Proposition 8:** Assuming positivity of all the probabilities $p_s$ and the validity of $(\sum_s P^s_2) \rho_2 = \rho_2$, if the distant mixture $\rho_1 = \sum_s p_s \rho^s_1$ (cf (2a)-(2c) changing what has to be changed) is orthogonal, then

$$\left( \sum_s P^s_2 \right) \rho_{12} = \rho_{12} = \left( \sum_s Q^s_1 \right) \rho_{12},$$

and

$$\forall s: \quad Q^s_1 \rho_{12} = P^s_2 \rho_{12}$$

are valid.

**Proof:** Let $Q_2$ be the range projector of the nearby state $\rho_2$. The relation $(\sum_s P^s_2) \rho_2 = \rho_2$ then implies $\left( \sum_s P^s_2 \right) Q_2 = Q_2$ (see Appendix A in [12]). Since
one can always write $\rho_{12} = Q_2 \rho_{12}$ (cf relation (12a) in [21]), the first equality in (34a) follows.

The relation (33b) and the fact that one can write $\rho_{12} = Q_1 \rho_{12}$, then make also the second equality in (34a) seen to be valid.

Next, we prove that

$$s \neq s' \implies Q_i^s P_i^{s'} \rho_{12} = 0. \quad (35)$$

For unequal index values one has $\text{tr}(Q_i^s P_i^{s'} \rho_{12}) = p_s \text{tr}(Q_i^s \rho_{12}^{s'}) = p_s \text{tr}(Q_i^s (Q_i^s \rho_{12}^{s'}) = 0$. Further, $0 = \text{tr}(Q_i^s P_i^{s'} \rho_{12}) = \text{tr}((Q_i^s P_i^{s'}) \rho_{12})(Q_i^s P_i^{s'})$, and $(Q_i^s P_i^{s'}) \rho_{12} = (Q_i^s P_i^{s'}) = 0$ is well known to ensue. Then, the Lemma of L"{u}ders ([10] or see FN 16 in [25]) entails the claimed relation (35).

Finally, utilizing relations (34a) and (35), one can argue as follows. $Q_i^s \rho_{12} = Q_i^s \left(\sum_s P_i^{s'}\right) \rho_{12} = Q_i^s P_i^{s'} \rho_{12} = P_i^s \left(\sum_s Q_i^s\right) \rho_{12} = P_i^s \rho_{12}$ as claimed in (34b). $\square$

If one defines a first-subsystem observable $B_1 \equiv \sum p_s Q_i^s$ with arbitrary but distinct nonzero detectable eigenvalues $\{b_s : \forall s\}$, then, according to Theorem 1 in [12] and the theorem on so-called twin observables (p. 05232-1-3 in [20]) imply that proposition 8, actually, gives one more necessary and sufficient condition for $(B_1, B_2)$ to be twin observables in $\rho_{12}$.

Twin observables have a number of remarkable properties (cf also [21] and the references therein). For this study an important implication is that $[B_i, \rho_j] = 0$, $i = 1, 2$ (cf the mentioned Theorem 1 in [12]).

Two obvious consequences on the elaborate subsystem entropy decomposition (28b), which is the basic object of this study, follow:

$$I_C(B_2, \rho_2) = 0 = I_C(B_1, \rho_1), \quad (36a)$$

and, on account of (7), $\delta_{A_2}(\rho_{12}) = I_C(B_2, \rho_{12})$.

Thus, in this case (28b) simplifies to

$$S_{12} = S_1 - I_{12} + S_2 = \left\{\sum_s p_s S(\rho_{12}^s) + H(p_s)\right\} - \left\{H(p_s) + I_C(B_2, \rho_{12}) + \sum_p I(\rho_{12})\right\} + \left\{H(p_s) + \sum p_s S(\rho_{12}^s)\right\},$$

where the mixing property is utilized for the orthogonal mixture $\rho_1 = \sum p_s \rho_{12}^s$.

If $I_C(B_2, \rho_{12}) > 0$, then we have the case of so-called correlations incompatibility (cf Section 6 in [12]), in which the discord equals the coherence or incompatibility information of $B_2$ in $\rho_{12}$. Besides, there is no quantumness in $\rho_2$ with respect to $B_2$. (One has global coherence without local coherence.)

The quantity of uncertainty $S(\rho_{2})$ of the nearby subsystem state now (possibly) exceeds the quantity of uncertainty $S(B_2, \rho_2)$ of the observable $B_2$ in $\rho_2$, which equals the pure information gain $H(p_s) = J_{B_2}$.

The assumption $\left(\sum_s P_i^{s'}\right) \rho_2 = \rho_2$ is satisfied for $B_2^{ess}$ due to the very definition of the indices $s$ (all detectable $l$ values of $A_2$ are used up in it). Besides, on account of the definition of $B_2^{ess}$, all probabilities $p_s$ are positive.
So far in this subsection we had in mind the special case when the distant mixture \( \rho_1 = \sum_s p_s \rho_1^s \) without repetition in the admixed states turns out orthogonal. Now we return to the general case and prove that there always exists a (possibly trivial) unique minimal coarsening \( C_2 = \sum_t c_t P_t^2 \) of \( B_2^{\infty} \), and, by consequence, of \( A_2 \), that gives an orthogonal distant mixture and, by a definition analogous to the above of \( B_1 \), an observable \( C_1 = \overline{\sum t c_t Q_1^t} \) that is its twin observable.

### 4.5 Minimal orthogonal coarsening of a mixture

Before we proceed, we first expound some relevant properties of mixtures as far as orthogonal coarsenings of them are concerned.

**Lemma 6:** For any two states \( \rho \) and \( \rho' \) one has \( \text{tr}(\rho \rho') \geq 0 \), and \( \text{tr}(\rho \rho') = 0 \) if and only if \( \rho \rho' = 0 \).

**Proof:** Always \( \text{tr}(\rho \rho') = \text{tr}(\rho^0 \rho' \rho^{1/2}) \) \( \geq 0 \) because \( \rho^0 \rho' \rho^{1/2} \) is a positive operator. Sufficiency of orthogonality for trace orthogonality is obvious. Necessity is seen as follows: \( \text{tr}(\rho \rho') = 0 \) implies \( \rho^0 \rho' \rho^{1/2} = 0 \), and this has, due to the Lemma of Lüders (\([10]\)), \( 0 = \rho^0 \rho^{1/2} = \rho' \rho \) as its consequence. \( \square \)

**Definition 3:** Let \( \rho = \sum_s w_s \rho_s \) be a mixture with positive weights and without repetitions of the admixed states \( \rho_s \). We say that the states \( \rho_k \) and \( \rho_{k'} \) are *linked* if \( \text{tr}(\rho_k \rho_{k'}) > 0 \). If \( \rho_k \) and \( \rho_{k'} \) are such that there exists an integer \( n, n = 1, 2, \ldots \), and there can be found a chain of admixed states \( \{ \rho_{k_i} : i = 1, 2, \ldots, n \} \) such that \( \rho_k = \rho_{k_1}, \rho_{k_n} = \rho_{k'} \), and any two neighboring states in the chain are linked, then we say that \( \rho_k \) and \( \rho_{k'} \) are *chained*, and we speak of m-chaining.

**Definition 4:** We say that a mixture \( \rho = \sum_s w_s \rho_s \) is a *coarsening* of another mixture \( \rho = \sum_s w_s \rho_s \), the latter being without repetition in the admixed states \( \rho_s \) and with positive weights, if the index set \( \{ s : \forall s \} \) is partitioned into m-classes \( T_i \): \( \{ s : \forall s \} = \sum_t T_t \) (the sum stands for the union of the non-overlapping classes), is enumerated by \( t \), and \( \rho_1 = \sum_{s \in T_t} \left( w_s/w_t \right) \rho_s \), where \( \forall t : w_t \equiv \left( \sum_{s \in T_t} w_s \right) \). In this case we speak of m-coarsening.

**Proposition 9:** Let \( \rho = \sum_s p_s \rho_s \) be a mixture with all weights positive and without repetition. Let, further, another mixture \( \rho = \sum_s w_s \rho_s \) be a coarsening of the former mixture, obtained by *chaining* (chained m-coarsening). Then the latter mixture is orthogonal, and it is minimal as such, i.e., if also \( \rho = \sum_s p_s \rho_s \) is an orthogonal coarsening of the initial mixture, then it is also a coarsening of its chained m-coarsening.

**Proof:** Orthogonality can be proved as follows. Let \( t \neq t' \), and let \( s \in T_t, s' \in T_{t'} \). We assume *ab contrario* that \( \text{tr}(\rho_s \rho_{s'}) > 0 \). Then, according to definitions 4 and 3, \( \rho_s \) and \( \rho_{s'} \) are linked, and hence belong to the same m-class \( T_t \) contrary to assumption. Hence, \( \rho^0 \rho_{s'} = 0 \) (cf lemma 6), implying \( \rho^0 \rho' = 0 \) (cf definition 4).

Minimality is proved in the following way. Let the partitioning \( \{ s : \forall s \} = \sum_u \mathcal{U}_u \) define an orthogonal coarsening \( \rho = \sum_u p_u \rho^u \) in analogy with definition 4. Considering the initial mixture \( \rho = \sum_s p_s \rho^s \), we assume that two distinct index values \( s, s' \) are m-linked (cf definition 3). Lemma 6 claims that \( \rho^s \) and \( \rho^{s'} \) then cannot be orthogonal; hence \( s \) and \( s' \) must belong to one and the same m-class \( \mathcal{U}_u \). Next, let \( s \) and \( s' \) be chained. Then any two neighboring index values in the chain belong to one and the same m-class \( \mathcal{U}_u \), entailing the fact that also \( s \) and \( s' \) belong to the same m-class. Thus, any m-class \( T_t \) is a subset of some m-class \( \mathcal{U}_u \). This means that the \( u \)-mixture is a coarsening of the \( t \)-mixture, and the latter is thus proved to be minimal.
4.6 The pure part of information gain

We return now to our investigation of an arbitrary bipartite state \( \rho_{12} \). We have defined \( B_{2}^{\text{oss}} = \sum b_i P_i^2 \) to eliminate redundant noise.

**Definition 5:** We define \( C_{2}^{tw} \equiv \sum_{t} c_t P_t^2 \) as a coarsening of \( B_{2}^{\text{oss}} \) that induces \( m \)-chaining (cf definitions 4 and 3) of the distant mixture \( \rho_1 = \sum_i p_i \rho_i^1 \) (induced by \( B_{2}^{\text{oss}} \)), and by a spectrum \( \{ c_t : \forall t \} \) consisting of any distinct nonzero real numbers.

If \( \{ Q_t^1 : \forall t \} \) are the range projectors of the distant admixed states \( \rho_i^1 \), then defining, further, \( C_{1}^{tw} \equiv \sum_{t} c_t Q_t^1 \), (the eigenvalues \( c_t \) any distinct and nonzero real numbers), then, according to proposition 8 and the discussion after its proof, one obtains **twin observables** \( (C_{1}^{tw}, C_{2}^{tw}) \).

**Corollary 2:** In case \( C_{2}^{tw} \) is nontrivial, one has two parallel orthogonal mixtures with the common index \( t \), the nearby one \( \rho_2 = \sum_i p_i \rho_i^2 \), and the distant one \( \rho_1 = \sum_i p_i \rho_i^1 \). In general, \( \rho_{12} \) is not a mixture of the global states \( \rho_{12}^2 \equiv P_i^2 \rho_{12} P_i^2/\rho_i^2 \), which give \( \rho_i^2 \) and \( \rho_i^1 \) as their reductions. The global states \( \rho_{12}^i \) are **biorthogonal**, i.e., \( t \neq t' \Rightarrow \rho_i^t \rho_i^{t'} = 0, i = 1, 2 \).

Since \( C_{2}^{tw} \) is a coarsening of \( B_{2}^{\text{oss}} \), the information gain \( J_{C_{2}^{tw}} \) of the former is not larger than that of the latter (see theorem 2), i.e.,

\[
J_{C_{2}^{tw}} \leq J_{B_{2}^{\text{oss}}} = J_{A_{2}} \tag{37a}
\]

and

\[
J_{C_{2}^{tw}} = H(p_1) \leq H(p_t) \leq H(p_i) \tag{37b}
\]

(cf (31b)). One should remember that \( H(p_i) \geq J_{A_{2}} \), due to (possible) essential noise.

Observable coarsening (or \( \omega \)-coarsening) \( \rho_2 \leq \rho_2^{\text{coars}} \) is the opposite relation to (state-dependent) observable refinement (or \( \omega \)-refinement) explained in section II. It is a reflexive and transitive binary relation, i.e., it is a partial order in the set of all observables. One has

\[
C_{2}^{tw} \rho_2^{\text{coars}} \leq B_{2}^{\text{oss}} \rho_2 \leq A_2 \tag{37c}
\]

parallel to (37a) and (37b).

Returning to the elaborate subsystem entropy decomposition (28b), and having the relations (37a) and (37b) in mind, one can write

\[
H(p_t) = S(A_2, \rho_2) = \left\{ H(p_i) - H(p_t) \right\} + \left\{ H(p_t) - J_{B_{2}^{\text{oss}}} \right\} + \left\{ J_{B_{2}^{\text{oss}}} - H(p_i) \right\} + H(p_t). \tag{38}
\]

The **physical interpretation** of (38) goes as follows. The entropy \( S(A_2, \rho_2) \) (quantifying the uncertainty) of the initial subsystem observable \( A_2 \) in the nearby local state \( \rho_2 \) consist of redundant noise \( \{ H(p_i) - H(p_t) \} \), of essential noise \( \{ H(p_t) - J_{B_{2}^{\text{oss}}} \} \), of **garbled information gain** \( \{ J_{B_{2}^{\text{oss}}} - H(p_i) \} \), and, finally of **pure information gain** \( H(p_t) \). Naturally, each of the terms is positive or zero, as the case may be. The latter occurs when the corresponding subsystem observable is trivial, i.e., when it has only one detectable eigenvalue (the probability of which is then, of course, one).
For different choices of $A_2$ one may come to different $C_2^{ss}$. One may wonder if there always exists one $C_2^{ss}$ for all choices of $A_2$, as refined as possible. This is not true in the general case. Taking as an example the well-known singlet pure bipartite state, it is easy to see that for any choice of a nontrivial $A_2$, one has $A_2 = B_2^{ss} = C_2^{ss}$, and one obtains thus infinitely many different $C_2^{ss}$ observables that are all complete.

The case of pure bipartite states deserves a separate discussion (see subsection II). But first we again need some more general theory.

### 4.7 Minimal compatible coarsening of an observable

We begin by some relevant theory on relation between observable and state.

**Definition 6:** We say that two (equal or distinct) index values $t$ and $t'$ of detectable eigenvalues $c_t$ and $c_{t'}$, respectively of a given observable $C = \sum_{t', c_{t'}} P_{t'}$ are linked with respect to a given state $\rho$ if $P_t \rho P_{t'} \neq 0$. When $t$ and $t'$ are such that there exists an integer $n$, $n = 1$, or $2$, or $\ldots$, and there can be found a chain of index values \( \{t_i : i = 1, 2, \ldots, n\} \) such that $t = t_1$, $t_n = t'$, and any two neighboring index values in the chain are linked, then we say that $t$ and $t'$ are chained. Occasionally, when it is desirable to make a distinction with respect to $m$-linking and $m$-chaining, we shall speak of $o$-linking (short for observable-linking) and $o$-chaining.

O-chaining includes $o$-linking, and it is reflexive, symmetric and transitive, i.e., an equivalence relation in the detectable part of the spectrum of the observable $C$.

**Definition 7:** Let $C = \sum c_t P_t$ be a given observable, and let the observable $D$ be a coarsening of $C$ defined by means of chaining in the detectable part of the spectrum of the latter with respect to a given state $\rho$ (cf definition 6):

$$D \equiv \sum_k d_k P_k, \quad \forall k : P_k \equiv \sum_{t \in C_k} P_t,$$

where

$$\{t : \forall t, \rho P_t \equiv \mathrm{tr}(\rho P_t) > 0\} = \sum_k C_k\quad (39a)$$

partitions the detectable part of the spectrum of $C$ into the equivalence classes $C_k$ obtained by $o$-chaining, and the eigenvalues of $D$ are arbitrary distinct nonzero real numbers. We call $D$ the chained coarsening of $C$ with respect to $\rho$.

**Proposition 10:** The chained coarsening $D \preceq C$ given in definition 7 is compatible with $\rho$, $[D, \rho] = 0$. It is the most refined coarsening of $C = \sum c_t P_t$ compatible with $\rho$, i.e., if $C \equiv \sum c_{ij} P_{ij}$, is a coarsening of $C$, $\forall j : P_{ij} \equiv \sum_{t \in C_j} P_t$ with arbitrary distinct nonzero real eigenvalues of $C$ and $\{t : \forall t, \rho P_t > 0\} = \sum_j C_j$ a partitioning of the detectable part of the spectrum of $C$, such that $\tilde{C}$ is compatible with $\rho$, $\tilde{C} \preceq C$, $[\tilde{C}, \rho] = 0$, then it is also a coarsening of $D$: $\tilde{C} \preceq D$.

**Proof:** Let $k, k'$ be two distinct index values of $D$, and let us keep in mind that $P_k \rho P_{k'} = \left( \sum_{t \in C_k} P_t \right) \rho \left( \sum_{t' \in C_{k'}} P_{t'} \right)$. Since $t$ and $t'$ are not $o$-chained by assumption, they are not $o$-linked either. Hence, each term is zero $P_t \rho P_{t'} = 0$ (cf definition 6), implying $P_k \rho P_{k'} = 0$. Thus, one can write $\rho = \left( \sum_k P_k \right) \rho \left( \sum_{k'} P_{k'} \right) = \sum_k (P_k \rho P_k)$, i.e., $\rho$ is compatible with each eigenprojector of $D$, hence also with $D$ itself.

---

Page 19
Let $\tilde{C}$ be an arbitrary coarsening of $C$ compatible with $\rho$ as given in the theorem. Let $t \in C_j$, $t' \in \tilde{C}_j$. Since $j \neq j' \Rightarrow P_j \rho P_{j'} = 0$, due to $[\tilde{C}, \rho] = 0$, multiplication from the left by $P_t$ and from the right by $P_{t'}$ gives $P_t \rho P_{t'} = 0$, i.e., $t, t'$ are not o-linked. Equivalently, if $t, t'$ are o-linked, then $j = j'$. In other words, o-linked index values, and hence also o-chained index values, belong to one and the same equivalence class $\tilde{C}_j$. Thus, $\forall k : \exists j : C_k \subseteq \tilde{C}_j$. This proves $\tilde{C} \subseteq D$. □

4.8 Twin observables for pure bipartite states

As it was stated, if $(C_1, C_2)$ are twin observables in any bipartite state $\rho_{12}$, we have the necessary local condition of compatibility of observable and state $[C_2, \rho_2] = 0$ (and symmetrically, $[C_1, \rho_1] = 0$). If one deals with a pure state $\rho_{12} = |\Phi\rangle_1 |\Phi\rangle_2$, then this condition is also sufficient [26], [27].

If the bipartite state is pure, then for finding $C^{tw}_{22}$ for a given local observable $B^{ess}_{22}$ one need not resort to the distant mixture $\rho_1 = \sum_s \rho_s \rho_s^{t'}$.

**Theorem 4:** If the bipartite state is pure, and a nearby local observable $A_2 = \sum_i a_i P_i^2$ is given, then the observable $C^{tw}_{22}$ (cf subsection F) is the chained coarsening of $A_2$ with respect to $\rho_2$ (cf definitions 6 and 7).

**Proof:** Proposition 10 makes it clear that those and only those coarsenings of a given observable with respect to a given state are compatible with the state that are coarsenings of the chained coarsening of the observable. Naturally, the latter is the most refined one. That is precisely what $C^{tw}_{22}$ is regarding $A_2$.

At first glance one might wonder why is $C^{tw}_{22}$ not the chained coarsening of $A_2$ with respect to $\rho_2$ for a general state $\rho_{12}$. The answer lies, of course, in the fact that there may exist coarsenings $C_2 = \sum_t p_t P_t^2$ of $A_2$ that are compatible with $\rho_2$ and that do not induce orthogonal distant mixtures $\rho_1 = \sum_t p_t \rho_t^{t'}$. This is so because compatibility of local observable with local state is in general, in contrast to the special, pure-state case, not sufficient for twin observables. Note that orthogonality of both mixtures $\rho_i = \sum_t p_t \rho_t^{t'}$, $i = 1, 2$ is a characteristic property of twin observables (cf the "measurement-theoretic condition" $\forall t : P_t^2 \rho_{12} P_t^2 = P_t^2 \rho_{12} P_t^2$, $i \neq i'$, $i, i' = 1, 2$ defining twin observables in [20]).

Let a subsystem observable be complete $A_2 = \sum_i a_i |l\rangle_2 \langle l|_2$, and let

$|\Phi\rangle_{12} = \sum_i a_i |l\rangle_2 \langle l|_2 \otimes |l\rangle_2 \tag{40b}$

be a (generalized) expansion of $|\Phi\rangle_{12}$ in the eigenbasis $\{|l\rangle_2 : \forall l\}$ of the subsystem observable $A_2$, omitting undetectable index values, and the $|l\rangle_1$ being unit vectors. (It is non-unique because the phase factors of the $|l\rangle_1$ vectors are not specified.) Then (40b) implies that

$\forall l : p_l = |a_l|^2, \quad \rho_1^{l} = |l\rangle_1 \langle l|_1$

in (40a) as easily seen.

Since for a pure state $|\Phi\rangle_{12}$ compatibility of local observable and local state is necessary and sufficient for possessing an (opposite-subsystem) twin observable, one can distinguish two kinds of choices for $A_2$. In the first, $A_2$ is incompatible with $\rho_2$. Then, if there exists a nontrivial orthogonal decomposition of the range of $\rho_2$ that is invariant both for $A_2$ and for $\rho_2$, then there exists a finest of this kind defining a nontrivial $C^{tw}_{22}$. In the second choice, $A_2$ is compatible with $\rho_2$. Then $B^{ess}_{22} =$
The string of inequalities (37c) can be continued in the general case. Namely, $I_C(C_{tw}^w, \rho_2) = 0$. Thus, $\rho_2$ has no quantumness with respect to $C_{tw}^w$, but, in general, there still is quantumness in the mutual information in the form of discord.

**4.9 Quasi-classical correlations**

Let us now return to the general case of a bipartite state $\rho_{12}$, and the initial subsystem observable $A_2 = \sum a_i P_{2i}$.

**Definition 8**: Let the observable $D_2^{qc} = \sum_{i} d_i P_i$ be the chained coarsening of $C_{tw}^w$ with respect to $\rho_{12}$ (with distinct nonzero real eigenvalues).

**Theorem 5**: The following strings of inequalities are valid:

$$D_2^{qc} \leq C_{tw}^w \leq B_2^{cas} \leq A_2.$$  \hspace{1cm} (41a)

$$J_{D_2^{qc}} \leq J_{C_{tw}^w} \leq J_{B_2^{cas}} = J_{A_2},$$  \hspace{1cm} (41b)

$$J_{D_2^{qc}} = H(p) \leq J_{C_{tw}^w} = H(p) \leq H(p) \leq H(p) \leq H(p).$$  \hspace{1cm} (41c)

**Proof**: The first "inequality" in (41a) follows directly from definition 8, and then, due to transitivity, the rest of them are consequences. The rest of inequalities in theorem 5 are implied by theorem 2 and (31b).

One can write

$$H(p) = S(A_2, \rho_2) = \{H(p) - H(p)\} + \{H(p) - J_{B_2^{cas}}\} + \{J_{B_2^{cas}} - H(p)\} + \{H(p) - H(p)\} + H(p).$$  \hspace{1cm} (42)

The last two terms in (42) are possibly positive pure information gain. It consists of a purely quantum term $\{H(p) - H(p)\}$, and a quasi-classical term $H(p)$.

**Corollary 3**: On account of the compatibility $[D_2^{qc}, \rho_{12}] = 0$, the subsystem observable $D_2^{qc}$, if non-trivial, induces not only the subsystem mixtures $\rho_i = \sum_k p_k \rho_{12}^k$, $i = 1, 2$, but also the global mixture $\rho_{12} = \sum_k p_k \rho_{12}^k$.

$$\rho_{12} = \sum_k p_k \rho_{12}^k.$$  \hspace{1cm} (43)

where $\forall k : \rho_{12}^k \equiv P_{2k} \rho_{12}/p_k$. The mixture is biorthogonal, i.e., $k \neq k' \Rightarrow p^k_i p^{k'}_i = 0$, $i = 1, 2.$

21
Proof: The claimed compatibility of $D^{qc}_2$ and $\rho_{12}$ follows from definition 8 and proposition 10. It implies compatibility of $D^{qc}_2$ and $\rho_2$. In general, when $A_2 = \sum_i a_i P^i_2$ is given, the corresponding nearby mixture is $\rho^i_2(A_2) \equiv \sum_i P^i_2 \rho P^i_2$, which, in this case amounts to the orthogonal mixture $\rho_2 = \sum_k p_k (P^k_2 \rho P^k_2)$. Orthogonality of the distant mixture is implied by the first "inequality" in (41a): any coarsening of an orthogonal mixture is orthogonal. The global mixture (43) itself is, of course, due to $[D^{qc}_2, \rho_{12}] = 0$.

When one writes down the elaborate subsystem entropy decomposition (28b) for $D^{qc}_2$ (changing what has to be changed), then $0 = I_C(D^{qc}_2, \rho_{12}) = I_C(D^{qc}_2, \rho_2) = \delta_{D^{qc}_2}$. Thus, if we "interrogate" $\rho_{12}$ by $D^{qc}_2$, then quantumness has disappeared not only in the nearby subsystem (analogously as due to $C^{tw}_2$), but also in the global system, i.e., it has disappeared completely. For this reason we call this last step "quasi-classical".

4.10 Measures of quantumness
We now complete inequalities (41b), (41c), and equality (42) by analogous relations for the measures of quantumness.

Theorem 6: A) The following strings of inequalities parallel (41a):

$$I_C(D^{qc}_2, \rho_{12}) \leq I_C(C^{tw}_2, \rho_{12}) \leq I_C(D^{ess}_2, \rho_{12}) \leq I_C(A_2, \rho_{12});$$

$$I_C(D^{qc}_2, \rho_2) \leq I_C(C^{tw}_2, \rho_2) \leq I_C(D^{ess}_2, \rho_2) \leq I_C(A_2, \rho_2);$$

$$\delta_{D^{qc}_2} \leq \delta_{C^{tw}_2} \leq \delta_{D^{ess}_2} \leq \delta_{A_2}.$$

B) The coherence informations satisfy also the straight-line relations:

$$I_C(A_2, \rho_i) = I_C(D^{qc}_2, \rho_i) + I_C(C^{tw}_2, \sum_k P^i_2 \rho P^i_2) +$$

$$I_C(D^{ess}_2, \sum_i \sum_k P^i_2 P^k_2 \rho_i \rho P^k_2 P^i_2) +$$

$$I_C(A_2, \sum_i \sum_k P^i_2 P^k_2 \rho_i \rho P^k_2 P^i_2), \quad i = 2, 12.$$

Proof: A) The first two strings of inequalities are an immediate consequence of the inequality proved in previous work [12] (Theorem 3 there, "Ec" is written instead of "Ic"). The third string of inequalities is an immediate consequence of theorem 2.

B) The straight-line relations are an immediate implication of Corollary 2 in [2].

5 Examples

5.1 Pure states

Example 1: Let $\{|i\rangle_2 : i = 1, 2, 3\}$ be an orthonormal set in the state space of the nearby subsystem, and let $\{|j\rangle_1 : j = 1, 2\}$ be an orthonormal set in that of the distant subsystem. We define

$$|\Phi\rangle_{12} \equiv \alpha_1 |j = 1\rangle_1 |i = 1\rangle_2 + \alpha_2 |j = 1\rangle_1 |i = 2\rangle_2 + \alpha_3 |j = 2\rangle_1 \otimes |i = 3\rangle_2, \quad (44a)$$
where, of course, \( \sum_{q=1}^{3} |\alpha_q|^2 = 1 \) is valid.

If \( \{ |i\rangle_2 : i = 1, 2, 3 \} \) is a subset of the eigenbasis of a complete subsystem observable \( A_2 \), then the corresponding distant state decomposition is

\[
p_1 = |\alpha_1|^2 |j = 1\rangle_1 \langle j = 1| + |\alpha_2|^2 |j = 1\rangle_1 \langle j = 2| + |\alpha_3|^2 |j = 2\rangle_1 \langle j = 2|,
\]

(cf (40b) and the next relation). It contains repetition in the admixed states. This makes the subsystem observables \( B_2^{\text{ss}} = C_2^{\text{sw}} \equiv b_1 P_2^{r=1} + b_2 |s = 2\rangle_2 \langle s = 2|_2 \), where \( P_2^{r=1} \equiv |i = 1\rangle_2 \langle i = 1| + |i = 2\rangle_2 \langle i = 2| \) and \( |s = 2\rangle_2 \equiv |i = 3\rangle_2 \), nontrivial. The corresponding pure information gain is

\[
J_{B_2^{\text{ss}}} = H(p_{r=1}, p_{s=2}) = S(B_2^{\text{ss}}, \rho_2) = S(C_2^{\text{sw}}, \rho_2) = -(|\alpha_1|^2 + |\alpha_2|^2) \log(|\alpha_1|^2 + |\alpha_2|^2) - |\alpha_3|^2 \log|\alpha_3|^2.
\]

A Schmidt expansion of \(|\Phi\rangle_{12}\) is

\[
|\Phi\rangle_{12} = r_1^{1/2} |j = 1\rangle_1 |r_1\rangle_2 + |\alpha_3| (e^{i\lambda_3} |j = 2\rangle_1) |i = 3\rangle_2,
\]

where \( r_1 = |\alpha_1|^2 + |\alpha_2|^2 \), \( e^{i\lambda_3} \) is the phase factor of \( \alpha_3 \), and

\[
|r_1\rangle_2 \equiv (\alpha_1 |i = 1\rangle + \alpha_2 |i = 2\rangle)/r_1^{1/2}.
\]

**Example 2:** We assume that all positive eigenvalues of \( \rho_2 \) of \(|\Psi\rangle_{12}\) are non-degenerate. Let \( \{ |q\rangle_2 : \forall q \} \) be the unique (up to phase factors) eigen-sub-basis of \( \rho_2 \) corresponding to its positive eigenvalues \( \{ r_q > 0 : \forall q \} \). Finally, let

\[
|\Psi\rangle_{12} = \sum_q r_q^{1/2} |q\rangle_1 \otimes |q\rangle_2
\]

be a Schmidt expansion of \(|\Psi\rangle_{12}\).

Let \( A_2^q = \sum_l |l\rangle_2 \langle l|_2 \) be a complete observable. Then the induced nearby mixture is

\[
\rho_2^q (A_2^q) \equiv \sum_l P_2^l \rho_2 P_2^l = \sum_l p_l |l\rangle_2 \langle l|_2
\]

Let the eigenbasis of \( A_2^q \) contain \( \{ |q\rangle_2 : \forall q \} \) as a subset. Then, as easily seen from theorem 4, \( B_2^{\text{ss}} = \sum_q b_q |q\rangle_2 \langle q|_2 \), \( C_2^{\text{sw}} = \sum_q c_q |q\rangle_2 \langle q|_2 \), \( Q_2 A_2^q = \sum_q a_q |q\rangle_2 \langle q|_2 \), and \( D_2^{\text{sw}} = Q_2 \), where \( Q_2 \) is the range projector of \( \rho_2 \). The observable \( D_2^{\text{sw}} \) is trivial because a pure state cannot be written as a nontrivial mixture - cf corollary 3.

Let us now take another complete observable \( A_2^q \) as follows. Let \( |l = 1\rangle_2 \equiv |q = 1\rangle_2 \) from above. Further, let \( \langle q\rangle_2 \langle q|_2 |l\rangle_2 \equiv 0 \) unless \( q = 1 \), and at least one of the index values \( l, l' \) is not equal to 1, when it is zero. Then, as easily seen, theorem 4 implies that \( C_2^{\text{sw}} = c_1 |q = l = 1\rangle_2 \langle q = l = 1|_2 + c_2 P_2^{r=2} \), where \( P_2^{r=2} \equiv \sum_{q \geq 2} |q\rangle_2 \langle q|_2 \).

**5.2 Mixed states**

**Example 3:** We assume that all vectors \(|q\rangle_1\) in (45a) are orthogonal to all \(|j\rangle_1\) in (44a), and symmetrically, that all \(|q\rangle_2\) in (45) are orthogonal to all \(|i\rangle_2\) in (44a). Then we take a mixture of the bipartite pure state vectors given by (44b) and (45a): \( \rho_{12} \equiv (1/2) |\Phi\rangle_{12} \langle \Phi|_{12} + (1/2) |\Psi\rangle_{12} \langle \Psi|_{12} \).

Further, we define \( A_2 \) to be complete and such that its eigenbasis contains all the mentioned orthonormal vectors for the nearby subsystem as subsets. Then all four
observables in (41a) are nontrivial: \( D^{\text{gc}}_2 = d_{k=1} P^{k=1}_2 + d_2 P^{k=2}_2 \), where \( P^{k=1}_2 \equiv \sum_{i=1}^3 |i\rangle_2 \langle i|_2 \), and \( P^{k=2}_2 \equiv \sum_q |q\rangle_2 \langle q|_2 \):

\[
C^{tw}_2 = B^{ass}_2 \equiv c_{t=1} |r_1\rangle_2 \langle r_1|_2 + c_{t=2} |i = 3\rangle_2 \langle i = 3|_2 + \\
\sum_q c_{t=q+3} |t = q+3\rangle_2 \langle t = q+3|_2
\]

(cf (44c)).

6 Has the Discord Disappeared in Measurement?

As it was stated, to extract the information gain \( J_{A_2} \) from \( \rho_{12} \), one measures \( A_2 \) locally on the nearby subsystem, and, by this very fact also \( 1 \otimes A_2 \) globally in \( \rho_{12} \). In general, one thus obtains \( S(A_2, \rho_2) = H(\rho_2) \), in which to \( J_{A_2} \) is inseparably added both the essential and the redundant noise, and \( J_{A_2} \) necessarily contains garbled information gain in the general case. If it contains a positive amount of pure information, this, in turn, consists of a quantum and a quasi-classical term.

As far as quantities are concerned, the results of the preceding section allow one to evaluate how much of each of the mentioned entities is contained in \( H(\rho_t) \). But qualitatively, when one deals with an ensemble \( \rho_{12} \) of individual bipartite systems in the laboratory, on each of which \( A_2 \) is measured, at first glance, one can do nothing in the way of separation of these entities. One can, of course, measure locally \( B^{tw}_2 \) (or \( C^{tw}_2 \)) instead of \( A_2 \). Actually, if the laboratory ensemble is sufficiently large, the thing to do is to measure the mentioned observables on subensembles, which, if randomly taken, also represent empirically the same bipartite state \( \rho_{12} \).

The simplest way to measure \( A_2 \) is the so-called ideal measurement, which, by definition, changes a state \( \rho \) into its Lüders mixture \( \rho_L \) [10]. Then relations (5) and (4a) imply \( I(\rho^I_{12}) = I(\rho_{12}) - \delta_{A_2} \). The discord has disappeared from the bipartite state. Hence the title of this section.

In addition to this disappearance, one has the following known fact.

Lemma 7: If \( \rho^f_{12} \equiv (U_1 \otimes U_2) \rho_{12} (U_1 \otimes U_2)^\dagger \), where \( \rho_{12} \) is an arbitrary bipartite state, \( U_i, \ i = 1, 2 \) are any unitary subsystem operators, and the suffix \( f \) denotes "final", then

\[
I(\rho^f_{12}) = I(\rho_{12}).
\]

Putting it in words, in any bipartite state, when it is dynamically closed and the two subsystems do not interact, the mutual information does not change.

Proof is straightforward.

To apply Lemma 7 to the case of ideal measurement of \( A_2 \) in \( \rho_{12} \), let the instrument that performs a measurement of the observable be subsystem 3. Subsystems 1 and \( (2 + 3) \) do not interact during the subsystem measurement, and the tripartite system is dynamically closed. Hence, according to Lemma 7, the mutual information between subsystems 1 and \( (2 + 3) \) does not change. Writing \( \rho^I_{i,23} \) for the state \( \rho^I_{123} \) of the bipartite system \( 1 + (2 + 3) \), we have

\[
I(\rho^i_{1,23}) = I(\rho^f_{1,23})
\]

(the suffix "i" denotes "initial"). Further, strong subadditivity of entropy requires that \( I_{12} \leq I_{1,23} \) be always valid (cf relation (7) in [14]). On the other hand, the initial state \( \rho^i_3 \) of the measuring apparatus is uncorrelated with the \( (1 + 2) \) system at the
beginning of measurement, i.e., $\rho_{123}^f \equiv \rho_{12} \otimes \rho_3^f$. Hence we have a case of equality in the strong subadditivity of entropy inequality:

$$I(\rho_{12}) = I(\rho_{1,23}^f)$$

[14] (see relation (8) there). Altogether,

$$I(\rho_{12}) = I(\rho_{1,23}^f). \quad (46a)$$

Thus, the amount of mutual information between subsystems 1 and 2 at the beginning of measurement is preserved as the amount of mutual information between subsystems 1 and $(2 + 3)$ at the end of ideal measurement.

Strong subadditivity of entropy requires $I(\rho_{12}^f) \leq I(\rho_{1,23}^f)$, where $\rho_{12}^f = \text{tr}_3 \rho_{123}^f$, and, in case of ideal measurement, it is seen from (4a) and (5) that we now have a proper inequality: $I(\rho_{12}^f) < I(\rho_{1,23}^f) = I(\rho_{12})$ in the general case.

The final mutual information $I(\rho_{1,23}^f)$ can be decomposed according to (4a) (changing what has to be changed) with respect to the same observable $A_2$:

$$I(\rho_{1,23}^f) = \sum_i \left( p_i^f \left(S(\rho_{1,23}^f) | \rho_i^f \right) - I_C(A_2, \rho_{1,23}^f) \right) + \sum_i \left( p_i^f I(\rho_{1,23}^f) \right), \quad (46b)$$

where the suffix $f$ denotes that the quantity is derived from the final state $\rho_{123}^f$, and $l$ stems from the eigenprojector $P_l^f$ of $A_2$. In particular, $\forall l: p_l^f \equiv \text{tr}(\rho_{123}^f P_l^f)$;

$$\forall l, p_l^f > 0: \quad \rho_{12}^{fl} \equiv \text{tr}_{23}(\rho_{123}^f P_l^f/P_l^f) = \text{tr}_{23}(P_l^f \rho_{123}^f P_l^f/P_l^f);$$

e etc.

To find out how each of the three terms changes from $\rho_{12}$ to $\rho_{1,23}^f$, i.e., from (4a) to (46b), we define the simplest measuring apparatus for ideal measurement:

The initial state of subsystem 3 is pure $\rho_3^f \equiv \langle \phi | \rangle_3 \langle \phi |_3$; the ”pointer observable” is a complete one $A_3 = \sum l | l \rangle_3 \langle l |$ (spectral form in terms of distinct eigenvalues - ”pointer positions”); finally, the interaction evolution goes as follows

$$\rho_{123}^f = U_{23}(\rho_{12} \otimes | \phi \rangle_3 \langle \phi |_3) U_{23}^\dagger, \quad (47a)$$

and it is such that

$$\forall l: \quad \text{tr}(\rho_{123}^f | l \rangle_3 \langle l |_3) = p_l \quad (47b)$$

(cf (2a)), and

$$\forall l, p_l > 0: \quad p_l^{-1} \text{tr}_3(\rho_{123}^f | l \rangle_3 \langle l |) = \rho_{12}^l \quad (47c)$$

(cf (2b)).

**Theorem 7:** Comparing (4a) and (46b), all corresponding quantities on the RHSs are equal. More precisely,

$$\forall l: \quad p_l = p_l^f, \quad (48a)$$

$$\forall l, p_l > 0: \quad \rho_l = \rho_l^{fl}, \quad (48b)$$

$$I_C(A_2, \rho_{12}) = I_C(A_2, \rho_{123}^f), \quad (48d)$$

$$I_C(A_2, \rho_2) = I_C(A_2, \rho_{23}^f), \quad (48c)$$

where $\rho_{23}^f = \text{tr}_1 \rho_{123}^f$, and finally

$$\forall l, p_l > 0: \quad I(\rho_{12}^f) = I(\rho_{1,23}^f). \quad (48f)$$
The theorem is proved in Appendix C.

For further use, we establish that $A_2$ and $A_3$ are twin observables in relation to $\rho_{12,3}$. Subsystem 3 is viewed as the nearby one, and the bipartite system $(1 + 2)$ as the distant one.

As it was shown in subsection IV.D and proposition 8, it is sufficient to point out that $\sum_l |l\rangle_3 \langle l| = 1$, and $\sum_l |l\rangle_3 \langle l| \rho_{12}^l = \rho_3^l$ (the undetectable $l$ values are omitted). It follows from (47c) that $\sum_l p_l \rho_{12}^l = \rho_{12}^l \equiv \text{tr}_3 \rho_{12,3}$, with all weights positive and the admixed states $\rho_{12}^l$ orthogonal because $\rho_{12}^l = P_3^l \rho_{12}^l P_2^l$ (cf. (28b)). Hence, according to proposition 8, $A_2$ and $A_3$ are twin observables in $\rho_{12,3}$.

It is also of interest to consider a relevant subsystem entropy decomposition for $\rho_{12,3}$, i.e., for the bipartite system $(1 + 2) + 3$ in the state $\rho_{123}$.

**Theorem 8:** The following subsystem entropy decomposition in terms of $A_2$ and $\rho_{12}$ entities is valid:

$$S(\rho_{12}^l) = S(\rho_{12}^l) - I(\rho_{12}^l, \rho_3) + S(\rho_3) =$$

$$\left\{S(\rho_{12}) + I_C(A_2, \rho_{12})\right\} - \left\{I_C(A_2, \rho_{12}) + H(p_l)\right\} + \left\{H(p_l)\right\}. \quad (49)$$

(It is understood that each large-brackets expression equals the corresponding entity in the preceding decomposition.)

**Proof:** Relation (47c) implies

$$\rho_{12}^l \equiv \text{tr}_3 (\rho_{123}^l) = \sum_l \left(\text{tr}_3 (\rho_{123}^l | l\rangle_3 \langle l|_3)\right) = \sum_l p_l \rho_{12}^l.$$

Hence,

$$\rho_{12}^l = \rho_{12}^l. \quad (50)$$

i.e., it is the Lüders mixture of the initial state $\rho_{12}$ with respect to $A_2$. Further, definition (1a) gives $I_C(A_2, \rho_{12}) \equiv S(\rho_{12}^l) - S(\rho_{12})$. Thus, the first large-brackets expression follows.

Next we prove the third large-brackets expression. The pointer observable $A_3$, being a twin observable, is necessarily compatible with $\rho_{12}^l$. Since it is also complete (by definition), its entropy coincides with the entropy of $\rho_{12}^l$: $S(A_3, \rho_{12}^l) = H(p_l) = S(\rho_{12}^l)$.

Finally, in view of the fact that the system $(1 + 2 + 3)$ is dynamically closed (isolated) during the measurement interaction, the total entropy is preserved: $S(\rho_{123}^l) = S(\rho_{12} \otimes \phi_3 (\phi|_3) = S(\rho_{12})$. The second large-brackets expression follows from this. □

We have to clarify how theorem 8 relates to the proved disappearance of the discord $\delta_{A_2}(\rho_{12}) = I_C(A_2, \rho_{12}) - I_C(A_2, \rho_2)$ (cf. (7)) in the measurement interaction. Since $\delta_{A_2}(\rho_{12})$ is a term in the mutual information (cf. (28b)), at first glance one would expect that $S(\rho_{12})$ increases by $\delta_{A_2}(\rho_{12})$ when $\rho_{12}$ goes over into $\rho_{12}^l$. But this is not so because, as seen in (28b), $I_C(A_2, \rho_2)$ actually cancels out in $S(\rho_{12})$. In $\rho_{12,3}$ the measured observable $A_2$ and the pointer observable $A_3$ are twin observables, and, as a consequence (cf. (36a)), one has compatibility $[A_2, \rho_{12}^l] = 0$, $[A_3, \rho_{12}^l] = 0$. Therefore, we can forget about the quantunness of $A_2$ in relation to $\rho_{12}^l$, and we do cancel $I_C(A_2, \rho_2)$ in $I(\rho_{12})$ and $S(\rho_2)$ in (28b). Thus, the increase in $S(\rho_{12})$ is $I_C(A_2, \rho_{12})$ (cf. (7)) in accordance with (49).
Let us write down next (28b) for $S(\rho'_{12,3})$ with respect to $A_3$:

$$S(\rho'_{12,3}) = S(\rho'_{12}) - I(\rho'_{12,3}) + S(\rho'_{1}) = \left\{ \sum_l p_l S(\rho'_{12}) + H(p_l) \right\} - \left\{ H(p_l) + I_C(A_3, \rho'_{123}) \right\} + \left\{ H(p_l) \right\}. \quad (51)$$

(One should note that $I_C(A_3, \rho'_{1}) = 0$, and that on account of $A_3$ being complete, the respective residual terms in $I_{12,3}$ and $S(\rho'_{1})$ are zero.)

Comparing (49) and (51), one infers that

$$I_C(A_2, \rho_{12}) = I_C(A_3, \rho'_{123}) = I_C(A_2, \rho'_{123}).$$

The last equality is a general property of twin observables: they have the same coherence information in the bipartite state as follows from (in our case) $\forall l : P_l^2 \rho'_{123} = |l\rangle_3 \langle l|_3 \rho'_{123}$, (cf (34b)), which is one of the equivalent definitions of twin observables [20].

Thus, we have proved

**Theorem 9:** It is not the discord $\delta_{A_2}(\rho_{12}) = I_C(A_2, \rho_{12}) - I_C(A_2, \rho_{12})$ (cf (7)), but only the non-negative global term in it that is preserved in the measurement interaction:

$$I_C(A_2, \rho_{12}) = I_C(A_2, \rho'_{123}). \quad (52)$$

**Corollary 4:** One has $A_3 = C_{1+2}^{tw}$ with respect to $\rho'_{123}$, i.e., $S(A_3, \rho'_{1}) = H(p_l)$ is pure information on the distant mixture $\rho'_{12} = \sum_l p_l \rho'_{12}$ (cf (50)), which is orthogonal.

This pure information is not the information at issue. The subject of our investigation is $J_{A_2}(\rho_{12})$, the information gain in the distant mixture $\rho_1 = \sum_l p_l \rho_1^l$ induced by $A_2$ in $\rho_{12}$. Thus, one should view $H(p_l)$ decomposed according to (42), which shows that it consists of a redundant-noise term $\left( H(p_l) - H(p_k) \right)$, an essential-noise term $\left( H(p_k) - J_{B_2^{ess}} \right)$, a term $\left(J_{B_2^{ess}} - H(p_l)\right)$ of garbled information (due to the overlap in the admixed states $\rho_1^l$), of a term $\left( H(p_l) - H(p_k) \right)$ of pure quantum information, and, finally, of a term $H(p_k)$ of pure quasi-classical information. Naturally, any of these terms can be zero.

The measurement interaction, or pre-measurement as it is called in the thorough measurement theory [28], is not the final step in measurement. It is collapse, objectification or reduction (cf also [11], which makes ample use of [29]), which turns $\rho'_{123}$ into the Lüders mixture

$$\sum_l p_l \left( \rho_{12}^l \otimes |l\rangle_3 \langle l|_3 \right) \quad (53)$$

of $\rho'_{123}$ with respect to $A_3$. The admixed Lüders states $\left( \rho_{12}^l \otimes |l\rangle_3 \langle l|_3 \right)$ correspond to the individual results $a_l$ of $A_2$ revealed by the pointer position $|l\rangle_3 \langle l|_3$ of the pointer observable $A_3$. Incidentally, the state (53) is a quasi-classical mixture, well known in laboratory measurements.

Both in the final state of premasurement $\rho'_{123}$ and in the final state of measurement given by (53) there are, in general, correlations in the subsystem $(1+3)$ though 1 and 3 have not interacted. Thus, subsystem 1 has simultaneous correlations with 2 and with 3, and so-called monogamy [35], expressing mutual restrictions in the two mentioned correlations, enters the scene. Koashi and Winter have quantified monogamy [35]. In one of their inequalities appears, as a measure of correlations, the
so-called entanglement of formation expressing the least expected entanglement of any ensemble of pure states realizing a given bipartite state [36]. Their inequality (6) can easily be rearranged to the effect that for (1+3) the entanglement of formation cannot exceed the minimal residual entropy

\[
\inf_{\{B_k\}} \sum_k p_k S(\rho^1_k)
\]

(cf (6) in this article with \(k\) instead of \(l\)). The latter quantity applies to subsystem 1 when all imaginable choices of the observable \(B_2 = \sum b_k P_k\) (all \(b_k\) distinct eigenvalues) with a view to be measured in the state \(\rho^1_{12} = \sum p_l \rho^1_{l12}\) of subsystem \((1 + 2)\) are taken into account. (Note that this is the common reduced state of \(\rho^1_{123}\) and of the state (53).)

7 Summing Up

The investigation reported in this article is restricted to von Neumann entropy, and von Neumann mutual information defined by the subsystem entropy decomposition

\[ S(\rho_{12}) = S(\rho_1) - I(\rho_{12}) + S(\rho_2), \]

where \(\rho_{12}\) is an arbitrary bipartite state, and \(\rho_i, i = 1, 2\) are its reductions. The approach is based on the use of coherence or incompatibility information \(I_C(A_2, \rho_i), i = 2, 12\) (cf (1a)-(1c)), which quantifies the quantumness in the relation of an observable and state.

Zurek’s idea of “interrogating” the quantum correlations of the composite state \(\rho_{12}\) through the choice of a local observable \(A_2\) is elaborated via the mentioned subsystem decomposition of entropy.

The first result (theorem 1 and (4a)) has introduced coherence information into mutual information through one of three relevant non-negative terms. It is Zurek’s discord [3], which turned out to be coherence-information excess (global minus local) (cf (7)). The other two terms are the information gain and the residual mutual information.

The notion of function of observable or its coarsening is made ample use of extending discord also to incomplete observables. The second result (theorem 2 and (9)) revealed that in refinement (opposite of coarsening) both information gain and discord are non-decreasing, and the residual mutual information is non-increasing. It is known from previous work [12] that coherence information is non-decreasing in refinement. It is somewhat surprising that also the (global minus local) coherence-information excess (the discord) is non-decreasing. (The finer observable “sees” more quantumness both locally and in the correlations; and the latter outweigh the former.)

The zero-discord problem was explored in detail. Two kinds of zero discord have been distinguished: strong, when both terms in the excess coherence information are zero, and weak, when they are nonzero, but equal. Necessary and sufficient conditions were given where possible. Desirable results that have not been obtained were pointed out.

A unique string of coarsenings of the "interrogating" observable \(A_2\) has been derived:

\[
D^{qc}_{2} \leq C^{w}_{2} \leq B^{cs}_{2} \leq A_{2},
\]

corresponding to (reading from right to left) redundant noise, essential noise, garbled information, pure quantum information and pure quasi-classical information respectively (see section IV.).

Finally, simplest possible measurement interaction for measuring \(A_2\) leading to a tripartite state \(\rho^1_{123}\), in which the measuring apparatus is subsystem 3, was considered. The entropy relations in this state were discussed. It was shown that all three terms in the mutual information of \(\rho_{12}\) are shifted to the bipartite system \((1 + 2 + 3)\) in \(\rho^1_{123}\) (theorem 7 in section VI.). Further, it was shown that the global coherence information \(I_C(A_2, \rho_{12})\) is shifted into the global coherence information \(I_C(A_2, \rho^1_{123})\).
in $\rho'_{123}$ (cf (52)).

**Appendix A.**

Proof of theorem 2: In (4a) each $I(\rho'_{12})$ in the last term can be further decomposed according to (4a) itself. Performing this and substituting the result for each $l$ value in (4a), one obtains the RHS of the claimed relation (9).

To prove that the expression in the first large brackets is the information gain, we write down the decomposition of $S(\rho_1)$ due to probing with $A'_2$ analogous to (6) in two versions: directly and as a two-step procedure.

\[
S(\rho_1) = \sum_{l,q} [p_l p_{l,q} S(\rho'_1 l||\rho_1)] + \sum_{l,q} [p_l p_{l,q} S(\rho'_1 q||\rho_1)].
\]  

(A.1)

\[
S(\rho_1) = \left\{ \sum_l [p_l S(\rho'_1 l||\rho_1)] + \sum_{l,q} [p_l p_{l,q} S(\rho'_1 q||\rho_1)] \right\} + \sum_{l,q} [p_l p_{l,q} S(\rho'_1 q)].
\]  

(A.2)

Comparison of (A.1) and (A.2) proves the claim of theorem 2 as far as the information gain with respect to $A'_2$ is concerned.

It is obvious in (9) that the last expression is the amount of inaccessible correlations. Since the LHS is the same in (9) and (4a), the expression in the second large brackets must be the quantum discord.

\[\square\]

**Appendix B.**

We prove now the last claim in proposition 7. We need auxiliary lemmata.

**Lemma A.1:** If in a mixture of pure states $\rho = \sum_{l=1}^m p_l \mid l \rangle \langle l \mid$ one has $p_{l=1} = r'_{max}$, where $r'_{max}$ is the maximal eigenvalue of $\rho$, then necessarily $\mid l = 1 \rangle$ is an eigenvector of $\rho$ corresponding to the eigenvalue $r'_{max}$.

**Proof:** It is known that for all $l$ values $1 = ||\rho^{1/2} \rho^{-1/2} \mid l \rangle||^2$, where $\rho^{-1/2}$ is the inverse of the restriction of $\rho^{1/2}$ to the range of $\rho$ ([30], see Theorem 1 there). This implies

\[p_l = \left(\langle l \mid \rho^{-1} \mid l \rangle\right)^{-1} \quad l = 1, \ldots, m.\]  

(A.3)

(The operator $\rho^{-1}$, by definition, inverts the restriction of $\rho$ to its range.)

Let us expand $\mid l \rangle = \sum_{k=1}^d \alpha_k^l \mid r_k \rangle$, $l = 1, \ldots, m$ where $\{r_k : k = 1, \ldots, d\}$ is the positive spectrum of $\rho$, and $\{\mid r_k \rangle : k = 1, \ldots, d\}$ is a corresponding orthonormal set of eigenvectors. Substituting this in (A.3), one obtains

\[p_l = \left(\sum_{k=1}^d |\alpha_k^l|^2 r_k^{-1}\right)^{-1} \quad l = 1, \ldots, m.\]  

(A.4)

Assuming now that $p_{l=1} = r'_{max}$, one can write $p_{l=1}^{-1} - r'_{max}^{-1} = 0$, entailing with the use of (A.4)

\[\sum_{k=1}^d |\alpha_k^1|^2 \left(r_k^{-1} - r'_{max}^{-1}\right) = 0.\]

All terms are nonnegative. This implies $r_k < r'_{max} \Rightarrow \alpha_k^1 = 0$. Hence, if $q$ enumerates the possible multiplicity in $r_{max}$, then

\[\mid l = 1 \rangle = \sum_{q} \alpha_q^1 \mid r_{max}, q \rangle.\]  

(A.5)
Lemma A.2: If \( \rho = \sum_{l=1}^{d} r_l |l\rangle \langle l | \) is a mixture, and the weights \( \{ r_l : \forall l \} \) coincide with the positive eigenvalues of \( \rho \) (with possible repetition in the latter), then also the state vectors coincide each with a corresponding eigenvector of \( \rho \). Naturally, \( \rho = \sum_{l=1}^{d} r_l |r_l\rangle \langle r_l | \) is a spectral form of \( \rho \).

Proof: We assume that in the mixture the weights are written in non-increasing order. Then, according to lemma A.1, \( \rho = \sum_{l=1}^{n} r_l |l\rangle \langle l | \) and \( \sum_{l=n+1}^{d} r_l |l\rangle \langle l | \). To apply total induction, we further assume that the demonstration has already been done up to \( n \) : \( \rho = \sum_{l=1}^{n} r_l |r_l\rangle \langle r_l | + \sum_{l=n+1}^{d} r_l |l\rangle \langle l | \), where \( 1 \leq n \leq (d - 1) \).

Let us introduce \( \beta_n \equiv \sum_{k=n+1}^{d} r_k \), \( \beta_n > 0 \). Then \( \rho' = \sum_{l=n+1}^{d} r_l \beta_n |l\rangle \langle l | \). The last equality follows from the spectral form of \( \rho \). It is a spectral form of \( \rho' \). Hence, \( r_{n+1}/\beta_n \) is its largest eigenvalue. On account of lemma A.1, \( |l = n + 1\rangle \rangle r_{n+1}/\beta_n \). Total induction then proves the claim of lemma A.2.

Lemma A.3: If \( \rho = \sum_{l=1}^{m} p_l |l\rangle \langle l | \) is a mixture, and the so-called mixing entropy equals the entropy of the state, i.e., \( H(p_l) = S(\rho) \), then \( m = d \), and \( \{ p_l = r_l : l = 1, \ldots, d \} \) is the positive spectrum of \( \rho \) (with possible repetition in the eigenvalues).

Proof: According to Theorem 3 in a remarkable article by Nielsen [31], the existence of the mixture in lemma A.3 implies that its probability distribution is majorized by the spectrum of \( \rho \). This means that when both \( \{ p_l = r_l : l = 1, \ldots, m \} \) and \( \{ r_k : k = 1, \ldots, m \} \) are written in non-increasing order (if \( m < d \), then \( m - d \) zeros are added at the end of the positive spectrum of \( \rho \) then \( \sum_{l=1}^{m} p_l \leq \sum_{l=1}^{m} r_l \), \( n = 1, \ldots, (m - 1) \).

Next, we assume that the state space of \( \rho \) is at least \( m \)-dimensional. (If it is not, we can orthogonally add a space to the null space of \( \rho \) without loosing generality of the argument.) We define \( \rho' = \sum_{k=1}^{m} p_k |l\rangle \langle l | \), where \( \{ |l\rangle = l = 1, \ldots, m \} \) is an arbitrary orthonormal set.

Ruch introduced the term ”mixing character” for the positive spectrum of \( \rho \) (with possible zeros) [32] (see also [33]), and ”larger” for the majorized spectrum. In a previous article by the present author [34] the concept ”strictly larger mixing character” (when ”larger” is not valid symmetrically for the given mixing characters) was treated, and it was shown that von Neumann entropy is strictly mixing-homomorphic. This means that if the mixing character of \( \rho' \) is strictly larger than that of \( \rho \), then \( S(\rho') > S(\rho) \).

As it was stated, thanks to Nielsen, we know that the mixing character of \( \rho' \) is larger than that of \( \rho \). Since the entropies \( S(\rho') = H(p_l) \) and \( S(\rho) \) are assumed to be equal, the former cannot be strictly larger. It must be equal. Mixing characters are equal if and only if the corresponding states have equal positive eigenvalues with equal multiplicities. Hence, \( \{ p_l = r_l : l = 1, \ldots, m \} \) and \( m = d \) (the number of positive eigenvalues of \( \rho \) with possible repetitions) as claimed.

Finally, we prove the last part of proposition 7 claiming that if any mixture \( \rho = \sum_{l=1}^{m} p_l \rho_l \) of a finite number of admixed states is given, and it has the property that \( S(\rho) = H(p_l) + \sum_{l=1}^{m} p_l S(\rho_l) \), then the mixture is orthogonal, i.e., \( l \neq l', p_l > 0 < p_{l'} \Rightarrow p_l p_{l'} = 0 \).

Let \( \rho = \sum_{l=1}^{m} p_l \rho_l \) be the given (initial) mixture. Let, further, \( \forall l : p_l = \sum_{k} \rho_k \rho_l \rho_k \) be spectral forms. Substitution in the initial mixture gives \( \rho = \sum_{l} \sum_{k} p_l \rho_k \rho_l \), with the mixing entropy \( H(p_l \rho_k \rho_l) = H(p_l) + \sum_{l} p_l H(\rho_k) \). Since by assumption \( S(\rho) = H(p_l) + \sum_{l} p_l S(\rho_l) \), and \( \forall l : S(\rho_l) = 30 \).
\[ H(r_\downarrow) \], it is clear that one has \( H(\rho \uparrow) = H(\rho \downarrow) \). Hence, lemma A.3 is applicable to the mixture \( \rho = \sum_k \sum_n \rho_{nk}^k \) |l_k\rangle\langle l_k| \), making \( \{ \rho_{nk}^k : \forall n, \forall k \} \) the positive spectrum of \( \rho \). Then lemma A.2 implies that the mixture is a spectral form of \( \rho \). This cannot be unless the initial mixture is orthogonal as claimed. \( \square \)

**Appendix C**

Before we prove theorem 7, we establish some facts. Since the operators at issue are twin observables, one has
\[ \forall l : \rho_{123}^{l}|l\rangle\langle l|_3 \] (A.6a)
(cf the adjoint of (34b) changing what has to be changed), and equivalently (cf p. 052321-3 in [20]),
\[ \forall l : P_2 \rho_{123}^{l} P_2 = |l\rangle\langle l|_3 \rho_{123}^{l}|l\rangle\langle l|_3 . \] (A.6b)
Besides, we need the following result.

If \( \rho_{12} \) is a bipartite density operator and \(| b \rangle_2 \) is a second-subsystem unit vector, then
\[ |b\rangle_2 |b\rangle_2 \rho_{12} |b\rangle_2 |b\rangle_2 = \left[ tr_2 (\rho_{12} |b\rangle_2 \langle b|_2 ) \right] \otimes |b\rangle_2 |b\rangle_2 \] (A.7)

(see the necessity part in the proof of proposition 3).

**Proof** of theorem 7:

a) The validity of (48a) is a consequence of (A.6a), (47b), and of the definition of \( p_l^l \) (see beneath (46b)).

b) Utilizing (2c), (47c) and (A.6a), and finally the definition of \( p_l^l \) (beneath (46b)), and (48a), one has
\[ \forall l, \ p_l > 0 : \ \rho_l^l \equiv tr_2 (\rho_{12}^l) = p_l^{-1} tr_3 (\rho_{123}^l P_2^l) = \rho_l^l. \]
This proves (48b).

c) Claim (48c) is an immediate consequence of definition (47a).

d) Making use of (1a), of the mixing property of entropy, and of (2b), one has
\[ I_C (A_2, \rho_{12}) = H(p_l) + \sum_l p_l S(P_2^l \rho_{12}^l P_2^l / p_l) - S(\rho_{12}) = H(p_l) + \sum_l p_l S(\rho_{12}^l) - S(\rho_{12}). \] (A.8)

On account of (47c), (A.7) changing what has to be changed, and (A.6b), one can write
\[ S(\rho_{12}^l) = S\left( tr_3 (\rho_{123}^l |l\rangle\langle l|_3 / p_l ) / p_l \right) = S\left( tr_3 (\rho_{123}^l |l\rangle\langle l|_3 / p_l ) \otimes |l\rangle_3 \langle l|_3 \right) = S\left( |l\rangle_3 \langle l|_3 \rho_{123}^l |l\rangle_3 \langle l|_3 / p_l \right) = S\left( P_2^l \rho_{123}^l P_2^l / p_l \right). \]

Substituting this in (A.8), making use of the mixing property of entropy, taking into account that
\[ S(\rho_{12}) = S(\rho_{12} \otimes |\phi\rangle_3 \langle \phi|_3 ) = S\left( U_{23} (\rho_{12} \otimes |\phi\rangle_3 \langle \phi|_3 ) U_{23}^\dagger \right) = S(\rho_{123}^l), \]
and utilizing (1a), one derives the RHS of (48d).
e) To prove (48e), we argue in analogy with the preceding item.

\[ I_C(A_2, \rho_2) = H(p_i) + \sum_i p_i S(P_2^l \rho_2 P_2^l / p_i) - S(\rho_2). \]  

(A.9a)

Further, (2b), (2c) and (47c), upon taking partial trace 1 of it, imply

\[ S(P_2^l \rho_2 P_2^l / p_i) = S \left( \frac{1}{p_i} \text{tr}_3 (\rho_{23}^l | l \rangle_3 \langle l |_3) \otimes | l \rangle_3 \langle l |_3 \right) \).

Further evaluation using (A.7) and (A.6b) after taking partial trace 1 in it gives

\[ S(P_2^l \rho_2 P_2^l / p_i) = S(P_2^l \rho_2^l P_2^l / p_i). \]

Making use of the mixing property of entropy once again, on account of tr(\rho_{23}^l P_2^l) = p_i = p_i (cf beneath (46) and (48a)), one obtains

\[ S \left( \sum_i P_2^l \rho_2 P_2^l \right) = S \left( \sum_i P_2^l \rho_{23}^l P_2^l \right). \]  

(A.9b)

Returning to the last term on the RHS of (A.9a), one can write

\[ S(\rho_2) = S \left( U_{23} (\rho_2 \otimes | \phi \rangle_3 \langle \phi |_3) U_{23}^\dagger \right) = S(\rho_{23}) \]

(cf (47a) upon taking tr_1 in it). Substituting this and (A.9b) in (A.9a), in view of (1a), the RHS of (48e) is derived.

f) Finally, to prove (48f), we write down the definitions

\[ I(\rho_{12}^l) \equiv S(\rho_1^l) + S(\rho_2^l) - S(\rho_{12}^l), \]  

\[ I(\rho_{12}^{l_1 l_2}) \equiv S(\rho_1^{l_1}) + S(\rho_2^{l_2}) - S(\rho_{12}^{l_1 l_2}). \]  

(A.10a)  

(A.10b)

The first terms on the RHSs coincide due to (48b). Further, utilizing tr_1 of (47c),

\[ S(\rho_2) = S \left[ \text{tr}_3 (\rho_{23}^l | l \rangle_3 \langle l |_3) \otimes | l \rangle_3 \langle l |_3 \right) = S(P_2^l \rho_{23}^l P_2^l / p_i) = S(\rho_{23}^l) \]

(cf (A.7), (A.6b), and the definitions beneath (46)). As to the third terms on the RHSs of (A.10a) and (A.10b), equality is established by a similar argument:

\[ S(\rho_{12}^l) = S \left[ \frac{1}{p_i} \text{tr}_3 (\rho_{123}^l | l \rangle_3 \langle l |_3) \otimes | l \rangle_3 \langle l |_3 \right) \]

(cf (47c)). This equals S(\rho_{123}^l) (cf (A.7) and (A.6b)).  

\[ \square \]

References

[1] C. H. Bennett, D. P. DiVincenzo, J. A. Smolin, and W. K. Wootters, Phys. Rev. A 54, 3824 (1996).
[2] F. Herbut, J. Phys. A: Math. Gen. 38, 2959 (2005); e-print quant-ph/0503077.
[3] W. H. Zurek, Ann. Phys. (Leipzig) 9, 5 (2000); e-print quant-ph/0011039.
[4] H. Ollivier and W. H. Zurek, Phys. Rev. Lett., 88, 17901 (2002); e-print quant-ph/0105072.
[5] W. H. Zurek, *Phys. Rev. A* **67**, 012320 (2003); e-print quant-ph/0301127.
[6] W. H. Zurek, *Rev. Mod. Phys.* **75**, 715 (2003); quant-ph/0105127.
[7] J. Oppenheim, M. Horodecki, P. Horodecki, and R. Horodecki, *Phys. Rev. Lett.* **89**, 180402 (2002); e-print quant-ph/0112074.
[8] M. Horodecki, P. Horodecki, R. Horodecki, J. Oppenheim, A. Sen(De), U. Sen, and B. Synak, e-print quant-ph/0410090.
[9] A. Uhlmann, *Lect. Notes Phys.* **539**, 93 (2000); e-print quant-ph/9901027.
[10] G. Lüders, *Ann. Phys.* (Leipzig) **8**, 322 (1951); A. Messiah, *Quantum Mechanics*, vol. I (North-Holland, Amsterdam, 1961), p. 333; C. Cohen-Tannoudji, B. Diu, and F. Laloe, *Quantum Mechanics*, vol. I (Wiley-Interscience, New York, 1977), p. 221 (Fifth Postulate).
[11] G. Auletta, *Foundations and Interpretation of Quantum Mechanics*, (World Scientific, Singapore, 2001).
[12] F. Herbut, *J. Phys. A: Math. Gen.* **36**, 8479 (2003); e-print quant-ph/0309181.
[13] A. Wehrl, *Rev. Mod. Phys.* **50**, 221 (1978).
[14] F. Herbut, *J. Phys. A: Math. Gen.* **37**, 3535 (2004); e-print quant-ph/0311193.
[15] G. Lindblad, *Commun. Math. Phys.* **40**, 147 (1975).
[16] A. Uhlmann, *Comm. Math. Phys.* **54**, 21 (1977).
[17] L. Henderson and V. Vedral, *J. Phys. A: Math. Gen.*, **34**, 6899 (2001); e-print quant-ph/0105028.
[18] E. H. Lieb, *Bull. Am. Math. Soc.* **81**, January (1975).
[19] C. Adami and N. J. Cerf, *Phys. Rev. A* **56**, 3470 (1997).
[20] F. Herbut, *Phys. Rev. A* **66**, 052321 (2002); e-print quant-ph/0305187.
[21] F. Herbut and M. Damnjanović, *J. Phys. A: Math. Gen.* **33**, 6023 (2000); e-print quant-ph/0004085.
[22] N. J. Cerf and C. Adami, *Phys. Rev. Lett.* **79**, 5194 (1997).
[23] I. Devetak and A. E. Staples, e-print quant-ph/0112166.
[24] B. Schumacher and M. A. Nielsen, *Phys. Rev. A* **54**, 2629 (1996); e-print quant-ph/9604022.
[25] F. Herbut, *Ann. Phys.* **NY** **55**, 271 (1969).
[26] F. Herbut and M. Vujčić, *Ann. Phys. NY* **96**, 382 (1976).
[27] M. Vujčić and F. Herbut, *J. Math. Phys.* **25**, 2253 (1984).
[28] P. Busch, P. J. Lahti, and P. Mittelstaedt, *The Quantum Theory of Measurement*, Lecture Notes in Physics, M2 (Springer-Verlag, Berlin, 1991).
[29] V. B. Braginsky and F. Y. Khalili, *Quantum Measurement*, (Cambridge University Press, Cambridge, 1992).
[30] G. Cassinelli, E. De Vito, and A. Leverre, *J. Math. Analysis and Appl.* **210**, 472 (1997).
[31] M. A. Nielsen, *Phys. Rev. A* **61**, 064301 (2000); e-print quant-ph/9909020.
[32] E. Ruch, *Theor. Chim. Acta* **38**, 167 (1975).
[33] B. Lesche, *J. Math. Phys.* **17**, 427 (1976).

[34] F. Herbut and I. D. Ivanović, *J. Phys. A: Math. Gen.* **15** 1775 (1982).

[35] M. Koashi and A. Winter, *Phys. Rev. A* **69** 022309 (2004); e-print quant-ph/0310037.

[36] C. H. Bennett, D. P. DiVincenzo, J. A. Smolin and W. K. Wootters, *Phys. Rev. A* **54** 3824 (1996); e-print quant-ph/9604024.