On a quantum version of Shannon's conditional entropy

R. Schrader*
Institut für Theoretische Physik
Freie Universität Berlin, Arnimallee 14
D-14195 Berlin, Germany

May 31, 2000

Abstract

In this article we propose a quantum version of Shannon's conditional entropy. Given two density matrices ρ and σ on a finite dimensional Hilbert space and with $S(\rho) = -\operatorname{Tr}\rho\ln\rho$ being the usual von Neumann entropy, this quantity $S(\rho|\sigma)$ is concave in ρ and satisfies $0 \le S(\rho|\sigma) \le S(\rho)$, a quantum analogue of Shannon's famous inequality. Thus we view $S(\rho|\sigma)$ as the entropy of ρ conditioned by σ. The second inequality is an equality if σ is a multiple of the identity. In contrast to the classical case, however, $S(\rho|\rho) = 0$ if and only if the non-vanishing eigenvalues of ρ are all non-degenerate. Also in general and again in contrast to the corresponding classical situation $S(\rho, \sigma) = S(\sigma) + S(\rho|\sigma)$ is not symmetric in ρ and σ even if they commute. We also show that there is no quantum version of conditional entropy in terms of two density matrices, which shares more properties with the classical case and which in particular reduces to the classical case when the two density matrices commute. As an alternative we propose to use spectral resolutions of the unit matrix instead of density matrices. We briefly compare this with the algebraic approach of Connes and Størmer and Connes, Narnhofer and Thirring.

1 Introduction

The concept of entropy plays a major role in thermodynamics and statistical mechanics. It serves to describe the behavior of macroscopic systems. The name "entropy" was introduced by Clausius (1865) and derives from $\epsilon\nu\tau\rho o\pi\iota\eta$ "transformation". It was von Neumann (1927 [**?**]), who generalized the classical expression of Boltzmann and Gibbs for the entropy to quantum mechanics by using the concept of what is now called a density matrix, also introduced quite generally by him in the same year [**?**]. In the special context of radiation damping the density matrix was discovered independently by L. Landau [**?**] and by F. Bloch [**?**], again in the same year (see also the citation in [**?**]). For a technical overview of the developments up to 1978 and with further historical references see [**?**]. For recent expositions see [**?**, **?**]. In the theory of dynamical systems entropy and the derived notion of topological entropy also plays an important role, see e.g. the contributions in [**?**].

In a seminal article Shannon (1948, [**?**]) introduced the concept of entropy into information theory. Roughly speaking a gain in information means a decrease in entropy. Shannon also provided the concept of conditional entropy. It is a measure how entropy

*e-mail: schrader@physik.fu-berlin.de, Supported in part by DFG SFB 288 "Differentialgeometrie und Quantenphysik"

is reduced given a preexisting knowledge. To the author's best knowledge the first construction in quantum mechanics coming close to such a notion is due to E. Lieb [?] (see also [?, ?]). It involves tensor product structures and it was called a relative entropy in [?] (but a conditional entropy in [?], p. 259). In view of recent developments in quantum computation and quantum coding (see [?, ?] for a concise account) it is highly desirable to have such a quantity at ones disposal. There is a construction of a non-commutative analogue of Shannon's conditional entropy by Connes and Størmer [?] and Connes, Narnhofer and Thirring [?](for an exposition and a discussion of further developments see e.g. [?, ?]). More recently attempts have been made to construct a mutual information analogous to Shannon's conditional entropy in the context of quantum error-correction. In two of these attempts [?, ?], made independently, yielded the same quantity. The first article exhibits necessary and sufficient conditions for quantum error-correction to be possible in terms of the mutual information like the quantity given there, and a conjecture is made on its connection with quantum channel capacity, explored in more detail [?]. The connection with channel capacity was also analyzed in [?]. In [?] its connection with entanglement is discussed. In yet another approach [?] the starting point is one density matrix on a tensor product. The conditioning is then obtained by looking at the two density matrices in the two sub-systems resulting by taking the corresponding partial traces. An alternative approach, which also starts with one density matrix in a tensor product is given in [?]. There the resulting conditional entropy may become negative.

In this article we will propose a different candidate for a quantum mechanical conditional entropy $S(\rho|\sigma) \geq 0$, a function of two density matrices ρ and σ in a same Hilbert space and having the interpretation of the entropy of ρ conditioned by the "knowledge" given by σ. For simplicity we will only discuss the finite dimensional case although an extension to the infinite dimensional case seems possible. If we view ρ as the analogue of X and σ the analogue of Y such that von Neumann's entropy $S(\rho)$ is the analogue of Shannon's entropy $H(X)$, then this conditional entropy shares several but not all properties of Shannon's conditional entropy $H(X|Y)$ (see section 3 for a brief recapitulation of Shannon's theory). In particular the "knowledge" of σ reduces the entropy, i.e. the inequality $S(\rho|\sigma) \leq S(\rho)$ holds. This corresponds exactly to Shannon's famous inequality $H(X|Y) \leq H(X)$ and was our main motivation for our construction. Also and again in analogy to the classical theory we wanted the conditioning to be given by a quantity on the same footing as the original density matrix, i.e. conditioning should also be given by a density matrix. If as in the classical case σ contains no information, i.e. if it is a multiple of the identity such that $S(\sigma)$ is maximal, then $S(\rho|\sigma) = S(\rho)$. In contrast to the classical case $H(X|X) = 0$, however, the relation $S(\rho|\rho) = 0$ holds if and only if the non-zero eigenvalues of ρ are non-degenerate. In particular $S(\rho|\rho) = 0$ if ρ is pure. We will not elaborate on the question, whether the failure of our $S(\rho|\sigma)$ to satisfy all corresponding classical properties, like this last property, is due to a fundamental difference of quantum and classical information theory. In particular we will not provide a more detailed quantum mechanical interpretation of $S(\rho|\sigma)$. Also so far we have not analyzed whether it may be used in the context of channel capacity. Rather we will argue that other quantum mechanical versions of conditional entropy, which share more properties with the classical counterpart $H(X|Y)$, do not exist.

The article is organized as follows. In section 2 we provide the construction of a quantum version $S(\rho|\sigma)$ of the conditional entropy and establish several properties. In section 3 and after a brief review of Shannon's theory we compare this with Shannon's conditional entropy. In section 4 we first present a list of desirable properties for a quantum version of conditional entropy given in terms of two density matrices. We then show that even parts of these desiderata can not be fulfilled simultaneously. In particular there is no version involving two density matrices and which reduces to the classical case, when these two density matrices commute. We will provide an alternative in terms of resolutions of the unit matrix in terms of orthogonal projections and which share more properties with the classical case. Briefly we will compare this ansatz with the algebraic constructions given by Connes and Størmer and Connes, Narnhofer and Thirring.

2 Construction of a quantum conditional entropy

Let ρ be a density matrix on a finite dimensional Hilbert space $\mathcal{H}$, i.e $\rho \geq 0$ and $\operatorname{Tr}\rho = 1$, where Tr denotes the canonical trace on $\mathcal{H}$. We write $\rho = \sum_i \rho_i \; P_i$ for the spectral representation of ρ where the projections $P_i \neq 0$ are pairwise orthogonal (i.e. $P_i P_j = \delta_{ij} P_i, P_i = P_i^\dagger$), such that $\rho_i \geq 0$, $\rho_i \neq \rho_j$ for $i \neq j$ and $\sum_i P_i = \mathbb{I}$, where $\mathbb{I}$ is the identity operator on $\mathcal{H}$. Thus $\operatorname{Tr}\rho = \sum_i \dim P_i \; \rho_i = 1$ with $\dim P = \operatorname{Tr} P = \dim P\mathcal{H}$ for any projection P. Here and in what follows projection operators are always understood to be orthogonal. With this notational convention the P_i are canonically defined in terms of ρ. Since this fact will be crucial in what follows, let us briefly recall a standard proof. The eigenvalues ρ_i (and their degeneracies (= $\dim P_i$)) are of course uniquely determined by ρ as solutions in λ of the secular equation $\det(\lambda\mathbb{I} - \rho) = 0$, a basis independent relation, such that $\det(\lambda\mathbb{I} - \rho) = \prod_i (\lambda - \rho_i)^{\dim P_i}$. Order the ρ_i in such a way that $1 \geq \rho_1 > \rho_2 > \rho_3 > \ldots$. Then $P_1 = \lim_{n\to\infty} (\rho/\rho_1)^n$, $P_2 = \lim_{n\to\infty} ((\rho - \rho_1 P_1)/\rho_2)^n$, etc.

The quantum mechanical entropy of ρ is given as $S(\rho) = -\sum_i \dim P_i \; \rho_i \ln \rho_i$, which is continuous and concave in ρ (for an account of sub-additivity and convexity properties of the entropy and related quantities see e.g. [?, ?, ?]). Let σ be another density matrix on the same space $\mathcal{H}$ with the spectral representation $\sigma = \sum_j \sigma_j Q_j$ again written in a canonical way. We define the conditional entropy by

$$
\begin{aligned}
S(\rho|\sigma) &= \sum_j \dim Q_j \; \sigma_j \; F(\rho, Q_j) \\
&= \sum_j \operatorname{Tr} Q_j \sigma \; F(\rho, Q_j) \qquad (1)
\end{aligned}
$$

where

$$
F(\rho, Q) = -\operatorname{Tr}(Q\rho Q \ln(Q\rho Q)) + \operatorname{Tr}(Q\rho Q) \ln \operatorname{Tr}(Q\rho Q) \qquad (2)
$$

for any orthogonal projection Q. Since the Q_j's and σ_j's are well defined in terms of σ and since trivially $Q\rho Q \geq 0$, $S(\rho|\sigma)$ is well defined. Also as usual in this context $A \ln A$ for any non-negative operator A is defined in terms of the spectral representation of A with the natural convention that $x \ln x|_{x=0} = 0$. If $Q\rho Q \neq 0$ then also $0 \neq \operatorname{Tr} Q\rho Q = \operatorname{Tr} Q\rho$ and then we may write

$$
F(\rho, Q) = \operatorname{Tr} Q\rho \cdot S(\rho_Q) \qquad (3)
$$

with

$$
\rho_Q = \frac{1}{\operatorname{Tr}(Q\rho)} \cdot Q\rho Q \qquad (4)
$$

being a density matrix. Actually we might use (**??**) instead of (**??**) as a definition for $F(\rho, Q)$ with the convention, usually made in similar contexts (see e.g. [?]), that 0 times something undefined is 0. Relation (**??**) shows that $F(\rho, Q) \geq 0$ for all ρ and Q. Using (**??**) we may rewrite $S(\rho|\sigma)$ as

$$S(\rho|\sigma) = \sum_{j\,:\,Q_j\rho Q_j \neq 0} \mathrm{Tr}\, Q_j\sigma \;\; \mathrm{Tr}\, Q_j\rho \; S(\rho_{Q_j}). \tag{5}$$

There is yet another way of writing $F(\rho, Q)$. It uses the relative entropy $0 \leq S_{rel}(A, B) = \mathrm{Tr}\, A(\ln A - \ln B) \leq \infty$, which is defined for any $A \geq 0$ and $B \geq 0$. The relative entropy is lower semi-continuous in A and jointly convex in A and B, see e.g. [**?**, **?**]. Obviously $S_{rel}(\lambda A, \lambda B) = \lambda S_{rel}(A, B)$ holds for any $\lambda > 0$ and we have

$$F(\rho, Q) = -S_{rel}(Q\,\rho\,Q, \mathrm{Tr}(Q\,\rho\,Q)\mathbb{I}) \tag{6}$$

such that

$$S(\rho|\sigma) = -\sum_j \mathrm{Tr}\, Q_j\sigma \cdot S_{rel}(Q_j\rho Q_j, \mathrm{Tr}(Q_j\rho Q_j)\mathbb{I}) \tag{7}$$

It is instructive to compare $S(\rho|\sigma)$ with $S(E_{\underline{Q}}(\rho))$ and which actually motivated our construction of $S(\rho|\sigma)$. $E_{\underline{Q}}$ is the linear map on the set of linear operators A on $\mathcal{H}$ given as $E_{\underline{Q}}(A) = \sum_j Q_j A Q_j$. The Q_j's are as above, i.e a any set $\underline{Q} = \{Q_j\}$ of pairwise orthogonal nonzero projection operators with $\sum_j Q_j = \mathbb{I}$ and which is called a resolution of the identity. $E_{\underline{Q}}$ is a conditional expectation (see e.g. [**?**]) with range being the $\star$-algebra consisting of all linear operators which commute with all Q_i. In particular $E_{\underline{Q}}$ maps density matrices into density matrices. More precisely, let $\mathcal{B} = \mathcal{B}(\mathcal{H})$ be the $\star$-algebra of all linear operators on $\mathcal{H}$, which is (isomorhic to) a full matrix-algebra. Then $E_{\underline{Q}}(\mathcal{B})$ is a $\star$-sub-algebra of $\mathcal{B}$ and the direct sum of the $\star$-sub-algebras $Q_j\mathcal{B}Q_j = \mathcal{B}(Q_j\mathcal{H})$, which are (isomorphic to) full matrix algebras. Although any finite dimensional $\star$-algebra is (isomorphic to) a direct sum of full matrix algebras, not all $\star$-sub-algebras of $\mathcal{B}$ are of the form $E_{\underline{Q}}(\mathcal{B})$ for a suitable $\underline{Q}$. As an example consider the algebra generated by $\mathbb{I}$ alone. It can easily be shown that any $\star$-sub-algebra is of this form if and only if it contains a maximal abelian sub-algebra. Also from $E_{\underline{Q}}(\mathcal{B})$ $\underline{Q}$ may be recovered. Indeed the Q_j's are just the minimal self-adjoint idempotents (i.e. the orthogonal projections) in $E_{\underline{Q}}(\mathcal{B})$ and which are central. Also on the set of all spectral resolutions of the identity we introduce a partial ordering $\leq$ by setting $\underline{P} \leq \underline{Q}$ if to each i there is $j(i)$ (which is unique) such that $P_i \leq Q_{j(i)}$. Note that each j is of the form $j = j(i)$ for at least one i. Then in particular all P_i commute with all Q_j. Also $\underline{P} \leq \{\mathbb{I}\}$ holds for all $\underline{P}$. It is easy to see that $\underline{P} \leq \underline{Q}$ if and only if $E_{\underline{P}}(\mathcal{B}) \subseteq E_{\underline{Q}}(\mathcal{B})$. With respect to these orderings $\underline{P}$ or equivalently $E_{\underline{P}}(\mathcal{B})$ is minimal if and only if each P_i is one-dimensional. $E_{\underline{P}}(\mathcal{B})$ is then commutative with dimension equal to $\dim \mathcal{H}$. To sum up, with respect to the partial ordering $\leq$ there is a unique maximal element but there are many minimal elements in the set of spectral resolutions $\underline{P}$.

Now one has the well known result $S(E_{\underline{Q}}(\rho)) \geq S(\rho)$ (see e.g. [**?**] for a direct proof and [**?**] for the special case when $\dim Q_j = 1$ for all j. It is a special case of Uhlmann's monotonicity theorem [**?**], see also [**?**]). It means that projective measurements increase entropy and compares with the inequality $S(\rho|\sigma) \leq S(\rho)$ to be proven below. Its interpretation is that of a projective measurement described by the family $\underline{Q}$ of projections on a system given by ρ, but where we never learn of the result of the measurement. In contrast $S(\rho|\sigma)$ is interpreted as a set of projective measurements given by the projections Q_j, each performed with the probability dim Q_j σ_j, and where we learn of each outcome $F(\rho, Q_j)$ separately. The sum in (**??**) and (**??**) then reflects the occurrence of a quantum decoherence. In other words one considers the family of density operators ρ_{Q_j}, $Q_j \rho Q_j \neq 0$, takes their von Neumannn entropy and then forms the linear combination with the non-negative coefficients $\mathrm{Tr}\, Q_j\sigma\ \mathrm{Tr}\, Q_j\rho$.

By definition we have

$$F(\rho, Q = \mathbb{I}) = S(\rho|\sigma = (1/\dim \mathbb{I}) \cdot \mathbb{I}) = S(\rho). \tag{8}$$

We consider this property to be necessary for any other sensible definition of a conditional entropy involving two density matrices. It holds for Shannon's conditional entropy $H(X|Y)$ in the form $H(X|Y) = H(X)$ when Y is the trivial partition (see section 3), which means that there is no gain in information, if Y contains no information. We will return to this point in section 3.

Some additional remarks are in order. Since the quantity $S(\rho|\sigma)$ is supposed to be a quantum mechanical mechanical analogue of Shannon's conditional entropy $H(X, Y)$, ρ corresponds to X and σ to Y. In analogy to the classical case, where X and Y may be considered to be stochastic variables living on the same space, here the density matrices ρ and σ also live on the same space. Unfortunately with this correspondence $S(\rho|\sigma)$ does not reduce to the classical case when ρ and σ commute (see (**??**) and its discussion in section 3). As matter of fact, we shall argue in section 4 that a quantum conditional entropy with this property does not exist.

By construction we have the obvious invariance under unitary automorphisms

$$F(\rho, Q) = F(U\rho U^{-1}, UQU^{-1}), \tag{9}$$

for any $U \in \mathcal{U}(\mathcal{H})$, the group of unitary operators in $\mathcal{H}$. This relation (**??**) immediately implies

$$S(U\rho U^{-1}|U\sigma U^{-1}) = S(\rho|\sigma) \tag{10}$$

for all U. Relation (**??**) reflects the fact that $S(\rho|\sigma)$ is defined intrinsically and is in particular basis independent. Therefore this invariance property should also hold for any alternative, sensible definition of a quantum mechanical conditional entropy defined in terms of two density matrices. We shall comment on the classical analogue to (**??**) in section 4.

The next observation is also important. It is easy to see that $F(\rho, Q)$ is continuous in ρ and Q by the same arguments used to prove continuity of $S(\rho)$. Therefore $S(\rho|\sigma)$ is also continuous in ρ for fixed σ. However, $S(\rho|\sigma)$ is not continuous in σ everywhere for all fixed ρ. It is continuous on the dense open subset where the eigenvalues of σ are non-degenerate.In fact, it is zero there(see below). So this lack of continuity occurs where σ has degenerate eigenvalues and is due to the fact that for $Q = Q' + Q''$ being the sum of two projections both $\neq 0$ and which are orthogonal to each other, i.e. $Q'Q'' = 0$, in general one has

$$\dim Q\, F(\rho, Q) \neq \dim Q'\, F(\rho, Q') + \dim Q''\, F(\rho, Q''). \tag{11}$$

To understand this consider the case when $\dim \mathcal{H} = 2$. Then $S(\rho|\sigma) = S(\rho)$ if $\sigma = 1/2\ \mathbb{I}$ and $S(\rho|\sigma) = 0$ otherwise. At the moment we do not know whether this lack of continuity of $S(\rho|\sigma)$ in σ is a desirable feature or not, i.e whether this can be understood quantum mechanically, when we interpret $S(\rho|\sigma)$ as the entropy of ρ conditioned by σ. Observe that a degeneracy typically occurs when a non-trivial symmetry is present. In other words there is then a non-trivial non-abelian subgroup $\mathcal{G} = \mathcal{G}(\sigma)$ of $\mathcal{U}(\mathcal{H})$ such that $U\sigma U^{-1} = \sigma$ for all $U \in \mathcal{G}$. Note that $\mathcal{G}$ always contains a subgroup isomorphic to the abelian group $U(N = \dim \mathcal{H})$. In this picture a removal of degeneracies is related to a breakdown of symmetry, a familiar phenomenon in physics.

To proceed further, $F(\rho, Q) = 0$ if $Q\rho Q = 0$, which can happen for $Q \neq 0$ only if ρ has zero as an eigenvalue, i.e. if ρ is not strictly positive. Then also $(\mathbb{I} - Q)\rho Q = Q\rho(\mathbb{I} - Q) = 0$. In fact, by Schwarz inequality for any $\psi, \psi' \in \mathcal{H}$ we have

$$| < \psi, Q\rho(\mathbb{I} - Q)\psi' > | \leq ||\rho^{1/2}Q\psi|| ||\rho^{1/2}(\mathbb{I} - Q)\psi'|| = 0.$$

This also shows that $Q\rho Q = 0$ is equivalent to $Q\rho = 0$, which in turn by the self-adjointness of ρ and Q is equivalent to $\rho Q = 0$. By the trivial identity

$$\rho = Q\rho Q + (\mathbb{I} - Q)\rho Q + Q\rho(\mathbb{I} - Q) + (\mathbb{I} - Q)\rho(\mathbb{I} - Q), \tag{12}$$

valid for all ρ, Q, we therefore also have $\rho = (\mathbb{I} - Q)\rho(\mathbb{I} - Q)$ whenever $Q\rho Q = 0$. Obviously (**??**) gives $\mathrm{Tr}\,\rho = \mathrm{Tr}\,Q\rho Q + \mathrm{Tr}(\mathbb{I} - Q)\rho(\mathbb{I} - Q)$ such that in particular the inequalities $0 \leq \mathrm{Tr}\,Q\rho Q \leq 1$ and $0 \leq \mathrm{Tr}(\mathbb{I} - Q)\rho(\mathbb{I} - Q) \leq 1$ hold for any ρ and Q. By relation (**??**) we also have $F(\rho, Q) \geq 0$ and hence $S(\rho|\sigma) \geq 0$ for all ρ, Q and σ. Now $S(\rho_Q) = 0, Q\rho Q \neq 0$ holds if and only if ρ_Q is a pure state, i.e. a one-dimensional projection. Also for $\dim Q = 1$ one always has $Q\rho Q = (\mathrm{Tr}\,Q\rho Q)Q$. We collect this observation in

Lemma 2.1. *$F(\rho, Q) = 0$ if and only if $Q\rho Q$ is a multiple of a one-dimensional projection.*

This multiple is allowed to be zero. To characterize such Q's fulfilling the conditions of the lemma, let $P(\rho) \neq 0$ be the projection operator onto the subspace corresponding to the non-zero eigenvalues, such that $P(\rho)\rho = \rho = \rho P(\rho)$ and in particular $P(\rho) = \mathbb{I}$ if $\rho > 0$. Using the spectral representation of ρ it is easy to see that $Q\rho Q$ is a multiple (possibly zero) of a one-dimensional projection if and only if Q may be written as $Q = Q' + Q''$ with $\dim Q' \leq 1$ and $P(\rho)Q'' = \rho Q'' = 0$.

More generally consider the case where $Q\rho Q = (\mathrm{Tr}(Q\rho Q)/\dim Q') \cdot Q', Q \neq 0$ holds for a suitable projection operator Q' such that in particular $0 \neq Q' \leq Q$ and Q' is unique whenever $Q\rho Q \neq 0$. Then $F(\rho, Q) = (\mathrm{Tr}\,Q\rho Q)\ln\dim Q'$ and $\rho_Q = (1/\dim Q')Q'$. This gives the

Lemma 2.2. *If all non-zero eigenvalues of σ are non-degenerate then $S(\rho|\sigma) = 0$ for all ρ. More generally if $Q_j\rho Q_j$ is a multiple (possibly zero) of some projection operator $Q'_j\,(\leq Q_j)$ for all j with $\sigma_j > 0$, then*

$$S(\rho|\sigma) = \sum_j \mathrm{Tr}\,Q_j\rho\ \ \mathrm{Tr}\,Q_j\sigma\ln\dim Q'_j. \tag{13}$$

Observe that $S(\rho|\sigma) = 0$ for all pure states σ and all ρ. If ρ is pure then $Q\rho Q$ is always a multiple of a pure state for all Q. Therefore $S(\rho|\sigma) = 0$ also holds for all σ whenever ρ is pure. Also if $\rho\sigma = 0$ which is equivalent to $\mathrm{Tr}\,\rho\sigma = 0$ and which can happen only if neither ρ nor σ is strictly positive, then again $S(\rho|\sigma) = 0$. Sufficient (but not necessary) for the condition of Lemma **??** to hold is that to each j with $\sigma_j > 0$ there is $i(j)$ with $Q_j \leq P_{i(j)}$. For these j's $Q'_j = Q_j$, $Q_j\rho Q_j = \rho_{i(j)}Q_j$ and hence $\mathrm{Tr}\,Q_j\rho = \rho_{i(j)}\dim Q_j$. This gives in particular

$$S(\rho|\rho) = \sum_i \rho_i^2(\dim P_i)^2\ln\dim P_i. \tag{14}$$

Therefore the relation $S(\rho|\rho) = 0$ holds if and only if all the non-zero eigenvalues of ρ are non-degenerate, the if part being a special case of Lemma **??**.

If in addition to the property $Q_j \leq P_{i(j)}$ the density matrix σ is such that

$$\sum_{j:i(j)=i} \sigma_j (\dim Q_j)^2 \ln \dim Q_j \leq \rho_i (\dim P_i)^2 \ln \dim P_i$$

holds for all i, then by (**??**) and (**??**) $S(\rho|\sigma) \leq S(\rho|\rho)$. Note that this last condition is satisfied if

$$\sum_{j:i(j)=i} \sigma_j \dim Q_j \leq \rho_i \dim P_i$$

holds since trivially $\dim Q_j \leq \dim P_{i(j)}$.

We return to a discussion of the general properties of $F(\rho, Q)$ and $S(\rho|\sigma)$. The first main result of this article shows that $S(\rho|\sigma)$ shares an important property with $S(\rho)$ (see e.g. [**?**, **?**] for the classical and the quantum entropy and [**?**] for Shannon's conditional entropy and derived quantities).

Theorem 2.1. *$F(\rho, Q)$ and $S(\rho|\sigma)$ are both concave in ρ.*

Again we consider this property to be necessary for any sensible definition of a quantum conditional entropy. Like for the entropy $S(\rho)$ itself it states that mixing (in ρ) increases (conditional) entropy. On the other hand the case $\dim \mathcal{H} = 2$ discussed above shows that in general $S(\rho|\sigma)$ for fixed ρ is neither convex nor concave in σ. Intuitively it would be desirable to have concavity with respect to σ since mixing the conditioning should increase conditional entropy.

The proof follows easily from the presentation (**??**) and (**??**) and the known convexity property of the relative entropy.

The second main result of this article shows in particular that $S(\rho|\sigma)$ satisfies Shannon's inequality.

Theorem 2.2. *The following inequalities hold for all density matrices ρ and σ in a fixed finite dimensional Hilbert space*

$$0 \leq S(\rho|\sigma) \leq S(\rho). \tag{15}$$

If $\rho > 0$ the last inequality is strict unless $\sigma = (1/\dim \mathbb{I}) \cdot \mathbb{I}$.

The above comparison of $S(\rho|\sigma)$ with $S(E_{\underline{Q}}(\rho)$ suggests another definition of conditional entropy with the conditioning not given in terms of a density matrix σ but rather only in terms of any resolution $\underline{Q}$ of the identity.

$$S(\rho|\underline{Q}) = \sum_j \frac{\dim Q_j}{\dim \mathbb{I}} F(\rho, Q_j). \tag{16}$$

By (**??**) below we have

$$0 \leq S(\rho|\underline{Q}) \leq S(\rho),$$

where the first inequality is an equality if $\dim Q_j = 1$ for all j and the second one an equality if the spectral resolution is trivial, i.e. if $\underline{Q} = \{\mathbb{I}\}$. We note that in (**??**) any sequence of numbers $\sigma'_j \geq 0$ (labeled in the same way as the Q_j's) with $\sum_j \sigma'_j = 1$ and replacing $\dim Q_j/\dim \mathbb{I}$ would do equally well. But then we may combine and encode these data $\underline{Q}$ and $\{\sigma\}$ in the density matrix $\sigma = \sum_j \sigma_j Q_j$ with $\sigma_j = \sigma_j'/\dim Q_j$. If in addition all the σ_j's are pairwise different, then by our discussion above they and the spectral resolution $\underline{Q}$ may be recovered from σ and we are back to our construction $S(\rho|\sigma)$.

Due to the relation $1 = \operatorname{Tr}\sigma = \sum_j \dim Q_j\ \sigma_j$ this second theorem is an immediate consequence of the following

Lemma 2.3. *For all ρ and Q the inequality*

$$F(\rho, Q) \leq S(\rho) \tag{17}$$

holds. If $\rho > 0$ this inequality is strict unless $Q = \mathbb{I}$.

Before we turn to a proof we make some remarks. We conjecture that in the general case $\rho \geq 0$, the inequality (**??**) is strict unless $Q\rho = \rho$. This would imply that the second inequality in (**??**) is strict unless $\sigma\rho = (\operatorname{Tr}\sigma\rho)\rho$, which means the following. Any σ with $\sigma\rho = (\operatorname{Tr}\sigma\rho)\rho$ is of the form $\sigma = (\operatorname{Tr}\sigma\rho)P(\rho) + \sigma'$ with $(\mathbb{I} - P(\rho))\sigma' = \sigma'$.

Instead of $F(\rho, Q)$ one might be tempted to consider instead the quantity (see (**??**))

$$\tilde{F}(\rho, Q) = -\operatorname{Tr}(Q\rho Q \ln Q\rho Q) \geq 0$$

and try to prove $\tilde{F}(\rho, Q) \leq S(\rho)$. Obviously we have $\tilde{F}(\rho, Q) \geq F(\rho, Q)$. Consider, however, the case where $\dim Q = 1$ and $\rho = P$, $\dim P = 1$ (i.e. ρ is pure) and with P chosen such that $QPQ = (\operatorname{Tr} QP)Q$ satisfies $0 < \operatorname{Tr} PQ < 1$. Then $0 = S(\rho = P) < \tilde{F}(\rho = P, Q)$. Furthermore one has $F(\rho, Q) \leq S(\rho_Q)$ when $0 < \operatorname{Tr} Q\rho Q (\leq 1)$. But it does not make sense to replace $F(\rho, Q)$ by $S(\rho_Q)$ as an alternative, since $S(\rho_Q)$ is only defined when $Q\rho Q \neq 0$. Even if $Q\rho Q \neq 0$, one does not have $S(\rho_Q) \leq S(\rho)$ in general. To see this we will consider an example. For any $0 \neq \psi \in \mathcal{H}$ let P_ψ be the 1-dim. projection onto the subspace spanned by ψ.

Example 2.1. *Let* $\dim \mathcal{H} = 4$ *with $\psi_1, \psi_2, \psi_3, \psi_4$ being an orthonormal basis. Let $Q = P_{\psi_1} + P_{\psi_2}$ be the 2-dim. projection onto the sub-space spanned by ψ_1 and ψ_2. Choose $\rho(\phi_1, \phi_2) = \rho_1 P_{\psi_1'} + \rho_2 P_{\psi_2'}$, $\rho_1 + \rho_2 = 1$ with*

$$\begin{aligned}
\psi_1' &= \cos\phi_1\, \psi_1 + \sin\phi_1\, \psi_3, \\
\psi_2' &= \cos\phi_2\, \psi_2 + \sin\phi_2\, \psi_4,\ \cos\phi_1 \neq 0 \neq \cos\phi_2.
\end{aligned}$$

Then

$$\rho(\phi_1, \phi_2)_Q = \frac{\cos^2\phi_1\,\rho_1}{\cos^2\phi_1\,\rho_1 + \cos^2\phi_2\,\rho_2} P_{\psi_1} + \frac{\cos^2\phi_2\,\rho_2}{\cos^2\phi_1\,\rho_1 + \cos^2\phi_2\,\rho_2} P_{\psi_2}.$$

Assume $0 < \rho_1 < 1$ such that $S(\rho(\phi_1, \phi_2)) \neq 0$ and choose ϕ_1 and ϕ_2 such that $\cos^2\phi_1\,\rho_1 = \cos^2\phi_2\,\rho_2$. This gives $\rho(\phi_1, \phi_2)_Q = 1/2Q$ with $S(\rho(\phi_1, \phi_2)_Q) = \ln 2 > S(\rho(\phi_1, \phi_2))$ whenever $\rho_1 \neq 1/2$. On the other hand, some easy estimates show that indeed $F(\rho(\phi_1, \phi_2), Q) \leq S(\rho(\phi_1, \phi_2))$ holds for all ϕ_1 and ϕ_2.

This example also shows that in general neither ρ_Q nor $S(\rho_Q)$ for $Q\rho Q = 0$ may be defined by a limiting procedure. In fact, we may let ϕ_1 and ϕ_2 tend to $\pi/2$ in such a way that $\cos^2\phi_2/\cos^2\phi_1$ tends to an arbitrary constant ≥ 0 showing that in the limit for $\rho(\phi_1, \phi_2)_Q$ we may obtain an arbitrary convex combination of P_{ψ_1} and P_{ψ_2} and hence an arbitrary value between 0 and $\ln 2$ for the entropy. By the convexity of the relative entropy we also have

$$F(\rho, Q) + F(\rho, \mathbb{I} - Q) \leq S(E_{\{Q, \mathbb{I}-Q\}}(\rho)).$$

On the other hand, in general $F(\rho, Q) + F(\rho, \mathbb{I} - Q)$ is in general not bounded above by $S(\rho)$. Indeed, consider the following

Example 2.2. *Let the set-up be as in Example 2.1. With repsect to this basis let*

$$\rho(\kappa)=\frac{1}{4}\begin{pmatrix}1&0&0&\kappa\\0&1&\kappa&0\\0&\kappa&1&0\\\kappa&0&0&1\end{pmatrix}$$

with $0\le\kappa\le 1$. *The two two-fold degenerate eigenvalues are* $1/4(1\pm\kappa)$. *This gives* $F(\rho,Q)+F(\rho,\mathbb{I}-Q)=\ln 2$ *whereas* $S(\rho(\kappa))=\ln 2-1/2((1+\kappa)\ln(1+\kappa)+(1-\kappa)\ln(1-\kappa))<\ln 2$, *whenever* $0<\kappa$.

The quantity

$$\Delta S(\rho)=S(\rho)-S(\rho|\rho)\ge-\sum_i \dim P_i\;\rho_i\ln(\dim P_i\;\rho_i) \tag{18}$$

is of special interest. The inequality is a consequence of $\dim P_i\;\rho_i\le 1$ and again implies that the right hand side is non-negative and equal to zero if and only if $\rho=(1/\dim\;\mathbb{I})\cdot\mathbb{I}$ such that $\Delta S(\rho)>0$ unless $\rho=(1/\dim\;\mathbb{I})\mathbb{I}$. In more detail the inequality in (**??**) may also be written as follows. Let $S_{cl}(\underline{p})\ge 0$ be the classical entropy for the probability distribution $\underline{p}=(p_1,p_2,...p_n),p_k\ge 0,\sum_k p_k=1$

$$S_{cl}(\underline{p})=-\sum_{k=1}^{n}p_k\;\ln p_k.$$

such that in particular

$$S(\rho)=S_{cl}(\underline{p}(\rho)) \tag{19}$$

with

$$\underline{p}(\rho)=(\underbrace{\rho_1,..,\rho_1}_{\dim P_1},\underbrace{\rho_2,..,\rho_2}_{\dim P_2},....). \tag{20}$$

(**??**) may now be rewritten as

$$0\le S_{cl}(\underline{\hat{p}}(\rho))\le\Delta S(\rho) \tag{21}$$

with

$$\underline{\hat{p}}(\rho)=(\dim P_1\;\rho_1,\dim P_2\;\rho_2,...)$$

and where $S_{cl}(\underline{\hat{p}})=0$ if and only if ρ is a pure state. $\Delta S(\rho)$ is easily shown to be continuous in ρ and is obviously bounded above by $\ln\dim\;\mathbb{I}=\ln\dim\mathcal{H}=S(\rho=(1/\dim\;\mathbb{I})\cdot\mathbb{I})$. It would be interesting to find its maximum in ρ for fixed dimension of $\dim\mathcal{H}$. Note also that

$$S(\rho)=S_{cl}(\underline{p}(\rho))=S_{cl}(\underline{\hat{p}}(\rho))+\sum_i\dim P_i\;\sigma_i\ln\dim P_i\ge S_{cl}(\underline{\hat{p}}(\rho)) \tag{22}$$

with equality if and only if $\dim P_i=1$ for all i with $\sigma_i>0$. We will discuss $\Delta S(\rho)$ below when we compare $S(\rho|\sigma)$ with Shannon's conditional entropy.

We turn to the proof of (**??**). First recall that $F(\rho,Q)$ is continuous in ρ (and Q). Hence it suffices to consider the case $\rho>0$ which implies that $Q\rho Q\ne 0$ for all $Q\ne 0$.

Since $F(\rho, Q) = 0$ for $Q = 0$ and $\dim Q = 1$ and since $F(\rho, \mathbb{I}) = S(\rho)$ it suffices to consider the case $1 < \dim Q < \dim \mathbb{I}$.

Now $\mathcal{U}(\mathcal{H})$ operates transitively and continuously on the Grassmannian of all n-dimensional subspaces of $\mathcal{H}$, $(1 \leq n \leq \dim \mathcal{H})$. For each n this space is therefore compact and homeomorphic to the set of all projections of dimension n. Obviously on this set $\mathcal{U}(\mathcal{H})$ operates, again continuously, via $U : P \mapsto UPU^{-1}$. By (**??**)

$$F_n(\rho) = \sup_{Q:\dim Q=n} F(\rho, Q) = \sup_{U:U\in\mathcal{U}(\mathcal{H})} F(\rho, UQ_0U^{-1}) = \sup_{U:U\in\mathcal{U}(\mathcal{H})} F(U\rho U^{-1}, Q_0), \quad (23)$$

which is finite for each n. Here Q_0 is any orthogonal projection with $\dim Q_0 = n$. In particular we may choose Q_0 such that $F_n(\rho) = F(\rho, Q_0)$. Consider the one-parameter unitary subgroup $U(t) = \exp(-itK)$, where K is an arbitrary self-adjoint operator on $\mathcal{H}$. Then we must have $f_K(t) = F(U(t)\rho U(-t), Q_0) \leq F(\rho, Q_0) = f_K(t = 0)$ for all t and all s.a. K. Now it is well known that for any one parameter family of strictly positive operators $A(t)$ which is differentiable in t one has

$$\frac{d}{dt}\operatorname{Tr}(A(t) \ln A(t)) = \operatorname{Tr}((\mathbb{I} + \ln A(t))\frac{d}{dt}A(t)).$$

Recalling the assumption $\rho > 0$ such that $Q_0\rho Q_0 > 0$ when restricted to the subspace $Q_0\mathcal{H}$, it is easy to see that $f_K(t)$ is also differentiable in t at $t = 0$ and

$$\begin{aligned}\tfrac{d}{dt} f_K(t)|_{t=0} \;&=\; -i\ \operatorname{Tr}((\mathbb{I} + \ln Q_0\rho Q_0)Q_0[K, \rho]Q_0) \\ &\qquad + i\ (1 + \ln \operatorname{Tr}(Q_0\rho Q_0)) \operatorname{Tr} Q_0[K, \rho]Q_0 \\ &=\; i\ \operatorname{Tr}([\rho, \ln \operatorname{Tr}(Q_0\rho Q_0) \cdot Q_0 - Q_0(\ln Q_0\rho Q_0)Q_0]K).\end{aligned} \quad (24)$$

By definition of Q_0 we must have $d/dt f_K(t = 0) = 0$ for all K. But then (**??**) implies that ρ commutes with $B = Q_0B = BQ_0$ given as

$$B = \ln \operatorname{Tr}(Q_0\rho Q_0) \cdot Q_0 - Q_0(\ln Q_0\rho Q_0)Q_0.$$

This in turn implies that ρ commutes with Q_0 itself, which is easy to see. Indeed, use the spectral representation $Q_0\rho Q_0 = \sum_k \rho'_k Q'_k$ with $Q'_k \leq Q_0, \dim Q'_k = 1$ and $\sum_k Q'_k = Q_0$ to write B as

$$B = \sum_k (\ln(\sum_l \rho'_l) - \ln \rho'_k)Q'_k.$$

Now write any $\psi \in Q_0\mathcal{H}$ as $\psi = \sum_k a_k\psi_k$, where ψ_k is a unit vector in $Q'_k\mathcal{H}$. Set

$$\phi = \sum_k \frac{a_k}{(\ln(\sum_l \rho'_l) - \ln \rho'_k)}\psi_k \in Q_0\mathcal{H}.$$

ϕ is well defined since $\sum_l \rho'_l \neq \rho'_k$ for every k. This follows from our assumption $n > 1$, the fact that $\ln x$ is strictly monotonic in x and that $\rho'_k > 0$ for all k, since $Q_0\rho Q_0$ when restricted to $Q_0\mathcal{H}$ is strictly positive. By construction $\psi = B\phi$ such that $\rho\psi = \rho B\phi = B\rho\phi = Q_0B\rho\phi \in Q_0\mathcal{H}$. Thus ρ leaves $Q_0\mathcal{H}$ invariant and hence commutes with Q_0, as was claimed. But then we have $\rho = Q_0\rho Q_0 + (\mathbb{I} - Q_0)\rho(\mathbb{I} - Q_0)$ which implies

$$S(\rho) = -\operatorname{Tr}(Q_0\rho Q_0 \ln Q_0\rho Q_0) - \operatorname{Tr}((\mathbb{I} - Q_0)\rho(\mathbb{I} - Q_0) \ln(\mathbb{I} - Q_0)\rho(\mathbb{I} - Q_0)).$$

This gives

$$\begin{aligned} S(\rho) \;&= F(\rho, Q_0) \;- \operatorname{Tr}((\mathbb{I} - Q_0)\rho(\mathbb{I} - Q_0) \ln(\mathbb{I} - Q_0)\rho(\mathbb{I} - Q_0)) \\ &\qquad - \operatorname{Tr} Q_0\rho Q_0 \ln \operatorname{Tr} Q_0\rho Q_0.\end{aligned} \quad (25)$$

The two last terms in (**??**), however, are non-negative. This concludes the proof of the claim (**??**). To prove the second part of Lemma **??**, we observe that the last two terms in (**??**) vanish exactly when $(\mathbb{I}-Q_0)\rho(\mathbb{I}-Q_0)=0$. But this contradicts the assumption $\rho>0$ and $\dim Q_0<\dim\ \mathbb{I}$, the case $Q=\mathbb{I}$ having been discussed previously. This completes the proof of Lemma **??**.

3 Comparison with the classical case

In this section we provide a comparison with the classical theory of Shannon (see [**?**] and for expositions e.g. [**?**, **?**, **?**]). For the convenience of the reader and in order to establish notation we recall the basic facts. Let $\{\Omega,\mu\}$ be a probability space. Furthermore let $X=\{X_\alpha\}$ and $Y=\{Y_\beta\}$ be any two partitions (up to measure zero) of Ω into disjoint subsets of non-zero measure. For simplicity we will assume these partitions to be finite, i.e. we choose the indices α and β to be in the range $1\le\alpha\le n, 1\le\beta\le m$. Set $\underline{p}(X)=\{p_\alpha\}$ with $p_\alpha=\mu(X_\alpha)>0$ and $\underline{p}(Y)=\{q_\beta\}$ with $q_\beta=\mu(Y_\beta)>0$ such that $\sum_\alpha p_\alpha=1$ and $\sum_\beta q_\beta=1$. Here and in what follows α is an index referring to X and β to Y. Then $H(X)=-\sum_\alpha p_\alpha\ln p_\alpha\ge 0$ and similarly $H(Y)=-\sum_\beta q_\beta\ln q_\beta\ge 0$ is Shannon's entropy. Actually Shannon used $\log_2$ instead of ln adapting to the situation where information is coded in bits, but this is not relevant for our purpose. Since $H(X)=S_{cl}(\underline{p}(X))$ this concept of information theory relates to the concept of entropy in classical statistical mechanics. Shannon's conditional entropy is now given as follows. Let

$$p_{\alpha|\beta}=\frac{\mu(X_\alpha\cap Y_\beta)}{\mu(Y_\beta)},\qquad q_{\beta|\alpha}=\frac{\mu(Y_\beta\cap X_\alpha)}{\mu(X_\alpha)}$$

be conditional probabilities associated to X and Y (i.e. $p_{\alpha|\beta}$ is the probability that X_α will happen, given that Y_β has happened). Obviously

$$p_{\alpha|\beta}\,q_\beta=q_{\beta|\alpha}\,p_\alpha(=\mu(A_\alpha\cap B_\beta)) \tag{26}$$

for all α,β, which is called Bayes rule for $p_{\alpha|\beta}$ and $q_{\beta|\alpha}$. Let $\underline{p}_\beta=(p_{1|\beta},p_{2|\beta},..,p_{n|\beta})$ and $\underline{q}_\alpha=(q_{1|\alpha},q_{2|\alpha},....q_{m|\alpha})$, such that

$$\underline{p}=\sum_{\beta=1}^{m}q_\beta\underline{p}_\beta,\qquad \underline{q}=\sum_{\alpha=1}^{n}p_\alpha\underline{q}_\alpha. \tag{27}$$

Shannon's conditional entropy is now defined as

$$H(X|Y)=\sum_{\beta=1}^{m}q_\beta S_{cl}(\underline{p}_\beta) \tag{28}$$

and it satisfies

$$0\le H(X|Y)\le H(X). \tag{29}$$

We observe that the second inequality, called Shannon's inequality, is a consequence of the concavity of the function $\underline{p}\mapsto S_{cl}(\underline{p})$ and (**??**). It states that on average information on X is gained if Y is known. Also $0\le H(X,Y)=H(Y)+H(X|Y)$ is symmetric in X and Y and satisfies

$$H(Y)\le H(X,Y)\le H(X)+H(Y). \tag{30}$$

Actually $H(X,Y) = H(X \vee Y)$, where $\vee$ denotes the join of two partitions. The inequalities in (**??**) and (**??**) turn into equalities if the following conditions hold. X and Y are said to be independent if $p_{\alpha|\beta} = p_\alpha$ holds for all α and β. This means that $\underline{p}_\beta$ is actually independent of β and equals $\underline{p}$ and $\underline{q}_\alpha$ is independent of α and equals $\underline{q}$. In particular $S_{cl}(\underline{p}_\beta) = S_{cl}(\underline{p})$ holds for all β and $S_{cl}(\underline{q}_\alpha) = S_{cl}(\underline{q})$ for all α. The second inequality in (**??**) and the second inequality in (**??**) (which are equivalent) are now equalities if and only if X and Y are independent. It follows from the fact that $S_{cl}(\underline{p})$ is strictly concave in $\underline{p}$. Secondly X is called a consequence of Y if to each α there is $\beta(\alpha)$ such that $p_{\alpha|\beta(\alpha)} = 1$. So this means that $p_{\alpha|\beta} = 0$ for all $\beta \neq \beta(\alpha)$ and hence $S_{cl}(\underline{p}_\beta) = 0$ for all β. Therefore the first inequality in (**??**) and equivalently the first inequality in (**??**) are equalities if and only if X is a consequence of Y. In particular

$$H(X|X) = 0, \tag{31}$$

i.e. $H(X,X) = H(X)$.

With this brief review of Shannon's theory we turn to a comparison with our quantum mechanical construction. Obviously (**??**) corresponds to (**??**) when we let X correspond to ρ and Y to σ. Note, however, the difference between (**??**) and (**??**). Moreover for the quantity $S(\rho,\sigma) = S(\sigma) + S(\rho|\sigma)$ we have the inequalities

$$S(\sigma) \leq S(\rho,\sigma) \leq S(\rho) + S(\sigma), \tag{32}$$

which correspond to (**??**). $S(\rho,\sigma)$ is in general not symmetric in ρ and σ . To see this consider commuting ρ and σ. Then we have

$$S(\rho|\sigma) = -\sum_{j,i} \dim\ Q_j\, \sigma_j\, \rho_i\, \dim(P_i Q_j) \ln \frac{\rho_i}{\mathrm{Tr}(\rho Q_j)}. \tag{33}$$

We remark that if $\mathrm{Tr}(\rho Q_j) = 0$ for a fixed j then $\mathrm{Tr}(P_i Q_j) = 0$ for all i. Also (**??**) is a special case of (**??**). (**??**) shows that even in the commutative case $S(\rho,\sigma)$ is not symmetric in ρ and σ. So this implies that in the commutative case $S(\rho|\sigma)$ does not reduce to $H(X|Y)$ for *any* choice of $X = X(\rho)$ and $Y = Y(\sigma)$ with $H(X) = S(\rho)$ and $H(Y) = S(\sigma)$. This lack of symmetry of $S(\rho,\sigma)$ is in contrast to the symmetry of its classical counterpart $H(X,Y)$, which has an important interpretation. The relation $H(X,Y) = H(Y,X)$ is equivalent to $H(Y)+H(X|Y) = H(X)+H(Y|X)$, a consequence of Bayes rule. But this means that on average the information on Y plus the information on X given Y is equal to the information on X plus the information on Y given X. It would be interesting to see whether this failure of symmetry for $S(\rho,\sigma)$ has a sensible interpretation in the context of the familiar *Alice and Bob* set-up in quantum information theory, see e.g. [**?**].

Finally consider

$$0 \leq S(\rho||\sigma) = S(\rho) + S(\sigma) - S(\rho,\sigma) = S(\rho) - S(\rho|\sigma) \leq S(\rho) \tag{34}$$

which corresponds to

$$0 \leq I(X||Y) = H(X) + H(Y) - H(X,Y) = H(X) - H(X|Y).$$

On average $0 \leq I(X||Y) \leq H(X)$ gives the information gain on X when knowing Y. Thus if there is no information content at all in Y, i.e. if Y is the trivial partition $\{\Omega\}$, then there is no information gain in X

$$I(X||Y = \{\Omega\}) = 0. \tag{35}$$

Thus (**??**) corresponds to (**??**) when rewritten as $S(\rho||\sigma = (1/\dim \mathbb{I})\mathbb{I}) = 0$. Therefore we also interpret the quantum mechanical analogue $S(\rho||\sigma)$ as a quantum information gain for ρ given σ and which by (**??**) can be at most $S(\rho)$. In particular the gain is maximal for all ρ, if all non-zero eigenvalues of σ are non-degenerate. The gain is also maximal if $\rho\sigma = 0$, since then $S(\rho|\sigma) = 0$, see Lemma **??** and the remark thereafter.

Finally $\Delta S(\rho)$ (see (**??**)) corresponds to $I(X||X)$ and describes the situation where ρ is conditioned on itself, $\sigma = \rho$. Then by (**??**) there is non-zero information gain unless ρ is pure (and then a gain is not necessary). In contrast to the classical situation, $I(X||X) = H(X)$, which gives complete information gain when X is conditioned on itself, there is complete information gain in the quantum case, $\Delta S(\rho) = S(\rho)$, if and only if all non-zero eigenvalues of σ are non-degenerate.

4 Attempts of alternative constructions

We conclude by addressing the natural question whether there is a quantity $S^?(\rho|\sigma)$ which shares more properties with Shannon's conditional entropy than the $S(\rho|\sigma)$ we have given. More precisely and by the arguments given in the preceding sections it would be desirable for $S^?(\rho|\sigma)$ to have (most of) the following properties

1. Invariance under the group $\mathcal{U}(\mathcal{H})$: $S^?(U\rho U^{-1}|U\sigma U^{-1}) = S^?(\rho|\sigma)$ for all $U \in \mathcal{U}(\mathcal{H})$ (compare (**??**)).
2. Bounds: $0 \leq S^?(\rho|\sigma) \leq S(\rho)$ for all ρ and σ with $S^?(\rho|\rho) = 0$ and $S^?(\rho|\sigma = (1/\dim \mathbb{I})\mathbb{I}) = S(\rho)$.
3. Classical equivalence with Shannon's conditional entropy.
4. Symmetry: $S^?(\rho,\sigma) = S(\sigma) + S^?(\rho|\sigma)$ is symmetric in ρ and σ.
5. Continuity of $S^?(\rho|\sigma)$ in ρ and in σ.
6. Concavity of $S^?(\rho|\sigma)$ in ρ and σ.

Note that $S(\rho|\sigma)$ fulfills condition 1, condition 2 apart from the property $S(\rho|\rho) = 0$, condition 5 up to a set of measure zero and condition 6 only with respect to ρ.

Both the equality requirements of condition 2 can never be satisfied simultaneously. Indeed, with the choice $\rho = \sigma = 1/\dim \mathcal{H}\,\mathbb{I}$ we should have both $S(1/\dim \mathcal{H}\,\mathbb{I}|1/dim\mathcal{H}\,\mathbb{I}) = 0$ and $S(1/\dim \mathcal{H}\,\mathbb{I}|1/dim\mathcal{H}\,\mathbb{I}) = \ln \dim \mathcal{H}$. Also the condition $S(\rho|\rho) = 0$ combined with $S(\rho|\sigma) \geq 0$ is incompatible with concavity of $S(\rho|\sigma)$ in ρ (condition 6). In fact, let $\rho = \lambda\,\rho_1 + (1-\lambda)\,\rho_2$, $0 < \lambda < 1$. But this gives $0 = S(\rho|\rho) \geq \lambda\, S(\rho_1|\rho) + (1-\lambda)S(\rho_2|\rho)$. Hence $S(\rho_1|\rho) = 0$ for all ρ_1 for which there is $\lambda > 0$ with $\lambda\,\rho_1 < \rho$. This condition is fulfilled for all ρ_1, whenever $\rho > 0$ (I owe these observations to H. Narnhofer).

Next let us look at the condition 3, by which we mean the situation where ρ and σ commute such that $S(\rho) = H(X)$, $S(\sigma) = H(Y)$ and $S^?(\rho|\sigma) = H(X|Y)$ holds for suitable $X = X(\rho)$ and $Y = Y(\sigma)$. Also the dependence of $X(\rho)$ and $Y(\sigma)$ on ρ and σ respectively should be non-trivial w.r.t. their eigenvalues. In particular condition 3 means that the symmetry condition 4 must hold at least when ρ and σ commute. In view of the destruction of quantum coherence when measurements are performed and due to the occurrence of the sum by which $S^?(\rho,\sigma)$ is defined, it is unclear to the author whether the symmetry condition 4 also should hold for non-commuting ρ and σ (see below , however, a construction of conditional entropy in terms of spectral resolutions of the identity below). It is natural to make the assumption on $X(\rho)$, that

$\mu(X_k) = p_k(\rho)$, see (**??**). Then it may be shown that the continuity condition and the classical equivalence condition are not compatible. The concavity condition in σ is at least intuitively desirable since taking convex combinations decreases conditioning, i.e. increases uncertainty, and hence should increase conditional entropy.

We would also like to point out another difference between the classical and the quantum case in the way we have presented it so far. In the classical case the conditioning Y is trivial when $Y = \{\Omega\}$, which means no information content and for which we have $H(Y) = 0$. Within the context of density matrices the only sensible candidate for a trivial conditioning is $\sigma = 1/\dim\mathcal{H}\,\mathbb{I}$, since this is the density matrix with no information content. Its von Neumann entropy, however, is maximal. Recall that we used this quantum notion of trivial conditioning in our discussion of the inequality $S(\rho|\sigma) \leq S(\rho)$ (see also the discussion following (**??**)). We note that several authors consider von Neumann's entropy not to be a good generalization of classical entropy (see e.g. [**?**], page 141). In fact, in classical theory finer partitions give rise to higher uncertainty and hence to larger classical entropy. This was the reason for the algebraic approach of Connes and Størmer and of Connes, Narnhofer and Thirring, in which a classical finer partitioning corresponds to a larger algebra. In particular the larger the algebra, the larger the entropy and similarly the larger the conditioning algebra the larger the conditional entropy.

We claim, however, that there is a way to reconcile this with von Neumann's entropy. Indeed, given a quantum system in the state ρ, the measurements one can perform without disturbing ρ are given by the observables (i.e. the self-adjoint operators) in $\mathcal{A}(\rho)$, which by definition is the $\star$-sub-algebra of $\mathcal{B}$ consisting of all elements in $\mathcal{B}$ which commute with ρ. In particular $\mathcal{A}(\rho = 1/dim\mathcal{H}\,\mathbb{I}) = \mathcal{B}$. In this sense again larger uncertainties correspond to larger algebras. In other words, the larger the entropy the more measurements on can perform without disturbing the system in the given state ρ. To be more precise, we introduce a partial ordering $\preceq$ on the set of all density matrices (which differs from the one introduced by Uhlmann, see e.g. [**?**]). By definition $\rho \preceq \sigma$ (σ is more mixed than ρ), if and only if a) $\underline{P} \leq \underline{Q}$ and b) $\operatorname{Tr} \rho\, Q_j = \operatorname{Tr} \sigma\, Q_j = \sigma_j \operatorname{Tr} Q_j$ holds for all j. It is easy to see that $\rho \preceq \sigma$ and $\sigma \preceq \tau$ implies $\rho \preceq \tau$ and that $\rho \preceq 1/\dim\mathcal{H}\,\mathbb{I}$ and $\rho \preceq \rho$ holds for all ρ. So whenever $\rho \preceq \sigma$ then condition a) implies $\mathcal{A}(\rho) \subseteq \mathcal{A}(\sigma)$ and a) and b) combined imply $S(\rho) \leq S(\sigma)$ by the concavity of the von Neumann entropy. Note, however, that the correspondence between ρ and $\mathcal{A}(\rho)$ is not one-to-one. In fact, $\mathcal{A}(\rho)$ only depends on the spectral resolution of the identity $\underline{P} = \underline{P}(\rho)$ associated to ρ and not on the eigenvalues ρ_i of ρ. Indeed, one has $\mathcal{A}(\rho) = E_{\underline{P}(\rho)}(\mathcal{B})$, as one may easily verify.

Returning to our discussion of conditions 1-6, there is a way out, however, if one considers spectral resolutions of the identity $\underline{P}$ instead of density matrices. It works as follows. First observe that the actual choice of the probability space $\{\Omega, \mu\}$ for Shannon's theory is irrelevant. What is relevant are the the sets of non-negative numbers $\underline{p} = \{p_\alpha\}$, $\underline{q} = \{q_\beta\}$, $\underline{p} \vee \underline{q} = \{p_{\alpha|\beta}\}$ and $\underline{q} \vee \underline{p} = \{q_{\beta|\alpha}\}$ subject to the following conditions of which the last one is Bayes rule

$$\sum_\alpha p_\alpha = \sum_\beta q_\beta = 1,\ \sum_\beta p_{\alpha|\beta} q_\beta = p_\alpha,\ \sum_\alpha q_{\beta|\alpha} p_\alpha = q_\beta,\ p_{\alpha|\beta}\, q_\beta = q_{\beta|\alpha}\, p_\alpha. \tag{36}$$

Note that then

$$\begin{aligned} \sum_\alpha p_{\alpha|\beta} &= \frac{1}{q_\beta} \sum_\alpha q_{\beta|\alpha} p_\alpha = 1 \\ \sum_\beta q_{\beta|\alpha} &= \frac{1}{p_\alpha} \sum_\beta p_{\alpha|\beta} q_\beta = 1. \end{aligned}$$

We consider these conditions (**??**), which mean independence of a particular realization of partitions X and Y on a probability space, the classical analogue of the relation (**??**). Setting $p_{\alpha,\beta} = p_{\alpha|\beta} q_\beta$ and $q_{\beta,\alpha} = q_{\beta|\alpha} p_\alpha$, Bayes rule gives $p_{\alpha,\beta} = q_{\beta,\alpha}$. We will therefore write $H(X|Y) = H(\underline{p}|\underline{q})$ by a slight abuse of notation since all the data $\underline{p}$, $\underline{q}$, $\underline{p} \vee \underline{q}$ and $\underline{q} \vee \underline{q}$ in (**??**) are necessary for a specification of $H(X|Y)$. But given these data it makes sense to say that $\underline{p}$ is a consequence of $\underline{q}$ or that $\underline{p}$ and $\underline{q}$ are independent.

Now let $\tau = 1/\dim \mathcal{H} \operatorname{Tr}$ denote the normalized trace, i.e. $\tau(\mathbb{I}) = 1$. For any two spectral resolutions $\underline{P}$ and $\underline{Q}$ let $p_i = \tau(P_i)$, $q_j = \tau(Q_j)$, $p_{i|j} = \tau(P_i Q_j)/\tau(Q_j)$, $q_{j|i} = \tau(Q_j P_i)/\tau(P_i)$. Note that by definition all P_i and all Q_j are non-zero projections. The conditions (**??**) are obviously satisfied. We then set $H(\underline{P}) = S_{cl}(\underline{p}), H(\underline{Q}) = S_{cl}(\underline{q})$,such that $H(\underline{Q} = \{\mathbb{I}\}) = \ln \dim \mathcal{H}$ and finally $H(\underline{P}|\underline{Q}) = H(\underline{p}|\underline{q})$, such that $0 \le H(\underline{P}|\underline{Q}) \le H(\underline{P})$ as desired. Note that now $H(\underline{P}|\underline{Q})$ is completely specified by $\underline{P}$ and $\underline{Q}$. Also $H(\underline{P}, \underline{Q}) = H(\underline{Q}) + H(\underline{P}|\underline{Q})$ is symmetric in $\underline{P}$ and $\underline{Q}$.

It is easy to see that $\underline{p}$ is a consequence of $\underline{q}$ if and only if $\underline{Q} \le \underline{P}$ such that $H(\underline{P}|\underline{Q}) = 0$ if and only if $\underline{Q} \le \underline{P}$. Similarly $\underline{p}$ and $\underline{q}$ are independent if and only if $\underline{P} = \{\mathbb{I}\}$ or $\underline{Q} = \{\mathbb{I}\}$. Therefore, whenever $H(\underline{P}) \neq 0$, $H(\underline{P}|\underline{Q}) = H(\underline{P})$ if and only if $\underline{Q} = \{\mathbb{I}\}$, which in this context is the trivial conditioning and for which the entropy is zero in contrast to our construction in terms of density matrices. Finally we set $Ad\, U \underline{Q} = \{U\, Q_i\, U^{-1}\}$ for any $\underline{Q}$ and any unitary U. Then obviously $H(Ad\, U \underline{P} | Ad\, U \underline{Q}) = H(\underline{P}|\underline{Q})$ (compare condition 1).

Since the P_i's and the Q_j's need not commute, this construction is a non-commutative version of Shannon's conditional entropy in (commutative) classical probability theory. Thus a classical partition X is replaced by a spectral resolution of the identity $\underline{P}$, which in turn corresponds to the $\star$-algebra $E_{\underline{P}}(\mathcal{B})$ and which is abelian if and only if each P_i is one-dimensional. The choice $\underline{Q} = \{\mathbb{I}\}$ giving maximal entropy $H(\underline{Q})$ and maximal conditional entropy $H(\underline{P}|\underline{Q})$ corresponds to the maximal algebra $E_{\underline{Q}=\{\mathbb{I}\}}(\mathcal{B}) = \mathcal{B}$. Our construction of $H(\underline{P}|\underline{Q})$ differs from the construction in [**?**, **?**].

We might have defined the conditional entropy of two density matrices ρ and σ by $H(\underline{P}(\rho)|\underline{Q}(\sigma))$. Conditions 1,2 and 4 are then satisfied but not condition 5 and condition 3, since the dependence on the eigenvalues of ρ and σ drops out. We conjecture that condition 6 is also not satisfied.

Acknowledgements: The author would like to thank M. Karowski, H. Narnhofer, M.A. Nielsen, M. Schmidt and E. Størmer for helpful remarks.

References

[1] H. Barnum, M.A. Nielsen and B.W. Schumacher, "Information Transmission through a Noisy Quantum Channel", *Phys. Rev. A* **57**, 4153 – 4175 (1998).

[2] F. Benatti, *Deterministic Chaos in Infinite Quantum Systems*, Trieste lecture notes in Physics, Springer, Berlin, 1993.

[3] F. Bloch, " Zur Strahlungsdämpfung in der Quantenmechanik", *Phys. Zeitschrift* **29**, 58 – 66 (1928).

[4] H.J. Borchers, "On the Structure of the Algebra of Field Operators", *Nuovo Cimento* **24**, 214 – 236 (1992).

[5] N.J. Cerf and C. Adami, "Negative Entropy and Information in Quantum Mechanics", *Phys. Rev. Lett.* **79**, 5194 – 5197 (1997).

[6] A. Connes, H. Narnhofer and W. Thirring, " Dynamic Entropy of $C^\star$-algebras and von Neumann Algebras", *Commun. Math. Phys.* **112**, 691 – 719 (1987).

[7] A. Connes and E. Størmer, "The Entropy for Automorphisms of II_1 von Neumann Algebras", *Acta Mathematica* **134**, 289 – 306 (1975).

[8] R.G. Gallagher, *Information Theory and Reliable Communication*, Wiley, New York, 1968.

[9] F.M. Goodman, P. de la Harpe and V.F.R. Jones, *Coxeter Graphs and Towers of Algebras*, Springer, New York, 1989.

[10] L.D. Landau and E.M. Lifschitz, *Quantum Mechanics*, 3^{rd} edition, Pergamon, Oxford (1992).

[11] L.D. Landau, "Das Dämpfungsproblem in der Wellenmechanik", *Zeitschrift für Physik* **45**, 430 – 441 (1927).

[12] L.B. Levitin: " Quantum Generalization of Conditional Entropy and Information" in: C.P.Williams (ed.): Quantum Computing and Quantum Communication, *Lecture Notes in Computer Science* **1509**, 269 –275 (1999).

[13] E. Lieb, "Some Convexity and Subadditive Properties of Entropy", *Bull. Am. Math. Soc.* **81**, 1 – 13 (1975).

[14] S. Lloyd, "The Capacity of the Noisy Quantum Channel", *Phys. Rev. A* **55**, 1613 – 1622 (1997).

[15] R.J. McEliece, *The Theory of Information and Coding*, Addison-Wesley, Reading, 1977.

[16] J. von Neumann, "Wahrscheinlichkeitstheoretischer Aufbau der Quantenmechanik", *Gött. Nachr.*, 245 – 272 (1927).

[17] J. von Neumann, "Thermodynamik quantenmechanischer Grössen", *Gött. Nachr.*, 273 – 291 (1927).

[18] M.A. Nielsen, " Quantum Information Theory", available from `http://theory.caltech.edu/ mnielsen/phd/`

[19] M.A. Nielsen and I.L. Chuang, *Quantum Computation and Quantum Information*, Cambridge University Press, Cambridge, 2000.

[20] M. Ohya and D. Petz, *Quantum Entropy and its Use*, Springer, Berlin, 1993.

[21] J. Preskill, *A Course on Quantum Computation*; available from `http://www.theory.caltech.edu/people/preskill/ph_229`.

[22] B.W. Schumacher and M.A. Nielsen, "Quantum Data Processing and Error Correction", *Phys. Rev. A* **54**, 2629 – 2635 (1996).

[23] C.E. Shannon, "A Mathematical Theory of Information", *Bell System Tech. J.* **27**, 379 – 423, 623 – 656 (1948), reprinted in: C.E. Shannon and W. Weaver, *The Mathematical Theory of Information*, The University of Illinois Press, Urbana, 1964.

[24] Y. Sinai (Ed.), *Dynamical Systems II, Ergodic Theory with Applications to Dynamical Systems and Statistical Mechanics*, Vol. 2 of *Encyclopaedia of Mathematical Sciences*, Editor-in-chief: R.V. Gamkrelidze, Springer, Berlin, 1989.

[25] F.Topsøe, *Informationstheorie*, B.G. Teubner, Stuttgart, 1974.

[26] A. Uhlmann, "Relative Entropy and the Wigner-Yanase-Dyson-Lieb Concavity in an Interpolation Theory", *Commun. Math. Phys.* **54**, 21 – 32 (1977).

[27] A. Wehrl, "General Properties of Entropy", *Rev. Mod. Phys.* **50**, 221 – 260 (1978).