Extended quantum conditional entropy and quantum uncertainty inequalities

Rupert L. Frank\textsuperscript{1} and Elliott H. Lieb\textsuperscript{1,2}

\textsuperscript{1}Department of Mathematics, Princeton University, Princeton, NJ 08544, USA
\textsuperscript{2}Department of Physics, Princeton University, P. O. Box 708, Princeton, NJ 08542, USA

Quantum states can be subjected to classical measurements, whose incompatibility, or uncertainty, can be quantified by a comparison of certain entropies. There is a long history of such entropy inequalities between position and momentum. Recently these inequalities have been generalized to the tensor product of several Hilbert spaces and we show here how their derivations can be shortened to a few lines and how they can be generalized. All the recently derived uncertainty relations utilize the strong subadditivity (SSA) theorem; our contribution relies on directly utilizing the proof technique of the original derivation of SSA.

PACS numbers: 03.67.-a, 03.67.Mn, 03.67.Hk, 03.65.Ta

A celebrated inequality of Maassen–Uffink\textsuperscript{1}, based on earlier work in\textsuperscript{2,3}, relates the ‘classical’ entropy of a quantum density matrix in two different bases and shows that although one of them could be very small, the sum of the two is bounded below by a positive constant. The word ‘classical’ refers to that although one of them could be very small, the sum of the two is bounded below by a positive constant. The analogous inequalities become more difficult mathematically because of the well-known entropy of the above spaces. If $\rho_{123}$ is defined on $\mathcal{H}_{123}$ then there is a natural $\rho_{12}$ on $\mathcal{H}_{12}$ given by $\rho_{12} = \text{Tr}_3 \rho_{123}$, where $\text{Tr}_3$ is the partial trace over $\mathcal{H}_3$. In general, entropy is defined by $S(\rho) = -\text{Tr}_\mathcal{H} \rho \ln \rho$, where $\mathcal{H}$ is one of the above spaces. If $\rho$ and $\mathcal{H}$ are understood, then $S_{12} = S(\rho_{12})$ and $S_1 = S(\rho_1)$. Conditional entropy is defined by

$$S(1|2) = S_{12} - S_2 = S(\rho_{12}) - S(\rho_2)$$

and, as shown in\textsuperscript{3}, this is a concave function of $\rho_{12}$. This concavity is mathematically equivalent to SSA, and we use it in the Lemma below.

A classical measurement on $\mathcal{H}_1$ is defined by a sequence of operators $A_1, A_2, \text{etc.}$ on $\mathcal{H}_1$ such that $\sum_j A_j^* A_j = 1$. We will need two measurements in our discussion, so we will need a second sequence $B_1, B_2, \text{etc.}$ on $\mathcal{H}_1$ such that $\sum_k B_k^* B_k = 1$. At the end of this paper we will let the indices $j$ and $k$ be continuous and the sums become integrals, but for simplicity we stay with sums for now.
A measurement $A$ is called ‘rank-one’ if every $A_j$ is a rank-one matrix, namely, $A_j = [a_j_j](a_j)$. The classical/quantum conditional entropy corresponding to $A$ and $\rho$ is

$$H(1^A|2) = - \sum_j \text{Tr}_2 \left( \text{Tr}_1 A_j \rho_{12} A_j^* \right) \ln \left( \text{Tr}_1 A_j \rho_{12} A_j^* \right) - S(\rho_2).$$

Note where the summation sits. (This is not the conditional entropy of $\sum_j A_j \rho A_j^*$. The quantity $H(1^B|2)$ is defined similarly.

**2–space theorem.**— For any $\rho_{12}$,

$$H(1^A|2) + H(1^B|2) \geq S(1|2) - 2 \ln c_1, \quad (1)$$

where $c_1 = \sup_{j,k} \sqrt{\text{Tr}_1 B_k A_j^* A_j B_k^*}$.

In the special case that $\rho_{12} = \rho_1 \otimes \rho_2$ is a product state one easily sees that (1) reduces to the 1–space theorems in [1, 7, 12], which extend the Maassen–Uffink theorem.

Coles et al. [6] prove (1) if either $A$ or $B$ is rank-one and if $c_1$ is replaced by the smaller, better, number

$$c_\infty = \sup_{j,k} \sqrt{\| B_k A_j^* A_j B_k^* \|_\infty}.$$ 

They state correctly that the theorem cannot hold for general $A$ and $B$ in the $c_\infty$ version. However, if either $A$ or $B$ is rank-one, then $c_1 = c_\infty$. So our improvement of Coles et al. consists of eliminating the rank-one condition and using $c_1$.

The norm $\| B_k A_j^* A_j B_k^* \|_\infty$ is the largest eigenvalue of $B_k A_j^* A_j B_k^*$, which is also the largest eigenvalue of $\sqrt{B_k A_j^* A_j B_k^*}$, and thus coincides with $\| A_j^* A_j \|_\infty^{1/2}$. This shows that our $c_\infty^2$ is, indeed, the same as that in [6, Eq. (57)].

**3–space theorem.**— For any $\rho_{123}$,

$$H(1^A|2) + H(1^B|3) \geq -2 \ln c_\infty.$$ 

We shall prove the 3–space theorem by first proving the following theorem which is not conveniently expressed in the $H(\cdot|\cdot)$ notation.

**$2^{1/2}$–space theorem.**— For any $\rho_{12}$,

$$H(1^A|2) - \sum_k \text{Tr}_{12} B_k \rho B_k^* \ln B_k \rho B_k^* - S(\rho_{12}) \geq -2 \ln c_\infty.$$ 

This looks like another 2–space theorem but, as we explain below, it is really the 3–space theorem applied to a pure state (i.e., rank-one) $\rho_{123}$. Thus our name.

The key inequality behind the proofs of all our three theorems is the three operator generalization of the Golden–Thompson inequality from [13], which was also the key ingredient in the proof of SSA in [9]. It states that for non-negative operators $X, Y, Z$,

$$\text{Tr} e^{\ln X - \ln Y + \ln Z} \leq \int_0^\infty dt \text{Tr} X(Y + t)^{-1} Z(Y + t)^{-1}. \quad (2)$$

Note that in the special case $Z = 1$ this reduces to the classical Golden–Thompson inequality. This inequality was a byproduct of the proof of concavity of the general-

alized Wigner–Yanase skew information.

We shall also need the Gibbs variational principle (equivalent to the Peierls–Bogolubov inequality),

$$\text{Tr} \rho h - S(\rho) \geq - \ln \text{Tr} e^{-h} \quad (3)$$

for any density matrix $\rho$ and any self-adjoint $h$, and the Davis operator Jensen inequality [14],

$$\sum_j A_j^* (\ln K_j) A_j \leq \ln \left( \sum_j A_j^* K_j A_j \right) \quad (4)$$

for any positive operators $K_j$ and any $A_j$ with $\sum_j A_j^* A_j = 1$.

**Proof of the 2–space theorem.** We use (1) for both $A$ and $B$ to bound

$$H(1^A|2) + H(1^B|2) - S(1|2) \geq \text{Tr}_{12} \rho_{12} h - S(\rho_{12}), \quad (5)$$

with the operator

$$h = \ln \rho_2 - \ln \sum_j A_j^* A_j \left( \text{Tr}_1 A_j \rho_{12} A_j^* \right) - \ln \sum_k B_k^* B_k \left( \text{Tr}_1 B_k \rho_{12} B_k^* \right).$$

We do not need to invoke (4) when $A$ and $B$ are rank-one; in that case (4) is an equality and the proof simplifies further.

Thus, by (5) and (3),

$$H(1^A|2) + H(1^B|2) - S(1|2) \geq - \ln \text{Tr}_{12} e^{-h},$$

and it remains to show that $\text{Tr}_{12} e^{-h} \leq c_\infty^2$. Now comes the crucial step! We use (2) to bound

$$\text{Tr}_{12} e^{-h} \leq \int_0^\infty dt \sum_{j,k} \text{Tr}_{12} C_{j,k}(t)$$

with

$$C_{j,k}(t) = A_j^* A_j \left( \text{Tr}_1 A_j \rho_{12} A_j^* \right) (\rho_2 + t)^{-1} \times B_k^* B_k \left( \text{Tr}_1 B_k \rho_{12} B_k^* \right) (\rho_2 + t)^{-1} = A_j^* A_j B_k^* B_k D_{j,k}(t).$$

Here, $D_{j,k}(t)$ is the operator on $\mathcal{H}_2$ given by

$$\left( \text{Tr}_1 A_j \rho_{12} A_j^* \right) (\rho_2 + t)^{-1} \left( \text{Tr}_1 B_k \rho_{12} B_k^* \right) (\rho_2 + t)^{-1}.$$
Thus,
\[ \text{Tr}_{12} C_{j,k}(t) = (\text{Tr}_1 A_{j}^{*} A_{j} B_{k}^* B_{k}^*) \text{Tr}_2 D_{j,k}(t) \leq c_{2}^{2} \text{Tr}_2 D_{j,k}(t). \]

We next note that \( \sum_j \text{Tr}_1 A_j \rho_{12} A_j^* = \sum_j \text{Tr}_1 A_j^* A_j \rho_{12} = \rho_2 \), and similarly for the \( k \) sum, and obtain
\[ \sum_{j,k} \text{Tr}_2 D_{j,k}(t) = \text{Tr}_2 \rho_2^2 (\rho_2 + t)^{-2}. \]

Thus, since \( \int_{0}^{\infty} dt (\rho_2 + t)^{-2} = \rho_2^{-1} \),
\[ \int_{0}^{\infty} dt \sum_{j,k} \text{Tr}_2 D_{j,k}(t) = \text{Tr}_2 \rho_2 = 1. \]

This completes the proof of the 2–space theorem. QED

We shall now show that the 3–space theorem is a corollary of the 2\( \frac{1}{2} \)–space theorem, so that the proof of the 2\( \frac{1}{2} \)–space theorem will finish everything. We use the following:

**Lemma.**—— \( H(1^{A}|2) \) is a concave function on the set of non-negative operators \( \rho_{12} \) on \( \mathcal{H}_{12} \).

**Proof of the Lemma.**—— The idea is to view the sum over \( j \) as the trace over an auxiliary space \( \mathcal{K} \) of a matrix that happens to be diagonal in this space, and to apply concavity of the conditional entropy \( \rho_j \) in \( \mathcal{H}_2 \otimes \mathcal{K} \). The details are as follows: Let \((\epsilon_j)\) be an orthonormal basis of \( \mathcal{K} \) and consider the operator \( \Gamma = \sum_j (\text{Tr}_1 A_j \rho_{12} A_j^*) \otimes \langle \epsilon_j | \epsilon_j \rangle \) on \( \mathcal{H}_2 \otimes \mathcal{K} \). As in the proof of the 2–space theorem we have \( \Gamma_2 = \text{Tr}_K \Gamma = \rho_2 \) and therefore
\[ H(1^{A}|2) = S(\Gamma) - S(\Gamma_2). \]

This is the conditional entropy of \( \Gamma \) with respect to \( \mathcal{H}_2 \). Since \( \rho_{12} \mapsto \Gamma \) is linear, the asserted concavity follows from the fact that conditional entropy is concave, as shown in \( \rho_j \) (Thm. 1). QED

**Proof of the 3–space theorem.**—— It follows from the Lemma that \( H(1^{A}|2) + H(1^{B}|3) \) is a concave function of \( \rho_{123} \). Thus, for the proof we may assume that \( \rho_{123} \) is a pure state (rank one). In that case \( S(\rho_1) = S(\rho_{12}) \) and, since \( B_k \rho_{123} B_k^* \) is pure as well,
\[ H(1^{B}|3) = - \sum_k \text{Tr}_2 B_k^* \rho_2 B_k \ln B_k \rho_2 B_k^* - S(\rho_{12}). \]

This reduces the inequality of the 3–space theorem to that of the 2\( \frac{1}{2} \)–space theorem. QED

**Proof of the 2\( \frac{1}{2} \)–space theorem.**—— The proof runs very parallel to that of the 2–space theorem. Namely, the right side in the desired inequality is bounded from below by \( \text{Tr}_{12} \rho_2 \hat{h} - S(\rho_{12}) \), where now
\[ \hat{h} = \ln \rho_2 - \ln \sum_j A_j^* A_j (\text{Tr}_1 A_j \rho_{12} A_j) \]
\[ - \ln \sum_k B_k^* B_k \rho_{12} B_k^* B_k. \]

Here we used \( \rho \). After applying \( \rho \) as before, everything is reduced to showing \( \text{Tr}_{12} e^{-\hat{h}} \leq c_{2}^{2} \). The crucial ingredient is again \( \rho \) which now leads to
\[ \text{Tr}_{12} e^{-\hat{h}} \leq \int_{0}^{\infty} dt \sum_{j,k} \text{Tr}_{12} \tilde{C}_{j,k}(t) \]

with
\[ \tilde{C}_{j,k}(t) = A_j^* A_j (\text{Tr}_1 A_j \rho_{12} A_j)^{-1} \]
\[ \times B_k^* B_k \rho_{12} B_k^* B_k (\rho_2 + t)^{-1}. \]

At this point the proof diverges somewhat from that of the 2–space theorem. Namely, by cyclicity of the trace we write
\[ \text{Tr}_{12} \tilde{C}_{j,k}(t) = \text{Tr}_{12} \left( B_k A_j^* A_j B_k^* \tilde{D}_{j,k}(t) \right) \]

with \( \tilde{D}_{j,k}(t) \) given by
\[ \left( \text{Tr}_1 A_j \rho_{12} A_j^* \right) (\rho_2 + t)^{-1} B_k \rho_{12} B_k^* (\rho_2 + t)^{-1}. \]

Since
\[ B_k A_j^* A_j B_k^* (\text{Tr}_1 A_j \rho_{12} A_j^*) \leq c_{2}^{2} (\text{Tr}_1 A_j \rho_{12} A_j^*) \]
and since \( (\rho_2 + t)^{-1} B_k \rho_{12} B_k^* (\rho_2 + t)^{-1} \geq 0 \), we have
\[ \text{Tr}_{12} \tilde{C}_{j,k}(t) \leq c_{2}^{2} \text{Tr}_{12} \tilde{D}_{j,k}(t). \]

From here, everything is as before:
\[ \sum_{j,k} \text{Tr}_{12} \tilde{D}_{j,k}(t) = \text{Tr}_2 \rho_2^2 (\rho_2 + t)^{-2} \]
and
\[ \int_{0}^{\infty} dt \sum_{j,k} \text{Tr}_{12} \tilde{D}_{j,k}(t) = \text{Tr}_2 \rho_2 = 1. \]

QED

At last, we turn to the continuous version, the most important application being the position-momentum uncertainty. We start with this case. Take \( \mathcal{H}_1 \) to be \( L^2(\mathbb{R}^d) \), the square-integrable functions on \( \mathbb{R}^d \). The spaces \( \mathcal{H}_2 \) and \( \mathcal{H}_3 \) can be anything. The measurement \( A_j^* A_j \) is \( \{a_j | a_j \} \) where \( a_j \) is the delta function \( \delta(x - x') \) for some \( x' \in \mathbb{R}^n \). The \( j \) becomes \( x' \), which is a continuous variable, the sum \( \sum_j \) becomes the integral \( \int dx' \delta(x - x') \), and the normalization condition \( \sum_j A_j^* A_j = 1 \) becomes \( \int dx' \delta(x - x') \delta(x'' - x') \) (We realize that the delta-function is not a function, but all of this can be made rigorous.) Similarly, the \( B_k^* B_k \)'s are \( \{b_k | b_k \} \) where the \( b_k \) will also not be square-integrable functions. They will be plane waves \( e^{ \pm i k \cdot x} \) for some \( k \) in \( \mathbb{R}^d \). Again, \( k \) is a continuous index and sums are integrals. The normalization condition \( \sum_k \) becomes \( \int d^d k e^{-2 \pi i k \cdot x} e^{2 \pi i k \cdot x'} = \delta(x - x') \).
Let $\rho_{12}$ be a density matrix on the Hilbert space $L^2(\mathbb{R}^d) \otimes \mathcal{H}_2$. For every fixed $x \in \mathbb{R}^d$ we can define $\langle x|\rho_{12}|x\rangle$ as an operator on $\mathcal{H}_2$. This really means the partial trace $\text{Tr}_1$ of $A_j^* \rho_{12} A_j$. Likewise we define $(k|\rho_{12}|k)$ to be $\text{Tr}_1 B_k^* \rho_{12} B_k$. We see that $\langle x|\rho_{12}|x\rangle$ is an operator-valued density in position space and $(k|\rho_{12}|k)$ is the corresponding density in momentum space.

Now we apply the 2-space theorem and infer that

$$H(1^A|2) + H(1^B|2) \geq S(1|2),$$

where

$$H(1^A|2) = -\int_{\mathbb{R}^d} d^d x \text{Tr}_2(\rho_{12}|x) \ln(\rho_{12}|x) - S(\rho_2)$$

and

$$H(1^B|2) = -\int_{\mathbb{R}^d} d^d k \text{Tr}_2(k|\rho_{12}|k) \ln(k|\rho_{12}|k) - S(\rho_2).$$

Here $c_1 = c_\infty = \sup_{x,k} |e^{2\pi ikx}| = 1$ and $\ln c_1 = \ln c_\infty = 0$. This is a generalization of the uncertainty principle in [7], which is what [6] reduces to when $\rho_{12} = \rho_1 \otimes \rho_2$ is a product state. The 3-space theorem and the $2^{1/2}$-space theorem obviously generalize in a similar way for the Fourier transform.

This example of a continuum version of our theorems has the obvious generalization to positive operator-valued measures (POVMs). The interested reader can work this out for him/herself, but we mention here one further specific generalization. Let us take $\mathcal{H}_1 = L^2(X, dx)$, where $X$ is some configuration space (e.g., $X = \mathbb{R}^d$ or $X$ is a torus or $X$ is a lattice) and $L^2(X, dx)$ are the square integrable functions on $X$ with respect to some measure $dx$. The measurements $A_j^* A_j$ are again given by rank one projections corresponding to the functions $\delta_{x'}(x)$ for some $x'$ in $X$. Now let $\mathcal{H}_1'$ be a second Hilbert space of the form $L^2(K, dk)$ and let $U$ be a unitary from $\mathcal{H}_1$ to $\mathcal{H}_1'$, which is given by a kernel $U(k, x)$. For each $k \in K$, we can think of $U(k, x)$ as a function of $x$ and we define the $B_k^* B_k$'s to be the rank-one projections onto these functions. In this way we obtain, as before, operators $\langle x|\rho_{12}|x\rangle$ and $(k|\rho_{12}|k)$ on $\mathcal{H}_2$. Now $H(1^A|2)$ and $H(1^B|2)$ are defined as in the Fourier transform case, except that the integration is over the sets $X$ and $K$, respectively.

Generalized Fourier transform theorem. — [1] is valid with

$$c_1 = \sup_{x,k} |U(k, x)|.$$

An example in which the classical entropies are simultaneously discrete and continuous is the following. Suppose $X = \mathbb{Z}$, i.e., the integers, like the sites in a tight binding model. An $L^2$ function is the wave function of an itinerant electron. The second space $\mathcal{H}_1'$ is the square integrable functions on $K = (-1/2, 1/2)$, the Brillouin zone. In this case the delta-functions, $\delta_{x'k}$, on $X$ are legitimate Kronecker deltas, and the (normalized) plane waves on $X$, parametrized by $k \in K$, are $e^{2\pi ikx}$. The unitary here is $U(k, x) = e^{2\pi ikx}$ and therefore $c_1 = 1$. Since $dx$ is counting measure on $\mathbb{Z}$ the expression of $H(1^A|2)$ is a sum $\sum_{x \in \mathbb{Z}}$, whereas $dk$ is ordinary Lebesgue measure on $(-1/2, 1/2)$ and $H(1^B|2)$ is an integral $\int_K dk$.

In conclusion, we have shown several things. (1) The entropy inequalities in [8, 10] can be proved in a few lines essentially by imitating the original proof of strong subadditivity of entropy. (2) We have carried at least one of these inequalities forward by utilizing the trace norm $c_1$ instead of the operator norm $c_\infty$. (3) We have shown how these inequalities, suitably interpreted, extend the 2-space entropy uncertainty principle to continuous bases such as position and momentum.

We thank Zhihao Ma for helpful correspondence. U.S. National Science Foundation grants PHY-1068285 (R.F.) and PHY-0965859 (E.L.) are acknowledged.