Lecture 1: Strong Subadditivity of Entropy

Abstract : We discuss motivation for strong subadditivity, and discuss its elementary derivation. We also discuss a special structure that arises when the strong subadditivity is satisfied with an equality.

1 Motivation

Strong subadditivity of entropy is arguably one of the most important inequalities in quantum information theory. It states that the following linear combination of entanglement entropies is always nonnegative:

$$S(AB) + S(BC) - S(B) - S(ABC) \geq 0, \tag{1}$$

where $S(X)$ is the von Neumann entropy of a reduced density matrix over a subsystem X; you can plug in AB, BC, B, and ABC into X. The von Neumann entropy is defined as $S(X) = -\text{tr}(\rho_X \log \rho_X)$, where ρ_X is the reduced desnity matrix over X.

Why is this useful? There are two reasons. First of all, it is because quantum information explores the fundamental limit of quantum information processing. This means that we are trying to understand whether certain information tasks can be done or not in principle. Strong subadditivity applies to any quantum states, so it is a suitable tool to study these problems. The second reason is that there is a sense in which the strong subadditivity of entropy is the strongest possible nontrivial statement. For inequalities involving entanglement entropies of three subsystems, any inequality that holds for all quantum states must reduce to strong subadditivity.

But more recently, strong subadditivity has gained a lot of interest in a broader context. They have found applications in condensed matter physics and high energy physics community. These results can be understood in a similar vain, in that the goal is to understand the fundamental constraints that are imposed on physical systems. I will not be able to review all of these results in this lecture, so I decided to pick one direction that seemed to have been quite fruitful. This is a class of monotonicity theorems that were derived for renormalization group flows.

2 Derivation

It is often said that there is no simple derivation of strong subadditivity. I think this is not true. If you know a few tricks, then the derivation becomes quite simple. What I will discuss below will have analogues in infinite dimension. This can be found in Witten's note. For quantum field theory, you do want to be careful about these subtleties. But I will take a point of view that sensible quantum field theories have a lattice regularization, so that we can justify using versions of strong subadditivity that are derived in finite dimensions.

2.1 Entanglement Trick

Here we will discuss a simple trick. Suppose we have operator A and B. Let's consider an unnormalized maximally entangled state between two Hilbert spaces:

$$|\Phi\rangle = \sum_{I=1}^{d} |I\rangle |I\rangle . \tag{2}$$

Then we have the following identity:

$$\langle\Phi| A \otimes B |\Phi\rangle = \text{tr}(AB^T). \tag{3}$$

This is easy to show.

$$\begin{aligned} \langle\Phi| A \otimes B |\Phi\rangle &= \sum_{I=1}^{d} A_{IJ} B_{JI} \\ &= \text{tr}(AB^T). \end{aligned} \tag{4}$$

2.2 Operator Convex Functions

There is a beautiful theory of operator convex functions, but I won't have time to discuss this subject. It really is an awesome subject, so if you have spare time, you should at least take a look at Eric Carlen's lecture note [1]. But for us, a much simpler analysis will suffice.

Let us recall that a function f is convex if it obeys Jensen's inequality:

$$f(\lambda x + (1-\lambda)y) \leq \lambda f(x) + (1-\lambda)f(y) \tag{5}$$

for $\lambda \in [0,1]$. We say that f is operator convex if it obeys Operator Jensen inequality:

$$f(\lambda X + (1-\lambda)Y) \leq \lambda f(X) + (1-\lambda)f(Y). \tag{6}$$

One might ask, how is $f(X)$ defined? We consider the case in which X is positive semidefinite. In this case, we can diagonlize X and get its eigenvalues $\{x_i\}$ and eigenvectors $|i\rangle$. We define $f(X)$ to be

$$f(X) = \sum_i x_i \left|i\right\rangle \left\langle i\right|. \tag{7}$$

Another question: what does the inequality mean for operators? It means that the following operator

$$\lambda f(X) + (1-\lambda)f(Y) - f(\lambda X + (1-\lambda)Y) \tag{8}$$

is positive semidefinite.

Dealing with general operators is difficult, and operator convex functions are one of the few methods that are easy to understand but yield powerful results. We are actually interested in one function: $f(x) = x\log x$. To see why this function is operator convex, note the following integral representation:

$$\begin{aligned} f(x) &= x\int_0^\infty \frac{1}{t+1} - \frac{1}{t+x} \\ &= \int_0^\infty \frac{x}{t+1} - 1 + \frac{t}{t+x}. \end{aligned} \tag{9}$$

The first two terms are convex by definition, so whether $x\log x$ is operator convex boils down to the question of whether a function $\frac{1}{t+x}$ is operator convex.

This is not too difficult to see. The operator convexity of the function $\frac{1}{t+x}$ follows from the following inequality.

$$\lambda\frac{1}{X} + (1-\lambda)\frac{1}{Y} \leq \frac{1}{\lambda X + (1-\lambda Y)}. \tag{10}$$

This holds because, after multiplying both sides by $X^{1/2}$ from the left and the right, we get

$$\lambda I + (1-\lambda)\frac{1}{Z} \leq \frac{1}{\lambda I + (1-\lambda)Z}, \tag{11}$$

where $Z = X^{-1/2}YX^{-1/2}$. Because I and Z commutes, one can write down this inequality in the eigenbasis of Z, which can be proved by the convexity of $\frac{1}{t+x}$ for *numbers*.

Recall that all the operators were assumed to be positive semidefinite, so we can reverse our action to deduce the operator convexity of $\frac{1}{t+x}$, which subsequently implies operator convexity of $f(x) = x\log x$. So we can conclude:

$$\lambda A\log A + (1-\lambda)B\log B \geq C\log C, \tag{12}$$

where $C = \lambda A + (1-\lambda)B$.

With a little extra effort, one can also show that

$$T_1 A\log A T_1^\dagger + T_2 B\log B T_2^\dagger \geq C\log C, \tag{13}$$

where $C = T_1 A T_1^\dagger + T_2 B T_2^\dagger$, $T_1T_1^\dagger + T_2T_2^\dagger = I$.

2.3 Perspective Function

Now we need one more tool: perspective function. Perspective function is defined as $g(t,s) = tf(s/t)$. One can view this as a collection of rays emanating from the origin that passes through $f(s)$ at $t=1$; see Fig.1 as an example. Hopefully this gives you an intuition on why this function is convex not just in s, and not just on t, but jointly on s and t.[1]

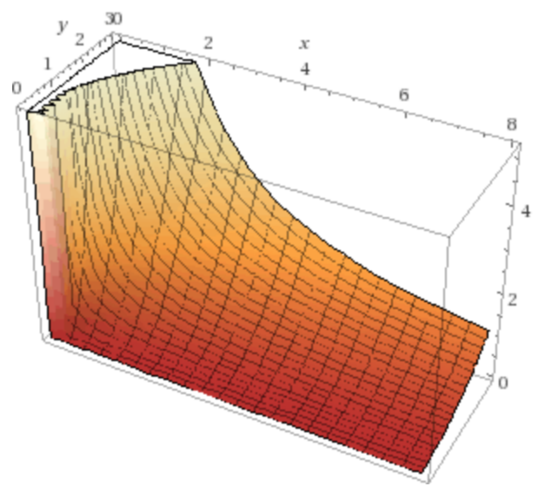

Figure 1: Plot of a perspective function $g(x,y) = x(y/x)^2$.

However, not all convex functions are operator convex. For example, $f(x) = x^4$ is a convex function, but it is *not* an operator convex function. Again, if you are interested in this subject, I advise to study Carlen' lecture note [1].

Effros [2] proved a beautiful theorem that, if $f(x)$ is an operator convex function, then $g(x,y) = xf(y/x)$ is jointly convex in x and y, provided that x commutes with y.[2] This means that if

$$[A_1, B_1] = [A_2, B_2] = 0, \tag{14}$$

then

$$g(tA_1 + (1-t)A_2, tB_1 + (1-t)B_2) \leq tg(A_1, B_1) + (1-t)g(A_2, B_2). \tag{15}$$

Here A_1, A_2, B_1, B_2 are matrices. This proof is due to Effros [2].

$$\begin{aligned}
g(A,B) &= Af(B/A) \\
&= A^{1/2} f(B/A) A^{1/2} \\
&= A^{1/2} f(A^{-1/2} B A^{-1/2}) A^{1/2} \\
&= A^{1/2} f(T_1 (B_1/A_1) T_1^\dagger + T_2 (B_2/A_2) T_2^\dagger) A^{1/2} \\
&\leq A^{1/2} T_1 f(B_1/A_1) T_1^\dagger A^{1/2} + A^{1/2} T_2 f(B_2/A_2) T_2^\dagger A^{1/2} \\
&= tg(A_1, B_1) + (1-t) g(A_2, B_2),
\end{aligned} \tag{16}$$

where we chose $T_1 = A^{-1/2}(tA_1)^{1/2}$ and $T_2 = A^{-1/2}((1-t)A_2)$.

2.4 Joint Convexity of relative entropy

Now all we need to do is to plug in an appropriate set of density matrices. Note that Eq.15 for any $g(A,B) = Af(B/A)$ for any operator convex function f, provided that A and B commute so to speak. Let us choose

[1]These facts can be easily checked by computing the Hessian.

[2]Under an appropriate definition of matrix inverse, this condition can be removed.

$f(x) = x \log x$, and A as $\rho \otimes I$ and B as $I \otimes \sigma^T$. Here σ^T is a transpose of σ. Now we see that

$$\langle\Phi| g(\rho,\sigma) |\Phi\rangle = \mathrm{tr}(\rho(\log\rho - \log\sigma)) =: D(\rho\|\sigma). \tag{17}$$

This is a quantity that is known as relative entropy. So we have arrived at an important conclusion. That $D(\rho\|\sigma)$ is jointly convex in its arguments.

2.5 Monotonicity

Joint convexity of relative entropy implies that relative entropy does not increase if under partial trace. That is,

$$D(\rho\|\sigma) \geq D(\rho_A\|\sigma_A). \tag{18}$$

This follows from what is known as the "twirling trick." Without loss of generality, suppose that the Hilbert space has a tensor product structure $\mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_B$. By joint convexity,

$$\begin{aligned} D(\rho\|\sigma) &= \int dU_B D(U_B\rho U_B^\dagger \| U_B\sigma U_B^\dagger) \\ &\geq D(\int dU_B U_B \rho U_B^\dagger \| \int dU_B U_B\rho U_B^\dagger), \end{aligned} \tag{19}$$

where $\int dU_B$ is a normalized integration over the Haar measure over the set of unitary operators acting on $\mathcal{H}_B$. The first line follows from the fact that the relative entropy is invariant under unitary transformation on both of its arguments.[3] The second line follows from the joint convexity argument.

With a little thought, one can show that

$$\int dU_B U_B\rho U_B^\dagger = \rho_A \otimes \frac{I_B}{d_B}, \tag{20}$$

where d_B is the dimension of B. Here is a heuristic argument. One can easily check that

$$[\int dU_B U_B \rho U_B^\dagger, I_A \otimes V_B] = 0 \tag{21}$$

for any unitary operator V_B, where I_A is the identity on $\mathcal{H}_A$. Because the set of unitary operator spans the space of operators, we conclude that $\int dU_B U_B\rho U_B^\dagger = X_A \otimes I_B$ for some operator X_A. This operator must furthermore be positive semi-definite, and its expectation values on A must be consistent with that of ρ. What else can it be other than ρ_A?

So we have

$$\begin{aligned} D(\rho\|\sigma) &\geq D(\rho_A \otimes I_B/d_B \| \rho_B \otimes I_B/d_B) \\ &= D(\rho_A\|\sigma_A). \end{aligned} \tag{22}$$

2.6 Strong subadditivity

We can choose $\rho = \rho_{ABC}$, $\sigma = \frac{I_A}{d_A} \otimes \rho_{BC}$. We have

$$\begin{aligned} D(\rho\|\sigma) &\geq D(\mathrm{tr}_C\rho \| \mathrm{tr}_C\sigma) \\ &= D(\rho_{AB}\| \frac{I_A}{d_A} \otimes \rho_B). \end{aligned} \tag{23}$$

Rearranging the terms, we have $S(AB) + S(BC) - S(B) - S(ABC) \geq 0$.

[3] Of course, the same unitary must be applied. Otherwise the relation does not hold.

3 Comment

Strong subadditivity is a nontrivial theorem. One needs to invoke some nontrivial theorems to prove it. People do have intuitions about these proofs, but they are often difficult to digest at first sights. The goal of this note was to emphasize the geometric aspect of these inequalities. Everything followed from the convexity of some operator-valued function. The core of the analysis lied on perspective function, which, in a sense, "boosts" convexity of an operator-valued function with single argument to a joint convexity of a function with two arguments. The rest of the derivation is relatively straightforward in comparison.

References

[1] E. Carlen, "Trace inequalities and quantum entropy: An introductory course," 2009.

[2] E. G. Effros, "A matrix convexity approach to some celebrated quantum inequalities," *Proceedings of the National Academy of Sciences*, vol. 106, no. 4, pp. 1006–1008, 2009.