Quantum Channels and Their Capacities, Homework 1

In this question you will prove the following two results. The first is used in the "converse" proof of the classical noisy channel coding theorem. The quantum version of the second is closely related to strong-subadditivity, and is used in the converse for degradable quantum channels (you'll learn about these in lectures 3 + 4):

$$H(X,Y) \leq H(X) + H(Y) \tag{1}$$

and

$$H(Y_1Y_2|X_1X_2) \leq H(Y_1|X_1) + H(Y_2|X_2). \tag{2}$$

You'll first prove Eq. (1), then use it to prove Eq. (2).

You're going to need some background information to do this. First, the definition of a convex function:

Definition A function f is convex on (a,b) if, for any $x, y \in (a,b)$ and $p \in (0,1)$, it satisfies

$$pf(x) + (1-p)f(y) \geq f\left(px + (1-p)y\right). \tag{3}$$

You'll also need these two facts about convex functions:

Fact 1 If f has a non-negative second derivative on (a,b), then it is convex on (a,b).

Fact 2 (Jensen's Inequality) If f is a convex function and X is a random variable, then $\langle f(X)\rangle \geq f(\langle X\rangle)$, where $\langle\rangle$ indicates expectation.

Also, don't forget that $H(X) = -\sum_x p(x)\log p(x)$.

Questions

A Show that $-\log x$ is convex on $(0,\infty)$.

B Using the definition of entropy, show that Eq. (1) is equivalent to

$$-\sum_{x,y} p(x,y)\log\left(\frac{p(x)p(y)}{p(x,y)}\right) \geq 0. \tag{4}$$

C Apply Jensen's inequality to prove Eq. (4), and thus Eq.(1).

D Starting from the definition of entropy, $H(X) = -\sum_x p(x)\log p(x)$, and of conditional entropy, $H(Y|X) = H(X,Y) - H(X)$, show that

$$H(Y|X) = \sum_x p(x)\sum_y -p(y|x)\log(p(y|x)). \tag{5}$$

As a result, we know that $H(Y|X)$ is the average over x of the entropies of random variables Y_x, distributed according to $p(y|x)$.

E Use part D to show that $I(X;Y|Z) := I(X;YZ) - I(X;Z)$ is an average (over Z) of the mutual information between Z-dependent random variables. By positivity of mutual information (which you showed in C), argue that

$$I(X;YZ) \geq I(X;Z). \tag{6}$$

F Show that Eq.(2) can be rephrased as

$$I(X_1;X_2) \leq I(X_1Y_1;X_2Y_2). \tag{7}$$

Use E to show this is true.