PHYS 7895: Quantum Information Theory Fall 2015

Lecture 19 — October 28, 2015

Prof. Mark M. Wilde *Scribe: Mark M. Wilde*

1 Overview

In the previous lecture, we introduced quantum entropy and some other variants of it, such as conditional entropy, mutual information, and conditional mutual information.

In this lecture, we discuss several entropy inequalities that play an important role in quantum information processing: the monotonicity of quantum relative entropy, strong subadditivity, the quantum data-processing inequalities, and continuity of quantum entropy.

2 Quantum Relative Entropy

The quantum relative entropy is one of the most important entropic quantities in quantum information theory, mainly because we can reexpress many of the entropies given in the previous sections in terms of it. This in turn allows us to establish many properties of these quantities from the properties of relative entropies. Its definition is a natural extension of that for the classical relative entropy. Before defining it, we need the notion of the support of an operator:

Definition 1 (Kernel and Support)**.** *The kernel of an operator* $A \in \mathcal{L}(\mathcal{H}, \mathcal{H}')$ *is defined as*

$$\ker(A) \equiv \{|\psi\rangle \in \mathcal{H} : A|\psi\rangle = 0\}. \tag{1}$$

The support of A *is the subspace of* $\mathcal{H}$ *orthogonal to its kernel:*

$$\operatorname{supp}(A) \equiv \{|\psi\rangle \in \mathcal{H} : A|\psi\rangle \neq 0\}. \tag{2}$$

If A *is Hermitian and thus has a spectral decomposition as* $A = \sum_{i:a_i \neq 0} a_i |i\rangle\langle i|$, *then* $\operatorname{supp}(A) = \operatorname{span}\{|i\rangle : a_i \neq 0\}$. *The projection onto the support of* A *is denoted by*

$$\Pi_A \equiv \sum_{i:a_i \neq 0} |i\rangle\langle i|. \tag{3}$$

Definition 2. *The quantum relative entropy* $D(\rho\|\sigma)$ *between a density operator* $\rho \in \mathcal{D}(\mathcal{H})$ *and a positive semi-definite operator* $\sigma \in \mathcal{L}(\mathcal{H})$ *is defined as follows:*

$$D(\rho\|\sigma) \equiv \operatorname{Tr}\{\rho[\log\rho - \log\sigma]\}, \tag{4}$$

if the following support condition is satisfied

$$\operatorname{supp}(\rho) \subseteq \operatorname{supp}(\sigma), \tag{5}$$

and it is defined to be equal to $+\infty$ *otherwise.*

This definition is consistent with the classical definition of relative entropy. However, we should note that there could be be several ways to generalize the classical definition to obtain a quantum definition of relative entropy. For example, one could take

$$D'(\rho\|\sigma) = \mathrm{Tr}\left\{\rho \log\left(\rho^{1/2}\sigma^{-1}\rho^{1/2}\right)\right\}, \tag{6}$$

as a definition and it reduces to the classical definition as well. In fact, it is easy to see that there are an infinite number of quantum generalizations of the classical definition of relative entropy. So how do we single out which definition is the right one to use? The definition given in (4) is the answer to a meaningful quantum information-processing task. Furthermore, this definition generalizes the quantum entropic quantities we have given in this chapter, which all in turn are the answers to meaningful quantum information-processing tasks. For these reasons, we take the definition given in (4) as *the* quantum relative entropy. Recall that it was this same line of reasoning that allowed us to single out the entropy and the mutual information as meaningful measures of information in the classical case.

Similar to the classical case, we can intuitively think of the quantum relative entropy as a distance measure between quantum states. But it is not strictly a distance measure in the mathematical sense because it is not symmetric and it does not obey a triangle inequality.

The following proposition justifies why we take the definition of quantum relative entropy to have the particular support conditions as given above:

Proposition 3. *Let $\rho \in \mathcal{D}(\mathcal{H})$ and $\sigma \in \mathcal{L}(\mathcal{H})$ be positive semi-definite. The quantum relative entropy is consistent with the following limit:*

$$D(\rho\|\sigma) = \lim_{\varepsilon \searrow 0} D(\rho\|\sigma + \varepsilon I). \tag{7}$$

One of the most fundamental entropy inequalities in quantum information theory is the monotonicity of quantum relative entropy. When the arguments to the quantum relative entropy are quantum states, the physical interpretation of this entropy inequality is that states become less distinguishable when noise acts on them. We defer a proof of this theorem until later, where we also establish a strengthening of it.

Theorem 4 (Monotonicity of Quantum Relative Entropy)**.** *Let $\rho \in \mathcal{D}(\mathcal{H})$, $\sigma \in \mathcal{L}(\mathcal{H})$ be positive semi-definite, and $\mathcal{N} : \mathcal{L}(\mathcal{H}) \to \mathcal{L}(\mathcal{H}')$ be a quantum channel. The quantum relative entropy can only decrease if we apply the same quantum channel $\mathcal{N}$ to ρ and σ:*

$$D(\rho\|\sigma) \geq D(\mathcal{N}(\rho)\|\mathcal{N}(\sigma)). \tag{8}$$

Theorem 4 then implies non-negativity of quantum relative entropy in certain cases.

Theorem 5 (Non-Negativity)**.** *Let $\rho \in \mathcal{D}(\mathcal{H})$, and let $\sigma \in \mathcal{L}(\mathcal{H})$ be positive semi-definite and such that $\mathrm{Tr}\{\sigma\} \leq 1$. The quantum relative entropy $D(\rho\|\sigma)$ is non-negative:*

$$D(\rho\|\sigma) \geq 0, \tag{9}$$

and $D(\rho\|\sigma) = 0$ if and only if $\rho = \sigma$.

Proof. The first part of the theorem follows from applying Theorem 4, taking the quantum channel to be the trace-out map. We then have that

$$D(\rho\|\sigma) \geq D(\operatorname{Tr}\{\rho\}\|\operatorname{Tr}\{\sigma\}) \tag{10}$$

$$= \operatorname{Tr}\{\rho\}\log\left(\frac{\operatorname{Tr}\{\rho\}}{\operatorname{Tr}\{\sigma\}}\right) \tag{11}$$

$$\geq 0. \tag{12}$$

If $\rho = \sigma$, then the support condition in (5) is satisfied and plugging into (4) gives that $D(\rho\|\sigma) = 0$. Now suppose that $D(\rho\|\sigma) = 0$. This means that the inequality above is saturated and thus $\operatorname{Tr}\{\sigma\} = \operatorname{Tr}\{\rho\} = 1$, so that σ is a density operator. Let $\mathcal{M}$ be an arbitrary measurement channel. From the monotonicity of quantum relative entropy (Theorem 4), we can conclude that $D(\mathcal{M}(\rho)\|\mathcal{M}(\sigma)) = 0$. The equality condition for the non-negativity of the classical relative entropy in turn implies that $\mathcal{M}(\rho) = \mathcal{M}(\sigma)$. Now since this equality holds for any possible measurement channel, we can conclude that $\rho = \sigma$. (For example, we could take $\mathcal{M}$ to be the optimal measurement for the trace distance, which would allow us to conclude that $\max_{\mathcal{M}} \|\mathcal{M}(\rho) - \mathcal{M}(\sigma)\|_1 = \|\rho - \sigma\|_1 = 0$, and hence $\rho = \sigma$.) □

2.1 Deriving Other Entropies from Quantum Relative Entropy

There is a sense in which the quantum relative entropy is a "parent quantity" for other entropies in quantum information theory, such as the von Neumann entropy, the conditional quantum entropy, the quantum mutual information, and the conditional quantum mutual information. The following exercises explore these relations. The main tool needed to solve some of them is the non-negativity of quantum relative entropy.

Exercise 6. *Let $P_A \in \mathcal{L}(\mathcal{H}_A)$ and $Q_B \in \mathcal{L}(\mathcal{H}_B)$ be positive semi-definite operators. Show that the following identity holds:*

$$\log(P_A \otimes Q_B) = \log(P_A) \otimes I_B + I_A \otimes \log(Q_B). \tag{13}$$

Exercise 7 (Mutual Information and Relative Entropy)**.** *Let $\rho_{AB} \in \mathcal{D}(\mathcal{H}_A \otimes \mathcal{H}_B)$. Show that the following identities hold:*

$$I(A;B)_\rho = D(\rho_{AB}\|\rho_A \otimes \rho_B) \tag{14}$$

$$= \min_{\sigma_B} D(\rho_{AB}\|\rho_A \otimes \sigma_B) \tag{15}$$

$$= \min_{\omega_A} D(\rho_{AB}\|\omega_A \otimes \rho_B) \tag{16}$$

$$= \min_{\omega_A,\sigma_B} D(\rho_{AB}\|\omega_A \otimes \sigma_B), \tag{17}$$

where the optimizations are with respect to $\omega_A \in \mathcal{D}(\mathcal{H}_A)$ and $\sigma_B \in \mathcal{D}(\mathcal{H}_B)$.

Exercise 8 (Conditional Entropy and Relative Entropy)**.** *Let $\rho_{AB} \in \mathcal{D}(\mathcal{H}_A \otimes \mathcal{H}_B)$. Show that the following identities hold:*

$$I(A\rangle B)_\rho = D(\rho_{AB}\|I_A \otimes \rho_B) \tag{18}$$

$$= \min_{\sigma_B \in \mathcal{D}(\mathcal{H}_B)} D(\rho_{AB}\|I_A \otimes \sigma_B). \tag{19}$$

Note that these imply that

$$H(A|B)_\rho = -D(\rho_{AB} \| I_A \otimes \rho_B) \tag{20}$$
$$= -\min_{\sigma_B \in \mathcal{D}(\mathcal{H}_B)} D(\rho_{AB} \| I_A \otimes \sigma_B). \tag{21}$$

Exercise 9 (Relative Entropy of Classical–Quantum States)**.** *Show that the quantum relative entropy between classical–quantum states ρ_{XB} and σ_{XB} is as follows:*

$$D(\rho_{XB} \| \sigma_{XB}) = \sum_x p_X(x) D(\rho_B^x \| \sigma_B^x), \tag{22}$$

where

$$\rho_{XB} \equiv \sum_x p_X(x) |x\rangle\langle x|_X \otimes \rho_B^x, \qquad \sigma_{XB} \equiv \sum_x p_X(x) |x\rangle\langle x|_X \otimes \sigma_B^x, \tag{23}$$

with p_X a probability distribution over a finite alphabet $\mathcal{X}$, $\rho_B^x \in \mathcal{D}(\mathcal{H}_B)$ for all $x \in \mathcal{X}$, and $\sigma_B^x \in \mathcal{L}(\mathcal{H}_B)$ positive semi-definite for all $x \in \mathcal{X}$.

3 Quantum Entropy Inequalities

Monotonicity of quantum relative entropy has as its corollaries many of the important entropy inequalities in quantum information theory.

Corollary 10 (Strong Subadditivity)**.** *Let $\rho_{ABC} \in \mathcal{D}(\mathcal{H}_A \otimes \mathcal{H}_B \otimes \mathcal{H}_C)$. The von Neumann entropy is strongly subadditive, in the following sense:*

$$H(AC)_\rho + H(BC)_\rho \geq H(ABC)_\rho + H(C)_\rho. \tag{24}$$

Equivalently, we have that

$$I(A; B|C)_\rho \geq 0. \tag{25}$$

Proof. Consider that

$$I(A; B|C)_\rho = H(AC)_\rho + H(BC)_\rho - H(ABC)_\rho - H(C)_\rho, \tag{26}$$

so that

$$I(A; B|C)_\rho = H(B|C)_\rho - H(B|AC)_\rho. \tag{27}$$

From Exercise 8, we know that

$$-H(B|AC)_\rho = D(\rho_{ABC} \| I_B \otimes \rho_{AC}), \tag{28}$$
$$H(B|C)_\rho = -D(\rho_{BC} \| I_B \otimes \rho_C). \tag{29}$$

Then

$$D(\rho_{ABC} \| I_B \otimes \rho_{AC}) \geq D(\mathrm{Tr}_A\{\rho_{ABC}\} \| \mathrm{Tr}_A\{I_B \otimes \rho_{AC}\}) \tag{30}$$
$$= D(\rho_{BC} \| I_B \otimes \rho_C). \tag{31}$$

The inequality is a consequence of the monotonicity of quantum relative entropy (Theorem 4), taking $\rho = \rho_{ABC}$, $\sigma = I_B \otimes \rho_{AC}$, and $\mathcal{N} = \mathrm{Tr}_A$. By (26)–(29), the inequality in (30)–(31) is equivalent to the inequality in the statement of the corollary. ☐

Corollary 11 (Joint Convexity of Quantum Relative Entropy)**.** *Let* p_X *be a probability distribution over a finite alphabet* $\mathcal{X}$*,* $\rho^x \in \mathcal{D}(\mathcal{H})$ *for all* $x \in \mathcal{X}$*, and* $\sigma^x \in \mathcal{L}(\mathcal{H})$ *be positive semi-definite for all* $x \in \mathcal{X}$*. Set* $\rho \equiv \sum_x p_X(x)\rho^x$ *and* $\sigma \equiv \sum_x p_X(x)\sigma^x$*. The quantum relative entropy is jointly convex in its arguments:*

$$D(\rho\|\sigma) \leq \sum_x p_X(x) D(\rho^x\|\sigma^x). \tag{32}$$

Proof. Consider classical–quantum states of the following form:

$$\rho_{XB} \equiv \sum_x p_X(x)|x\rangle\langle x|_X \otimes \rho_B^x, \tag{33}$$

$$\sigma_{XB} \equiv \sum_x p_X(x)|x\rangle\langle x|_X \otimes \sigma_B^x. \tag{34}$$

Then

$$\sum_x p_X(x) D(\rho_B^x\|\sigma_B^x) = D(\rho_{XB}\|\sigma_{XB}) \geq D(\rho_B\|\sigma_B). \tag{35}$$

The equality follows from Exercise 9, and the inequality follows from monotonicity of quantum relative entropy (Theorem 4), where we take the channel to be the partial trace over the system X. □

Corollary 12 (Unital Channels Increase Entropy)**.** *Let* $\rho \in \mathcal{D}(\mathcal{H})$ *and let* $\mathcal{N} : \mathcal{L}(\mathcal{H}) \to \mathcal{L}(\mathcal{H})$ *be a unital quantum channel. Then*

$$H(\mathcal{N}(\rho)) \geq H(\rho). \tag{36}$$

Proof. Consider that

$$H(\rho) = -D(\rho\|I), \tag{37}$$

$$H(\mathcal{N}(\rho)) = -D(\mathcal{N}(\rho)\|I) = -D(\mathcal{N}(\rho)\|\mathcal{N}(I)), \tag{38}$$

where in the last equality, we have used that $\mathcal{N}$ is a unital quantum channel. The inequality in (36) is a consequence of the monotonicity of quantum relative entropy (Theorem 4) because $D(\rho\|I) \geq D(\mathcal{N}(\rho)\|\mathcal{N}(I))$. □

3.1 Quantum Data Processing

The quantum data-processing inequalities discussed below are similar in spirit to the classical data-processing inequality. Recall that the classical data-processing inequality states that processing classical data reduces classical correlations. The quantum data-processing inequalities state that processing *quantum* data reduces *quantum* correlations.

One variant applies to the following scenario. Suppose that Alice and Bob share some bipartite state ρ_{AB}. The coherent information $I(A\rangle B)_\rho$ is one measure of the quantum correlations present in this state. Bob then processes his system B according to some quantum channel $\mathcal{N}_{B\to B'}$ to produce some quantum system B' and let $\sigma_{AB'}$ denote the resulting state. The quantum data-processing inequality states that this step of quantum data processing reduces quantum correlations, in the sense that

$$I(A\rangle B)_\rho \geq I(A\rangle B')_\sigma. \tag{39}$$

Theorem 13 (Quantum Data Processing for Coherent Information)**.** *Let $\rho_{AB} \in \mathcal{D}(\mathcal{H}_A \otimes \mathcal{H}_B)$ and let $\mathcal{N} : \mathcal{L}(\mathcal{H}_B) \to \mathcal{L}(\mathcal{H}_{B'})$ be a quantum channel. Set $\sigma_{AB'} \equiv \mathcal{N}_{B\to B'}(\rho_{AB})$. Then the following quantum data-processing inequality holds*

$$I(A\rangle B)_\rho \geq I(A\rangle B')_\sigma. \tag{40}$$

Proof. This is a consequence of Exercise 8 and Theorem 4. By Exercise 8, we know that

$$\begin{aligned} I(A\rangle B)_\rho &= D(\rho_{AB} \| I_A \otimes \rho_B), & (41)\\ I(A\rangle B')_\sigma &= D(\sigma_{AB'} \| I_A \otimes \sigma_{B'}) & (42)\\ &= D(\mathcal{N}_{B\to B'}(\rho_{AB}) \| I_A \otimes \mathcal{N}_{B\to B'}(\rho_B)) & (43)\\ &= D(\mathcal{N}_{B\to B'}(\rho_{AB}) \| \mathcal{N}_{B\to B'}(I_A \otimes \rho_B)). & (44) \end{aligned}$$

The statement then follows from the monotonicity of quantum relative entropy by picking $\rho = \rho_{AB}$, $\sigma = I_A \otimes \rho_B$, and $\mathcal{N} = \mathrm{id}_A \otimes \mathcal{N}_{B\to B'}$ in Theorem 4. □

Theorem 14 (Quantum Data Processing for Mutual Information)**.** *Let $\rho_{AB} \in \mathcal{D}(\mathcal{H}_A \otimes \mathcal{H}_B)$, $\mathcal{N} : \mathcal{L}(\mathcal{H}_A) \to \mathcal{L}(\mathcal{H}_{A'})$ be a quantum channel, and $\mathcal{M} : \mathcal{L}(\mathcal{H}_B) \to \mathcal{L}(\mathcal{H}_{B'})$ be a quantum channel. Set $\sigma_{A'B'} \equiv (\mathcal{N}_{A\to A'} \otimes \mathcal{M}_{B\to B'})(\rho_{AB})$. Then the following quantum data-processing inequality applies to the quantum mutual information:*

$$I(A;B)_\rho \geq I(A';B')_\sigma. \tag{45}$$

Proof. From Exercise 7, we know that

$$\begin{aligned} I(A;B)_\rho &= D(\rho_{AB} \| \rho_A \otimes \rho_B), & (46)\\ I(A';B')_\sigma &= D(\sigma_{A'B'} \| \sigma_{A'} \otimes \sigma_{B'}) & (47)\\ &= D((\mathcal{N}_{A\to A'} \otimes \mathcal{M}_{B\to B'})(\rho_{AB}) \| \mathcal{N}_{A\to A'}(\rho_A) \otimes \mathcal{M}_{B\to B'}(\rho_B)) & (48)\\ &= D((\mathcal{N}_{A\to A'} \otimes \mathcal{M}_{B\to B'})(\rho_{AB}) \| (\mathcal{N}_{A\to A'} \otimes \mathcal{M}_{B\to B'})(\rho_A \otimes \rho_B)). & (49) \end{aligned}$$

The statement then follows from the monotonicity of quantum relative entropy by picking $\rho = \rho_{AB}$, $\sigma = \rho_A \otimes \rho_B$, and $\mathcal{N} = \mathcal{N}_{A\to A'} \otimes \mathcal{M}_{B\to B'}$ in Theorem 4. □

4 Continuity of Entropy

An important theorem below, the Alicki–Fannes–Winter (AFW) inequality, states that conditional quantum entropies are close as well. This statement does follow directly from the Fannes–Audenaert inequality, but the main advantage of the AFW inequality is that the upper bound has a dependence only on the dimension of the first system in the conditional entropy (no dependence on the conditioning system). The AFW inequality also finds application in a proof of a converse theorem in quantum Shannon theory.

Theorem 15 (AFW Inequality)**.** *Let $\rho_{AB}, \sigma_{AB} \in \mathcal{D}(\mathcal{H}_A \otimes \mathcal{H}_B)$. Suppose that*

$$\frac{1}{2} \|\rho_{AB} - \sigma_{AB}\|_1 \leq \varepsilon, \tag{50}$$

for $\varepsilon \in [0,1]$. *Then*

$$|H(A|B)_\rho - H(A|B)_\sigma| \leq 2\varepsilon \log \dim(\mathcal{H}_A) + (1+\varepsilon) h_2\left(\frac{\varepsilon}{1+\varepsilon}\right). \tag{51}$$

If ρ_{XB} *and* σ_{XB} *are classical–quantum and have the following form:*

$$\rho_{XB} = \sum_x p(x)|x\rangle\langle x|_X \otimes \rho_B^x, \tag{52}$$

$$\sigma_{XB} = \sum_x q(x)|x\rangle\langle x|_X \otimes \sigma_B^x, \tag{53}$$

where p *and* q *are probability distributions defined over a finite alphabet* $\mathcal{X}$, $\{|x\rangle\}$ *is an orthonormal basis, and* $\rho_B^x, \sigma_B^x \in \mathcal{D}(\mathcal{H}_B)$ *for all* $x \in \mathcal{X}$, *then*

$$|H(X|B)_\rho - H(X|B)_\sigma| \leq \varepsilon \log \dim(\mathcal{H}_X) + (1+\varepsilon) h_2\left(\frac{\varepsilon}{1+\varepsilon}\right), \tag{54}$$

$$|H(B|X)_\rho - H(B|X)_\sigma| \leq \varepsilon \log \dim(\mathcal{H}_B) + (1+\varepsilon) h_2\left(\frac{\varepsilon}{1+\varepsilon}\right). \tag{55}$$

Proof. The bounds trivially hold when $\varepsilon = 0$, so henceforth we assume that $\varepsilon \in (0,1]$. All of the upper bounds are monotone non-decreasing with ε, so it suffices to assume that $\frac{1}{2}\left\|\rho_{AB} - \sigma_{AB}\right\|_1 = \varepsilon$. Let $\rho_{AB} - \sigma_{AB} = P_{AB} - Q_{AB}$ be a decomposition of $\rho_{AB} - \sigma_{AB}$ into its positive part $P_{AB} \geq 0$ and its negative part $Q_{AB} \geq 0$. Let $\Delta_{AB} \equiv P_{AB}/\varepsilon$. Since $\text{Tr}\{P_{AB}\} = \frac{1}{2}\left\|\rho_{AB} - \sigma_{AB}\right\|_1$ (recall this from the development with trace distance), it follows that Δ_{AB} is a density operator. Now consider that

$$\rho_{AB} = \sigma_{AB} + (\rho_{AB} - \sigma_{AB}) \tag{56}$$

$$= \sigma_{AB} + P_{AB} - Q_{AB} \tag{57}$$

$$\leq \sigma_{AB} + P_{AB} \tag{58}$$

$$= \sigma_{AB} + \varepsilon \Delta_{AB} \tag{59}$$

$$= (1+\varepsilon)\left(\frac{1}{1+\varepsilon}\sigma_{AB} + \frac{\varepsilon}{1+\varepsilon}\Delta_{AB}\right) \tag{60}$$

$$= (1+\varepsilon)\,\omega_{AB}, \tag{61}$$

where we define

$$\omega_{AB} \equiv \frac{1}{1+\varepsilon}\sigma_{AB} + \frac{\varepsilon}{1+\varepsilon}\Delta_{AB}. \tag{62}$$

Now let

$$\Delta'_{AB} \equiv \frac{1}{\varepsilon}\left[(1+\varepsilon)\,\omega_{AB} - \rho_{AB}\right]. \tag{63}$$

It follows from (56)–(61) that Δ'_{AB} is positive semi-definite. Furthermore, one can check that $\text{Tr}\{\Delta'_{AB}\} = 1$, so that Δ'_{AB} is a density operator. One can also quickly check that

$$\omega_{AB} = \frac{1}{1+\varepsilon}\rho_{AB} + \frac{\varepsilon}{1+\varepsilon}\Delta'_{AB} = \frac{1}{1+\varepsilon}\sigma_{AB} + \frac{\varepsilon}{1+\varepsilon}\Delta_{AB}. \tag{64}$$

Now consider that

$$H(A|B)_\omega = -D(\omega_{AB} \| I_A \otimes \omega_B) \tag{65}$$
$$= H(\omega_{AB}) + \mathrm{Tr}\{\omega_{AB} \log \omega_B\} \tag{66}$$
$$\leq h_2\left(\frac{\varepsilon}{1+\varepsilon}\right) + \frac{1}{1+\varepsilon} H(\rho_{AB}) + \frac{\varepsilon}{1+\varepsilon} H(\Delta'_{AB}) \tag{67}$$
$$+ \frac{1}{1+\varepsilon} \mathrm{Tr}\{\rho_{AB} \log \omega_B\} + \frac{\varepsilon}{1+\varepsilon} \mathrm{Tr}\{\Delta'_{AB} \log \omega_B\} \tag{68}$$
$$= h_2\left(\frac{\varepsilon}{1+\varepsilon}\right) - \frac{1}{1+\varepsilon} D(\rho_{AB} \| I_A \otimes \omega_B) \tag{69}$$
$$- \frac{\varepsilon}{1+\varepsilon} D(\Delta'_{AB} \| I_A \otimes \omega_B) \tag{70}$$
$$\leq h_2\left(\frac{\varepsilon}{1+\varepsilon}\right) + \frac{1}{1+\varepsilon} H(A|B)_\rho + \frac{\varepsilon}{1+\varepsilon} H(A|B)_{\Delta'}. \tag{71}$$

The first equality follows from Exercise 8, and the second equality follows from the definition of quantum relative entropy. The first inequality follows because $H(AB) \leq H(Y) + H(AB|Y)$ for a classical–quantum state on systems Y and AB, here taking the state as

$$\frac{1}{1+\varepsilon} |0\rangle\langle 0|_Y \otimes \rho_{AB} + \frac{\varepsilon}{1+\varepsilon} |1\rangle\langle 1|_Y \otimes \Delta'_{AB}. \tag{72}$$

The third equality follows from algebra and the definition of quantum relative entropy. The last inequality follows from Exercise 8. From concavity of the conditional entropy, we have that

$$H(A|B)_\omega \geq \frac{1}{1+\varepsilon} H(A|B)_\sigma + \frac{\varepsilon}{1+\varepsilon} H(A|B)_\Delta. \tag{73}$$

Putting together the upper and lower bounds on $H(A|B)_\omega$, we find that

$$H(A|B)_\sigma - H(A|B)_\rho \leq (1+\varepsilon)\, h_2\left(\frac{\varepsilon}{1+\varepsilon}\right) + \varepsilon \left[H(A|B)_{\Delta'} - H(A|B)_\Delta\right] \tag{74}$$
$$\leq (1+\varepsilon)\, h_2\left(\frac{\varepsilon}{1+\varepsilon}\right) + 2\varepsilon \log \dim(\mathcal{H}_A), \tag{75}$$

where the second inequality follows from a dimension bound for the conditional entropy.

The statements for classical–quantum states follow because the density operator Δ is classical–quantum in this case and we know that $H(X|B)_\Delta, H(B|X)_\Delta \geq 0$. □