Lecture 2

Max-relative entropy and conditional min-entropy

In this lecture we will first define the *max-relative entropy* and observe some of its properties. We will then define the *conditional min-entropy* in terms of the quantum max-relative entropy, derive an alternative characterization of this quantity, and consider the conditional min-entropy of a few example classes of states.

Before proceeding to the definition of the max-relative entropy, it will be helpful to consider the ordinary quantum relative entropy and its relationship to the conditional quantum entropy as a source of inspiration. Recall that the quantum relative entropy is defined as follows for all density operators ρ and all positive semidefinite operators Q acting on the same complex Euclidean space:

$$\mathrm{D}(\rho\|Q) = \begin{cases} \mathrm{Tr}(\rho\log\rho) - \mathrm{Tr}(\rho\log Q) & \mathrm{im}(\rho) \subseteq \mathrm{im}(Q) \\ \infty & \mathrm{im}(\rho) \not\subseteq \mathrm{im}(Q). \end{cases} \tag{2.1}$$

We can define this function more generally for any positive semidefinite operator P in place of the density operator ρ, but our focus will be on the case where the first argument is a density operator.

One way to think about the quantum relative entropy is that it represents the *loss of efficiency*, measured in bits, that is incurred when one plans ahead for Q but receives ρ instead. This is highly informal, and should not be taken too seriously, but we will allow this intuitive description to suggest some useful terminology: we will refer to the second argument Q in the quantum relative entropy as the *model*, and to the first argument ρ as the *actual state*, for the sake of convenience.

Irrespective of how we choose to interpret the quantum relative entropy function, there is no denying its enormous utility as a "helper function," through which fundamental entropic quantities may be defined and analyzed. In particular, the conditional quantum entropy and the quantum mutual information are defined in

terms of the quantum relative entropy as follows:

$$\begin{aligned} \mathrm{H}(\mathsf{X}\,|\,\mathsf{Y})_\rho &= -\,\mathrm{D}(\rho\|\mathbb{1}_{\mathcal{X}}\otimes\rho[\mathsf{Y}]), \\ \mathrm{I}(\mathsf{X}:\mathsf{Y})_\rho &= \mathrm{D}(\rho\|\rho[\mathsf{X}]\otimes\rho[\mathsf{Y}]), \end{aligned} \tag{2.2}$$

for all $\rho \in \mathrm{D}(\mathcal{X}\otimes\mathcal{Y})$. Then, though properties of the quantum relative entropy, one may establish important properties of the conditional quantum entropy and quantum mutual information. For example, through the *joint convexity* of quantum relative entropy,

$$\mathrm{D}(\lambda\rho_0 + (1-\lambda)\rho_1\|\lambda Q_0 + (1-\lambda)Q_1) \leq \lambda\,\mathrm{D}(\rho_0\|Q_0) + (1-\lambda)\,\mathrm{D}(\rho_1\|Q_1), \tag{2.3}$$

one may prove the critically important *strong subadditivity* property of von Neumann entropy, which may be expressed as

$$\mathrm{H}(\mathsf{X}|\mathsf{Y},\mathsf{Z})_\rho \leq \mathrm{H}(\mathsf{X}|\mathsf{Y})_\rho \tag{2.4}$$

for every $\rho \in \mathrm{D}(\mathcal{X}\otimes\mathcal{Y}\otimes\mathcal{Z})$.

The quantum relative entropy, and the entropic quantities it defines, tell us a great deal about the so-called *i.i.d. limit*, where an increasing number of independent copies of a given state are made available. In contrast, when our interest is in the so-called *one-shot* setting, where our concern is primarily with a single copy of a given state, the quantum relative entropy and the quantities it defines have limited value.

2.1 Quantum max-relative entropy

The *quantum max-relative entropy* (or just max-relative entropy for short) offers an alternative to the ordinary quantum relative entropy that is relevant in the one-shot setting. While it is a different function from the quantum relative entropy, it does possess some of the same general characteristics that make the quantum relative entropy function useful. As we will see in a couple of lectures, the ordinary quantum relative entropy can in fact be recovered from the max-relative entropy (or, to be more precise, a *smoothed* version of max-relative entropy) by applying it in the i.i.d. limit.

Definition 2.1 (Quantum max-relative entropy)**.** For a density operator ρ and a positive semidefinite operator Q acting on the same complex Euclidean space, the *quantum max-relative entropy* of ρ with respect to Q is defined as follows:

$$\mathrm{D}_{\max}(\rho\|Q) = \inf\{\lambda\in\mathbb{R} \,:\, \rho \leq 2^{\lambda}Q\}. \tag{2.5}$$

Remark 2.2. The same definition may be used when ρ is any positive semidefinite operator, and not necessarily a density operator. It is common, in particular, that ρ is taken to be a sub-normalized state, meaning $\rho \geq 0$ and $\mathrm{Tr}(\rho) \leq 1$. In this course, however, we will focus on the case that ρ is normalized.

Let us first observe the equivalence of the following statements:

1. $\mathrm{D}_{\max}(\rho \| Q) < \infty$
2. $\mathrm{im}(\rho) \subseteq \mathrm{im}(Q)$ (or, equivalently, $\ker(Q) \subseteq \ker(\rho)$)
3. $\mathrm{D}(\rho \| Q) < \infty$

In particular, the max-relative entropy is finite if and only if the ordinary quantum relative entropy is finite.

We may also observe that the max-relative entropy can be expressed through a semidefinite program. More specifically, the max-relative entropy is the *logarithm* of the optimal value of the following semidefinite program (where $\mathcal{X}$ is the complex Euclidean space upon which ρ and Q act).

Problem 2.1 (SDP for max-relative entropy)

Primal problem		Dual problem	
minimize:	η	*maximize:*	$\langle \rho, X \rangle$
subject to:	$\rho \leq \eta Q$	*subject to:*	$\langle Q, X \rangle \leq 1$
	$\eta \geq 0$		$X \in \mathrm{Pos}(\mathcal{X})$

Alternatively, the max-relative entropy is the *negative logarithm* of the optimal value of the following semidefinite program.

Problem 2.2 (Reciprocal SDP for max-relative entropy)

Primal problem		Dual problem	
maximize:	μ	*minimize:*	$\langle Q, Y \rangle$
subject to:	$\mu\rho \leq Q$	*subject to:*	$\langle \rho, Y \rangle \geq 1$
	$\mu \geq 0$		$Y \in \mathrm{Pos}(\mathcal{X})$

Notice that all four of the problems just suggested are strictly feasible when $\mathrm{im}(\rho) \subseteq \mathrm{im}(Q)$. Slater's theorem therefore implies that strong duality holds under this assumption for both semidefinite programs, with optimal values always being achieved in all four problems. Strong duality also holds when $\mathrm{im}(\rho) \not\subseteq \mathrm{im}(Q)$; in this case the optimal value of both the primal and dual forms in Optimization Problem 2.1 is positive infinity, while the optimal value of both the primal and dual forms in Optimization Problem 2.2 is zero.

Two alternative characterizations of max-relative entropy

We will now take moment to observe two alternative characterizations of the max-relative entropy. For the first, observe that if $\operatorname{im}(\rho) \subseteq \operatorname{im}(Q)$, then the condition $\rho \leq 2^{\lambda} Q$ is equivalent to

$$\left\| \sqrt{Q^{+}} \rho \sqrt{Q^{+}} \right\| \leq 2^{\lambda}. \tag{2.6}$$

Therefore, we have

$$\mathrm{D}_{\max}(\rho \| Q) = \begin{cases} \log\left(\left\| \sqrt{Q^{+}} \rho \sqrt{Q^{+}} \right\| \right) & \operatorname{im}(\rho) \subseteq \operatorname{im}(Q) \\ \infty & \operatorname{im}(\rho) \not\subseteq \operatorname{im}(Q). \end{cases} \tag{2.7}$$

We may alternatively write

$$\mathrm{D}_{\max}(\rho \| Q) = \log\left(\left\| Q^{-1/2} \rho Q^{-1/2} \right\| \right), \tag{2.8}$$

with the somewhat informal understanding that the expression evaluates to ∞ in case $\operatorname{im}(\rho) \not\subseteq \operatorname{im}(Q)$.

The second alternative characterization of the max-relative entropy begins with the observation that the condition $\rho \leq 2^{\lambda} Q$ is equivalent to $\langle \rho, Z \rangle \leq 2^{\lambda} \langle Q, Z \rangle$ for all positive definite operators Z. Therefore, assuming $Q \neq 0$, we find that

$$\mathrm{D}_{\max}(\rho \| Q) = \sup_{Z > 0} \log\left(\frac{\langle \rho, Z \rangle}{\langle Q, Z \rangle} \right). \tag{2.9}$$

Interpretation of max-relative entropy

One simple and intuitive way to think about the max-relative entropy $\mathrm{D}_{\max}(\rho \| Q)$ is as follows. Suppose that one attempts to express Q as a nonnegative linear combination of ρ along with any other collection of positive semidefinite operators. We can amalgamate the other positive semidefinite operators and associated nonnegative scalars into a single positive semidefinite operator R, for the sake of focusing on the relationship between ρ and Q, and we obtain an expression like this:

$$Q = \eta \rho + R \qquad (\text{where } R \geq 0). \tag{2.10}$$

The largest that the value η can be, assuming we are free to choose R however we wish, is precisely $2^{-\mathrm{D}_{\max}(\rho \| Q)}$.

If $Q = \sigma$ is itself a density operator, then necessarily $\eta \in [0, 1]$, and we may think of this value as being a probability. The simple fact that $\eta \leq 1$ immediately yields a variant of Klein's inequality for the max-relative entropy:

$$\mathrm{D}_{\max}(\rho \| \sigma) \geq 0, \tag{2.11}$$

with equality if and only if $\rho = \sigma$. If, on the other hand, σ is "highly dissimilar" to ρ, then any convex combination involving ρ and yielding σ must take the probability η associated with ρ to be small, so $\mathrm{D}_{\max}(\rho\|\sigma)$ must be large. In the extreme case that $\mathrm{im}(\rho) \not\subseteq \mathrm{im}(Q)$, then any expression of Q taking the form (2.10) must have $\eta = 0$, which is consistent with $\mathrm{D}_{\max}(\rho\|\sigma) = \infty$.

Monotonicity of max-relative entropy

Next let us observe that the max-relative entropy is monotonic with respect to the action of channels, meaning that

$$\mathrm{D}_{\max}(\Phi(\rho)\|\Phi(Q)) \leq \mathrm{D}_{\max}(\rho\|Q) \tag{2.12}$$

for all $\rho \in \mathrm{D}(\mathcal{X})$, $Q \in \mathrm{Pos}(\mathcal{X})$, and $\Phi \in \mathrm{C}(\mathcal{X}, \mathcal{Y})$. In fact, complete positivity is not required; the inequality (2.12) holds for all Φ positive and trace preserving.

Before we prove that the max-relative entropy is monotonic in the sense just described, let us noted that we cannot follow a similar route to this fact that we followed when proving the analogous fact for the ordinary quantum relative entropy in CS 766/QIC 820—which was through the joint convexity of quantum relative entropy. This is because *the max-relative entropy is not jointly convex*—and this is a sense in which it differs from the ordinary quantum relative entropy. The max-relative entropy is, however, *jointly quasi-convex*:

$$\mathrm{D}_{\max}\left(\sum_{k=1}^{n} p_k \rho_k \,\middle\|\, \sum_{k=1}^{n} p_k Q_k\right) \leq \max_{k\in\{1,\ldots,n\}} \mathrm{D}_{\max}(\rho_k\|Q_k). \tag{2.13}$$

The fact that the max-relative entropy is monotonic with respect to the action of channels, however, is not only true but is almost immediate from the definition of the max-relative entropy. Specifically, if we have $\rho \leq 2^{\lambda} Q$ for some choice of λ, then $\Phi(\rho) \leq 2^{\lambda}\Phi(Q)$ by the positivity of Φ, from which (2.12) follows. The assumption that Φ preserves trace implies that $\mathrm{Tr}(\Phi(\rho)) = 1$, so that it is a suitable first argument to the max-relative entropy—but this assumption can be dropped altogether, provided that we're willing to allow $\Phi(\rho)$ as a first argument to the max-relative entropy function.

Max-relative entropy upper-bounds relative entropy

One can prove that the max-relative entropy is at least as large as the ordinary quantum relative entropy, meaning

$$\mathrm{D}(\rho\|Q) \leq \mathrm{D}_{\max}(\rho\|Q) \tag{2.14}$$

for all density operators ρ and all positive semidefinite operators Q.

One way to prove this is to use the fact that the logarithm is an *operator monotone* function: for all positive definite operators P and Q with $P \leq Q$, it is the case that $\log(P) \leq \log(Q)$. This is not a trivial fact to prove, but it is well-known, and you should have no trouble finding a proof if you search for one.

Now, suppose that λ satisfies $\rho \leq 2^{\lambda} Q$, or equivalently $2^{-\lambda}\rho \leq Q$. We then have

$$\mathrm{D}(\rho\|Q) = \mathrm{Tr}(\rho\log\rho) - \mathrm{Tr}(\rho\log Q) \leq \mathrm{Tr}(\rho\log\rho) - \mathrm{Tr}\big(\rho\log\big(2^{-\lambda}\rho\big)\big) = \lambda, \quad (2.15)$$

and the relation (2.14) follows by minimizing over λ.

Max-relative entropy of tensor products and block operators

The max-relative entropy is additive with respect to tensor products:

$$\mathrm{D}_{\max}\big(\rho_0 \otimes \rho_1 \big\| Q_0 \otimes Q_1\big) = \mathrm{D}_{\max}(\rho_0\|Q_0) + \mathrm{D}_{\max}(\rho_1\|Q_1). \quad (2.16)$$

(The ordinary relative entropy is also additive with respect to tensor products in the same way.) The characterization

$$\mathrm{D}_{\max}(\rho\|Q) = \begin{cases} \log\Big(\Big\| \sqrt{Q^{+}}\,\rho\,\sqrt{Q^{+}}\Big\|\Big) & \mathrm{im}(\rho) \subseteq \mathrm{im}(Q) \\ \infty & \mathrm{im}(\rho) \not\subseteq \mathrm{im}(Q) \end{cases} \quad (2.17)$$

offers an easy route to a proof of this fact. Observe in particular that this implies that, for every choice of ρ, Q, and a positive integer n, we have

$$\mathrm{D}_{\max}\big(\rho^{\otimes n}\big\|Q^{\otimes n}\big) = n\,\mathrm{D}_{\max}(\rho\|Q). \quad (2.18)$$

The max-relative entropy also obeys the following identity, for any choice of density operator $\rho_1, \ldots, \rho_n$, positive semidefinite operators $Q_1, \ldots, Q_n$, and a probability vector $(p_1, \ldots, p_n)$:

$$\mathrm{D}_{\max}\left(\sum_{k=1}^{n} p_k |k\rangle\langle k| \otimes \rho_k \,\middle\|\, \sum_{k=1}^{n} p_k |k\rangle\langle k| \otimes Q_k\right) = \max_{k\in\{1,\ldots,n\}} \mathrm{D}_{\max}(\rho_k\|Q_k). \quad (2.19)$$

Using the formula $\mathrm{D}_{\max}(\rho\|\eta Q) = \mathrm{D}_{\max}(\rho\|Q) - \log(\eta)$, we obtain this formula for the situation in which the probabilities $p_1, \ldots, p_n$ are not included in the blocks of the second operator:

$$\begin{aligned} &\mathrm{D}_{\max}\left(\sum_{k=1}^{n} p_k |k\rangle\langle k| \otimes \rho_k \,\middle\|\, \sum_{k=1}^{n} |k\rangle\langle k| \otimes Q_k\right) \\ &\qquad = \max_{k\in\{1,\ldots,n\}} \big(\mathrm{D}_{\max}(\rho_k\|Q_k) + \log(p_k)\big). \end{aligned} \quad (2.20)$$

In contrast, the ordinary quantum relative entropy obeys this equation:

$$\mathrm{D}\left(\sum_{k=1}^{n} p_k |k\rangle\langle k| \otimes \rho_k \,\middle\|\, \sum_{k=1}^{n} p_k |k\rangle\langle k| \otimes Q_k\right) = \sum_{k=1}^{n} p_k \,\mathrm{D}(\rho_k \| Q_k). \tag{2.21}$$

Using the equation $\mathrm{D}(\rho\|\eta Q) = \mathrm{D}(\rho\|Q) - \log(\eta)$, we then conclude that

$$\mathrm{D}\left(\sum_{k=1}^{n} p_k |k\rangle\langle k| \otimes \rho_k \,\middle\|\, \sum_{k=1}^{n} |k\rangle\langle k| \otimes Q_k\right) = \sum_{k=1}^{n} p_k \,\mathrm{D}(\rho_k \| Q_k) - \mathrm{H}(p). \tag{2.22}$$

2.2 Conditional min-entropy

As was already mentioned at the beginning of the lecture, the ordinary conditional quantum entropy is given by the formula

$$\mathrm{H}(\mathsf{X}|\mathsf{Y})_\rho = -\mathrm{D}(\rho\|\mathbb{1}_{\mathcal{X}} \otimes \rho[\mathsf{Y}]). \tag{2.23}$$

We may also observe that

$$\mathrm{D}(\rho\|\mathbb{1}_{\mathcal{X}} \otimes \rho[\mathsf{Y}]) = \inf_{\sigma\in\mathrm{D}(\mathcal{Y})} \mathrm{D}(\rho\|\mathbb{1}_{\mathcal{X}} \otimes \sigma); \tag{2.24}$$

the infimum value is always obtained when $\sigma = \rho[\mathsf{Y}]$. With this fact in mind, we define the conditional min-entropy as follows.

Definition 2.3. Let X and Y be registers and let $\rho \in \mathrm{D}(\mathcal{X}\otimes\mathcal{Y})$ be a state of these registers. The *conditional min-entropy* of X given Y for the state ρ is defined as

$$\mathrm{H}_{\min}(\mathsf{X}|\mathsf{Y})_\rho = -\inf_{\sigma\in\mathrm{D}(\mathcal{Y})} \mathrm{D}_{\max}(\rho\|\mathbb{1}_{\mathcal{X}} \otimes \sigma). \tag{2.25}$$

Remark 2.4. It is not, in general, the case that the infimum in (2.25) is achieved when $\sigma = \rho[\mathsf{Y}]$.

By expanding the definition of the max-relative entropy, one may alternatively express the conditional min-entropy in the following way:

$$\begin{aligned} 2^{-\mathrm{H}_{\min}(\mathsf{X}|\mathsf{Y})_\rho} &= \inf\{\eta \geq 0 \,:\, \rho \leq \eta\mathbb{1}_{\mathcal{X}} \otimes \sigma,\, \sigma \in \mathrm{D}(\mathcal{Y})\} \\ &= \inf\{\mathrm{Tr}(Y) \,:\, \rho \leq \mathbb{1}_{\mathcal{X}} \otimes Y,\, Y \in \mathrm{Pos}(\mathcal{Y})\}. \end{aligned} \tag{2.26}$$

The conditional min-entropy is always at most the ordinary conditional quantum entropy: $\mathrm{H}_{\min}(\mathsf{X}|\mathsf{Y})_\rho \leq \mathrm{H}(\mathsf{X}|\mathsf{Y})_\rho$. This fact follows from the fact that the max-relative entropy is at least the ordinary quantum relative entropy, for then we have

$$\mathrm{D}_{\max}(\rho\|\mathbb{1}_{\mathcal{X}} \otimes \sigma) \geq \mathrm{D}(\rho\|\mathbb{1}_{\mathcal{X}} \otimes \sigma) \tag{2.27}$$

for all density operators σ, implying the claimed inequality.

Semidefinite program for conditional min-entropy

It is evident from (2.26) that the conditional min-entropy can be expressed as a semidefinite program. In particular, the quantity $\mathrm{H}_{\min}(\mathsf{X}|\mathsf{Y})_\rho$ is the negative logarithm of the optimal value of the following semidefinite program.

Problem 2.3 (SDP for conditional min-entropy)

Primal problem		Dual problem	
maximize:	$\langle \rho, X \rangle$	*minimize:*	$\mathrm{Tr}(Y)$
subject to:	$\mathrm{Tr}_{\mathcal{X}}(X) = \mathbb{1}_{\mathcal{Y}}$	*subject to:*	$\mathbb{1}_{\mathcal{X}} \otimes Y \geq \rho$
	$X \in \mathrm{Pos}(\mathcal{X} \otimes \mathcal{Y})$		$Y \in \mathrm{Herm}(\mathcal{Y})$

The dual problem is clearly consistent with the expression (2.26), whereas the primal problem corresponds (essentially) to an optimization of a linear function (represented by the state ρ) over all channels $\Phi \in \mathrm{C}(\mathcal{Y}, \mathcal{X})$. There is a useful and intuitive way to think about this optimization, but first we will take a moment to introduce a useful concept, the *transpose* of a channel.

Definition 2.5. Let $\mathcal{X}$ and $\mathcal{Y}$ be complex Euclidean spaces and let $\Phi \in \mathrm{T}(\mathcal{Y}, \mathcal{X})$. The *transpose* of Φ is the unique map $\Phi^{\mathsf{T}} \in \mathrm{T}(\mathcal{X}, \mathcal{Y})$ satisfying the equation

$$\bigl(\Phi^{\mathsf{T}} \otimes \mathbb{1}_{\mathrm{L}(\mathcal{X})}\bigr)\bigl(\mathrm{vec}(\mathbb{1}_{\mathcal{X}})\,\mathrm{vec}(\mathbb{1}_{\mathcal{X}})^*\bigr) = \bigl(\mathbb{1}_{\mathrm{L}(\mathcal{Y})} \otimes \Phi\bigr)\bigl(\mathrm{vec}(\mathbb{1}_{\mathcal{Y}})\,\mathrm{vec}(\mathbb{1}_{\mathcal{Y}})^*\bigr). \tag{2.28}$$

Equivalently, the map $\Phi^{\mathsf{T}} \in \mathrm{T}(\mathcal{X}, \mathcal{Y})$ is the (uniquely determined) map whose Choi representation is given by

$$J(\Phi^{\mathsf{T}}) = \bigl(\mathbb{1}_{\mathrm{L}(\mathcal{Y})} \otimes \Phi\bigr)\bigl(\mathrm{vec}(\mathbb{1}_{\mathcal{Y}})\,\mathrm{vec}(\mathbb{1}_{\mathcal{Y}})^*\bigr). \tag{2.29}$$

Here is a short list of facts concerning this notion, all of which are straightforward to prove.

1. $(\Phi^{\mathsf{T}})^{\mathsf{T}} = \Phi$.
2. The map $\Phi \mapsto \Phi^{\mathsf{T}}$ from $\mathrm{T}(\mathcal{Y}, \mathcal{X})$ to $\mathrm{T}(\mathcal{X}, \mathcal{Y})$, is linear, one-to-one, and onto.
3. $\Phi^{\mathsf{T}} \in \mathrm{CP}(\mathcal{X}, \mathcal{Y})$ if and only if $\Phi \in \mathrm{CP}(\mathcal{Y}, \mathcal{X})$.
4. Φ^{T} is unital if and only if Φ preserves trace.

Finally, one may observe that Φ^{T} is (as you might have guessed) the map that is obtained by taking any Kraus representation of Φ and transposing the Kraus operators.

Returning to Optimization Problem 2.3, let us consider the set $\mathcal{A}$ of primal feasible operators, which can be expressed in multiple ways based on the facts about the transpose of a map just listed:

$$\begin{aligned}
\mathcal{A} &= \{X \in \mathrm{Pos}(\mathcal{X} \otimes \mathcal{Y}) \,:\, \mathrm{Tr}_{\mathcal{X}}(X) = \mathbb{1}_{\mathcal{Y}}\} \\
&= \big\{\big(\Phi \otimes \mathbb{1}_{\mathrm{L}(\mathcal{Y})}\big)\big(\mathrm{vec}(\mathbb{1}_{\mathcal{Y}})\,\mathrm{vec}(\mathbb{1}_{\mathcal{Y}})^*\big) \,:\, \Phi \in \mathrm{C}(\mathcal{Y},\mathcal{X})\big\} \\
&= \big\{\big(\mathbb{1}_{\mathrm{L}(\mathcal{X})} \otimes \Phi^{\mathsf{T}}\big)\big(\mathrm{vec}(\mathbb{1}_{\mathcal{X}})\,\mathrm{vec}(\mathbb{1}_{\mathcal{X}})^*\big) \,:\, \Phi \in \mathrm{C}(\mathcal{Y},\mathcal{X})\big\} \\
&= \big\{\big(\mathbb{1}_{\mathrm{L}(\mathcal{X})} \otimes \Psi\big)\big(\mathrm{vec}(\mathbb{1}_{\mathcal{X}})\,\mathrm{vec}(\mathbb{1}_{\mathcal{X}})^*\big) \,:\, \Psi \in \mathrm{CP}(\mathcal{X},\mathcal{Y}),\, \Psi(\mathbb{1}_{\mathcal{X}}) = \mathbb{1}_{\mathcal{Y}}\big\}.
\end{aligned} \tag{2.30}$$

The optimal value of the semidefinite program is $2^{-\mathrm{H}_{\min}(\mathsf{X}|\mathsf{Y})_\rho}$, so

$$\begin{aligned}
&2^{-\mathrm{H}_{\min}(\mathsf{X}|\mathsf{Y})_\rho} \\
&= \sup\big\{\big\langle \rho, (\mathbb{1}_{\mathrm{L}(\mathcal{X})} \otimes \Psi)\big(\mathrm{vec}(\mathbb{1}_{\mathcal{X}})\,\mathrm{vec}(\mathbb{1}_{\mathcal{X}})^*\big)\big\rangle \,:\, \Psi \in \mathrm{CP}(\mathcal{X},\mathcal{Y}),\, \Psi(\mathbb{1}_{\mathcal{X}}) = \mathbb{1}_{\mathcal{Y}}\big\} \\
&= \sup\big\{\big\langle (\mathbb{1}_{\mathrm{L}(\mathcal{X})} \otimes \Psi^*)(\rho), \mathrm{vec}(\mathbb{1}_{\mathcal{X}})\,\mathrm{vec}(\mathbb{1}_{\mathcal{X}})^*\big\rangle \,:\, \Psi \in \mathrm{CP}(\mathcal{X},\mathcal{Y}),\, \Psi(\mathbb{1}_{\mathcal{X}}) = \mathbb{1}_{\mathcal{Y}}\big\} \\
&= \sup\big\{\big\langle (\mathbb{1}_{\mathrm{L}(\mathcal{X})} \otimes \Xi)(\rho), \mathrm{vec}(\mathbb{1}_{\mathcal{X}})\,\mathrm{vec}(\mathbb{1}_{\mathcal{X}})^*\big\rangle \,:\, \Xi \in \mathrm{C}(\mathcal{Y},\mathcal{X})\big\}.
\end{aligned} \tag{2.31}$$

That is,

$$2^{-\mathrm{H}_{\min}(\mathsf{X}|\mathsf{Y})_\rho} = n \cdot \sup_{\Xi \in \mathrm{C}(\mathcal{Y},\mathcal{X})} \big\langle \big(\mathbb{1}_{\mathrm{L}(\mathcal{X})} \otimes \Xi\big)(\rho), \tau \big\rangle \tag{2.32}$$

where

$$\tau = \frac{1}{n} \sum_{a,b=1}^{n} |a\rangle\langle b| \otimes |a\rangle\langle b| \quad \text{and} \quad n = \dim(\mathcal{X}). \tag{2.33}$$

In words, $2^{-\mathrm{H}_{\min}(\mathsf{X}|\mathsf{Y})_\rho}$ is equal to $\dim(\mathcal{X})$ times the maximum squared-fidelity, over all channels $\Xi \in \mathrm{C}(\mathcal{Y},\mathcal{X})$, between the state $(\mathbb{1}_{\mathrm{L}(\mathcal{X})} \otimes \Xi)(\rho)$ and the canonical maximally entangled state $\tau \in \mathrm{D}(\mathcal{X} \otimes \mathcal{X})$.

2.3 Examples

We will conclude the lecture by considering the conditional min-entropy of a few classes of states.

Example 2.6. Suppose Y is trivial (i.e., one-dimensional), so that $\rho \in \mathrm{D}(\mathcal{X})$. We then find that

$$\begin{aligned}
\mathrm{H}_{\min}(\mathsf{X}|\mathsf{Y})_\rho &= -\inf_{\sigma \in \mathrm{D}(\mathcal{Y})} \mathrm{D}_{\max}(\rho \| \mathbb{1}_{\mathcal{X}} \otimes \sigma) \\
&= -\mathrm{D}_{\max}(\rho \| \mathbb{1}_{\mathcal{X}}) \\
&= -\log \lambda_1(\rho).
\end{aligned} \tag{2.34}$$

Naturally, we omit the register Y from this notation when it is trivial:

$$\mathrm{H}_{\min}(\mathsf{X})_\rho = \mathrm{H}_{\min}(\rho) = -\log \lambda_1(\rho). \tag{2.35}$$

Example 2.7. Suppose $\rho = \sigma \otimes \xi$ for $\sigma \in \mathrm{D}(\mathcal{X})$ and $\xi \in \mathrm{D}(\mathcal{Y})$. Through a similar calculation to the previous example, we find that

$$\begin{aligned} \mathrm{H}_{\min}(\mathsf{X}|\mathsf{Y})_\rho &= -\inf_{\xi' \in \mathrm{D}(\mathcal{Y})} \mathrm{D}_{\max}(\sigma \otimes \xi \| \mathbb{1}_{\mathcal{X}} \otimes \xi') \\ &= -\mathrm{D}_{\max}(\sigma \| \mathbb{1}_{\mathcal{X}}) \\ &= \mathrm{H}_{\min}(\mathsf{X})_\sigma. \end{aligned} \tag{2.36}$$

This is natural: if the registers X and Y are completely uncorrelated, the conditional min-entropy of X given Y is simply the min-entropy of X.

Example 2.8. Next, suppose that we have a separable state: $\rho \in \mathrm{SepD}(\mathcal{X} : \mathcal{Y})$. Then, for any channel $\Xi \in \mathrm{C}(\mathcal{Y}, \mathcal{X})$ we have

$$(\mathbb{1}_{\mathrm{L}(\mathcal{X})} \otimes \Xi)(\rho) \in \mathrm{SepD}(\mathcal{X} : \mathcal{X}); \tag{2.37}$$

applying a channel locally to one part of a separable state always results in a separable state. The inner-product between any separable state and the canonical maximally entangled state τ is at most $1/n$ (as we proved in CS 766/QIC 820), and therefore

$$2^{-\mathrm{H}_{\min}(\mathsf{X}|\mathsf{Y})_\rho} = n \cdot \sup_{\Xi \in \mathrm{C}(\mathcal{Y}, \mathcal{X})} \left\langle \left(\mathbb{1}_{\mathrm{L}(\mathcal{X})} \otimes \Xi\right)(\rho), \tau \right\rangle \leq n \cdot \frac{1}{n} = 1. \tag{2.38}$$

The conditional min-entropy of every separable state is therefore nonnegative.

By similar reasoning, for every PPT state $\rho \in \mathrm{PPT}(\mathcal{X} : \mathcal{Y}) \cap \mathrm{D}(\mathcal{X} \otimes \mathcal{Y})$ it is the case that $\mathrm{H}_{\min}(\mathsf{X}|\mathsf{Y})_\rho \geq 0$.

Example 2.9. Suppose that the τ can be recovered perfectly by applying a channel locally to Y for the state $\rho \in \mathrm{D}(\mathcal{X} \otimes \mathcal{Y})$. This is equivalent to ρ taking the form

$$\rho = (\mathbb{1}_{\mathcal{X}} \otimes V)(\tau \otimes \xi)(\mathbb{1}_{\mathcal{X}} \otimes V)^* \tag{2.39}$$

for some choice of a density operator $\xi \in \mathrm{D}(\mathcal{Z})$ and an isometry $V \in \mathrm{U}(\mathcal{X} \otimes \mathcal{Z}, \mathcal{Y})$. Then we have

$$\mathrm{H}_{\min}(\mathsf{X}|\mathsf{Y})_\rho = -\log(n), \tag{2.40}$$

which is the minimum possible value for the conditional min-entropy.

Example 2.10. Finally, suppose that ρ is a classical-quantum state:

$$\rho = \sum_{a=1}^{n} p(a) |a\rangle\langle a| \otimes \xi_a. \tag{2.41}$$

We find that

$$\begin{aligned} 2^{-\mathrm{H}_{\min}(\mathsf{X}|\mathsf{Y})_\rho} &= n \cdot \sup_{\Xi \in \mathrm{C}(\mathcal{Y},\mathcal{X})} \left\langle \left(\mathbb{1}_{\mathrm{L}(\mathcal{X})} \otimes \Xi\right)(\rho), \tau \right\rangle \\ &= \sup_{\Xi \in \mathrm{C}(\mathcal{Y},\mathcal{X})} \sum_{a=1}^{n} p(a) \langle a| \, \Xi(\xi_a) \, |a\rangle, \end{aligned} \tag{2.42}$$

with the simplification to the second line being possible because ρ is a classical-quantum state. This has the following intuitive meaning: $\mathrm{H}_{\min}(\mathsf{X}|\mathsf{Y})_\rho$ is the negative logarithm of the optimal correctness probability to identify a state chosen randomly according to the ensemble corresponding to ρ.