Contents

Chapter 4

Quantifying information

In Chapter **??** we saw an example of a communication task between Alice and Bob in which the goal was to exchange a message in a way that the transmitted message is hidden from any eavesdropper, Eve. In this scenario we saw the importance of using a large key K, that is secretly shared between Alice and Bob but looks completely random from Eve's perspective.

An important goal of the next few chapters is to identify a method for generating such a key using quantum information. The first step in doing so, that we accomplish in this chapter, is to learn about ways to quantify quantum information.

4.1 When are two quantum states almost the same?

We introduce two important measures for when two quantum states are close to each other.

4.1.1 Trace distance

The first measure of closeness is called the trace distance. This measure is essential in quantum cryptography, but also in quantum computing in general, such as for the design of quantum circuits.

Let us suppose that we would like to implement a protocol or algorithm that produces a certain quantum state ρ_{ideal}. Unfortunately, due to imperfections in the design or execution of the protocol we actually prepared a different state ρ_{real}. If we now use this protocol or algorithm as a subroutine in a much larger protocol or computation, how is this larger protocol affected if it is executed on top of the state ρ_{real} instead of ρ_{ideal}?

Intuitively, it is clear that if the states ρ_{real} and ρ_{ideal} are nearly impossible to distinguish with respect to any measurement, then it should also not matter much in the large protocol which one we use. Indeed, if the large protocol itself made the difference, then we could think of it as a measurement that distinguishes the two states.

We would thus like a distance measure that is directly related to how well it is possible to distinguish the two states by making any measurement on them. So suppose that we really don't know whether we have the real or ideal state: we are given ρ_{real} and ρ_{ideal} each with probability $1/2$, and we are challenged to decide which is the case. To this end, we can perform a two-outcome measurement with POVM elements M_{real} and $M_{\text{ideal}} = \mathbb{I} - M_{\text{real}}$. If we perform this measurement, the probability of giving the right answer is, on average,

$$p_{\text{succ}} = \frac{1}{2}\operatorname{tr}\left(M_{\text{real}}\rho_{\text{real}}\right) + \frac{1}{2}\operatorname{tr}\left(M_{\text{ideal}}\rho_{\text{ideal}}\right) = \frac{1}{2} + \frac{1}{2}\operatorname{tr}\left(M_{\text{real}}\left(\rho_{\text{real}} - \rho_{\text{ideal}}\right)\right) . \tag{4.1}$$

To find the best measurement, we can optimize over the choice of the term M_{real} above. Recall from the definition of a POVM that the only constraints are that M_{real} is Hermitian and that $0 \leq M_{\text{real}} \leq \mathbb{I}$, i.e. all eigenvalues of M_{real} are real and lie between 0 and 1. This means that we can write the maximum probability of successfully distinguishing between the two states as

$$p_{\text{succ}}^{\max} = \frac{1}{2} + \frac{1}{2}\max_{0\leq M\leq\mathbb{I}}\operatorname{tr}\left(M\left(\rho_{\text{real}} - \rho_{\text{ideal}}\right)\right) . \tag{4.2}$$

What is the operator M that maximizes the trace quantity $\operatorname{tr}\left(M\left(\rho_{\text{real}} - \rho_{\text{ideal}}\right)\right)$ appearing on the right-hand side? To find this out, consider the diagonalized form of the linear operator $\rho_{\text{real}} - \rho_{\text{ideal}} = \sum_j \lambda_j |u_j\rangle\langle u_j|$, where $\{\lambda_j\}$ are the eigenvalues and $\{|u_j\rangle\}$ the eigenvectors. Expanding the trace on the right-hand side of (**??**) and using cyclicity we get

$$p_{\text{succ}}^{\max} = \frac{1}{2} + \frac{1}{2}\max_{0\leq M\leq\mathbb{I}}\sum_j \lambda_j \langle u_j| M |u_j\rangle . \tag{4.3}$$

Note that for any M such that $0 \leq M \leq \mathbb{I}$, for any unit vector $|u\rangle$ we have that $0 \leq \langle u| M |u\rangle \leq 1$. Then it is clear that, at best, we should choose M in (4.3) such that $\langle u_j| M |u_j\rangle = 0$ whenever $\lambda_j \leq 0$, and $\langle u_j| M |u_j\rangle = 1$ whenever $\lambda_j > 0$. Both conditions are satisfied by choosing M as the projector onto the positive eigenspace of the matrix $\rho_{\text{real}} - \rho_{\text{ideal}}$: if we introduce the set $S_+ = \{j|\lambda_j > 0\}$ then an optimal M is given by

$$M_{\text{opt}} = \sum_{j\in S_+} |u_j\rangle\langle u_j| . \tag{4.4}$$

You can verify that this M_{opt} satisfies $0 \leq M_{\text{opt}} \leq \mathbb{I}$ (since it is a projector), so it is a valid solution to (4.2); furthermore, by the reasoning above it is an optimal solution.

The trace distance is a distance measure that exactly captures the maximum success probability in the state distinguishing task. Formally, it is defined as follows.

Definition 4.1.1 (Trace distance)**.** *The* trace distance *between two quantum states ρ_0 and ρ_1 of the same dimension is given by*

$$D(\rho_0, \rho_1) = \max_{0 \leq M \leq \mathbb{I}} \operatorname{tr}\left(M\left(\rho_0 - \rho_1\right)\right) \ . \tag{4.5}$$

The trace distance can also be expressed as

$$D(\rho_0, \rho_1) = \frac{1}{2} \operatorname{tr}\left(\sqrt{A^\dagger A}\right) \ , \tag{4.6}$$

where $A = \rho_0 - \rho_1$.

To see why (4.5) is correct, observe that for a Hermitian matrix A, $\operatorname{tr}(\sqrt{A^\dagger A})$ is exactly the sum of the singular values of A, i.e. the sum of the absolute values of its eigenvalues. If $\operatorname{tr}(A) = 0$ (for example if A is the difference of two density matrices) then $\operatorname{tr}(\sqrt{A^\dagger A})$ is exactly twice the sum of the positive eigenvalues of A. Therefore, (4.5) follows from the optimal choice of M in (4.5) given by (4.4).

Note that in the literature, you will also see the trace distance written using the following notation

$$D(\rho_0, \rho_1) = \|\rho_0 - \rho_1\|_{\text{tr}} = \frac{1}{2}\|\rho_0 - \rho_1\|_1 \ . \tag{4.7}$$

By definition, if two states are close in trace distance then there exists no measurement — no process in the universe — that can tell them apart very well. Going back to our example of preparing ρ_{real} instead of ρ_{ideal}, it also means that as long as the two are close in trace distance then we can safely conclude that also any surrounding larger protocol will not make much of a difference between the two states. Otherwise, it would be possible to tell two states apart, but by definition we know that this is impossible.

Definition 4.1.2 (Closeness in terms of trace distance)**.** *We say that two quantum states ρ and σ are ϵ-close if $D(\rho, \sigma) \leq \epsilon$. We also write this as $\rho \approx_\epsilon \sigma$.*

Proposition 4.1.1. *The trace distance is a metric, that is, a proper distance measure that corresponds to our intuitive notions of distance. More precisely, we have the following properties for all states ρ, σ, τ:*

1. *Non-negative:* $D(\rho,\sigma) \geq 0$*, where equality is achieved if and only if* $\rho = \sigma$.
2. *Symmetric:* $D(\rho,\sigma) = D(\sigma,\rho)$.
3. *Triangle inequality:* $D(\rho,\sigma) \leq D(\rho,\tau) + D(\tau,\sigma)$.
4. *Convexity: for all* (p_i,ρ_i) *such that* $p_i \geq 0$, $\sum_i p_i = 1$*, and* ρ_i *is a density matrix, it holds that* $D(\sum_i p_i\rho_i,\sigma) \leq \sum_i p_i D(\rho_i,\sigma)$.

Example 4.1.1. *Consider* $\rho_1 = |0\rangle\langle 0|$ *and* $\rho_2 = |+\rangle\langle +|$*. To evaluate* $D(\rho_1,\rho_2)$ *we first calculate*

$$\rho_1 - \rho_2 = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} - \frac{1}{2}\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} = \frac{1}{2}\begin{pmatrix} 1 & -1 \\ -1 & -1 \end{pmatrix}. \tag{4.8}$$

Therefore, the trace distance is equal to

$$D(\rho_1,\rho_2) = \frac{1}{2}\cdot\frac{1}{2}\,\mathrm{tr}\sqrt{\begin{pmatrix} 1 & -1 \\ -1 & -1 \end{pmatrix}^2} = \frac{1}{2}\cdot\frac{1}{2}\,\mathrm{tr}\sqrt{\begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}} = \frac{1}{\sqrt{2}}. \tag{4.9}$$

Another way to do so is to first consider the diagonalization of $\rho_1 - \rho_2$*, which can be done by first calculating its eigenvalues, solving the following equation:*

$$\det\begin{pmatrix} \frac{1}{2}-\lambda & -\frac{1}{2} \\ -\frac{1}{2} & -\frac{1}{2}-\lambda \end{pmatrix} = 0. \tag{4.10}$$

The solutions are given by $\lambda = \pm\frac{1}{\sqrt{2}}$*. One can also find the eigenvector* $|e_+\rangle = (x \quad y)^T$ *corresponding to* $\lambda = \frac{1}{\sqrt{2}}$*,*

$$\frac{1}{2}\begin{pmatrix} 1 & -1 \\ -1 & -1 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \frac{1}{\sqrt{2}}\begin{pmatrix} x \\ y \end{pmatrix} \quad\Longrightarrow\quad \frac{x}{y} = \frac{-1}{\sqrt{2}-1}. \tag{4.11}$$

On the other hand, normalization condition gives $x^2 + y^2 = 1$*, and the solution is found to be*

$$x = \cos\frac{\pi}{8},\;\; y = \sin\frac{\pi}{8}. \tag{4.12}$$

The optimal measurement operator that distinguishes ρ_1,ρ_2 *is then given by* $M_{\mathrm{opt}} = |e_+\rangle\langle e_+|$*, while*

$$\mathrm{tr}\left[M_{\mathrm{opt}}\left(\rho_1 - \rho_2\right)\right] = \frac{1}{\sqrt{2}}. \tag{4.13}$$

■

Since states which are $\varepsilon-$close to each other cannot be distinguished well, it will later be convenient to have the notion of a set of states which are all $\varepsilon-$close to a particular state ρ. This is often called the $\varepsilon-$ball of ρ.

Definition 4.1.3 ($\varepsilon-$ball around ρ)**.** *Given any density matrix ρ, the $\varepsilon-$*ball around *ρ is defined as the set of all states ρ' which are $\varepsilon-$close to ρ in terms of trace distance, i.e.*

$$\mathcal{B}^{\varepsilon}(\rho) := \{\rho' \mid \rho' \geq 0, \mathrm{tr}(\rho') = 1, D(\rho, \rho') \leq \varepsilon\}. \tag{4.14}$$

4.1.2 Fidelity

A second common measure for closeness of states is known as the fidelity, which for pure states is directly related to their inner product.

The fidelity has an intuitive interpretation that applies to a situation where we want to verify whether we have managed to produce a desired target state $|\Psi\rangle$. Suppose that we want to build a machine that produces $|\Psi\rangle\langle\Psi|$, yet we are only able to produce some state ρ. Let us suppose that, having prepared ρ, we perform a measurement on it to check whether we have prepared the correct state $|\Psi\rangle$. We can do this (theoretically) by performing the two-outcome measurement

$$M_{\mathrm{succ}} = |\Psi\rangle\langle\Psi| \,, \tag{4.15}$$

$$M_{\mathrm{fail}} = \mathbb{I} - |\Psi\rangle\langle\Psi| \,. \tag{4.16}$$

The success probability of this measurement is directly related to the fidelity between the actual output state ρ and the target state $|\Psi\rangle$ as

$$\mathrm{tr}\left(M_{\mathrm{succ}}\rho\right) = \langle\Psi|\,\rho\,|\Psi\rangle = F(|\Psi\rangle\,, \rho)^2 \,. \tag{4.17}$$

More generally, we define the fidelity between two density matrices as follows.

Definition 4.1.4 (Fidelity)**.** *Given density matrices ρ_1 and ρ_2, the* fidelity *between ρ_1 and ρ_2 is*

$$F(\rho_1, \rho_2) = \mathrm{tr}\left[\sqrt{\sqrt{\rho_1}\rho_2\sqrt{\rho_1}}\right] \,. \tag{4.18}$$

For pure states $\rho_1 = |\Psi_1\rangle\langle\Psi_1|$ and $\rho_2 = |\Psi_2\rangle\langle\Psi_2|$ the fidelity takes on a simplified form:

$$F(\rho_1, \rho_2) = |\langle\Psi_1|\Psi_2\rangle| \,. \tag{4.19}$$

If only one of the states $\rho_1 = |\Psi_1\rangle\langle\Psi_1|$ is pure, we have

$$F(\rho_1, \rho_2) = \sqrt{\langle\Psi_1| \rho_2 |\Psi_1\rangle}\,. \tag{4.20}$$

As a comment we note that another way to write the fidelity is as

$$\max_{|\rho_{AP}\rangle, |\sigma_{AP}\rangle} |\langle\rho_{AP}|\sigma_{AP}\rangle|\,, \tag{4.21}$$

where $|\rho_{AP}\rangle$ and $|\sigma_{AP}\rangle$ are purifications of the states ρ_A and σ_A using a purifying system P.

Proposition 4.1.2. *For any two quantum states ρ, σ, the fidelity satisfies the following properties*

1. *Normalization:* $0 \leq F(\rho, \sigma) \leq 1$.
2. *Symmetric:* $F(\rho, \sigma) = F(\sigma, \rho)$.
3. *Multiplicative under tensor product:* $F(\rho_1 \otimes \rho_2, \sigma_1 \otimes \sigma_2) = F(\rho_1, \sigma_1) \cdot F(\rho_2, \sigma_2)$.
4. *Invariant under unitary operations:* $F(\rho, \sigma) = F(U\rho U^\dagger, U\sigma U^\dagger)$.
5. *Relation to trace distance:* $1 - F(\rho, \sigma) \leq D(\rho, \sigma) \leq \sqrt{1 - F^2(\rho, \sigma)}$. *Conversely, we also have that* $1 - D(\rho, \sigma) \leq F(\rho, \sigma) \leq \sqrt{1 - D^2(\rho, \sigma)}$.

4.2 What it means to be ignorant

Suppose that we designed a quantum protocol such that Alice's outcome is a classical n-bit string k that will ultimately form part of a secret key, such as for use in the one-time pad. (In particular, we may also want another party, Bob, to obtain the same k; for the time being we concentrate on Alice.) If Alice obtains the string k, we can represent it as a quantum state $|k\rangle\langle k|_K$. More generally, if we know that Alice obtains k with probability p_k, then the state of her key is $\sum_k p_k |k\rangle\langle k|_K$.

However, due to possible eavesdropping during the protocol it is not necessarily the case that Alice's string is perfectly uniformly random from the point of view of a third party Eve, who may have obtained side information during the protocol. In general we can model Eve's side information as a quantum state ρ_k^E that depends on k. Then, the joint state of Alice's string x and Eve's side information about it can be expressed as

$$\rho_{KE} = \sum_{k \in \{0,1\}^n} p_k |k\rangle\langle k|_K \otimes \rho_k^E\,. \tag{4.22}$$

Given such a state, how do we quantify the "security" of the string k in system K, in terms of how safely k can be used as a secret key? Informally, we want the classical string in system K to be uniformly random and uncorrelated with Eve's system. Before we arrive at a formal definition, let us first consider a few examples, where for simplicity we consider just a single bit of key.

Example 4.2.1. *Consider the state*

$$\rho_{KE} = \frac{1}{2} \sum_{k \in \{0,1\}} |k\rangle\langle k|_K \otimes |k\rangle\langle k|_E \,. \tag{4.23}$$

Clearly, we have $\rho_K = \mathrm{tr}_E(\rho_{KE}) = \mathbb{I}_K/2$. This means that if we look only at the key by itself, then it is uniform. But clearly Eve knows everything about the key: whenever K is in the state $|k\rangle\langle k|$, then so is E! In this example the information that Eve has is simply an exact classical copy of k. States of the form (4.23) *are also called classically maximally correlated states. Both systems are diagonal in the standard basis, and both systems are prepared precisely in the same state $|k\rangle\langle k|$ with some probability.* ∎

Example 4.2.2. *Consider the state $\rho_{KE} = |0\rangle\langle 0|_K \otimes \rho_E$. In this case Eve is clearly uncorreated with the key: the state is a tensor product. However, ρ_K is certainly not uniform! In fact, the only possible key here is $k = 0$, which is easy to guess for anyone. Completely insecure!* ∎

Example 4.2.3. *Consider the maximally entangled state*

$$\left|\psi^+\right\rangle_{KE} = \frac{1}{\sqrt{2}} \left(|0\rangle_K |0\rangle_E + |1\rangle_K |1\rangle_E\right)$$

and let $\rho_{KE} = |\psi^+\rangle\langle\psi^+|_{KE}$. As you know, here $\rho_K = \mathrm{tr}_E(\rho_{KE}) = \mathbb{I}/2$. That is, the key K is uniform. But is it uncorrelated from Eve? Clearly it is not: no matter what basis we measure K in, there always exists a corresponding measurement on E that yields the same outcome. This is because for all unitaries U,

$$\begin{aligned}
U_K \otimes U_E^* \left|\psi^+\right\rangle_{KE} &= (U_K \otimes \mathbb{I}_E)(\mathbb{I}_K \otimes U_E^*) \left|\psi^+\right\rangle_{KE} && (4.24)\\
&= (U_K \otimes \mathbb{I}_E)((U_K^*)^T \otimes \mathbb{I}_E) \left|\psi^+\right\rangle_{KE} && (4.25)\\
&= (U_K \otimes \mathbb{I}_E)(U_K^\dagger \otimes \mathbb{I}_E) \left|\psi^+\right\rangle_{KE} && (4.26)\\
&= (U_K U_K^\dagger \otimes \mathbb{I}_E) \left|\psi^+\right\rangle_{KE} && (4.27)\\
&= \left|\psi^+\right\rangle_{KE}, && (4.28)
\end{aligned}$$

where in the second equality, we have made used of a special property that holds for the maximally entangled state $|\psi^+\rangle_{KE}$: for any U,

$$(\mathbb{I}_K \otimes U_E)\,|\Psi\rangle_{KE} = (U_K^T \otimes \mathbb{I}_E)\,|\Psi\rangle_{KE}\ .$$

Therefore, the corresponding measurement on E is simply to measure in the basis defined by U_E^* (i.e. the basis in which U_E^* is diagonalized). ∎

Based on intuition gained from these examples, we give the following definition.

Definition 4.2.1 (Ignorant)**.** *Let ρ_{KE} be a cq state, where K is an n-bit string. We say that Eve (holding system E) is* ignorant *about the key K if and only if*

$$\rho_{KE} = \rho_{KE}^{\text{ideal}} = \frac{1}{2^n}\mathbb{I}_K \otimes \rho_E. \tag{4.29}$$

That is, the key is uniform and *uncorrelated from Eve.*

In any actual implementation, we can never hope to attain the perfection given by the state in Eq. (4.29). However, we can hope to get close to such an ideal state, motivating the following definition.

Definition 4.2.2 (Almost ignorant)**.** *Let ρ_{KE}^{real} be a cq state, where K is an n-bit string, and $\varepsilon \geq 0$. We say that Eve (holding system E) is* ε-almost ignorant *about the key K if*

$$D\left(\rho_{KE}^{\text{real}}, \rho_{KE}^{\text{ideal}}\right) \leq \epsilon\ , \tag{4.30}$$

where $\rho_{KE}^{\text{ideal}} = \frac{1}{2^n}\mathbb{I}_K \otimes \rho_E$.

Why is this be a good definition? Recall from the previous section that the trace distance measures how well it is possible to optimally distinguish between two scenarios. We saw that if two states are ϵ-close in trace distance, then no measurement can tell them apart with an advantage more than $\epsilon/2$: if we are given one of the two states with equal probability, then *any measurement* allowed by quantum mechanics will only return a correct guess with probability at most $1/2 + \epsilon/2$.

This has important consequences if we want to later use the key in another protocol, for example, in an an encryption protocol such as the one-time pad. Recall from Chapter **??** that an encryption scheme is secret/secure if and only if for all prior distributions $p(m)$ over the messages, and for all messages m, we have $p(m) = p(m|c)$, where c denotes the ciphertext. Such a secrecy can be achieved

using the one-time pad, if Eve is completely ignorant about the key. You may think of the one-time pad scheme as a type of measurement to distinguish ρ_{KE}^{ideal} and ρ_{KE}^{real}. If the security of the protocol was very different if we used ρ_{KE}^{real} instead of the ideal ρ_{KE}^{ideal}, then any "attack" by an adversary in the protocol would give a meas to distinguish the two states. But this is precisely ruled out if the states are close in trace distance!

4.3 Measuring uncertainty: the min-entropy

In the previous section we measured the quality of Alice's key k through the distance between the state ρ_{XE}^{real} and an "ideal state" ρ_{KE}^{ideal}. In general though, this distance may be extremely large. For example, suppose that X is 100 bits long, and Eve has the first two bits of K only as her side information. Then, this situation is almost perfectly distinguishable from the "ideal" situation where Eve has no information at all. However, we'd still like to say that, in some respects, we are "in good shape": sure, Eve knows two bits, but there are still 98 on which she has no information at all! This motivates us to try and find a more refined measure of uncertainty about the classical string K. To clarify that we now consider general distributions over strings, and not necessarily strings that we plan to use right away as key, we use the less suggestive letter X to represent the classical information, instead of K in the previous section.

4.3.1 The min-entropy

Let us first consider a simpler scenario where Eve has no side information, but Alice's string is not necessarily uniform either. That is, we look at the state

$$\rho_X = \sum_x p_x |x\rangle\langle x|_X \ , \tag{4.31}$$

which is our way of representing the probability distribution $\{p_x\}$ over strings x. How could we measure the uncertainty inherent in ρ_X?

When discussing communication tasks, the most useful measure is the Shannon entropy (also called von Neumann entropy in the context of quantum information) $H(X) = -\sum_x p_x \log p_x$. Is this quantity also a useful measure in the context of cryptography?

To think about this question, consider the following scenario: Suppose Alice purchased a box (possibly from Eve!) which generates a string $x = x_1, \ldots, x_n$ when she presses the "ON" button. If the string was uniformly random, then for all x, $p_x = 1/2^n$ and the Shannon entropy is $H(X) = n$. Suppose now that while we

are promised that x is uncorrelated from Eve, the distribution p_x is *not* uniform. However, we are guaranteed that the entropy is still $H(X) \approx n/2$. But suppose that we know nothing else about the box, except for this fact about the entropy of the distribution under which it generates a string. Would you still be willing to use x as an encryption key?

At first sight the situation might not look too bad. After all, while the string does not have maximum entropy $H(X) = n$, it still has half as much entropy, which for very large n is still large. Intuitively, this should mean that there is a lot of uncertainty for Eve, shouldn't it?

Let us consider the following distribution as an example:

$$p_x = \begin{cases} \frac{1}{2} & \text{for } x = 11\ldots1 \\ \frac{1}{2} \cdot \frac{1}{2^n - 1} & \text{otherwise .} \end{cases} \tag{4.32}$$

Exercise 4.3.1. *Show that the entropy of X with distribution $\{p_x\}$ is $H(X) \approx n/2$.* ■

So the entropy is large. But is there a lot of uncertainty for Eve? Note that the probability that the box generates the string $x = 11\ldots1$ is $1/2$, independent of the length of the string. This means that if we attempt to use x as an encryption key, Eve will be able to guess the key, and thus decrypt any message encrypted with it, with probability $1/2$. The problem here is that this particular string has high probability of being returned by the box, which is not secure at all.

We thus see that the von Neumann/Shannon entropy is not a good measure for cryptography. Luckily, there exists an alternate entropy which is much more useful for our purposes.

Definition 4.3.1 (Min-entropy)**.** *Given any probability distribution $\{p_x\}_x$, the* min-entropy $\mathrm{H}_{\min}$ *is defined as* $\mathrm{H}_{\min}(X) = -\log \max_x p_x$.

In our example above, we see that $\mathrm{H}_{\min}(X) = -\log 1/2 = 1$. That is, the min-entropy is tiny, which reflects our observation on lack of security. Looking at it more closely, note that the min-entropy precisely captures our intuitive idea of what it means for Eve to be uncertain about x: here, Eve could guess the string output by the box with probability $1/2$. This is a constant (independent of n) and so the min-entropy is constant. In general, the best guessing strategy for Eve is to guess the most likely string, so the maximum success probability that she has in guessing the output of the box is precisely $P_{\text{guess}}(X) = \max_x p_x$. The min-entropy thus has as very neat operational interpretation as

$$\mathrm{H}_{\min}(X) = -\log P_{\text{guess}}(X) \, . \tag{4.33}$$

Remark 4.3.1. *You may wonder why the min-entropy is not also the right measure of uncertainty for communication tasks. Note that for communication one generally considers the case where the states that are communicated take the form $\rho^{\otimes n}$, where n is reasonably large. That is, Shannon considered what happens when the users can repeatedly use the same communication channel a large number of times. In this setting, Shannon's idea for the right way to measure uncertainty was to consider $i(x) := -\log p_x$ as a measure of "surprisal", that is, a measure of the amount of information gained by observing x. This led him to introduce the Shannon entropy as a measure of the* average *surprisal $H(X) = \sum_x p_x i(x)$. When doing cryptography, however, we are always interested in the worst case, not the average case. The min-entropy $\mathrm{H}_{\min}(X) = \min_x i(x)$ is precisely this " smallest surprisal". Fig.4.1 shows the difference between these quantities, for a binary random variable.*

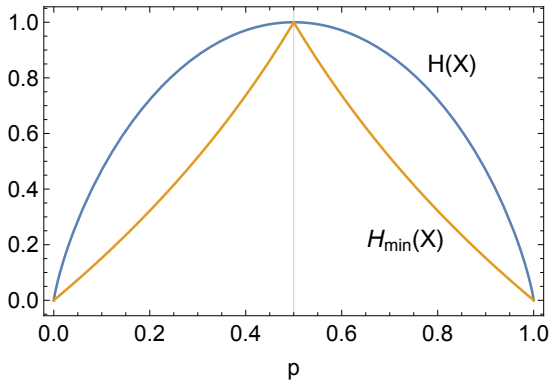

Figure 4.1: For a binary random variable $X = \{0, 1\}$, the comparison between Shannon entropy $H(X)$ and its min-entropy $\mathrm{H}_{\min}(X)$.

Exercise 4.3.2. *Show that the min-entropy satisfies the following bounds:*

$$0 \leq \mathrm{H}_{\min}(X) \leq H(X) \leq \log |X|. \tag{4.34}$$

■

4.3.2 The conditional min-entropy

Let's now consider the general case, where Eve has access to a quantum state ρ_x^E that is correlated with x. How can we quantify the uncertainty about X *given* the extra quantum register E? It turns out that just like for the von Neumann entropy ,

the min-entropy has a conditional variant $\mathrm{H}_{\min}(X|E)$. The easiest way to think about the conditional min-entropy is in terms of the probability that Eve manages to guess X given access to her quantum register E: given ρ_x^E with probability p_x, what is Eve's best chance in guessing x by making a measurement on E? This is precisely the problem of distinguishing quantum states that we considered earlier.

Definition 4.3.2 (Conditional min-entropy)**.** *Consider a bipartite cq-state ρ_{XE}, where X is classical. The* conditional min-entropy $\mathrm{H}_{\min}(X|E)$ *is defined as*

$$\mathrm{H}_{\min}(X|E)_{\rho_{XE}} := -\log P_{\mathrm{guess}}(X|E)\ , \tag{4.35}$$

where $P_{\mathrm{guess}}(X|E)$ is the probability that Eve guesses x, maximized over all possible measurements:

$$P_{\mathrm{guess}}(X|E) := \max_{\{M_x\}_x} \sum_x p_x \operatorname{tr}\left(M_x \rho_x^E\right)\ , \tag{4.36}$$

where the maximization is taken over all POVMS $\{M_x \geq 0 \mid \sum_x M_x = \mathbb{I}\}$ on E. In this context, E is also called side information *about X. When it is clear from context, we omit the subscript ρ_{XE}, i.e. we write $\mathrm{H}_{\min}(X|E)_{\rho_{XE}} = \mathrm{H}_{\min}(X|E)$.*

Note that the definition of the min-entropy involves a maximization over all possible POVM. In general, this could be hard to compute! Nevertheless, when $x \in \{0,1\}$ takes on only two values it is easy to find the optimal measurement, and the guessing probability P_{guess} is directly related to the distinguishability of reduced states ρ_0^E and ρ_1^E, i.e. the trace distance $D(\rho_0^E, \rho_1^E)$.

Example 4.3.1. *Consider the state $\rho_{XE} = \frac{1}{2}|0\rangle\langle 0|_X \otimes |0\rangle\langle 0|_E + \frac{1}{2}|1\rangle\langle 1|_X \otimes |+\rangle\langle +|_E$. Then the conditional min-entropy $\mathrm{H}_{\min}(X|E) = -\log P_{\mathrm{guess}}(X|E)$ where*

$$\begin{aligned}
P_{\mathrm{guess}}(X|E) &= \max_{\substack{M_1,\, M_2 \geq 0 \\ M_1 + M_2 = \mathbb{I}}} \left[\frac{1}{2}\operatorname{tr}\left(M_0|0\rangle\langle 0|_E\right) + \frac{1}{2}\operatorname{tr}\left(M_1|+\rangle\langle +|_E\right)\right] \\
&= \max_{0 \leq M \leq \mathbb{I}} \left[\frac{1}{2}\operatorname{tr}\left(M|0\rangle\langle 0|_E\right) + \frac{1}{2}\operatorname{tr}\left(|+\rangle\langle +|_E\right) - \frac{1}{2}\operatorname{tr}\left(M|+\rangle\langle +|_E\right)\right] \\
&= \frac{1}{2} + \frac{1}{2} \max_{0 \leq M \leq \mathbb{I}} \operatorname{tr}[M(|0\rangle\langle 0|_E - |+\rangle\langle +|_E)] \\
&= \frac{1}{2} + \frac{1}{2}\, D(|0\rangle\langle 0|_E, |+\rangle\langle +|_E).
\end{aligned}$$

■

However, if x can take more than two possible values, then it is in general difficult to compute $P_{\text{guess}}(X|E)$ by hand. However, the optimal success probability can be expressed as a *semidefinite program* (SDP). An SDP is a convex program that generalizes linear programs (LP). In particular, the optimum of an SDP can in general be approximated efficiently, in time polynomial in the dimension of the states ρ_x^E. Programming languages oriented towards linear algebra, such as Matlab or Julia, generally have packages that allow you to do this.

Exercise 4.3.3. *Show that for any cq-state ρ_{XE} we have*

$$0 \leq \mathrm{H}_{\min}(X|E) \leq \log |X| \,. \tag{4.37}$$

Furthermore, show (this is a little harder!) that

$$\mathrm{H}_{\min}(X|E) \geq \mathrm{H}_{\min}(X) - \log |E| \,, \tag{4.38}$$

where we used $|E|$ to denote the dimension of the system E (i.e. in case of qubits, the number of qubits of $|E|$ is $\log |E|$). ∎

Smoothed min-entropy. As we saw seen earlier, in our discussion of the trace distance, due to imperfections in a protocol or algorithm we often do not exactly produce the state ρ_{XE} that we want: rather, we may only manage to produce a state ρ'_{XE} which is close, and we may not even know the exact form of ρ'_{XE} (other than the fact that it is $\varepsilon-$close to ρ_{XE}). Due to this uncertainty it is usually more physically relevant to look at the *smoothed min-entropy*, which gives the maximum value of $\mathrm{H}_{\min}(X|E)$ over all states $\rho'_{AE} \in \mathcal{B}^\varepsilon(\rho_{AE})$.

Definition 4.3.3 (Smoothed conditional min-entropy)**.** *Consider a bipartite cq-state ρ_{XE}, where X is classical. The* smoothed conditional min-entropy $\mathrm{H}^\varepsilon_{\min}(X|E)$ *is defined as*

$$\mathrm{H}^\varepsilon_{\min}(X|E)_\rho = \max_{\rho' \in \mathcal{B}^\varepsilon(\rho)} \mathrm{H}_{\min}(X|E)_{\rho'}. \tag{4.39}$$

A general quantum conditional min-entropy. In full generality we may also be interested in quantifying uncertainty about *quantum* information, i.e. in a setting where the system X can also be quantum. To make the distinction explicit, we use A, instead of X, to label such a quantum system. How should we measure the conditional min-entropy $\mathrm{H}_{\min}(A|E)$? To gain intuition on how such a quantity could be defined, think of the guessing probability as a way of quantifying how close it is possible for Eve, holding system E, to put herself in a state that is *maximally correlated* with the classical system X, by guessing it correctly and holding

on to her guess. A natural quantum extension of this idea is to get as close as possible to the maximally entangled state between A and E while only allowing to perform operations on E. Here is a more formal definition.

Definition 4.3.4 (Quantum conditional min-entropy)**.** *For any bipartite density matrix ρ_{AE}, let $|A|$ be dimension of A and define*

$$\mathrm{Dec}(A|E) = \max_{\Lambda_{E\to A'}} F((\mathbb{I}_A \otimes \Lambda_{E\to A'})\rho_{AE}, |\Phi\rangle\langle\Phi|_{AA'})^2 \ ,$$

where A' is a system that has the same dimention as A and

$$|\Phi\rangle_{AA'} = \frac{1}{\sqrt{|A|}} \sum_{i=1}^{|A|} |a_i\rangle_A \otimes |a_i\rangle_{A'}$$

is the maximally entangled state between A and A' and the maximization is performed over all quantum channels Λ mapping system E to A', and the function F is the fidelity that we have seen in Definition 4.1.4. Then the conditional min-entropy of A, conditioned on E, is

$$\mathrm{H}_{\min}(A|E) = -\log\left(|A| \cdot \mathrm{Dec}(A|E)\right) \ .$$

As a remark, we note that an alternative way to express the quantum conditional min-entropy is as

$$\mathrm{H}_{\min}(A|E) = \max_{\sigma_B} \sup\left\{\lambda \in \mathbb{R} \middle| \rho_{AB} \leq 2^{-\lambda}\mathbb{I}_A \otimes \sigma_B\right\} \ . \tag{4.40}$$

The equivalence between the two definition is not obvious, but it can be shown using duality of semidefinite programming.

4.4 Uncertainty principles: a bipartite guessing game

With all the definitions in place, we start studying simple "games" that can be used as a means to certify that a certain string x generated by Alice must be at least partially unknown to an eavesdropper Eve. As a first step, we consider eavesdroppers that only have the ability to store and process classical information. Because Eve has less power, the setting will be easier to analyze. We will see in the next section how to handle the general setting of a quantum eavesdropper.

The essential property of quantum mechanics that will allow us to carry out our security arguments is called the *uncertainty principle*. Very informally, this principle allows us to limit how well Eve can predict the outcomes of different incompatible measurements on Alice's state.

To set things up, consider the following guessing game. For now, we study the game for its own sake, but later, we will see how the analysis can be used to show security of a more complex cryptographic protocol for secret key distribution.

Definition 4.4.1 (Guessing game — Alice and Eve)**.** *Suppose two parties, Alice and Eve, play the following game.*

1. *Eve prepares a qubit in an arbitrary state ρ_A and sends it to Alice.*
2. *Alice chooses a bit $\Theta \in \{0, 1\}$ uniformly at random.*
3. *If $\Theta = 0$, then Alice measures ρ_A in the standard basis. If $\Theta = 1$, then she measures in the Hadamard basis. She obtains and records a measurement outcome $X \in \{0, 1\}$.*
4. *Alice announces Θ to Eve.*
5. *Eve wins if she can guess the bit X.*

Suppose that you play the role of Eve in the guessing game, and that Alice plays exactly as described in the game. How should you choose the state ρ_A to maximize your chances of success? Note how tricky this is: for example, you could choose $\rho_A = |0\rangle\langle 0|$. Then, if Alice chooses $\Theta = 0$ you can predict the outcome — it is simply "0". But if she chooses $\Theta = 1$, her outcome is uniformly random. And if you choose $\rho_A = |+\rangle\langle +|$ then the situation is inverted.

Intuitively, the difficulty is that because the two measurements that Alice can form are in "incompatible" bases (their vectors all make an angle $0 < \varphi < \frac{\pi}{2}$), there does not seem to be a state that Eve can prepared such that she would be able to know a priori what the outcome should be in both bases.

Let's analyze formally what is the best that Eve can do in the game. In general, assuming that Alice chooses her measurement basis Θ uniformly at random her maximum success probability is

$$\begin{aligned} P_{\text{guess}}(X|\Theta) &= p(\Theta = 0) \cdot P_{\text{guess}}(X|\Theta = 0) + p(\Theta = 1) \cdot P_{\text{guess}}(X|\Theta = 1) \\ &= \frac{1}{2} \cdot \left[P_{\text{guess}}(X|\Theta = 0) + P_{\text{guess}}(X|\Theta = 1)\right] \ , \end{aligned} \tag{4.41}$$

where the second equality holds if

Note that Eve has to make a guess for each case, $\Theta = 0$ and $\Theta = 1$, when she is asked. Since we assumed that Eve has only classical information, we can record her guess for each of the two possibilities ahead of time. By symmetry, we can assume that her guess in each case is "$X = 0$". Thus her maximum success

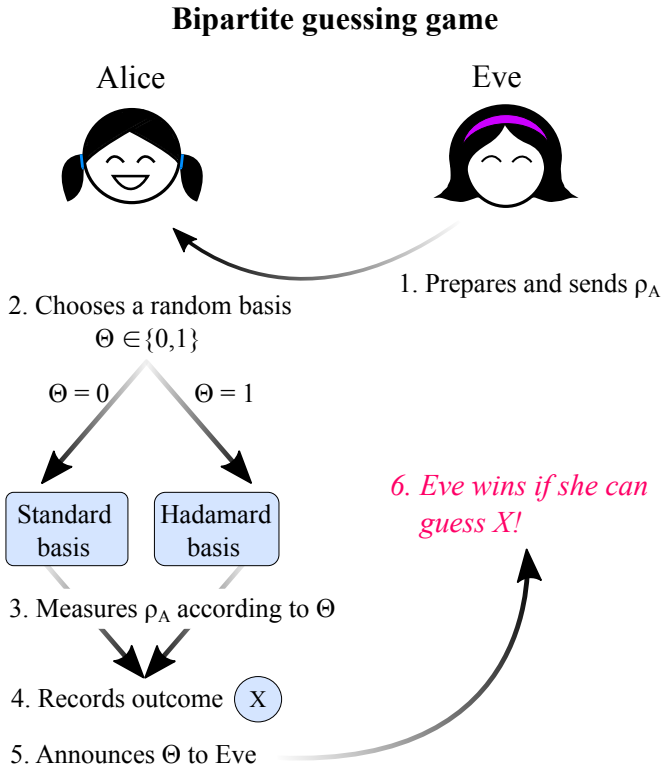

Figure 4.2: The guessing game between Alice and Eve, where Eve prepares a quantum state and sends it to Alice, who choses randomly to measure in the standard basis or in the Hadamard basis. Eve then tries to guess Alice's measurement outcome, given the basis she chosen.

probability is the maximum over all ρ_A of the chance that "$X = 0$" is actually the correct guess. In other words, continuing from (4.41) we get

$$P_{\text{guess}}(X|\Theta) = \frac{1}{2} \cdot [\text{tr}(\rho_A|0\rangle\langle 0| + \text{tr}(\rho_A|+\rangle\langle +|))] \tag{4.42}$$

$$= \frac{1}{2} \cdot \text{tr}\left[\rho_A(|0\rangle\langle 0| + |+\rangle\langle +|)\right]. \tag{4.43}$$

In the last equation the optimal choice of ρ_A is a pure state corresponding to the eigenvector of $|0\rangle\langle 0| + |+\rangle\langle +|$ with the largest eigenvalue. It is not hard to do the computation and obtain that $\lambda_{\max} = 1 + \frac{1}{\sqrt{2}}$. Therefore, we have that

$P_{\text{guess}}(X|\Theta) = \frac{1}{2} + \frac{1}{2\sqrt{2}} < 1$.

Exercise 4.4.1. *Write down explicitly the state ρ_A that Eve should prepare in order to succeed in the guessing game with probability exactly $\frac{1}{2} + \frac{1}{2\sqrt{2}}$ in the game.* ∎

Let's now consider a more general scenario where Eve may keep classical information about ρ_A. In other words, we allow Eve to prepare an arbitrary cq-state $\rho_{AC} = \sum_c p_c \rho_c^A \otimes |c\rangle\langle c|_C$ according to some distribution $\{p_c\}_c$, and send the qubit in A to Alice while keeping the classical system C. Let us convince ourselves that this scenario does not make any difference: it does not help Eve win with higher probability in the game. Indeed, the guessing probability conditioned on C is given by the average

$$P_{\text{guess}}(X|\Theta C)_{\rho_{AC}} = \sum_c p_c \, p_{\text{guess}}(X|\Theta)_{\rho_c^A} \ , \tag{4.44}$$

where we maximize over all possible $\{p_c, \rho_c^A\}_c$. But we have previously already shown the maximum possible value of $p_{\text{guess}}(X|\Theta)_{\rho_c^A}$, over all possible ρ_c^A! Therefore, we get that here also

$$P_{\text{guess}}(X|\Theta C)_{\rho_{AC}} = \frac{1}{2}\left(1 + \frac{1}{\sqrt{2}}\right) \approx 0.85. \tag{4.45}$$

The quantity $P_{\text{guess}}(X|\Theta C)$ allows us to directly compute the conditional min-entropy of Alice's outcome X, since

$$H_{\min}(X|\Theta C) = -\log P_{\text{guess}}(X|\Theta C) \approx 0.22 \ .$$

Next, let us make one more step and give Eve yet more power. Suppose that can now create an arbitrary quantum state ρ_{AE}, possibly entangled, and send only the qubit in A to Alice.

Exercise 4.4.2. *Show that if Eve can keep entanglement then there is a strategy that allows her to always win the game with probability* 1. ∎

The exercise shows that our assumption that Eve only keeps classical information was not only for convenience: it is also necessary for security, as as soon as we allow Eve to maintain entanglement with ρ_A then she may be able to guess Alice's outcome X perfectly.

How do we get around this? Here is the key: remember from Chapter **??** that entanglement is monogamous! In order to limit Eve's knowledge about Alice's measurement outcomes we will use *two* aspects of quantum mechanics:

- Uncertainty: If Eve has no (or little) entanglement with Alice, then she cannot predict the outcomes of two incompatible measurements (very well). In particular, this means it is difficult for here to guess Alice's measurement outcomes, i.e., $P_{\text{guess}}(X|E\Theta) < 1$.

- Monogamy: If we ensure that there is a large amount of entanglement between Alice and Bob, then we know that Eve can have only very little entanglement with either Alice or Bob.

4.5 Extended uncertainty relation principles: A tripartite guessing game

in order to make use of the monogamous property of entanglement we consider an extension of the guessing game from the previous section. This time we impose no constraint on the presence or absence of entanglement between Alice and Eve. Instead, we introduce a third party, Bob, whom Alice trusts. In particular, to show security against Eve, Alice and Bob may join forces to make an estimate of the amount of entanglement that Eve may be sharing with Alice. To do so, they perform an "entanglement test" between Alice and Bob to ensure that they share entanglement between themselves, in a way that will then guarantee that the entanglement between Alice and Eve is small. Here is the new formulation for the guessing game.

Definition 4.5.1 (Tripartite guessing game - Alice, Bob and Eve)**.** *Suppose three parties, Alice, Bob and Eve, play the following game.*

1. *Eve prepares an arbitrary state ρ_{ABE} such that A and B are both qubits. She sends qubit A to Alice and qubit B to Bob.*

2. *Alice chooses a bit $\Theta \in \{0, 1\}$ uniformly at random.*

3. *If $\Theta = 0$, then Alice measures A in the standard basis; if $\Theta = 1$, then she measures in the Hadamard basis. She obtains a measurement outcome $X \in \{0, 1\}$ and records it.*

4. *Alice announces Θ to both Bob and Eve.*

5. *Given Θ, Bob measures ρ_B in the basis Θ and obtains an outcome $\tilde{X}$. Likewise, Eve measures ρ_E and makes a guess X_E.*

6. *Bob and Eve win the game if $X_E = X = \tilde{X}$.*

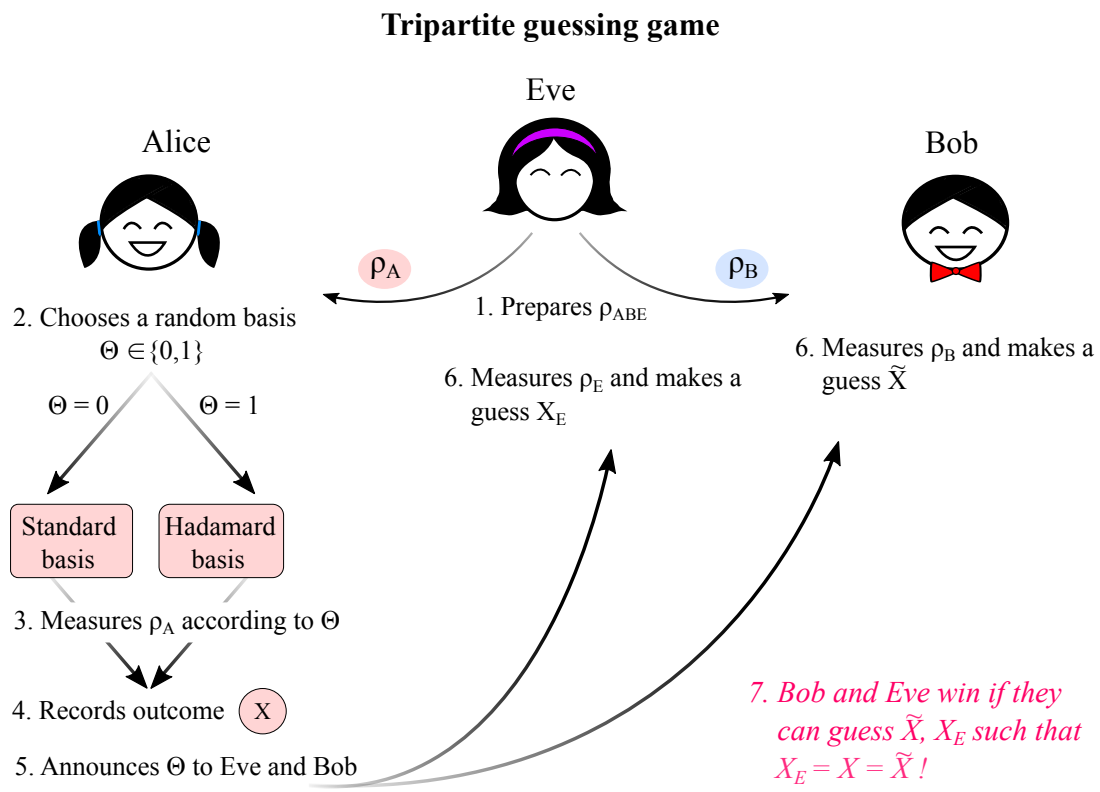

Figure 4.3: A tripartite guessing game where Eve gets to prepare the global state ρ_{ABE}. She send the qubits A and B to Alice and Bob respectively, where Alice measures randomly in either the standard or Hadamard basis. Bob and Eve both provide guesses $\tilde{X}, X_E$, and we say that they win the game if $X_E = X = \tilde{X}$.

Suppose that you try to design the best strategy for Eve in this game. How well can you do? We measure her success probability as

$$p_{\text{Tripartite}} = p\big(X = \tilde{X} = X_E\big) = \sum_{\Theta\in\{0,1\}} p_\Theta p(X = \tilde{X} = X_E|\Theta)$$
$$= \frac{1}{2}\sum_{\theta\in\{0,1\}} \text{tr}\left(\rho_{ABE}\Big(\sum_{x\in\{0,1\}} |x\rangle\langle x|^A_\theta \otimes |x\rangle\langle x|^B_\theta \otimes M^E_{x|\theta}\Big)\right), \quad (4.46)$$

where we used superscripts A, B and E to denote the systems on which we perform the measurements, and $|x\rangle_\theta$ to denote basis element x in the basis θ. That

is, $|0\rangle_0 = |0\rangle$, $|1\rangle_1 = |1\rangle$, and $|0\rangle_1 = |+\rangle$, $|1\rangle_1 = |-\rangle$. Of course, the difficulty is that we don't know anything a priori about the state ρ_{ABE} or the measurement $\{M^E_{x|\theta}\}_x$ with outcomes x that Eve will perform on E depending on the basis θ. We only know that this must be a quantum state, and Eve can only make measurements that are allowed by the laws of quantum mechanics. Since it is known that all POVMs can be realized as projective measurements using a potentially larger ancilla , we can without loss of generality assume that Eve's measurements are projective. Given her access to a smaller space only makes things more difficult for Eve and in a security analysis we are always allowed to make the adversary more (but not less!) powerful.

How do we bound the expression in (4.46)? In the previous section, when we considered a purely classical Eve, we were able to express the optimum as a simple eigenvalue problem. Here, if we fix Eve's measurements then again we obtain an eigenvalue problem

$$\max_{\rho_{ABE}} \operatorname{tr}\left(\rho_{ABE}\left(\frac{1}{2}\sum_{\Theta}\Pi_\Theta\right)\right), \tag{4.47}$$

where

$$\Pi_\Theta = \sum_{x\in\{0,1\}} |x\rangle\langle x|^A_\Theta \otimes |x\rangle\langle x|^B_\Theta \otimes M^E_{x|\Theta}\,. \tag{4.48}$$

Note that Π_Θ is a projector because for any Θ, $|x\rangle\langle x|_\Theta$ are orthogonal projectors for $x \in \{0,1\}$, and so are $M^E_{x|\Theta}$ by the assumption that Eve's measurements are projective.

Two tools from linear algebra

To make progress we we use two little tricks from linear algebra. To write them down, let us first introduce a shorthand for the maximization problem above. In general, the *operator norm* of an operator O is

$$\|O\|_\infty = \max_\rho \operatorname{tr}\left[\rho O\right]\,, \tag{4.49}$$

where the maximization is taken over all ρ such that $\operatorname{tr}[\rho] \leq 1$. When O is Hermitian we can just maximize over all quantum states ρ, that is, ρ satisfying $\rho \geq 0$ and $\operatorname{tr}(\rho) = 1$. Note that this means we can reduce the maximization problem above to studying

$$\left\|\frac{1}{2}\sum_{\theta\in\{0,1\}}\Pi_\theta\right\|_\infty. \tag{4.50}$$

For simplicity in the remainder of the section we omit the subscript ∞ and simply write $\|O\| = \|O\|_\infty$. Here are the two facts we will use:

1. For any two projectors Π_0 and Π_1 [1], we have

$$\|\Pi_0 + \Pi_1\| \leq \max\{\|\Pi_0\|, \|\Pi_1\|\} + \|\Pi_0\Pi_1\|. \tag{4.51}$$

2. If $\Pi_0 \leq P$ and $\Pi_1 \leq Q$ [2], then $\|\Pi_0\Pi_1\| \leq \|PQP\|$.

Assuming these two facts from elementary linear algebra, let us see how we can bound Eve's probability of winning in the tripartite guessing game. Using the first item,

$$\max_{M^E} \Big\| \frac{1}{2} \sum_{\theta\in\{0,1\}} \Pi_\theta \Big\|_\infty = \max_{M^E} \frac{1}{2} \Big\| \sum_{\theta\in\{0,1\}} \Pi_\theta \Big\|_\infty \tag{4.52}$$

$$\leq \frac{1}{2}\Big(1 + \max_{M^E} \|\Pi_0\Pi_1\|\Big)\,, \tag{4.53}$$

where we have used that $\|\Pi_0\|, \|\Pi_1\| \leq 1$ for any measurements M^E that Eve could make. It remains to analyze $\|\Pi_0\Pi_1\|$. For this we use the second item, for some smart choice of P and Q. Note that since all measurement operators $M^E_{x|\theta} \leq \mathbb{I}$ and $|x\rangle\langle x|_\theta \leq \mathbb{I}$, we have that

$$\Pi_0 \leq \sum_{x\in\{0,1\}} |x\rangle\langle x|^A_0 \otimes |x\rangle\langle x|^B_0 \otimes \mathbb{I}^E \tag{4.54}$$

$$\Pi_1 \leq \sum_{x\in\{0,1\}} |x\rangle\langle x|^A_1 \otimes \mathbb{I}^B \otimes M^E_{x|1} \tag{4.55}$$

Let P and Q be the operators on the right-hand side above, respectively. Using the fact that $\langle x|y\rangle = 0$ if $x \neq y$ in the same basis, and that $\sum_y M^E_{y|1} = \mathbb{I}$ for any

[1] Recall that a projector Π is a Hermitian operator such that $\Pi^2 = \Pi$.

[2] Recall that $A \leq B$ means that $B - A \geq 0$, i.e. $B - A$ is a positive semidefinite matrix.

quantum measurement Eve may make, we get

$$\begin{aligned} PQP &= \sum_{x,y,z} |x\rangle\langle x|_0^A |y\rangle\langle y|_1^A |z\rangle\langle z|_0^A \otimes |x\rangle\langle x|_0^B |z\rangle\langle z|_0^B \otimes M_{y|1}^E \\ &= \sum_{x,y} \frac{1}{2} |x\rangle\langle x|_0^A \otimes |x\rangle\langle x|_0^B \otimes M_{y|1}^E \\ &= \frac{1}{2} \sum_x |x\rangle\langle x|_0^A \otimes |x\rangle\langle x|_0^B \otimes \sum_y M_{y|1}^E \\ &= \frac{1}{2} \sum_x |x\rangle\langle x|_0^A \otimes |x\rangle\langle x|_0^B \otimes \mathbb{I}^E \, . \end{aligned}$$

This gives $\|PQP\| \leq 1/2$. Using the second trick and plugging into Eq. (4.53) we get

$$p_{\text{Tripartite}} \leq \frac{1}{2}\left(1 + \frac{1}{\sqrt{2}}\right) = \frac{1}{2} + \frac{1}{2\sqrt{2}} \, , \tag{4.56}$$

which is the same number that we obtained for the much simpler game in which Eve was all classical!

Exercise 4.5.1. *Identify explicitly Eve's optimal strategy in the game: find a state ρ_{ABE} and measurement operators $M_{x|\Theta}$ that give her a success probability of $\frac{1}{2} + \frac{1}{2\sqrt{2}}$.* ∎

As an additional remark we note that using even more linear algebra it is possible to show that also when playing the game n times "in parallel", i.e. allowing Eve to prepare a single ρ_{ABE} where A and B are n qubits, and Alice and Bob both measure all their qubits individually using an independent random basis choice for each qubit, then

$$p_{\text{Tripartite}}^{\text{n rounds}} \leq \left(\frac{1}{2} + \frac{1}{2\sqrt{2}}\right)^n \, , \tag{4.57}$$

and this bound can be achieved by Eve by preparing a tensor product state $\rho_{ABE}^{(n)} = (\rho_{ABE}^{(1)})^{\otimes n}$, where $\rho_{ABE}^{(1)}$ is an optimal choice of state for the single-qubit version of the game.

4.6 Important identities for calculations

Trace distance

$$D(\rho_{\text{real}}, \rho_{\text{ideal}}) := \max_{0 \leq M \leq \mathbb{I}} \text{tr}\left[M\left(\rho_{\text{real}} - \rho_{\text{ideal}}\right)\right] \tag{4.58}$$

$$= \frac{1}{2}\,\text{tr}\left[\sqrt{A^\dagger A}\right] \ , \qquad A = \rho_{\text{real}} - \rho_{\text{ideal}}. \tag{4.59}$$

Properties:

1. $D(\rho, \rho') \geq 0$ with equality iff $\rho = \rho'$.
2. $D(\rho, \rho') = D(\rho', \rho)$.
3. $D(\rho, \rho') + D(\rho', \rho'') \geq D(\rho, \rho'')$.
4. $D(\sum_i p_i \rho_i, \sigma) \leq \sum_i p_i D(\rho_i, \sigma)$.

Fidelity

$$F(\rho, \rho') := \text{tr}\left[\sqrt{\sqrt{\rho}\rho'\sqrt{\rho}}\right] \ . \tag{4.60}$$

If $\rho = |\psi\rangle\langle\psi|$ and $\rho' = |\psi'\rangle\langle\psi'|$, then $F(|\psi\rangle\langle\psi|, |\psi'\rangle\langle\psi'|) = \sqrt{\langle\Psi_1|\,\rho_2\,|\Psi_1\rangle}$.
Relation to trace distance: $1 - F \leq D \leq \sqrt{1 - F^2}$.

Min-entropy

Unconditional : $\text{H}_{\min}(X) = \text{H}_{\min}(\rho_X) = -\log \max_x p_x$.
Conditional : For a cq-state ρ_{XE}, $\text{H}_{\min}(X|E) := -\log P_{\text{guess}}(X|E)$, where

$$P_{\text{guess}}(X|E) := \max_{\{M_x\}_x} \sum_x p_x\,\text{tr}\left[M_x \rho_x^E\right], \{M_x \geq 0 \mid \sum_x M_x = \mathbb{I}\}.$$

Properties:

1. $0 \leq \text{H}_{\min}(X|E) \leq \text{H}_{\min}(X) \leq \log|X|$, but only for cq states! For quantum register X, $\text{H}_{\min}(X|E)$ can be negative.
2. $\text{H}_{\min}(X|E) \geq \text{H}_{\min}(X) - \log|E|$.

A secret key

A key K is secret from Eve iff it is *uniform and uncorrelated* from Eve, i.e. the joint state ρ_{KE} is of the form

$$\rho_{KE} = \frac{\mathbb{I}_K}{d_K} \otimes \rho_E. \tag{4.61}$$