1 Distinguishing between distributions of states

Let X, Y be two classical distributions over $\{0,1\}^n$ with x_i being the probability that X outputs i. Let $\{|\Psi_i\rangle \mid i \in \{0,1\}^n\}$, be an orthonormal basis. Define two quantum state probability distributions E_1, E_2.

$$E_1 = \{|\Psi_i\rangle @ \sqrt{x_i}\}$$
$$E_2 = \{|\Psi_i\rangle @ \sqrt{y_i}\}$$

How distinguishable are E_1, E_2 from one another?

We want to compute $TD(\rho_x, \rho_y)$ where

$$\rho_x = \sum_{i \in \{0,1\}^n} x_i |\Psi_i\rangle\langle\Psi_i|$$
$$\rho_y = \sum_{i \in \{0,1\}^n} y_i |\Psi_i\rangle\langle\Psi_i|$$

As these are two orthonormal bases, we know there exists a unitary U which maps $U|\Psi_i\rangle \to |i\rangle$. From the properties of trace distance, we know that $TD(\sigma, \rho) \geq TD(\mathcal{E}(\sigma), \mathcal{E}(\rho))$ so we can apply this with $\mathcal{E}$ being the unitary application of U.

$$TD(\rho_x, \rho_y) \geq TD(U\rho_x U^\dagger, U\rho_y U^\dagger)$$

We can apply $U^\dagger$ again to see that we haven't lost any distinguishing advantage.

$$TD(\rho_x, \rho_y) \geq TD(U\rho_x U^\dagger, U\rho_y U^\dagger) \geq TD(U^\dagger(U\rho_x U^\dagger)U, U^\dagger(U\rho_y U^\dagger)U) = TD(\rho_x, \rho_y)$$

From the above we see that

$$TD(U\rho_x U^\dagger, U\rho_y U^\dagger) = TD(\rho_x, \rho_y).$$

Now let's try to bound $TD(U\rho_x U^\dagger, U\rho_y U^\dagger)$.

$$U\rho_x U^\dagger = U(\sum_i x_i |\Psi_i\rangle\langle\Psi_i|)U^\dagger = \sum_i x_i U|\Psi_i\rangle\langle\Psi_i|U^\dagger = \sum_i x_i |i\rangle\langle i|$$
$$U\rho_y U^\dagger = U(\sum_i y_i |\Psi_i\rangle\langle\Psi_i|)U^\dagger = \sum_i y_i U|\Psi_i\rangle\langle\Psi_i|U^\dagger = \sum_i y_i |i\rangle\langle i|$$

Since these two matrices are diagonal, their difference is also diagonal. Thus we can find the trace distance between them

$$TD(\rho_x, \rho_y) = TD(U\rho_x U^\dagger, U\rho_y U^\dagger) = \frac{1}{2}\sum_i |x_i - y_i| = SD(X, Y)$$

2 Trace distance between two arbitrary states

2.1 If the states are orthogonal

Two orthogonal states $|\Psi\rangle, |\Phi\rangle$ should be perfectly distinguishable. To show this is the case, we know there exists some basis B with basis vectors b_i which contains both of these vectors with $b_0 := |\Psi\rangle, b_1 := |\Phi\rangle$. We know there also exists a unitary U such that $U|b_i\rangle \to |i\rangle$.

$$TD(|\Psi\rangle\langle\Psi|, |\Phi\rangle\langle\Phi|) \geq TD(U|\Psi\rangle\langle\Psi|U^\dagger, U|\Phi\rangle\langle\Phi|U^\dagger)$$

From the construction of U we know that $U|\Psi\rangle\langle\Psi|U^\dagger$ and $U|\Phi\rangle\langle\Phi|U^\dagger$ are both diagonal matrices with a singular 1. They are also on different places in the main diagonal as they would not be mapped to the same element due to orthogonality.

$$TD(U|\Psi\rangle\langle\Psi|U^\dagger, U|\Phi\rangle\langle\Phi|U^\dagger) = \frac{1}{2}(|1| - |-1|) = 1$$

2.2 If the states are not orthogonal

For simplicity let's look at the single qubit case. We can express $|\Phi\rangle = \alpha|\Psi\rangle + \beta|b\rangle$ for some normal $|b\rangle$ which is orthogonal to $|\Psi\rangle$. Again to keep things simpler we will assume α, β are real numbers.

There exists a U which maps this new basis to the computational basis.

$$U|\Psi\rangle \to |0\rangle$$
$$U|b\rangle \to |1\rangle$$

And so $U|\Phi\rangle = \alpha|0\rangle + \beta|1\rangle = \begin{pmatrix}\alpha\\ \beta\end{pmatrix}$. We can write out $U|\Phi\rangle\langle\Phi|U^\dagger$ as.

$$U|\Phi\rangle\langle\Phi|U^\dagger = \begin{pmatrix}\alpha^2 & \alpha\beta\\ \alpha\beta & \beta^2\end{pmatrix}$$

And use it to bound the trace distance of $|\Psi\rangle, |\Phi\rangle$.

$$TD(|\Phi\rangle\langle\Phi|, |\Psi\rangle\langle\Psi|) =^* TD(U|\Phi\rangle\langle\Phi|U^\dagger, U|\Psi\rangle\langle\Psi|U^\dagger) = \frac{1}{2}|\begin{pmatrix}\alpha^2 - 1 & \alpha\beta\\ \alpha\beta & \beta^2\end{pmatrix}|$$

$(*)$ - *This is an equality, not an inequality because we can apply $U^\dagger$ to receive $TD(|\Phi\rangle\langle\Phi|, |\Psi\rangle\langle\Psi|)$ again, thus the trace distance did not go down.*

Let e, f be the eigenvalues of the difference matrix. It is a known fact that the trace of a matrix is equal to the sum of its eigenvalues. Here the trace is $\alpha^2 - 1 + \beta^2 = 0$. This holds using the fact that $\alpha^2 + \beta^2 = 1$ from the norm of $|\Phi\rangle$. Thus we know that $e + f = 0 \Rightarrow e = -f$.

It is also a known fact that the determinant of a matrix is the product of its eigenvalues. Here the determinant is $(\alpha^2 - 1)\beta^2 - \alpha^2\beta^2 = -\beta^2$. Thus $e \cdot f = -\beta^2$. Since $e = -f$, $-f^2 = -\beta^2$. We now have two solutions for f.

$$f_1 = \beta, e_1 = -\beta$$
$$f_2 = -\beta, e_2 = \beta$$

In any case, we know that the trace distance is half of the sum of the absolute values of the eigenvalues, and thus $TD(|\Phi\rangle\langle\Phi|, |\Psi\rangle\langle\Psi|) = \frac{|\beta|+|-\beta|}{2} = |\beta|$.

3 QOTP without 0-keys

We have seen how the quantum one-time pad (QOTP) when used with a uniformly random key is secure. That is, with key register distribution $\rho_K := \sum_{i\in\{0,1\}^n} \frac{1}{2^n}|i\rangle\langle i|$, we can take any input and hide it so that without the key, all inputs look the same. So for any density operators ρ_A, ρ_B of equal dimension defined by the key, $\mathcal{E}(\rho_K \otimes \rho_A) = \mathcal{E}(\rho_K \otimes \rho_B) = \frac{1}{2^n} I$.

Note: our exact security definition was even stricter, see lecture slides if interested.

Previously we have had an exercise where a person avoids the all zeros key for a one-time pad because then the ciphertext is just the plaintext message, and that feels insecure to them. What would happen if the zero key was avoided in QOTP? How distinguishable are two different messages when using QOTP with this key distribution - what is $TD(\mathcal{E}(\rho'_K \otimes \rho_A), \mathcal{E}(\rho'_K \otimes \rho_B))$?

As the zero-key is no longer used, the key distribution is no longer ρ_K, but instead we define a $\rho'_K := \sum_{i\in\{0,1\}^n\setminus 0\ldots 0} \frac{1}{2^n-1}|i\rangle\langle i|$ which is uniform does not include the all zeros keys. Understanding the shapes of the ρ_K, ρ'_K matrices is useful in a later step. They are both diagonal matrices, but ρ'_K has no value on the $|0\ldots0\rangle\langle 0\ldots0|$ entry, and has larger values on the rest of the diagonal entries as all density operators have diagonal elements with a sum of 1.

$$\rho_K = \begin{pmatrix} \frac{1}{2^n} & 0 & \ldots & 0 \\ 0 & \frac{1}{2^n} & \ldots & 0 \\ \vdots & \vdots & \ddots & 0 \\ 0 & 0 & 0 & \frac{1}{2^n} \end{pmatrix} \qquad \rho'_K = \begin{pmatrix} 0 & 0 & \ldots & 0 \\ 0 & \frac{1}{2^n-1} & \ldots & 0 \\ \vdots & \vdots & \ddots & 0 \\ 0 & 0 & 0 & \frac{1}{2^n-1} \end{pmatrix}$$

Now we can compute how noticeable avoiding the zero key is by finding $TD(\mathcal{E}(\rho_K \otimes \rho_A), \mathcal{E}(\rho'_K \otimes \rho_A))$ - that is, the difference between using QOTP on some input ρ_A using the normal key distribution ρ_K, and the zeroless one ρ'_K.

$$TD(\mathcal{E}(\rho_K \otimes \rho_A), \mathcal{E}(\rho'_K \otimes \rho_A)) \leq TD(\rho_K \otimes \rho_A, \rho'_K \otimes \rho_A) = TD(\rho_K, \rho'_K)$$

The properties used for the above line can be found under the knowlet TD-Props in the lecture notes. As we have explicit constructions of ρ_K, ρ'_K, we can compute their trace. For the following computation, it helps to visualize ρ_K, ρ'_K and $|\rho_K - \rho'_K|$.

$$\rho_K - \rho'_K = \begin{pmatrix} \frac{1}{2^n} & 0 & \ldots & 0 \\ 0 & \frac{1}{2^n} - \frac{1}{2^n-1} & \ldots & 0 \\ \vdots & \vdots & \ddots & 0 \\ 0 & 0 & 0 & \frac{1}{2^n} - \frac{1}{2^n-1} \end{pmatrix}$$

$$
\begin{aligned}
TD(\rho_K, \rho'_K) &= \frac{1}{2} tr| \sum_{i\in\{0,1\}^n} \frac{1}{2^n}|i\rangle\langle i| + \sum_{i\in\{0,1\}^n\setminus 0\ldots 0} \frac{1}{2^n-1}|i\rangle\langle i|| \\
&=^1 \frac{1}{2}(|\frac{1}{2^n}| + \sum_{i\in\{0,1\}^n\setminus 0\ldots 0} |\frac{1}{2^n} - \frac{1}{2^n-1}|) \\
&=^2 \frac{1}{2}(|\frac{1}{2^n} + (2^n-1)|\frac{1}{2^n} - \frac{1}{2^n-1}|) \\
&= \frac{1}{2}(|\frac{1}{2^n} + (2^n-1)\frac{|2^n-1-2^n|}{2^n(2^n-1)}) \\
&= \frac{1}{2}(|\frac{1}{2^n} + |\frac{-1}{2^n}|) = \frac{1}{2^n}
\end{aligned}
$$

In the above, $=^1$ follows from the fact that anything in the form of $|i\rangle\langle i|$ is on the diagonal, and for a diagonal matrix, the trace of the absolute value of the matrix is the sum of the absolute values of main diagonal elements. In $=^2$ we use the fact that the sum $\sum_{i\in\{0,1\}^n\setminus 0\ldots 0}$ has 2^n-1 identical summands in it, thus it is just multiplying.

We have now shown that for any input ρ_A, using the zeroless key distribution is $\frac{1}{2^n}$-far from the secure QOTP using all keys uniformly. This also now allows us to show what is the distinguishing advantage when QOTP is used on two arbitrary inputs - the $TD(\mathcal{E}(\rho'_K \otimes \rho_A), \mathcal{E}(\rho'_K \otimes \rho_B))$ we wanted to find originally.

We denote by $A \approx^\varepsilon B$ that $TD(A, B) \leq \varepsilon$, meaning A and B are ε-distinguishable.

$$
\mathcal{E}(\rho'_K \otimes \rho_A) \approx^{\frac{1}{2^n}} \mathcal{E}(\rho_K \otimes \rho_A) = \mathcal{E}(\rho_K \otimes \rho_B) \approx^{\frac{1}{2^n}} \mathcal{E}(\rho'_K \otimes \rho_B)
$$

Her the first $\approx$ means that using QOPT with the zeroless key distribution on input ρ_A is $\frac{1}{2^n}$-far from using QOTP on it with the right key distribution. The middle equality holds as with the correct key distribution, QOTP provides complete indistinguishability for every input. And then the second $\approx$ is again comparing the proper key distribution vs the zeroless key distribution, but on input ρ_B.

Using the triangle inequality on this chain gives us

$$
TD(\mathcal{E}(\rho'_K \otimes \rho_A), \mathcal{E}(\rho'_K \otimes \rho_B)) \leq \frac{2}{2^n}
$$