Quantum Fingerprints

Recall from the previous lecture that *quantum fingerprinting* [BCWdW01] is a mechanism that allows us to encode $2^{\mathrm{O}(n)}$ bits into n qubits "in a useful manner". The fingerprints are constructed as follows.

Let $m = 2^{\mathrm{O}(n)}$, and consider any m-bit string x. We map $x \in \{0,1\}^m$ to (a larger string) $\hat{x} \in \{0,1\}^{\hat{m}}$, where $\hat{m} = \mathrm{O}(m)$, and $\{\hat{x}\}$ lie in a (classical) error correcting-code $\mathcal{C} \subseteq \{0,1\}^{\hat{m}}$. Figure 1 illustrates the mapping.

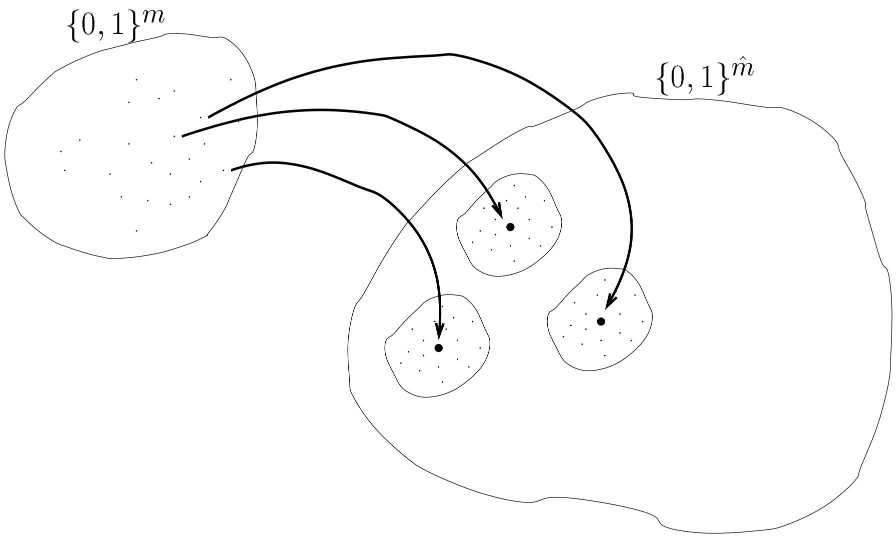

Figure 1: A (classical) error-correcting code.

We choose an error-correcting code such that the Hamming distance

$$\mathrm{d}(\hat{A}, \hat{B}) \geq \delta \hat{m}$$

for all $\hat{A}, \hat{B} \in \mathcal{C}$ and $\hat{A} \neq \hat{B}$, and a constant $0 < \delta < 1/2$. It is possible to choose such a code due to the following fact.

Fact : There exists a code $\mathcal{C}$ of size $|\mathcal{C}| \geq 2^{(1-H(\delta))\hat{m}}$, where $\mathcal{C}$ is an error-correcting code with minimum distance $\delta\hat{m}$, and

$$H(\delta) = \delta \times \log_2\left(\frac{1}{\delta}\right) + (1-\delta) \times \log_2\left(\frac{1}{1-\delta}\right)$$

is called the *binary entropy function*.

Figure 2 shows the shape of the binary entropy function. At $\delta = 0$ and $\delta = 1$, $H(\delta)$ is at its minimum value of 0. At $\delta = \frac{1}{2}$, $H(\delta)$ reaches its maximum value of 1. □

We construct the quantum fingerprints by associating every string x with a unique $\hat{x} \in \mathcal{C}$ using an injective map (denoted by $\mapsto$).

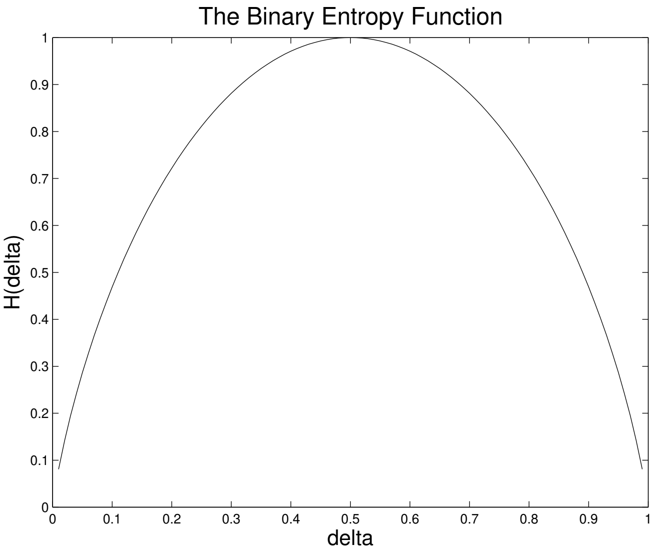

Figure 2: The binary entropy function.

$$
\begin{aligned}
x &\mapsto \hat{x} \in C \\
&\mapsto \sum_{i=1}^{\hat{m}} \frac{|i, \hat{x}_i\rangle}{\sqrt{\hat{m}}} = |\phi_x\rangle \in \mathbb{C}^{2\hat{m}}
\end{aligned}
$$

In the above, $\hat{x}_i$ is the ith bit of $\hat{x}$. Then $|i, \hat{x}_i\rangle$ is a quantum state consisting of the counter i and the bit $\hat{x}_i$ (which is normalized by $\sqrt{\hat{m}}$). Now $|\phi_x\rangle$ is a state over $\log_2(\hat{m}) + 1$ qubits which encodes m bits. Next, consider two strings $x, y \in \{0,1\}^m$ which we encode as $|\phi_x\rangle$ and $|\phi_y\rangle$ using the above method.

We see that

$$
\begin{aligned}
\langle\phi_x|\phi_y\rangle &= \frac{\left\langle \sum_{i=1}^{\hat{m}} |i, \hat{x}_i\rangle \middle| \sum_{j=1}^{\hat{m}} |j, \hat{y}_j\rangle \right\rangle}{\hat{m}} \\
&= \frac{1}{\hat{m}} \sum_{i=1}^{\hat{m}} \langle i|i\rangle\langle\hat{x}_i|\hat{y}_j\rangle \text{ (which is equal to the number of places where } \hat{x} \text{ and } \hat{y} \text{ are equal)} \\
&= \frac{\hat{m} - \mathrm{d}(\hat{x}, \hat{y})}{\hat{m}} \\
&= \begin{cases} 1 & \text{if } x = y\text{; and} \\ \leq \frac{\hat{m} - \delta\hat{m}}{\hat{m}} = 1 - \delta & \text{if } x \neq y. \end{cases}
\end{aligned}
$$

Thus, if the strings x and y are identical, then $\langle\phi_x|\phi_y\rangle = 1$, whereas if they are different, the inner product of their encoding has an absolute value of at most $(1 - \delta)$. (Such a set of states are an instance of *spherical codes*.)

Note that we have $2^m \geq 2^{(1-H(\delta))\hat{m}}$ such vectors in a space of dimension $2\hat{m}$. Thus, for any constant $\delta < \frac{1}{2}$, we have an exponential number of *nearly* orthogonal vectors in a space of dimension $2\hat{m}$.

The above property immediately suggests an application of quantum fingerprints in *equality testing*. Suppose that Alice and Bob are each in possession of an m-bit string, respectively $x, y \in \{0,1\}^m$. They want to determine if the two strings are equal, without revealing their strings to each other. Instead, they each send the fingerprint of his or her string to a third party called the *referee*, who follows a protocol to determine if the two strings x and y are equal. This set-up, using what is called the *simultaneous message passing model* of communication [Yao79], is illustrated in Figure 3.

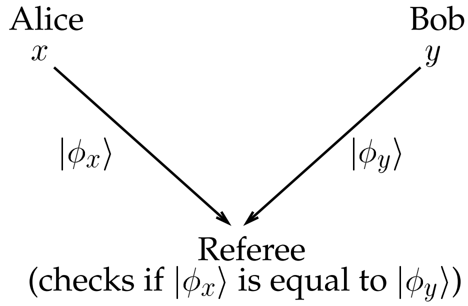

Figure 3: Equality-testing using the simultaneous messaging model.

Fact : Equality testing for m-bit strings has classical communication complexity $\Theta(\sqrt{m})$ in the simultaneous messages model. $\square$ However, the quantum protocol we describe below uses only $O(\log(m))$ qubits.

Given the fingerprints $|\phi_x\rangle$ and $|\phi_y\rangle$, the referee must determine whether they are identical or have an inner product of absolute value at most $(1-\delta)$. The referee has no information about the orientation of the two states relative to the standard basis (since he does not know x or y). Figure 4 shows a possible scenario.

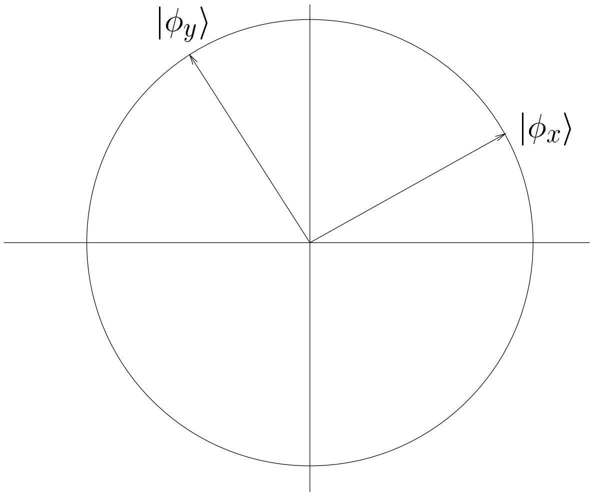

Figure 4: Comparing quantum fingerprints.

The referee has the state

$$|\psi_0\rangle = |\phi_x\rangle|\phi_y\rangle \in \mathbb{C}^{2\hat{m}} \otimes \mathbb{C}^{2\hat{m}}.$$

If he exchanges the qubits in the two fingerprints, the referee gets the possibly different state

$$|\psi_1\rangle = |\phi_y\rangle|\phi_x\rangle \in \mathbb{C}^{2\hat{m}} \otimes \mathbb{C}^{2\hat{m}}.$$

In fact, these two states are either the same or nearly orthogonal: $\langle\phi_1|\psi_2\rangle = |\langle\phi_x|\phi_y\rangle|^2$. The problem of distinguishing $|\phi_x\rangle$ and $|\phi_y\rangle$ then reduces to the problem of distinguishing $|\psi_0\rangle$ and $|\psi_1\rangle$. If $|\psi_0\rangle = |\psi_1\rangle$, then we can conclude that $|\phi_x\rangle = |\phi_y\rangle$, and hence $x = y$. While this may look like the same problem of having to distinguishing two states which are either exactly identical or "very different", in fact the problem has been simplified greatly. The referee's task is to check if the state $|\psi_0\rangle$ remains unchanged under the swap operation.

It is possible to decompose the tensor product space $\mathcal{H} \otimes \mathcal{H}$ of dimension n^2 into two orthogonal subspaces, the *symmetric subspace* S_n and the *alternating subspace* A_n, as follows:

$$\mathcal{H} \otimes \mathcal{H} \cong S_n \oplus A_n$$

, where the subspace S_n is invariant under the swap operation.

Let $\{|e_i\rangle\}_i$ be an orthonormal basis for $\mathcal{H}$. There are n such vectors. Then we can derive a orthonormal basis for $\mathcal{H} \otimes \mathcal{H}$, namely $\{|e_i\rangle|e_j\rangle\}_{i,j}$. Clearly, $|e_i\rangle \otimes |e_i\rangle \in S_n$ for each i. (We may abbreviate $|e_i\rangle \otimes |e_i\rangle$ as $|e_i\rangle|e_i\rangle$ or $|e_i, e_i\rangle$.)

Also, consider the vectors

$$\frac{|e_i\rangle|e_j\rangle + |e_i\rangle|e_j\rangle}{\sqrt{2}} \in S_n$$

for $i \neq j$. Together with $\{|e_i\rangle\}_i$, these form an orthonormal basis of size $n + \binom{n}{2}$ for S_n.

The following remaining $\binom{n}{2}$ vectors

$$\frac{|e_i\rangle|e_j\rangle - |e_i\rangle|e_j\rangle}{\sqrt{2}} \in S_n$$

for $i \neq j$ form an orthonormal basis for A_n.

Note that it immediately follows that S_n is swap invariant while A_n is not. In fact, the swap operation *reflects* any vector about the symmetric subspace S_n.

Let us denote by P_n the orthogonal projection onto S_n. We leave it to the reader to verify that

$$\| P_n |\phi_x\rangle|\phi_y\rangle \|^2 = \begin{cases} 1 & \text{if } x = y\text{; and} \\ \frac{1}{2}(1 + |\langle\phi_x|\phi_y\rangle|^2) & \text{if } x \neq y. \end{cases}$$

If $x \neq y$, then $|\langle\phi_x|\phi_y\rangle| \leq (1-\delta)$. Thus, if the referee uses the measurement $\{P_n, I - P_n\}$ to guess "=" or "$\neq$", then his probability of making an error is

$$\Pr(\text{error}) \begin{cases} = 0 & \text{if } x = y\text{; and} \\ \leq \frac{1}{2}(1 + (1-\delta)^2) & \text{if } x \neq y. \end{cases}$$

Note that $(1-\delta) < 1$, and we may reduce the probability of error in the standard way (by repeating the protocol in parallel, and taking the majority outcome).

Holevo Bound

From the fingerprinting example, one might think that we can somehow compress n classical bits into many fewer than n qubits. It turns out that there is a fundamental theorem that states that this is impossible *in general*. In its weakest form, Holevo's bound says:

"We need n quantum bits to encode n classical bits if we require that all n bits be recoverable simultaneously from the quantum state, with probability 1."

A simple way to infer this is as follows.

$$\begin{array}{ccc} \text{A(lice)} & & \text{B(ob)} \\ x \in \{0,1\}^n & \longmapsto & |\phi_x\rangle \in \mathbb{C}^{2^m} \end{array}$$

Figure 5: Pictorial explanation of the Holevo bound.

Suppose Alice encodes n bits x into an m-qubit state $|\phi_x\rangle$, and sends it to Bob. Bob measures the state according to the orthogonal measurement $\{P_y\}$, and on outcome y, guesses Alice's string to be y. If Bob's guess is correct with probability 1, then certainly, $\| P_x|\phi_x\rangle \|^2 = 1$, and $\| P_y|\phi_x\rangle \|^2 = 0$, for all $y \neq x$. Since the projection operators P_x are orthogonal, the states $|\phi_x\rangle$ themselves are orthogonal. There are 2^n such mutually orthognal states. Therefore, the dimension of the encoding space is at least 2^n. Thus, $m \geq n$.

References

[BCWdW01] Harry Buhrman, Richard Cleve, John Watrous, and Ronald de Wolf. Quantum fingerprinting. *Physical Review Letters*, 87(16):Article 167902, 2001.

[Yao79] Andrew Chi-Chih Yao. Some complexity questions related to distributive computing (Preliminary Report). In *Proceedings of the eleventh annual ACM Symposium on the Theory of Computing (STOC 79)*, pages 209–213, Atlanta, Georgia, United States, 1979. ACM.