CS181A Notes 9 Quantum Information

Dirac notation We follow the convention for column and row vectors known as Dirac's notation. Here, we denote column vectors as $|u\rangle$ (called "ket" u). The row vectors are obtained by taking the Hermitian transpose of the kets: $\langle u| = |u\rangle^\dagger$. This notation originated from the bracket notation used for the inner product $\langle u|v\rangle$ of vectors u and v.

States We imagine using quantum states for storing our information (bits). The basic building blocks are the two binary states

$$|0\rangle = \begin{bmatrix}1\\0\end{bmatrix} \quad \text{and} \quad |1\rangle = \begin{bmatrix}0\\1\end{bmatrix}$$

which are two-dimensional column vectors. Our actual *quantum bit* (qubit) is represented as

$$|\psi\rangle = a_0|0\rangle + a_1|1\rangle, \quad \text{where } a_b \in \mathbb{C} \text{ and } \sum |a_b|^2 = 1 \text{ and } b \in \{0,1\}$$

Despite its appearance, $|\psi\rangle$ is merely a two-dimensional complex vector of unit length.

Operators The laws of quantum mechanics allow us to apply an operation U to $|\psi\rangle$, and obtain $|\psi'\rangle$, as long as the result $|\psi'\rangle$ is also of unit length and that U is *invertible*. In short, U must be a *unitary* matrix; that is, $U^{-1} = U^\dagger$, where the Hermitian transpose is defined as $U^\dagger = \overline{U^T}$. A set of standard unitary operations (or gates) is given by the following 2×2 matrices:

$$I = \begin{bmatrix}1 & 0\\0 & 1\end{bmatrix}, \; X = \begin{bmatrix}0 & 1\\1 & 0\end{bmatrix}, \; Z = \begin{bmatrix}+1 & 0\\0 & -1\end{bmatrix}, \; Y = \begin{bmatrix}0 & -i\\+i & 0\end{bmatrix} = iXZ.$$

The above is also called the set of Pauli matrices. Other important gates include the *Hadamard* gate, the phase gate, and the $\pi/8$ gate:

$$H = \frac{1}{\sqrt{2}}\begin{bmatrix}+1 & +1\\+1 & -1\end{bmatrix}, \; S = \begin{bmatrix}1 & 0\\0 & i\end{bmatrix}, \; T = \begin{bmatrix}1 & 0\\0 & e^{i\pi/4}\end{bmatrix}.$$

Measurements In classical computation, the output of a computational process can be read out as is. If we would like to observe a quantum bit, the laws of quantum mechanics requires us to define a *measurement* operator $\mathcal{M}$ and then apply it to our qubit $|\psi\rangle$. Unlike the unitary requirement on U, the measurement operator $\mathcal{M}$ is required to be a *Hermitian* matrix (that is, $\mathcal{M}^\dagger = \mathcal{M}$); this guarantees that $\mathcal{M}$ has *real* eigenvalues.

In the simplest case, a measurement $\mathcal{M}_{B_{\text{rect}}}$ using the (rectilinear) basis $B_{\text{rect}} = \{|0\rangle, |1\rangle\}$ of $|\psi\rangle = a_0|0\rangle + a_1|1\rangle$ produces a probabilistic outcome

$$\mathcal{M}_{B_{\text{rect}}}|\psi\rangle = \begin{cases} |0\rangle & \text{with probability } |a_0|^2 \\ |1\rangle & \text{with probability } |a_1|^2 \end{cases}$$

We may also measure using a different basis, say, the diagonal basis $B_{\text{diag}} = \{|+\rangle, |-\rangle\}$, where

$$|\pm\rangle = \frac{1}{\sqrt{2}}|0\rangle \pm \frac{1}{\sqrt{2}}|1\rangle$$

In this case, we note

$$\mathcal{M}_{B_{\text{diag}}}|0\rangle = \begin{cases} |+\rangle & \text{with probability } 1/2 \\ |-\rangle & \text{with probability } 1/2 \end{cases}$$

but $\mathcal{M}_{B_{\text{diag}}}|\pm\rangle = |\pm\rangle$.

In a more general setting, suppose that the Hermitian operator $\mathcal{M}$ has as eigenvalues the real numbers $\lambda_1, \ldots, \lambda_r$. We consider the spectral decomposition of $\mathcal{M}$ given by

$$\mathcal{M} = \sum_{i=1}^{r} \lambda_i P_i$$

where P_i is the projector corresponding to the λ_i eigenspace. The laws of measurement stipulates that applying the measurement operator $\mathcal{M}$ to the quantum bit $|\psi\rangle$ produces the probabilistic outcome:

$$\mathcal{M}|\psi\rangle = \frac{P_i|\psi\rangle}{\sqrt{\langle\psi|P_i|\psi\rangle}}, \quad \text{with probability } ||P_i|\psi\rangle||^2 = \langle\psi|P_i|\psi\rangle$$

For a binary string $\beta \in \{0,1\}^k$, where $\beta = b_1 \ldots b_k$, we use the notation $|\beta\rangle$ to mean

$$|\beta\rangle = |b_1\rangle \ldots |b_k\rangle = |b_1\rangle \otimes \ldots \otimes |b_k\rangle.$$

A k-qubit $|\psi\rangle$ is given by

$$|\psi\rangle = \sum_{\beta\in\{0,1\}^k} a_\beta|\beta\rangle, \quad \text{where } \textstyle\sum_{\beta\in\{0,1\}^k} |a_\beta|^2 = 1$$

A very useful 2-qubit operator is the Controlled-NOT (CNOT) gate defined by the following unitary matrix:

$$CNOT = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

The action of CNOT on the computational basis states $\{|00\rangle, |01\rangle, |10\rangle, |11\rangle\}$ may be described as $CNOT(a, b) = (a, b \oplus a)$. If the input a is viewed as a control bit, then the second input is negated if the control bit is on.

Quantum key distribution We now describe the BB84 quantum protocol for key exchange due to Bennett and Brassard. There are two channels available to Alice and Bob: a one-way quantum channel from Alice (sender) to Bob (receiver) and a two-way classical authenticated channel (or broadcast). Alice and Bob agreed ahead of time on the number of particles N used in the protocol.

1. Alice selects a sequence of N random basis choices $\alpha \in \{\text{rect}, \text{diag}\}^N$.
2. Bob selects a sequence of N random basis choices $\beta \in \{\text{rect}, \text{diag}\}^N$.
3. Alice creates a sequence of N random quantum states $|\psi_i\rangle \in B_{\alpha_i}$, for each $i = 1, \ldots, N$. She then sends $|\psi_1\rangle, \ldots, |\psi_N\rangle$ to Bob.
4. Bob measures each incoming $|\psi_i\rangle$ using $\mathcal{M}_{B_{\beta_i}}$, for $i = 1, \ldots, N$.
5. Alice announces α to Bob.
6. Bob announces β to Alice.
7. Alice and Bob compute $S = \{i : \alpha_i = \beta_i\}$. Then, for each $i \in S$, they agree on a "secret" key $s \in \{0, 1\}^{|S|}$, where

$$s_i = \begin{cases} 0 & \text{if } |\psi_i\rangle \in \{|0\rangle, |+\rangle\} \\ 1 & \text{if } |\psi_i\rangle \in \{|1\rangle, |-\rangle\} \end{cases}$$

Alice and Bob will agree on M bits, where $M \sim N/2$ bits with high probability. This can be fomalized using Hoeffding bounds which we state below.

Lemma 1. *(Hoeffding bounds)*
Let $X_1, \ldots, X_m \in [a, b]$ be a sequence of independent identically distributed random variables. If $S_m = \sum_{i=1}^{m} X_i$, then for any $\varepsilon > 0$ we have

$$\mathbb{P}[|S_m - \mathbb{E}[S_m]| \geq \tau] \leq 2\exp(-2\tau^2/m).$$

Out of the M shared bits, Alice and Bob may agree to use half of these bits (say, a randomly chosen half) to check publicly (using the broadcast channel) that both of their halves match. Suppose this subset of positions is denoted $I \subseteq S$, where $|I| = |S|/2$. Then, Alice and Bob exchanged the identity of their j-th particle for each $j \in I$. If there is a mismatch at position $i \in I$, this implies that Eve had:

1. Measured $|\psi_i\rangle$ in the wrong basis (different from what Alice and Bob used)
2. Replaced $|\psi_i\rangle$ with an impostor particle $|\psi'\rangle$

We would like to consider the event X where Alice and Bob discover that Eve is eavesdropping. Conditioned on the event that Eve measured the i-th particle in basis B and sending Bob her measured particle, event X occurs if:

- Alice and Bob agreed on the same basis $B' \neq B$ for the i-th particle
- Alice and Bob agreed to use position i to perform the equality testing
- Bob obtained the "wrong" probabilistic outcome when he measured using B'

Classical Post-processing Suppose that Alice and Bob used QKD to *share* an n-bit string. At the end of the QKD protocol, Alice and Bob obtained $\alpha, \beta \in \{0,1\}^n$, respectively, which are not necessarily identical. They now typically perform the following two classical public post-processing protocols:

1. *Secret-key distillation*:
   This is a protocol that takes Alice's input α and Bob's input β and creates a string γ that both Alice and Bob agree on. It is likely that the amount of information that Eve knows about γ is non-zero (and possibly non-negligible).

   As an example of secret-key distillation, suppose Alice and Bob agree on a $[n, k]$-linear code $\mathcal{C}$. Suppose H is the parity check matrix of $\mathcal{C}$.

(a) Alice computes the syndrome $s_A = H\alpha$ and sends this to Bob.

(b) Bob computes the $(n-k)$-bit syndrome $s_B = H\beta$ and computes the difference $s = s_B - s_A = H(\beta - \alpha)$.

(c) Bob computes the unique codeword c for which $Hc = s$ and resets his string to $\beta' = \beta - c$.

If $\mathcal{C}$ can correct up to t errors and the weight of $\beta - \alpha$ is at most t, then there is a unique codeword c for which $Hc = s$. In fact, that codeword c must be equal to $\beta - \alpha$.

2. *Privacy amplification*:
   This is a protocol that Alice and Bob perform to reduce Eve's information about γ. The goal is for Alice and Bob to publicly agree on another string ω based on γ but for which Eve almost has no information about.

   Several methods commonly used here include universal hash functions (see the Leftover Hash Lemma described in the appendix), randomness extractors, and interactive parity check protocols.

Quantum teleportation We describe a protocol where Alice can send to Bob an *arbitrary* quantum bit, say $|\psi\rangle = a|0\rangle + b|1\rangle$, by sending only two classical bits provided they share some prior entanglement. Note due to the No-Cloning Theorem and other restrictions, this seems improbable as the identity of the arbitrary quantum bit $|\psi\rangle$ requires the specification of the two complex numbers a and b.

Consider the *Bell* state

$$|\beta_{00}\rangle = \frac{1}{\sqrt{2}}|00\rangle + \frac{1}{\sqrt{2}}|11\rangle,$$

which is also called an EPR (Einstein-Podolsky-Rosen) or *entangled* pair. We may create the above entangled pair by using the Hadamard and Controlled-NOT gates.

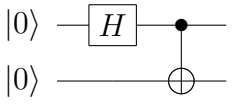

Figure 1: A quantum circuit to create the Bell state $|\beta_{00}\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$.

Let $\mathrm{CNOT}_{1,2}(a,b) = (a, a \oplus b)$ be the Controlled-NOT gate with input 1 serving as the control bit and input 2 being the result bit. Then,

$$\begin{aligned} |0\rangle \otimes |0\rangle \quad &\rightarrow_{H\otimes I} \quad \left[\frac{1}{\sqrt{2}}|0\rangle + \frac{1}{\sqrt{2}}|1\rangle\right] \otimes |0\rangle = \frac{1}{\sqrt{2}}|0\rangle \otimes |0\rangle + \frac{1}{\sqrt{2}}|1\rangle \otimes |0\rangle \\ &\rightarrow_{\mathrm{CNOT}_{1,2}} \quad \frac{1}{\sqrt{2}}|0\rangle \otimes |0\rangle + \frac{1}{\sqrt{2}}|1\rangle \otimes |1\rangle \end{aligned}$$

We are now ready to describe the quantum teleportation protocol.

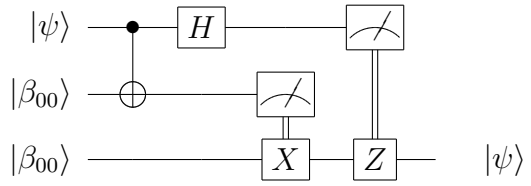

Figure 2: A quantum teleportation circuit.

Let $|\psi\rangle = a|0\rangle + b|1\rangle$ be Alice's quantum message and let $|\beta_{00}\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$ be the Bell (EPR) state shared between Alice and Bob. The initial state of the joint system is $|\phi_0\rangle = |\psi\rangle \otimes |\beta_{00}\rangle$. First, Alice perform two steps on her 2-qubit system; but we will follow the effects of Alice's operations on the whole system:

$$\begin{aligned} |\phi_1\rangle &= (\mathrm{CNOT}_{1,2} \otimes I)|\phi_0\rangle \\ |\phi_2\rangle &= (H \otimes I \otimes I)|\phi_1\rangle \\ |\phi_3\rangle &= (M_B \otimes M_B \otimes I)|\phi_2\rangle \end{aligned}$$

Here M_B is a measurement in the basis $B = \{|0\rangle, |1\rangle\}$. Alice obtains two classical bits $x, y \in \{0,1\}$ from these two measurements. We consider a more careful derivation of the above steps:

$$\begin{aligned} |\phi_0\rangle &= \frac{1}{\sqrt{2}}\left[a|0\rangle(|00\rangle + |11\rangle) + b|1\rangle(|00\rangle + |11\rangle)\right] \\ |\phi_1\rangle &= \frac{1}{\sqrt{2}}\left[a|0\rangle(|00\rangle + |11\rangle) + b|1\rangle(|10\rangle + |01\rangle)\right] \\ |\phi_2\rangle &= \frac{1}{2}\left[a(|0\rangle + |1\rangle)(|00\rangle + |11\rangle) + b(|0\rangle - |1\rangle)(|10\rangle + |01\rangle)\right] \\ &= \frac{1}{2}\left[|00\rangle(a|0\rangle + b|1\rangle) + |01\rangle(a|1\rangle + b|0\rangle) + |10\rangle(a|0\rangle - b|1\rangle) + |11\rangle(-b|0\rangle + a|1\rangle)\right] \\ &= \frac{1}{2}\left[|00\rangle|\psi\rangle + |01\rangle X|\psi\rangle + |10\rangle Z|\psi\rangle + |11\rangle XZ|\psi\rangle\right] \end{aligned}$$

The crux of this protocol is this: when Alice measures her two qubits and sends the results to Bob, this provides Bob with information about the state of his qubit. The state of Bob's qubit is either $|\psi\rangle$ (if the measurement of the first two qubits resulted in 00), $X|\psi\rangle$ (if the measurement outcome is 01), $Z|\psi\rangle$ (if the measurement outcome is 10), or $XZ|\psi\rangle$ (if the measurement outcome is 11). Knowing this information, Bob can *recover* $|\psi\rangle$ by applying a sequence of correct operators (note $X^2 = I$ and $Z^2 = I$). This implies that Bob knows what $|\psi\rangle$ is exactly; but in the process, Alice and Bob lost their entangled pairs.

References

[1] S. Loepp and W. Wootters, *Protecting Information: From Classical Error Correction to Quantum Cryptography*, Cambridge University Press, 2006.

[2] M. Nielsen and I. Chuang, *Quantum Computation and Quantum Information*, Cambridge University Press, 2000.

A Randomness extraction

We describe an interesting result about extracting randomness from a weak random source with the aid of universal hash functions. But first we define the requisite terminology.

For a distribution $\mathcal{D}$ over a finite set S, the *collision probability* of $\mathcal{D}$ is $\sum_{a \in S} \mathcal{D}(a)^2$. This gives the probability that two elements drawn according to $\mathcal{D}$ are identical. Note the uniform distribution $\mathcal{U}$ on S has a collision probability of $1/|S|$. The ***Rényi entropy*** of distribution $\mathcal{D}$ is defined as

$$H_R(\mathcal{D}) = -\log_2 \left(\sum_{a \in S} \mathcal{D}(a)^2 \right).$$

A distribution $\mathcal{D}$ on S is called δ-*uniform* if for each $A \subseteq S$ we have

$$|\mathcal{D}(A) - \mathcal{U}(A)| < \delta,$$

where $\mathcal{U}(A) = |A|/|S|$. This captures the idea that no statistical test that uses a binary partitioning on S can distinguish $\mathcal{D}$ from $\mathcal{U}$ with an advantage that is better than δ.

A family $\mathcal{H}$ of hash functions mapping n-bit strings to m-bit strings is called *almost-universal* if for each $x, y \in \{0,1\}^n$, where $x \neq y$, we have

$$\mathbb{P}[h(x) = h(y)] \leq \frac{1}{2^m} + \frac{1}{2^n}. \tag{1}$$

where h is chosen uniformly from $\mathcal{H}$. A simple example of such a family may be obtained from the set of linear maps $\{ax + b \bmod p : a, b \in \mathbb{Z}_p\}$ over integers modulo p.

The idea of the next lemma is as follows: if we have a random string with a certain amount of entropy (but we don't necessarily know where the randomness is or whether is concentrated somewhere or not), we can "extract" this randomness and localize it by "hashing" (but by also including the description of the randomly chosen hash functions which is viewed as a small investment towards extracting the randomness). So, if x is the random string which contain a good but unwieldy entropy, then the string $(h, h(x))$ is a higher quality random string, given that h is a random hash function (with *universal* property).

Lemma 2. *(Leftover Hash Lemma)*
Let $\mathcal{D}$ be a distribution over $\{0,1\}^n$ with collision probability at most 2^{-m}. Let $t > 0$ and let $\mathcal{H}$ be an almost universal family of hash functions which map n bits to $m - 2t$ bits.

Then, the distribution $(h, h(x))$ is $(1/2^t)$-uniform on $\mathcal{H} \times \{0,1\}^{m-2t}$, where the hash function h is chosen uniformly at random from $\mathcal{H}$ and x is randomly chosen according to $\mathcal{D}$.

Proof. (Impagliazzo, Rackoff)
First, note that any distribution $\mathcal{D}$ over S whose collision probability is at most

$$\frac{1+2\delta^2}{|S|},$$

is δ-uniform on S. This holds since if $\mathcal{D}$ is not δ-uniform on S, there is a subset $A \subseteq S$ for which $|\mathcal{D}(A) - \mathcal{U}(A)| \geq \delta$, where $\mathcal{U}(A) = |A|/|S|$. In fact, we may assume without loss of generality that $\mathcal{D}(A) = \delta + \mathcal{U}(A)$; otherwise, we may use $\delta' = \mathcal{D}(A) - \mathcal{U}(A)$ instead. Then,

$$\begin{aligned}
\sum_a \mathcal{D}(a)^2 &= \sum_{a\in A} \mathcal{D}(a)^2 + \sum_{a\in A^c} \mathcal{D}(a)^2 \\
&\geq \frac{\mathcal{D}(A)^2}{|A|} + \frac{\mathcal{D}(A^c)^2}{|S|-|A|}, \quad \text{by Cauchy-Schwarz inequality} \\
&= \frac{1}{|S|}\left[1 + \left(\frac{|S|}{|A|} + \frac{|S|}{|S|-|A|}\right)\delta^2\right] \geq \frac{1}{|S|}[1+4\delta^2].
\end{aligned}$$

Next, we consider the collision probability of the distribution $(h, h(x))$ where h is chosen uniformly from $\mathcal{H}$ and $x \in \{0,1\}^n$ is chosen according to D:

$$\begin{aligned}
\mathbb{P}[(h_1, h_1(x)) = (h_2, h_2(y))] &= \mathbb{P}[h_1 = h_2, h_1(x) = h_2(y)] \\
&= \frac{1}{|\mathcal{H}|}\,\mathbb{P}[h(x) = h(y)] \\
&\leq \frac{1}{|\mathcal{H}|}\,(\mathbb{P}[h(x) = h(y) | x \neq y] + \mathbb{P}[x = y]) \\
&= \frac{1}{|\mathcal{H}|}\,\left(2^{-n} + 2^{-m+2t} + 2^{-m}\right) \\
&\leq \frac{1+2(2^{-t})^2}{|H|2^{m-2t}}.
\end{aligned}$$

Both of these yield the lemma. □