COS 533: Advanced Cryptography — Princeton University
Lecture 20 (April 15, 2021)
Lecturer: Mark Zhandry — Scribe: Atanas Dinev

Notes for Lecture 20

1 Quantum Computing and its intersection with Cryptography

Quantum computers are believed to be able to solve certain problems faster than classical computers. It turns out that most of the problems that quantum computers can solve faster are problems that are used as a basis for cryptography. These are problems like Integer Factorization and Discrete Logarithm. In this lecture will motivate what quantum computers are and that they could be useful in cryptography. The behaviour of quantum mechanics will allow us to do new cryptographic tasks, which were not possible before. Before, we motivate what quantum computers are it is useful to start with a simpler non-quantum concept - a Stochastic Process.

2 Stochastic Process

Suppose we have n slots as shown Figure 1 and a ball that comes in from the top slot at time 0. Then it exits through one of the slots on the other side of the first wall. On this wall there is a randomized process that determines the trajectory of the ball. For simplicity, we will assume that the process just creates a probability distribution on the output. Suppose the ball has 0.5 probability of going to the top slot of the next chamber, 0.1 of going to the second, 0.2 of going to the third, and 0.2 of going to the fourth. The requirement is that the probabilities sum to one $0.5 + 0.1 + 0.2 + 0.2 = 1$ so that the ball doesn't get lost. We can also add a third wall with slots. There will be a transition from each slot of the second wall to each slot on the third (i.e. n^2 in total). We can generalize this construction and imagine that the ball moves to the next wall every second. We can represent the state of this ball at any time as a probability vector with the probabilities that the ball passes through each slot.

Definition 1 *(Probability vector) A vector $v \in \mathbb{R}^n$ is a probability vector if $v_i \geq 0$ for all i and*

$$\sum_{i=1}^{n} v_i = 1$$

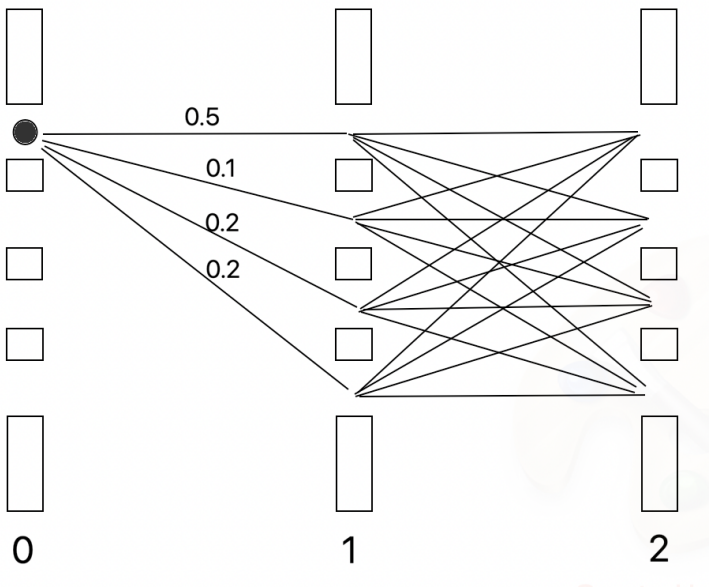

Figure 1: Stochastic Model with $n = 4$ slots

For example, the state of the ball at time 1 is just $v_1 = \begin{pmatrix} 0.5 \\ 0.1 \\ 0.2 \\ 0.2 \end{pmatrix}$. The transition function consists of a matrix $T_i \in \mathbb{R}^{n \times n}$ with all of the transition probabilities from time $i-1$ to time i. Here, we have that $(T_i)_{kl}$ is the probability that the ball transitions from slot l at time $i-1$ to slot k at time i. In the case above T_i will be a 4×4 matrix. Let v_i be the probability vector at time i, then it is clear that

$$v_i = T_i v_{i-1}$$

or more generally by iterating the above we get $v_i = \prod_{j=1}^{i} T_j v_0$ where v_0 is the starting state. The matrices T_i can be arbitrary but they need to preserve the fact that v_i is a probability vector. Those are called stochastic matrices

Definition 2 *(Stochastic matrix) A matrix $T \in \mathbb{R}^{n \times n}$ is a stochastic matrix if and only if $T_{ij} \geq 0$ for all i, j and its columns sum to 1. I.e. for every j*

$$\sum_{i=1}^{n} T_{ij} = 1$$

Note: It is easy to see that a matrix T_i is a stochastic matrix if and only if for all probability vectors v, $T_i v$ is a probability vector.

We introduce a couple more notions:

- **Measurement** Imagine we we look at where the ball is at time i. Then, we don't have any uncertainly as we know where the ball is exactly. We will see that ball at position $k \in [1, n]$ exactly with probability $e_k^T v_i = (v_i)_k$. Thus, our state vector becomes e_k with probability $e_k^T v_i$.

- **Paths/Trajectories**

  Definition 3 *(Path) A path is a tuple $P = (k_0, \ldots, k_t)$ and it represents that the ball is at slot k_i at time i. The probability of a path P is defined as the probability that the ball takes the path P, which is easily seen to be*

  $$Pr[P] = (v_0)_{k_0} \prod_{i=1}^{t} (T_i)_{k_i k_{i-1}}$$

  Note that by the law of total probability we have

  $$Pr[\text{ball exits at } j] = \sum_{P\text{: P ends in } j} Pr[P]$$

  Suppose we peek at time i and we consider slot k. Then we have by the law of total probability

  $$Pr[\text{ball passes through } k \text{ at time } i \text{ and exits through } j] = \sum_{P: P_i = k,\ P \text{ ends in } j} Pr[P]$$

  By linearity of the probability we have

  $$Pr[\text{ball exists at } j] = \sum_{i=1}^{n} Pr[\text{ball passes through } k \text{ at time } i \text{ and exits through } j]$$

  Therefore, looking at the ball at time k doesn't change its trajectory. In other words, looking at the fact that the ball was at slot k at time i does not affect the total probability that it exits at j, however it affects the conditional probability. We will soon see that this is not the case in the Quantum model.

3 Double Slit Experiment

Suppose we have a flash light, and we shine it towards a wall with a single opening. There is another wall that has detectors (which are like slots in the stochastic model). If you shine the light through the opening you will see a hump in the detectors. If you discretize it you will see a similar behaviour.
Suppose now, we have second opening on the first wall. If we model light as two balls passing though openings we will expect to see the superposition of two humps. Thus, the result is still expected to be hump. However, it was observed that you get peaks and troughs in a periodic fashion. This is not what we would expect from our model. There is a simple classical explanation: light does not behave as a particle, but it behaves as a wave.
We have

$$\text{Magnitute of light } \approx (\text{Electric field})^2$$

When you have 2 slits you get cancellations in the electric field, i.e. a peak from the first opening will cancel with a trough from from the second opening. This ends up causing, spots where light disappears. If the distance between the two openings is an integer multiple of the wavelength, the peaks and troughs will not cancel, while if it is a half-integer multiple of the wavelength they will cancel.

There were other experiments which discovered that light is composed of particles called photons. This, leads to the wave-particle duality of light. If light is a particle what would happen if we send exactly one particle in each slot? It turns out that we will see the above wave pattern due to interference. This became the origins of quantum mechanics. We can imagine one particle simultaneously entering both of the slits and it interferes with itself.

This will lead us to upgrade our stochastic model of the ball, and we will try to have a quantum version of our model, that will capture the fact that the particle is behaving like a wave.

4 Quantum model

Suppose again that we have a ball that passes through the top slot on the first wall as in Figure 1. Now, we think of the ball as a particle and also as an electric field corresponding to a wave. If we think of the ball as an electric field, then we will have transitions. These transitions are not longer stochastic transitions. The analogue of a probability vector or a state in the quantum model is called an *amplitude vector*. Below we define a complex conjugate and an amplitude vector

Definition 4 *Let $z = a + ib$ be a complex number, where $a, b \in \mathbb{R}$. Its conjugate is defined as $\overline{z} = a - ib$. The absolute of a complex number is defined as*

$$|z| = \sqrt{a^2 + b^2} = \sqrt{z\overline{z}}$$

Definition 5 *(Amplitude vector) A vector $v \in \mathbb{C}^n$ is an amplitude vector if*

$$v^T \overline{v} = \sum_{i=1}^{n} |v_i|^2 = 1$$

The quantity $v^T \overline{v}$ is also called the L_2 norm of v.

We think of the L_2 norm of the amplitude vector as the total energy. Our transition values are arbitrary complex values. Let U be a transition matrix. Then again we have

$$v_i = U_i v_{i-1}$$

The only difference is that U_i must preserve the L_2 norm (i.e. the energy) of the amplitude vectors. It turns out that this is equivalent to U being a unitary matrix and the proof is not hard. We introduce some more definitions:

Definition 6 *(Conjugate transpose) Let $U \in \mathbb{C}^{n \times n}$. The conjugate transpose of U is defined as $U^\dagger \in \mathbb{C}^{n \times n}$, such that $U^\dagger_{ij} = \overline{U_{ji}}$ for all $i, j \in [1, n]$*

Definition 7 *(Unitary matrix) Let $U \in \mathbb{C}^{n \times n}$. U is unitary, if $UU^\dagger = I_n$, where I_n is the identity matrix on $\mathbb{C}^{n \times n}$.*

Why do we use complex numbers?: For a complex number z, $|z|$ represents the magnitude of the electric field, and the relationship between the real and the imaginary part represents the phase of the wave. There is mathematical justification for that.

Now, let's consider the same setting as in the stochastic model. We have our ball that comes in the top slot. In the last wall we have a detector at each slot and we ask what is the probability of observing the ball coming through each of the detectors. We have that

$$Pr[\text{ball exits at } j] \propto \text{ amount of energy at } j$$

More specifically, if v_m is the amplitude vector at the last time m, we have that

$$Pr[\text{ball exits at } j] = |e_j^T v_m|^2 = |(v_m)_j|^2$$

Notice the since v_m is an amplitude vector

$$\sum_{j=1}^{n} Pr[\text{ball exits at } j] = \sum_{j=1}^{n} |(v_m)_j|^2 = 1$$

and thus we get a well defined probability distribution. Similarly to the stochastic model we can define

Definition 8 *(Path) A path is a tuple $P = (k_0, \ldots, k_t)$ and it represents that the ball is at slot k_i at time i. The weight of a path P is defined as*

$$w(P) = (v_0)_{k_0} \prod_{i=1}^{t} (U_i)_{k_i k_{i-1}}$$

Then, it is not hard to see by above that

$$Pr[\text{ball exits at } j] = |\sum_{\text{P: P ends at } j} w(P)|^2$$

Now, suppose that we observe the ball at an intermediate time at slot k at time i. Then, clearly

$$Pr[\text{observe ball at slot } k \text{ at time } i \text{ and it exits through } j] = |\sum_{\text{P: } P_i = k \text{ and } P \text{ ends at } j} w(P)|^2$$

What is interesting is that

$$\sum_{k=1}^{n} Pr[\text{observe ball at slot } k \text{ and it exits through } j] \neq Pr[\text{ball exits at } j]$$

i.e. sums don't commute with the squaring operation. This is known as

"The Observer effect": observing the ball at slot k changed its behaviour

i.e. the mere fact that we looked at the ball changed its behaviour. This, in contrast with the stochastic case, where if we see an object we will update our conditional probabilities but it will not affect our true total probabilities. Whereas, here the true total probabilities are affected.

Question: Should the above sum in the left hand side be greater or equal to the right hand side?

Answer: It can go both ways. We can think about it in the following way: we have paths which are interfereing with each other. The interference could be both constructive and destructive. It is possible that for some j all of the positive paths will cancel with all of the negative paths and we will get 0. But, now if we observe the ball at position j, we have partitioned the paths, and for this partition the sum of the squares of the values might end up being non-zero even the sum of the values is zero. On the other hand you can imagine that you have some positive weights, by Jensen's inequality you will end up with a higher amplitude if you first square and and then sum, than if you first sum and then square.

5 No cloning

Suppose we are given an amplitude vector state $\begin{pmatrix}\alpha\\ \beta\end{pmatrix}$ that was generated through some process. The normalization means that $|\alpha|^2 + |\beta|^2 = 1$.

Goal: Have a process that is independent of α and β, which clones $\begin{pmatrix}\alpha\\ \beta\end{pmatrix}$.

Before, we define cloning we introduce the concept of a joint system.

Definition 9 *(Joint system) Suppose we have two independent amplitude vectors (systems)* $v_0 = \begin{pmatrix}\alpha\\ \beta\end{pmatrix}$, *and* $v_1 = \begin{pmatrix}\gamma\\ \delta\end{pmatrix}$. *Their joint system is defined as the tensor product*

$$v_0 \otimes v_1 = \begin{pmatrix}\alpha\gamma\\ \alpha\delta\\ \beta\delta\\ \beta\delta\end{pmatrix}$$

As an analogue to the stochastic model we can think of the joint system of v_0 and v_1 as their joint distribution as two independent random variables. Then, indeed the probability of each pair of outcomes is the product of the respective probabilities.

What the goal of cloning translates to is that we want to have a process which converts $\begin{pmatrix}\alpha\\ \beta\end{pmatrix}$ into

$$\begin{pmatrix}\alpha\\ \beta\end{pmatrix} \otimes \begin{pmatrix}\alpha\\ \beta\end{pmatrix}$$

There is a caveat here. Our transition matrices need to be unitary, i.e. they need to be square and full rank. Therefore, U must preserve the dimensions. However, any process of the above type goes from a system of dimension 2 to a system of dimension 4. To fix that, we will start from a dummy system $\begin{pmatrix}1\\ 0\end{pmatrix}$. Thus, our process will take $\begin{pmatrix}\alpha\\ \beta\end{pmatrix} \otimes \begin{pmatrix}1\\ 0\end{pmatrix}$ as input and it will produce $\begin{pmatrix}\alpha\\ \beta\end{pmatrix} \otimes \begin{pmatrix}\alpha\\ \beta\end{pmatrix}$ as output. However, it turns out that this is actually impossible.

Theorem 10 *(**No-cloning theorem**) There is no process with the above properties that goes from $v \otimes e_1$ to $v \otimes v$ for all amplitude vectors v.*

Proof: Suppose the contrary. I.e. we have a process that goes from $v \otimes e_1$ to $v \otimes v$. It turns out we can compute inner products. Let v_1, v_2 be arbitrary vectors. We have

$$v_1 \otimes e_1 \to v_1 \otimes v_1$$

and

$$v_2 \otimes e_1 \to v_2 \otimes v_2$$

Now, let's look at the inner product

$$(v_1 \otimes e_1)^\dagger \cdot (v_2 \otimes e_1) = ((v_1)\dagger \cdot v_2)(e_1^\dagger \cdot e_1) = (v_1)^\dagger \cdot v_2$$

On the other hand, we have that

$$(v_1 \otimes v_1)^\dagger \cdot (v_2 \otimes v_2) = (v_1^\dagger \cdot v_2)^2$$

We know that unitary matrices must preserve inner products. This, means that the inner product of the inputs $v_1 \otimes e_1$ and $v_2 \otimes e_1$ has to be equal to the inner product of the outputs $v_2 \otimes e_1$ and $v_2 \otimes v_2$. This combined with above means

$$((v_1)^\dagger \cdot v_2)^2 = (v_1)^\dagger \cdot v_2$$

This implies that $(v_1)^\dagger \cdot v_2$ is 0 or 1. If the inner product is 0, we have that $v_1 \perp v_2$, and if it is 1, we must have that $v_1 = v_2$. So as long as the input vectors to the task

are orthogonal, there exists a cloning procedure. However, when the inputs are not orthogonal, then cloning is impossible, and in particular the cloning task cannot work for arbitrary input vectors.

Remarks:

- No-cloning also holds for probability vectors. Indeed, all need in the above proof of no-cloning is linearity, which holds for probability vectors. However, there is a fundamental difference between what it means to not be able to clone a probability vector versus what it means to not be able to clone an amplitude vector. No cloning for probability vectors is equivalent to

  Given one sample from distribution D it is impossible to produce two independent samples.

  Intuitively, all you can do is to produce the same sample. The only case in which this would be possible is if D was the singleton distribution. The cases in which D corresponds to the singleton distribution are exactly the the cases in which the probability vectors are orthogonal. However, if we have non-trivial distributions, the probability vectors for these distributions won't be orthogonal, and thus you won't be able to do cloning.

- **What makes quantum different?** - The amplitude vectors no longer represent uncertainty. More rigorously, you can actually check if the state is in a vector v . You can check if a given state is represented by an amplitude vector v. The way you can do this is:
  - Define a unitary matrix U such that $U \cdot v = e_1$.
  - Apply U to v and observe the output (measure), and see that the ball passes through slot 1.
  On the other hand, in the stochastic model checking if the probability vector is v turns out to be impossible. This is equivalent to the following task

  Given a sample, check if the sample came from a given distribution D

  This is clearly impossible. In some cases, if the sample is outside of the support of the distribution, you can succeed, but if not then it is impossible to check. This captures the fact that no-cloning for quantum means that we are not able to clone something "real" (since we can verify it), unlike in the stochastic model.

- This gives us Quantum Money. Here the banknotes will be quantum states. By the No-cloning theorem, we cannot clone them. You can also verify banknotes by using the process above by constructing the matrix U. This, gives you something that is not clone-able, but it is verifiable. We will go into more detail later in the course about Quantum Money. This should be a signal that interesting things are happening with quantum cryptography, that were not possible before.

Question Could it be that we have some v' that is pretty close to v in terms of α and β, but when we multiply it with U in ends up being pretty close to e_1, but maybe it is $e_1 + 0.01e_2$? Then there is a small probability that we get outcome 2, even though outcome 1 is much more likely and even though our state is not v exactly
Answer: You are right. This process for checking the state will potentially pass other states that are not v, and any state that is close to v will map to something close to e_1. When you try to observe it you will see a significant weight that the ball passes through 1. That is OK. We haven't formally defined the checking task. You could get a bigger advantage that what you get for stochastic. Here we are accepting the correct states with probability 1 and we also accept other states, but the probability of doing that decays are we get further and further than the one we care about. If we apply this to quantum money, you will accept banknotes that weren't truly the valid banknote, but we hope to ensure a no-cloning property that will make the probability of that very small.

Question: Once we measure the banknote, it is essentially lost, right? We cannot measure it twice?
Answer: Yes, that is correct. We will be more precise about this later. We will see that once you apply U and measure, and you get that ball passes through 1 you can construct an equivalent banknote by applying the inverse of U to e_1. So it is lost in some sense but you can also reconstruct it.

6 Next Time

- Introduce some more quantum notation
- Explore the relation between quantum computing and cryptography further.