CS 258: Quantum Cryptography (Fall 2025)
Homework 4 (100 points)

Recall the definition of a collapsing hash function:

Definition 1. *A hash function H is collapsing if, for all QPT adversaries $\mathcal{A}$, there exists a negligible function ϵ such that $|\Pr[W_0(\lambda)] - \Pr[W_1(\lambda)]| \leq \epsilon(\lambda)$, where $W_b(\lambda)$ is the event that $\mathcal{A}$ outputs 1 in the following:*

- *$\mathcal{A}$ produces a state $\sum_{x,z} \alpha_{x,z}|x,z\rangle$*
- *If $b = 1$, measure x; if $b = 0$, measure $H(x)$. Recall that "measuring $H(x)$" means to apply U_H to write $H(x)$ into a new register, which is then measured and discarded.*
- *Return the resulting state back to $\mathcal{A}$.*
- *$\mathcal{A}$ outputs a guess b'.*

1 Problem 1 (20 points)

Show that a hash function that is collapsing is also collision resistant. *[Hint: suppose toward contradiction that it is not collision-resistant. How can you take a collision and build a state that allows you to distinguish whether the input or output is measured?]*

2 Problem 2 (20 points)

Here, you will prove two facts that will be useful when we move to Problem 3. Recall the definition of a 2-universal hash family from Lecture 3.

Definition 2. *A family $\mathcal{H}$ of functions $h : \{0,1\}^n \to \{0,1\}^m$ is* 2-universal *if, for all $x, x' \in \{0,1\}^n$, $x \neq x'$, we have that $\Pr_{h \leftarrow \mathcal{H}}[h(x) = h(x')] = 2^{-m}$.*

- **Part (a) 10 points.** For a 2-universal hash family $\mathcal{H}$ with $m \geq 3n$, prove that, except with probability at most $2^{-O(n)}$, $h \leftarrow \mathcal{H}$ is injective. *[Hint: use a union bound.]*
- **Part (b). 10 points.** Let f_0, f_1 be two functions over the same domain, say $\{0,1\}^n$. Suppose we have the guarantee that, for all pairs $x, x' \in \{0,1\}^n$, $f_0(x) = f_0(x')$ if and only if $f_1(x) = f_1(x')$. In other words, the pre-image sets of f_0 and f_1 are the same, even if the images might be completely different.

  Let $|\psi\rangle = \sum_{x,z} \alpha_{x,z}|x,z\rangle$ be a quantum state. Consider the following two processes:

  (1) Use U_{f_0} to compute $\sum_{x,z} \alpha_{x,z}|x,z,f_0(x)\rangle$, measure $f_0(x)$, and discard the result. Let $|\psi'\rangle$ be the resulting state.

  (2) Use U_{f_1} to compute $\sum_{x,z} \alpha_{x,z}|x,z,f_1(x)\rangle$, measure $f_1(x)$, and discard the result. Let $|\psi'\rangle$ be the resulting state.

  Show that the distributions over $|\psi'\rangle$ in (1) and (2) are exactly the same.

3 Problem 3 (40 points)

In class, we saw one way to build collapsing hash function. Here, you will explore another way using *lossy functions*.

Definition 3. *A lossy function is a triple of algorithms* $(\mathsf{GenInj}, \mathsf{GenLossy}, \mathsf{Eval})$ *such that:*

- $\mathsf{GenInj}(1^\lambda), \mathsf{GenLossy}(1^\lambda)$ *are a PPT algorithms which each sample a key* k*.*
- $\mathsf{Eval}(k, x)$ *is a deterministic function which takes as input a key* k *(from* GenInj *or* $\mathsf{GenLossy}$*) and an* $x \in \{0,1\}^\lambda$*. It outputs a* y*.*

There are two correctness requirements:

- Injectivity in injective mode*: for* $k \leftarrow \mathsf{GenInj}(1^\lambda)$*,* $x \mapsto \mathsf{Eval}(k, x)$ *is injective.*
- Lossiness in lossy mode*: for* $k \leftarrow \mathsf{GenLossy}(1^\lambda)$*,* $|\mathsf{Eval}(k, \cdot)| \leq 2^{\lambda/3}$*. Here,* $|\mathsf{Eval}(k, \cdot)|$ *is the number of possible outputs of* $\mathsf{Eval}(k, x)$ *as* x *ranges over* $\{0,1\}^\lambda$*.*

Here, the "lossy mode" ($k \leftarrow \mathsf{GenLossy}(1^\lambda)$) means that the output of $\mathsf{Eval}(k, x)$ loses information about x, since there are only $2^{\lambda/3}$ outputs but $2^\lambda \gg 2^{\lambda/3}$ inputs.

Notice that in the "injective mode" ($k \leftarrow \mathsf{GenInj}(1^\lambda)$, $|\mathsf{Eval}(k, \cdot)| = 2^\lambda$. Thus the injective and lossy mode functions $\mathsf{Eval}(k, \cdot)$ are very different. However, the security of lossy functions will be that the two modes are indisitnguishable:

Definition 4. *A lossy function* $(\mathsf{GenInj}, \mathsf{GenLossy}, \mathsf{Eval})$ *is secure if, for all QPT adversary* $\mathcal{A}$*, there exists a negligible funciton* ϵ *such that*

$$|\Pr[1 \leftarrow \mathcal{A}(k) : k \leftarrow \mathsf{GenInj}(1^\lambda)] - \Pr[1 \leftarrow \mathcal{A}(k) : k \leftarrow \mathsf{GenLossy}(1^\lambda)]| \leq \epsilon(\lambda)$$

You will explore how to construct lossy functions in the next problem. Here, you will see how to construct a collapsing hash assuming a lossy funciton.

The construction is the following: let $k \leftarrow \mathsf{GenLossy}(1^\lambda)$. Assume the outputs of $\mathsf{Eval}(k, \cdot)$ lie in $\{0,1\}^\ell$. Then let $h \leftarrow \mathcal{H}$ where $\mathcal{H}$ is a 2-universal hash family (for definition, see Lecture 3), such that $h : \{0,1\}^\ell \to \{0,1\}^{\lambda-1}$. Then define $H(x) = h(\mathsf{Eval}(k, x))$, which takes λ bit inputs to $\lambda - 1$ bit outputs.

Now, in the proof, we will consider 5 different experiments:

- W_0: This is the case $b = 0$ in the collapsing experiment, where $k \leftarrow \mathsf{GenLossy}(1^\lambda)$ and we measure $H(x)$.
- V_0: Here, we sample $k \leftarrow \mathsf{GenLossy}(1^\lambda)$, but now we measure $\mathsf{Eval}(k, x)$ instead of $H(x)$
- V_1: Now we sample $k \leftarrow \mathsf{GenInj}(1^\lambda)$, but still measure $\mathsf{Eval}(k, x)$.
- V_2: We still sample $k \leftarrow \mathsf{GenInj}(1^\lambda)$, but now measure x instead of $\mathsf{Eval}(k, x)$.
- W_1: Now we sample $k \leftarrow \mathsf{GenLossy}(1^\lambda)$, but still measure x. This matches the case $b = 1$ in the collapsing experiment where we measure x.
- **Part (a). 10 points.** Show that $|\Pr[W_0] - \Pr[V_0]|$ is negligible. *[Hint: Use the facts proved in Problem 2. We can think of the function* h *as being restricted to the set of images of* $\mathsf{Eval}(k, \cdot)$*]*
- **Part (b). 10 points.** Show that $|\Pr[V_0] - \Pr[V_1]|$ is negligible, assuming the lossy function is secure.
- **Part (c). 10 points.** Show that $|\Pr[V_1] - \Pr[V_2]|$ is negligible.
- **Part (d). 10 points.** Show that $|\Pr[V_2] - \Pr[W_1]|$ is negligible, assuming the lossy function is secure.

Putting parts (a) through (d) together shows that H is collapsing.

4 Problem 4 (20 Points)

Now we will construct a lossy function from LWE. The function is simple: the key will consist of a matrix $\mathbf{B} \in \mathbb{Z}_q^{\ell \times m}$ for $\ell \gg m$. Then we have that

$$\mathsf{Eval}(\mathbf{B}, \mathbf{x} \in \{0,1\}^m) = \lfloor \mathbf{B} \cdot \mathbf{x} \rceil_{q/4}$$

GenInj samples $\mathbf{B}$ uniformly, while $\mathsf{GenLossy}$ samples $\mathbf{B}$ as $\mathbf{B} = \mathbf{S} \cdot \mathbf{A} + \mathbf{E} \bmod q$, where $\mathbf{A} \in \mathbb{Z}_q^{n \times m}$ and $\mathbf{S} \in \mathbb{Z}_q^{\ell \times n}$ are chosen uniformly, and $\mathbf{E} \in \mathbb{Z}^{\ell \times m}$ is sampled from the discrete Gaussian of width σ. In other words, the injective mode $\mathbf{B}$ is a random matrix, while the lossy mode $\mathbf{B}$ is close to a matrix $\mathbf{SA}$ of rank n.

- **Part (a). 10 points.** Prove that the security of $(\mathsf{GenInj}, \mathsf{GenLossy}, \mathsf{Eval})$, assuming the LWE assumption holds. To do so, prove that the distributions of $\mathcal{B}$ in GenInj and $\mathsf{GenLossy}$ are computational indistinguishable. This is accomplished by defining a set of hybrid experiments where the first i rows of $\mathbf{B}$ are as in $\mathsf{GenLossy}$, but the remaining $\ell - i$ rows are chosen as in GenInj. Show that the i and $i-1$ case are indistinguishable, by LWE.
- **Part (b). 10 points.** Suppose $\mathbf{E}$ did not exist, and $\mathsf{GenLossy}$ sampled $\mathbf{B}$ as $\mathbf{S} \cdot \mathbf{A}$. What is an upper bound on the size of the image of $\mathsf{Eval}(\mathbf{B}, \cdot)$?

We've actually already seen that GenInj is injective, for appropriately large ℓ (In Lecture 12, when we were trying to apply Kuperberg's algorithm to break LWE). As for showing that $\mathsf{GenLossy}$ is lossy, the hope is that the rounding causes the output of Eval to depend only on $\mathbf{S} \cdot \mathbf{A} \cdot \mathbf{x}$, and the rounding eliminates the part that depends on the error $\mathbf{E}$. This will be true for "most" inputs, but unfortunately is not true in general, since some fraction of the inputs will be close to the rounding boundary, causing different outputs. With some care, however, it is possible to extend the construction above to a full lossy function.