

COL872: QUANTUM AND POST-QUANTUM CRYPTOGRAPHY

PROBLEM SET 1

Due date: January 18th, 2023

INSTRUCTIONS

1. Assignments must be done in groups of size at most three. Each group must upload one submission, and mention the names of all group members.
2. You are welcome to discuss with other classmates and instructor, as well as refer to resources online. But if you do, please mention who all you collaborated with, or the online resources used.
3. You can use the following fact for this assignment:

Fact 1 (Leftover Hash Lemma - another simplified version). *Let n, m, q, t be integers such that $t \geq n \cdot m \cdot \log q + n$. The following distributions are statistically indistinguishable:*

$$\mathcal{D}_1 = \left\{ (\mathbf{A}_{i,b})_{i \in [t], b \in \{0,1\}} : \begin{array}{l} \mathbf{r} \leftarrow \{0,1\}^t \\ \mathbf{A}_{i,b} \leftarrow \mathbb{Z}_q^{n \times m} \text{ for all } i \in [t-1], b \in \{0,1\} \\ \mathbf{A}_{t,1-\mathbf{r}_t} \leftarrow \mathbb{Z}_q^{n \times m} \\ \mathbf{A}_{t,\mathbf{r}_t} = -\sum_{i < t} \mathbf{A}_{i,\mathbf{r}_i} \bmod q \end{array} \right\}$$
$$\mathcal{D}_2 = \left\{ (\mathbf{A}_{i,b})_{i \in [t], b \in \{0,1\}} : \mathbf{A}_{i,b} \leftarrow \mathbb{Z}_q^{n \times m} \text{ for all } i \in [t], b \in \{0,1\} \right\}$$

QUESTIONS

Question 1. Lossy Encryption (5 points)

A Secure Lossy Encryption Scheme for message space \mathcal{M} consists of four algorithms: (Setup, Setup-Lossy, Enc, Dec) with the following syntax-

- Setup(1^n): The (standard) setup algorithm takes as input the security parameter n and outputs public key pk and secret key sk .

Not much context for this question; it's meant to be a warm-up for the course. Also, it will be needed for Question 2.

-
- $\text{Setup-Lossy}(1^n)$: The lossy setup algorithm takes as input the security parameter n and outputs a lossy public key pk .
 - $\text{Enc}(\text{pk}, m)$: The encryption algorithm takes as input a public key pk (either standard or lossy) and message $m \in \mathcal{M}$, and outputs a ciphertext ct .
 - $\text{Dec}(\text{sk}, \text{ct})$: The decryption algorithm takes as input a secret key sk (output by the standard setup) and a ciphertext, and outputs $y \in \mathcal{M} \cup \{\perp\}$.
Here \perp simply denotes a special failure symbol, or a way to denote the algorithm aborts.

These algorithms must satisfy the following properties:

- **Correctness:** For all messages $m \in \mathcal{M}$, $(\text{pk}, \text{sk}) \leftarrow \text{Setup}(1^n)$,

$$\text{Dec}(\text{sk}, \text{Enc}(\text{pk}, m)) = m.$$
- **Indistinguishability of Modes:** The following distributions are computationally indistinguishable:
 - $\{\text{pk} : (\text{pk}, \text{sk}) \leftarrow \text{Setup}(1^n)\}$
 - $\{\text{pk} : \text{pk} \leftarrow \text{Setup-Lossy}(1^n)\}$
- **Statistical Indistinguishability in Lossy Mode:** For any messages m_0, m_1 , the following distributions are statistically indistinguishable:
 - $\{\text{pk}, \text{Enc}(\text{pk}, m_0) : \text{pk} \leftarrow \text{Setup-Lossy}(1^n)\}$
 - $\{\text{pk}, \text{Enc}(\text{pk}, m_1) : \text{pk} \leftarrow \text{Setup-Lossy}(1^n)\}$

1. Show that if $\mathbb{E} = (\text{Setup}, \text{Setup-Lossy}, \text{Enc}, \text{Dec})$ is a lossy encryption scheme satisfying all the above properties, then $\mathbb{E}' = (\text{Setup}, \text{Enc}, \text{Dec})$ is also a correct and semantically secure public key encryption scheme.

Question 2. A Lossy Encryption Scheme based on LWE (5 points)

In this problem, you will have to construct a lossy encryption mode for Regev encryption. The algorithms Setup , Enc , Dec are defined as in class (see Lecture Notes). You must define the Setup-Lossy algorithm, and then show that it is a secure lossy encryption scheme.

Easy problem. The aim of this problem is to get comfortable with Regev's scheme.

1. First, define the Lossy Setup algorithm Setup-Lossy formally.
2. Show that the public keys output by Setup and Setup-Lossy are computationally indistinguishable.
3. Finally, argue that it satisfies statistical indistinguishability in lossy mode.

Question 3. Small Secrets LWE - Matrix Version (5 points)

Consider the following two distributions:

$$\mathcal{D}_1 = \left\{ (\mathbf{A}, \mathbf{B}) : \begin{array}{l} \mathbf{A} \leftarrow \mathbb{Z}_q^{n \times m} \\ \mathbf{S} \leftarrow \chi^{n \times n} \\ \mathbf{E} \leftarrow \chi^{n \times m} \\ \mathbf{B} = \mathbf{S} \cdot \mathbf{A} + \mathbf{E} \end{array} \right\} \quad \mathcal{D}_2 = \left\{ (\mathbf{A}, \mathbf{B}) : \mathbf{A}, \mathbf{B} \leftarrow \mathbb{Z}_q^{n \times m} \right\}$$

Should follow from an easy hybrid argument; will be needed in Question 6, and also later in the course.

1. Prove that the above distributions are computationally indistinguishable, assuming $\text{ss-LWE}_{n,m,q,\chi}$ is a hard computational problem.
2. Let $q = 2^{\sqrt{n}}$ and $\chi \equiv \text{Unif}_{[-\sqrt{q}, \sqrt{q}]}$. Would this problem remain hard if even \mathbf{A} is chosen from $\chi^{n \times m}$ in both the distributions?
3. Recall, in the Learning-with-errors problem, if $m = n$, then the two distributions are statistically indistinguishable. Does the same hold true in the small-secrets setting?

Question 4. Full Rank Matrices vs Noisy Low Rank Matrices (5 marks)

Let q be a sufficiently large prime (say $q = O(2^{\sqrt{n}})$) and let $\chi = \text{Unif}_{[-\sqrt{q}, \sqrt{q}]}$. A uniformly random matrix $\mathbb{Z}_q^{2n \times 2n}$ will be a full rank matrix with overwhelming probability. Consider the following distribution:

Should follow from an easy hybrid argument.

$$\mathcal{D} = \left\{ (\mathbf{A}, \mathbf{A} \cdot \mathbf{B} + \mathbf{E}) : \mathbf{A} \leftarrow \mathbb{Z}_q^{2n \times n}, \mathbf{B} \leftarrow \mathbb{Z}_q^{n \times 2n}, \mathbf{E} \leftarrow \chi^{2n \times 2n} \right\}$$

Note that this distribution consists of noisy low-rank matrices.

1. Show that \mathcal{D} is computationally indistinguishable from the uniform distribution over $\mathbb{Z}_q^{2n \times n} \times \mathbb{Z}_q^{2n \times 2n}$.

Question 5. Random-looking t matrices with a special structure (8 points)

Let $q = 2^{\sqrt{n}}$ and $\chi \equiv \text{Unif}_{[-\sqrt{q}, \sqrt{q}]}$. We want to define an efficiently samplable distribution \mathcal{D} over t matrices of dimension $n \times 2n$ with the following properties:

Will be used in Question 6.

- The distribution \mathcal{D} is computationally indistinguishable from the uniform distribution over $(\mathbb{Z}_q^{n \times 2n})^t$.
- Let $t = \text{poly}(n)$. A matrix is said to have **small entries** if there is some entry in the range $[0, q^{0.75}] \cup [q - q^{0.75}, q]$. We require the probability of the following event to be 0:

$$\left[\exists \text{ nonempty set } S \subseteq [t] \text{ s.t. } \mathbf{A} = \sum_{i \in S} \mathbf{A}_i \text{ has small entries} : (\mathbf{A}_i)_{i \in [t]} \leftarrow \mathcal{D} \right]$$

1. Define distribution \mathcal{D} that satisfies the above properties.

Hint: Consider the matrix $\mathbf{B} = [\mathbf{A} \mid \mathbf{A} \cdot \mathbf{S} + \mathbf{E} + \mathbf{D}]$ where \mathbf{S} and \mathbf{E} are drawn from an appropriate error distribution, and \mathbf{D} is of some special form. If \mathbf{A} has non-small entries, then \mathbf{B} also has non-small entries. If not, then argue that $\mathbf{A} \cdot \mathbf{S} + \mathbf{E}$ will also have smallish entries, and therefore the \mathbf{D} will ensure that your overall matrix has non-small entries.

Question 6. Code Obfuscation (12 points)

The goal of code obfuscation is to compile a program so that the compiled code can be evaluated by everyone, but it ‘hides’ the source code. In this exercise, we will develop code obfuscation for point functions, based on LWE.

FUNCTION CLASS: Let $\mathcal{F}_t = \{f_z : \{0,1\}^t \rightarrow \{0,1\}\}_{z \in \{0,1\}^t}$, where

$$f_z(x) = \begin{cases} 1 & \text{if } x = z. \\ 0 & \text{otherwise.} \end{cases}$$

This will be a building block for more elaborate obfuscation schemes that we might see later in the course. Hopefully this exercise will also illustrate the usefulness of LWE as an assumption (even if we ignore the post-quantum aspect).

An obfuscation scheme for \mathcal{F}_t consists of algorithms Obf, Eval with the following syntax.

- Obf($1^n, f_z$): The obfuscation algorithm is randomized; it takes as input the security parameter n and a point function $f_z \in \mathcal{F}_t$. It outputs enc, an encoding of f_z .
- Eval(enc, x): The evaluation algorithm is deterministic; it takes as input an encoding enc, and an input $x \in \{0,1\}^t$, and outputs 0/1.

CORRECTNESS: For correctness, we require that for every n , $f_z \in \mathcal{F}_t$, and all $x \in \{0,1\}^t$, $\Pr[\text{Eval}(\text{Obf}(1^n, f_z), x) = f_z(x)] = 1$.

SECURITY: Intuitively, the security definition states that the obfuscation of a random function from \mathcal{F}_t is indistinguishable from a uniformly random string. Let ℓ denote the length of obfuscation of $f_z \in \mathcal{F}_t$. More formally, we say that the obfuscation scheme is secure if the following two distributions are computationally indistinguishable:

$$\mathcal{D}_1 = \left\{ \text{enc} : \begin{array}{l} z \leftarrow \{0,1\}^t \\ \text{enc} \leftarrow \text{Obf}(1^n, f_z) \end{array} \right\} \quad \mathcal{D}_2 = \text{Unif}_{\{0,1\}^\ell}$$

ATTEMPT 1: The obfuscation of f_z consists of $2t$ integers $\{a_{i,b}\}_{i \in [t], b \in \{0,1\}}$ chosen as follows. For all $i < t$, the integers $a_{i,b}$ are chosen uniformly at random from \mathbb{Z}_q . The integer a_{t,z_t} is set to $-\sum_{i < t} a_{i,z_i} \pmod{q}$, while $a_{t,1-z_t}$ is uniformly random.

Evaluation is defined as follows: on input x , compute $\sum_i a_{i,x_i}$. If the sum is 0, then output 1, else output 0. It is easy to check that $\text{Eval}(\text{Obf}(f_z), x) = 1$ if $x = z$.

One can show that, using Fact 1, \mathcal{D}_1 and \mathcal{D}_2 are statistically indistinguishable if t is large enough.

1. What is the issue with this attempt?

ATTEMPT 2: We will still follow the same approach as Attempt 1, but will use matrices instead of integers. The obfuscation function will use the bit string z to sample $2t$ matrices such that

- these matrices look like uniformly random matrices (this will follow from small-secrets LWE).
- the matrices corresponding to string z , when added together, will have small entries.
- for any string $x \neq z$, when the corresponding matrices are added together, there will have at least one non-small entry.

You should use the ideas developed in Question 5.

2. Describe Obf and Eval.

3. Show that your scheme satisfies correctness.

4. Prove security of your scheme, assuming ss-LWE. Note that t must be large for this proof to work. How large must t be?

This is version 2 of the assignment, updated on 15th January.