Tutorial 3: PRG and Symmetric encryption schemes

Exercise 1. *Learning with errors is back.*

Definition 1 (Learning with Errors). *Let $\ell < k \in \mathbb{N}$, $n < m \in \mathbb{N}$, $q = 2^k$, $B = 2^\ell$, $\mathbf{A} \hookleftarrow U(\mathbb{Z}_q^{m\times n})$. The Learning with Errors (*LWE*) distribution is defined as follows: $D_{\text{LWE},\mathbf{A}} = (\mathbf{A}, \mathbf{A}\cdot\mathbf{s} + \mathbf{e} \bmod q)$ for $\mathbf{s} \hookleftarrow U(\mathbb{Z}_q^n)$ and $\mathbf{e} \hookleftarrow U\left(\left[-\frac{B}{2}, \frac{B}{2} - 1\right]^m \cap \mathbb{Z}^m\right)$.*

The *LWE assumption* states that, given suitable parameters k, ℓ, m, n, it is computationally hard to distinguish $D_{\text{LWE},\mathbf{A}}$ from the distribution $(\mathbf{A}, U(\mathbb{Z}_q^m))$.
Let us consider the private-key encryption scheme below, which works under the following public parameters: $k, \ell, m, n, \mathbf{A}$, for which the LWE assumption holds.
Note. Here, $a \bmod q$ denotes the representative of the class of a in $[-\frac{q}{2}, \frac{q}{2} - 1] \cap \mathbb{Z}$ and not in $[0, q - 1] \cap \mathbb{Z}$ (this will ease the description of the scheme).

Keygen(1^λ)**:** from 1^λ, this algorithm outputs a random vector $\mathbf{s} \hookleftarrow U(\mathbb{Z}_q^n)$ as a secret key.

Enc$_\mathbf{s}(\mathfrak{m})$**:** from the secret key $\mathbf{s}$ and a message $\mathfrak{m} \in \{0,1\}^m$, the algorithm Enc samples a random vector $\mathbf{e} \hookleftarrow U\left(\left[-\frac{B}{2}, \frac{B}{2}\right]^m \cap \mathbb{Z}^m\right)$ and outputs $\mathbf{c} = \mathbf{A}\mathbf{s} + \mathbf{e} + \frac{q}{2}\mathfrak{m} \bmod q$ as a ciphertext.

Dec$_\mathbf{s}(\mathbf{c})$**:** from the secret key $\mathbf{s}$ and a ciphertext $\mathbf{c}$, the decryption algorithm computes $\mathbf{v} = \mathbf{c} - \mathbf{A}\cdot\mathbf{s}$. Then Dec constructs the message $\mathfrak{m}'$ from $\mathbf{v}$: for each component v_i of $\mathbf{v}$, sets the corresponding component of $\mathfrak{m}'$ as follows: 0 if $-q/4 \le v_i < q/4$, and 1 otherwise.

1. Prove the correctness of this cipher.
2. Show that this cipher is computationally secure.

If you take a look at this cipher, you can view it as a one-time pad on $\frac{q}{2}\mathfrak{m}$, which means that the message is hidden in the most significant bit of $\mathbf{e} + \frac{q}{2}\mathfrak{m}$.
Now, if one wants to hide the message in the least significant bit of the OTP, one solution is to encrypt a message as: $\mathbf{c} = 2\cdot(\mathbf{A}\cdot\mathbf{s} + \mathbf{e}) + \mathfrak{m} \bmod q$.

3. Construct a "decryption" algorithm that does not use the secret key to recover $\mathfrak{m}$ (i.e., show that this scheme is not secure).
4. Why is it also a bad idea to encrypt as $\mathbf{c} = \mathbf{A}\cdot\mathbf{s} + 2\mathbf{e} + \mathfrak{m}$?

Exercise 2. *Symmetric encryption scheme from a PRG.*
Let $G : \{0,1\}^s \to \{0,1\}^n$ be a pseudo-random generator. We define a symmetric encryption scheme (KeyGen, Enc, Dec) by

- Keygen(1^s) : outputs a uniform element $k \in \{0,1\}^s$;
- Enc$(k, m) = m \oplus G(k)$, where $m \in \{0,1\}^n$ and $k \in \{0,1\}^s$;
- Dec$(k, c) = c \oplus G(k)$, where $c \in \{0,1\}^n$ and $k \in \{0,1\}^s$,

where $\oplus$ denotes a xor performed component wise.

1. Show that this scheme is correct.

   Let $\mathcal{A}$ be a PPT algorithm such that there exist two messages m_0 and m_1 such that
   $$\text{Adv}(\mathcal{A}) := \Pr_{k,\beta}(\mathcal{A}(\text{Enc}(k, m_\beta)) = \beta) \ge 1/2 + \varepsilon,$$
   where the randomness is over the uniform choices of k, β and the internal randomness of $\mathcal{A}$.

2. Show that there exists a PPT adversary $\mathcal{A}'$ such that

$$\mathrm{Adv}(\mathcal{A}') = \Pr_{\substack{b \leftarrow U(\{0,1\}) \\ x \leftarrow D_b}} (\mathcal{A}'(x) = b) \geq 1/2 + \varepsilon/2,$$

where $D_0 = G(U(\{0,1\}^s))$ and $D_1 = U(\{0,1\}^n)$. What does it prove about the security of the encryption scheme?

Exercise 3. *Arbitrary length encryption.*

Toward Arbitrary-length Encryption. An arbitrary-length encryption scheme is a triple of algorithms $(KeyGen, Enc, Dec)$ such that $KeyGen$ outputs a key $k \in \{0,1\}^n$, Enc takes as inputs a key k, and a message M of arbitrary length ℓ and returns a ciphertext C, and Dec takes as inputs a key k and a ciphertext C and returns a plaintext M. We require that for all k, and for all M of arbitrary length, we have $Dec(k, Enc(k, M)) = M$.

We define one-time CPA-security for arbitrary-length messages with the following two games. For each $b \in \{0,1\}$, Game$_b$ starts by the challenger generating a key k uniformly at random. Then, the adversary should send two distinct plaintexts M_0 and M_1 of equal lengths to the challenger. The challenger encrypts M_b and sends the corresponding ciphertext. Then, the adversary outputs a bit b'. The scheme is considered secure if, for any probabilistic polynomial-time adversary, the difference between the probabilities that $b' = 1$ in Game$_0$ and Game$_1$ is negligible with respect to n.

1. Why do we need to assume that the challenge plaintexts are of equal lengths?

Let G^* be an arbitrary-length PRG. Define $Enc(k, M)$ as follows: $Enc(k, M) = G^*(k, 1^\ell) \oplus M$, with ℓ the length of M.

2. Propose a corresponding decryption algorithm Dec. Is this scheme still secure if we use the same key to encrypt two different messages?

3. Prove that if G^* is a secure arbitrary-length PRG, then $(KeyGen, Enc, Dec)$ is a one-time CPA-secure arbitrary-length encryption scheme.

Exercise 4. *Increasing the advantage of an attacker.*

Let G be a pseudo-random generator from $\{0,1\}^s$ to $\{0,1\}^n$ for some integers s and n. Let $i \in \{1, \cdots, n\}$ and let $\mathcal{A}$ be a PPT algorithm such that, for all $k \in \{0,1\}^s$, we have

$$\Pr[\mathcal{A}(G(k)_{1\cdots i-1}) = G(k)_i] \geq \frac{1}{2} + \varepsilon,$$

where the probability runs over the randomness of $\mathcal{A}$. Note that unlike the definition of the advantage seen in class, here we consider only the probability over the randomness of $\mathcal{A}$ and not over the random choice of k (we will see why later).

Our objective is to construct a new attacker $\mathcal{A}'$ with an advantage arbitrarily close to 1 (for instance $\Pr[\mathcal{A}(G(k)_{1\cdots i-1}) = G(k)_i] \geq 0.999$ for all $k \in \{0,1\}^s$).

1. Propose a method to improve the success probability of $\mathcal{A}$.

   Let m be some integer to be determined. Let $\mathcal{A}'$ be an algorithm that evaluates $\mathcal{A}$ on $G(k)_{1\cdots i-1}$ $2m+1$ times, to obtain $2m+1$ bits $b_1, \cdots, b_{2m+1}$ and then outputs the bit that appeared the most (i.e. at least $m+1$ times).

2. Give a lower bound on $\Pr[\mathcal{A}'(G(k)_{1\cdots i-1}) = G(k)_i]$, for all $k \in \{0,1\}^s$. We recall Hoeffding's inequality for Bernoulli variables: let $X_1, \cdots, X_{2m+1}$ be independent Bernoulli random variables, with $\Pr(X_i = 1) = 1 - \Pr(X_i = 0) = p$ for all i, and let $S = X_1 + \cdots + X_{2m+1}$. Then, for all $x > 0$, we have

$$\Pr[|S - \mathbb{E}(S)| \geq x\sqrt{2m+1}] \leq 2e^{-2x^2}.$$

3. What should be the value of m (depending on ε) if we want that $\Pr[\mathcal{A}'(G(k)_{1\cdots i-1}) = G(k)_i] \geq 0.999$ for all k? It may be useful to know that $e^{-8} \leq 0.0005$.

4. Do we have $\mathrm{Adv}_{\mathrm{unpredictability}}(\mathcal{A}') \geq 0.999$ if $\Pr[\mathcal{A}'(G(k)_{1\cdots i-1}) = G(k)_i] \geq 0.999$ for all k?

5. What condition on ε do we need to ensure that $\mathcal{A}'$ runs in polynomial time?

   Let now $\mathcal{A}$ be an attacker such that

$$\mathrm{Adv}(\mathcal{A}) = \Pr_{k \leftarrow U(\{0,1\}^s)}[\mathcal{A}(G(k)_{1\cdots i-1}) = G(k)_i] \geq \frac{1}{2} + \varepsilon.$$

   Note that we are now looking at the definition of advantage given in class, where the probability also depends on the uniform choice of k. We want to show that in this case, we cannot always amplify the success probability of the attacker by repeating the computation.

   In the following, we write $\Pr[\mathcal{A}(G(k)_{1\cdots i-1}) = G(k)_i]$ when we only consider the probability over the internal randomness of $\mathcal{A}$ (and k is fixed) and $\Pr_{k \leftarrow U(\{0,1\}^s)}[\mathcal{A}(G(k)_{1\cdots i-1}) = G(k)_i]$ when we consider the probability over the choice of k and the internal randomness of $\mathcal{A}$.

   Suppose that $s \geq 2$ and define

$$G(k) = \begin{cases} 00\cdots 0 & \text{if } k_0 = k_1 = 0 \\ G_0(k) & \text{otherwise,} \end{cases}$$

   where G_0 is a secure PRG from $\{0,1\}^s$ to $\{0,1\}^n$.

6. Show that there exists a PPT attacker $\mathcal{A}$ with non negligible advantage (for the unpredictability definition) against G.

7. Show on the contrary that there is no PPT attacker $\mathcal{A}$ with $\mathrm{Adv}(\mathcal{A}) \geq 7/8$ (assuming that G_0 is a secure PRG).