TD 4 Advanced cryptography 2024

Léo Colisson Palais

Exercice 1: Security of Regev's encryption

Solve the questions from the slides, i.e. show that for appropriately chosen parameters, the LWE-based encryption seen in the course is secure assuming the hardness of LWE.

Correction. We assume that LWE is secure, i.e.:

We want to prove that Regev's encryption is IND-CPA secure, i.e.:

$$\boxed{\begin{array}{c}\text{LWE-R}\\ \hline \begin{array}{l}(p_k, s_k) \leftarrow \mathsf{Gen}(1^\lambda)\\ \textsc{eavesdrop}(m_L, m_R): \\ \hline \quad \textbf{return } \mathsf{Enc}_{p_k}(m_L)\end{array}\end{array}} \approx \boxed{\begin{array}{c}\text{LWE-R}\\ \hline \begin{array}{l}(p_k, s_k) \leftarrow \mathsf{Gen}(1^\lambda)\\ \textsc{eavesdrop}(m_L, m_R): \\ \hline \quad \textbf{return } \mathsf{Enc}_{p_k}(m_R)\end{array}\end{array}} \tag{1}$$

□

Exercice 2: Worst case to average case reduction

One may be worried that the LWE problem may be hard in the worst case ("there exists a few hard instances"), but that it is still insecure on average ("if I pick a random LWE instance, it is easy to break on average"). In this exercise, we will show that if the LWE problem is hard in the worst case, it is also hard on average. Or said differently, if we can break LWE on average, we can break *all* LWE instances. In the following, $A_{\mathbf{s},\chi}$ will denote the LWE distribution sampling $\mathbf{a} \xleftarrow{\$} \mathbb{Z}_q^n$, $e \leftarrow \chi$ and returning $(\mathbf{a}, \mathbf{a}^T\mathbf{s} + e)$, where n and q are integers and χ is a distribution on $\mathbb{Z}_q$. For any set S, U_S will denote the uniform distribution on S, and we define $U \coloneqq U_{\mathbb{Z}_q^n \times \mathbb{Z}_q}$.

1. For any $\mathbf{t} \in \mathbb{Z}_q^n$, let $f_\mathbf{t} \colon \mathbb{Z}_q^n \times \mathbb{Z}_q \to \mathbb{Z}_q^n \times \mathbb{Z}_q$ be defined as $f_t(\mathbf{a}, b) \coloneqq (\mathbf{a}, b + \mathbf{a}^T\mathbf{t})$. For any $\mathbf{t}$ and $\mathbf{s}$, show that $f_\mathbf{t}(A_{\mathbf{s},\chi}) = A_{\mathbf{s}+\mathbf{t},\chi}$, i.e. that the distribution obtained by applying $f_\mathbf{t}$ to a sample of $A_{\mathbf{s},\chi}$ is statistically indistinguishable from a sample of $A_{\mathbf{s}+\mathbf{t},\chi}$.

   Correction. For any $x \in \mathbb{Z}_q^n \times \mathbb{Z}_q$, we have:

$$\Pr\left[\, f_\mathbf{t}(A_{\mathbf{s},\chi}) = x \,\right] \tag{2}$$

$$= \Pr_{\substack{\mathbf{a} \xleftarrow{\$} Z_q^n \\ e \leftarrow \chi}} \left[\, f_\mathbf{t}(\mathbf{a}, \mathbf{a}^T\mathbf{s} + e) = x \,\right] \tag{3}$$

$$= \Pr_{\substack{\mathbf{a} \xleftarrow{\$} Z_q^n \\ e \leftarrow \chi}} \left[\, (\mathbf{a}, \mathbf{a}^T\mathbf{s} + e + \mathbf{a}^T\mathbf{t}) = x \,\right] \tag{4}$$

$$= \Pr_{\substack{\mathbf{a} \xleftarrow{\$} Z_q^n \\ e \leftarrow \chi}} \left[\, (\mathbf{a}, \mathbf{a}^T(\mathbf{s} + \mathbf{t}) + e) = x \,\right] \tag{5}$$

$$= \Pr\left[\, A_{\mathbf{s}+\mathbf{t}} = x \,\right] \tag{6}$$

   □

2. Show that the image of the uniform distribution on $\mathbb{Z}_q^n \times \mathbb{Z}_q$ by $f_\mathbf{t}$ is the uniform distribution on $\mathbb{Z}_q^n \times \mathbb{Z}_q$, i.e. $f_\mathbf{t}(U_{\mathbb{Z}_q^n \times \mathbb{Z}_q}) = U_{\mathbb{Z}_q^n \times \mathbb{Z}_q}$.

Correction. For any $(x_0, x_1) \in \mathbb{Z}_q^n \times \mathbb{Z}_q$, we have:

$$\Pr\left[\, f_{\mathbf{t}}(U) = (x_0, x_1) \,\right] \tag{7}$$

$$= \Pr_{\substack{\mathbf{a} \overset{\$}{\leftarrow} Z_q^n \\ b \overset{\$}{\leftarrow} Z_q}} \left[\, f_{\mathbf{t}}(\mathbf{a}, b) = (x_0, x_1) \,\right] \tag{8}$$

$$= \Pr_{\substack{\mathbf{a} \overset{\$}{\leftarrow} Z_q^n \\ b \overset{\$}{\leftarrow} Z_q}} \left[\, (\mathbf{a}, \mathbf{a}^T\mathbf{t} + b) = (x_0, x_1) \,\right] \tag{9}$$

$$= \Pr_{\substack{\mathbf{a} \overset{\$}{\leftarrow} Z_q^n \\ b \overset{\$}{\leftarrow} Z_q}} \left[\, \mathbf{a} = x_0 \wedge \mathbf{a}^T\mathbf{t} + b = x_1 \,\right] \tag{10}$$

$$= \Pr_{\substack{\mathbf{a} \overset{\$}{\leftarrow} Z_q^n \\ b \overset{\$}{\leftarrow} Z_q}} \left[\, \mathbf{a} = x_0 \wedge b = x_1 - \mathbf{a}^T\mathbf{t} \,\right] \tag{11}$$

$$= \frac{1}{q^n} \times \frac{1}{q} \tag{12}$$

$$= \Pr\left[\, U = (x_0, x_1) \,\right] \tag{13}$$

□

3. We assume that there exists a polynomial-time algorithm W that distinguishes $U_{\mathbb{Z}_q^n \times \mathbb{Z}_q}$ from $A_{\mathbf{s},\chi}$ on average, i.e. such that there exists a set $S \subseteq \mathbb{Z}_q^n$ of elements where W guesses reasonably correctly, i.e. such that for any $s \in S$ and large enough n,

$$|\Pr_{x \leftarrow A_{\mathbf{s},\chi}}\left[\, W(x) = 1 \,\right] - \Pr_{x \overset{\$}{\leftarrow} \mathbb{Z}_q^n \times \mathbb{Z}_q}\left[\, W(x) = 1 \,\right]| \geq \frac{1}{n^c} \tag{14}$$

for some integer $c > 1$, and such that S covers a non-negligible fraction of $\mathbb{Z}_q^n \times \mathbb{Z}_q$, i.e. there exists some integer $c' > 1$ such that, for large enough n:

$$\frac{|S|}{q^{n+1}} \geq \frac{1}{n^{c'}} \tag{15}$$

Our goal is to use W to build an adversary able to break LWE for all instances.

(a) Without loss of generality, we can assume that W outputs either 0 or 1. Show that

$$\Pr_{x \leftarrow D}\left[\, W(x) = 1 \,\right] = \mathop{\mathbb{E}}_{x \leftarrow D}\left[\, W(x) \,\right] \tag{16}$$

Correction. We have by definition of the esperance and from the fact that W outputs either 0 or 1:

$$\mathop{\mathbb{E}}_{x \leftarrow D}\left[\, W(x) \,\right] = 0 \times \Pr_{x \leftarrow D}\left[\, W(x) = 0 \,\right] + 1 \times \Pr_{x \leftarrow D}\left[\, W(x) = 1 \,\right] = \Pr_{x \leftarrow D}\left[\, W(x) = 1 \,\right] \tag{17}$$

□

(b) First, show that for any distribution D, one can efficiently estimate the quantity $\Pr_{x \leftarrow D}\left[\, W(x) = 1 \,\right]$, i.e. that there exists a procedure $\mathsf{Estimate}(D)$ running in time polynomial in n, such that the probability to have:

$$|\mathsf{Estimate}(D) - \Pr_{x \leftarrow D}\left[\, W(x) = 1 \,\right]| \geq \frac{1}{10n^c} \tag{18}$$

is negligible in n.

Hint: For this, you may want to use the inequality of Hoeffding, that (in particular) states that if N independent random variables $V_1, \ldots, V_N$ are bounded by 0 and 1, then for any $t \geq 0$,

$$\Pr\left[\left| \sum_i V_i - \mathbb{E}\left[\sum_i V_i \right] \right| \geq t \right] \leq 2 \exp\left(-\frac{2t^2}{N} \right) \tag{19}$$

Correction. We define $\mathsf{Estimate}(D)$ as the operation that averages D over N (to be determine later) independent samples, i.e.:

$$\mathsf{Estimate}(D) \coloneqq \frac{1}{N}\sum_{i=1}^{N} D \tag{20}$$

We compute now the precision of this function:

$$\Pr\left[|\mathsf{Estimate}(D) - \Pr_{x\leftarrow D}[W(x)=1]| \geq \frac{1}{10n^c}\right] \tag{21}$$

$$= \Pr\left[|\frac{1}{N}\sum_{i=1}^{N} D - \Pr_{x\leftarrow D}[W(x)=1]| \geq \frac{1}{10n^c}\right] \tag{22}$$

$$= \Pr\left[|\sum_{i=1}^{N} D - N\Pr_{x\leftarrow D}[W(x)=1]| \geq \frac{N}{10n^c}\right] \quad = \Pr\left[|\sum_{i=1}^{N} D - \mathop{\mathbb{E}}_{x\leftarrow D}[NW(x)]| \geq \frac{N}{10n^c}\right] \tag{See eq. (16)}$$

We can now apply the inequality of Hoeffding with $V_1 \coloneqq \cdots = V_N \coloneqq D$ (all independent copies of D, which are indeed bounded by 0 and 1 if we consider that W outputs either 0 or 1), and $t = \frac{N}{10n^c}$:

$$\Pr\left[|\sum_{i=1}^{N} D - \mathop{\mathbb{E}}_{x\leftarrow D}[NW(x)]| \geq \frac{N}{10n^c}\right] \tag{23}$$

$$\leq 2\exp\left(-\frac{2\left(\frac{N}{10n^c}\right)^2}{N}\right) \tag{Hoeffding}$$

$$\leq 2\exp\left(-\frac{2N}{(10n^c)^2}\right) \tag{24}$$

If we define for instance $N = n(10n^c)^2$, then the probability of having a poor estimate is negligible:

$$\Pr\left[|\mathsf{Estimate}(D) - \Pr_{x\leftarrow D}[W(x)=1]| \geq \frac{1}{10n^c}\right] \leq 2\exp(-2n) = \mathsf{negl}(n) \tag{25}$$

□

(c) Show that with overwhelming[1] probability:

$$|\mathsf{Estimate}(U) - \mathsf{Estimate}(U)| \leq \frac{2}{10n^c} \tag{26}$$

and explain why this does not trivially simplify to 0.

Correction. $\mathsf{Estimate}(U)$ is not a deterministic function, hence this does not trivially simplify to 0. Using the triangle inequality, we know that

$$|\mathsf{Estimate}(U) - \mathsf{Estimate}(U)| \tag{27}$$

$$= \left|\mathsf{Estimate}(U) - \Pr_{x\leftarrow U}[W(x)=1] + \Pr_{x\leftarrow U}[W(x)=1] - \mathsf{Estimate}(U)\right| \tag{28}$$

$$\leq 2\left|\mathsf{Estimate}(U) - \Pr_{x\leftarrow U}[W(x)=1]\right| \tag{29}$$

$$\leq \frac{2}{10n^c} \tag{30}$$

where the last inequality holds with overwhelming probability based on eq. (18). □

(d) Show that for any $\mathbf{s} \in S$, with overwhelming probability:

$$|\mathsf{Estimate}(A_{\mathbf{s},\chi}) - \mathsf{Estimate}(U)| \geq \frac{8}{10n^c} \tag{31}$$

[1] I.e. the probability of not having this equation true is negligible over the randomness involved in $\mathsf{Estimate}$.

Correction. For simplicity, we denote by $a := \mathsf{Estimate}(A_{\mathbf{s},\chi})$, $u := \mathsf{Estimate}(U_{\mathbb{Z}_q^n \times \mathbb{Z}_q})$, $a' := \Pr\left[\, W(x) = 1 \;\middle|\; x \leftarrow A_{\mathbf{s},\chi} \,\right]$ and $u' := \Pr\left[\, W(x) = 1 \;\middle|\; x \leftarrow U_{\mathbb{Z}_q^n \times \mathbb{Z}_q} \,\right]$. By assumptions and based on the previous question, we know that:

- since $\mathbf{s} \in S$, we have $|a' - b'| \geq \frac{1}{n^c}$
- with overwhelming probability, $|a - a'| \geq \frac{1}{10n^c}$
- with overwhelming probability, $|b - b'| \geq \frac{1}{10n^c}$

Since we want to lower bound rather than upper bound $|a - b|$, we cannot directly use the triangle inequality on $|a - b|$, so instead we apply it to $|a' - b'|$ in order to make $|a - b|$ appear on the other side, before reordering the term to isolate $|a - b|$. More precisely, the triangle inequality gives:

$$|a' - b'| = |a' - a + a - b + b - b'| \leq |a' - a| + |a - b| + |b - b'| \tag{32}$$

hence

$$|a - b| \geq |a' - b'| - |a' - b| - |b - b'| \geq \frac{1}{n^c} - \frac{1}{10n^c} - \frac{1}{10n^c} = \frac{8}{10n^c} \tag{33}$$

□

(e) We define now the algorithm $W'(D)$ where D is a distribution on $\mathbb{Z}_q^n \times \mathbb{Z}_q$ given as a black-box oracle to W', that will internally calls W and try to guess if D is the uniform distribution or a LWE distribution. More precisely, $W'(D)$ will repeat M times (M being a polynomial in n to be determined later) the following procedure: it will pick a random $\mathbf{t} \overset{\$}{\leftarrow} \mathbb{Z}_q$, compute $|\mathsf{Estimate}(f_{\mathbf{t}}(D)) - \mathsf{Estimate}(U)|$: if this value is greater than $\frac{1}{2n^c}$, it will output "LWE", otherwise it continues the loop. If at the end it has not returned before, it will output "$\mathsf{Uniform}$".

i. Show that W' is correct when the input is a uniform distribution and when $M \geq 1$, i.e. with overwhelming probability, $W'(U) =$ "$\mathsf{Uniform}$".

Correction. We have seen in question 2 that $f_{\mathbf{t}}(U) = U$, so $\mathsf{Estimate}(f_{\mathbf{t}}(U)) = \mathsf{Estimate}(U)$, and with question 3.b we know that with overwhelming probability, $|\mathsf{Estimate}(U) - \mathsf{Estimate}(U)| \leq \frac{2}{10n^c} < \frac{1}{2n^c}$, so after the first round (exists since $M \geq 1$) with overwhelming probability we return "$\mathsf{Uniform}$". □

ii. Let $\mathbf{s} \in \mathbb{Z}_q^n$. Show that the probability that $W'(A_{\mathbf{s},\chi})$ returns "LWE'' after 1 iteration of its inner loop is lower bounded by $\frac{8}{10n^c}$.

Correction. Let $\mathbf{s} \in \mathbb{Z}_q^n$. We first determine the probability of returning "LWE" for a given iteration of the loop of $W'(A_{\mathbf{s},\chi})$. First, we saw in question 1 that $f_{\mathbf{t}}(A_{\mathbf{s},\chi}) = A_{\mathbf{s}+\mathbf{f},\chi}$. Since we know that W is good to answer correctly when the distribution is some $A_{\mathbf{s}',\chi}$ with $\mathbf{s}' \in S$, we want to compute the probability of having $\mathbf{s}' := \mathbf{s} + \mathbf{f} \in S$:

$$\Pr[\, \mathbf{s}' \in S \,] = \Pr_{\mathbf{t} \overset{\$}{\leftarrow} \mathbb{Z}_q^{n+1}} [\, \mathbf{s} + \mathbf{t} \in S \,] \tag{34}$$

$$= \Pr_{\mathbf{t} \overset{\$}{\leftarrow} \mathbb{Z}_q^{n+1}} [\, \mathbf{t} \in \{ \mathbf{x} - \mathbf{s} \mid \mathbf{x} \in S \} \,] \tag{35}$$

$$= \frac{|\{ \mathbf{x} - \mathbf{s} \mid \mathbf{x} \in S \}|}{|\mathbb{Z}_q^{n+1}|} = \frac{|S|}{q^{n+1}} \geq \frac{1}{n^{c'}} \tag{36}$$

where the last inequality is from eq. (15). Now, we return "LWE" iff $|\mathsf{Estimate}(f_{\mathbf{t}}(A_{\mathbf{s},\chi})) -$

$\mathsf{Estimate}(U)| \geq \frac{1}{2n^c}$ (remember that by assumption the input $D = A_{\mathbf{s},\chi}$), so

$$\begin{aligned}
&\Pr\left[\text{Return "}\mathsf{LWE}\text{" (one iteration)}\right] && (37)\\
&= \Pr\left[|\mathsf{Estimate}(f_{\mathbf{t}}(D)) - \mathsf{Estimate}(U)| \geq \frac{1}{2n^c}\right] && \text{(Definition } W\text{)}\\
&= \Pr\left[|\mathsf{Estimate}(A_{\mathbf{s}',\chi}) - \mathsf{Estimate}(U)| \geq \frac{1}{2n^c}\right] && \text{(Question 1)}\\
&= \Pr[\mathbf{s}' \in S]\Pr\left[|\mathsf{Estimate}(A_{\mathbf{s}',\chi}) - \mathsf{Estimate}(U)| \geq \frac{1}{2n^c} \mid \mathbf{s}' \in S\right]\\
&\quad + \Pr[\mathbf{s}' \notin S]\Pr\left[|\mathsf{Estimate}(A_{\mathbf{s}',\chi}) - \mathsf{Estimate}(U)| \geq \frac{1}{2n^c} \mid \mathbf{s}' \notin S\right] && (38)\\
&\geq \Pr[\mathbf{s}' \in S]\Pr\left[|\mathsf{Estimate}(A_{\mathbf{s}',\chi}) - \mathsf{Estimate}(U)| \geq \frac{1}{2n^c} \mid \mathbf{s}' \in S\right]\\
& && \text{(Lower bound second term by 0)}\\
&\geq \frac{1}{n^{c'}} \times \frac{8}{10n^c} = \frac{8}{10n^{c+c'}} && \text{(See eq. (36) and eq. (31))}
\end{aligned}$$

Therefore, the probability of correctly returning after one iteration is $\frac{8}{10n^{c+c'}}$. After □

iii. Show that if we choose M large enough (but polynomial), W' is correct (with overwhelming probability) when the input is any LWE distribution, i.e. **for any s** $\in \mathbb{Z}_q$, with overwhelming probability on the randomness of W', $W'(A_{\mathbf{s},\chi}) =$ "LWE''. What value can we choose for M?

Hint: you may need to use the fact that $1 + x \leq e^x$.

Correction. We saw in the previous question that each iteration has probability $\frac{8}{10n^{c+c'}}$ of (correctly) returning "LWE". Therefore, W returns an incorrect value if none of the M iterations returns, i.e.:

$$\begin{aligned}
\Pr\left[W(A_{\mathbf{s},\chi}) \neq \text{"}\mathsf{LWE}''\right] &= \left(1 - \frac{8}{10n^{c+c'}}\right)^M && (39)\\
&\leq \left(e^{-\frac{8}{10n^{c+c'}}}\right)^M && (1 + x \leq e^x \text{ with } x = -\tfrac{8}{10n^{c+c'}})\\
&= e^{-\frac{8M}{10n^{c+c'}}} && (40)
\end{aligned}$$

If we choose for instance $M = n^{c+c'+1}$, we have

$$\Pr\left[W(A_{\mathbf{s},\chi}) \neq \text{"}\mathsf{LWE}''\right] \leq e^{-8n/10} = \mathsf{negl}(n) \tag{41}$$

i.e. the probability of W' outputting an incorrect value is negligible. □

iv. Conclude by showing that W' runs in polynomial time and that it therefore solves efficiently the LWE problem in the worst case by computing its advantage.

Correction. Trivial: each operation runs in polynomial time since W' is efficient, and the number of repetition M is polynomial, hence W' runs in polynomial time. We already shown that W' was correct (with overwhelming probability on the internal randomness of W') both when it has oracle access to the uniform distribution and to $A_{\mathbf{s},\chi}$ for any $\mathbf{s}$, hence its advantage is:

$$\begin{aligned}
&|\Pr\left[W'(A_{\mathbf{s},\chi}) = \text{"}\mathsf{LWE}''\right] - \Pr\left[W'(U) = \text{"}\mathsf{LWE}''\right]| && (42)\\
&= |(1 - \mathsf{negl}(n)) - (0 + \mathsf{negl}(n))| && (43)\\
&= 1 - \mathsf{negl}(n) && (44)
\end{aligned}$$

which is clearly not negligible since it is lower bounded for instance by the constant $1/2$. □