MAD 4471: Introduction to Cryptography and Coding Theory **Fall 2022**

Lecture 14: The Learning With Error (LWE) Problem

Lecturer: Jean-François Biasse *TA: William Youmans*

In this lecture, we introduce the Learning With Errors problem which is at the hear of the lattice-based cryptosystems that we present in this course.

14.1 Intuition: learning without errors

To understand the Learning With Error (LWE) problem (and why it is computationally difficult), we start with something easy: solving linear systems. The (easy) problem we propose to solve is parametrized by the following values:

- A modulus $q \in \mathbb{Z}_{>0}$.
- A dimension $n > 0$.
- A number of samples $m > 0$.

An instance of the problem is given by a m vectors $\mathbf{a}_i \in \mathbb{Z}_q^n$ (here we use the notation $\mathbb{Z}_q = \mathbb{Z}/q\mathbb{Z}$), and the values

$$b_1 = \langle \mathbf{s}, \mathbf{a}_1 \rangle, \dots, b_m = \langle \mathbf{s}, \mathbf{a}_m \rangle,$$

for a secret vector $\mathbf{s} \in \mathbb{Z}_q^n$. Can you find $\mathbf{s}$?

If enough samples $b_i = \langle \mathbf{s}, \mathbf{a}_i \rangle$ are given, the answer is yes. It is simple linear algebra. Let $A \in \mathbb{Z}_q^{m \times n}$ be the matrix whose m rows are the vectors $\mathbf{a}_i$, then the problem is to find $\mathbf{s} \in \mathbb{Z}_q^n$ such that

$$A\mathbf{s} = \mathbf{b} \quad \text{for } \mathbf{b} = (b_1, \dots, b_m) \in \mathbb{Z}_q^m.$$

To simplify, we saw that the number m of samples is "large enough" if there is a matrix $U \in \mathbb{Z}_q^{m \times m}$ such that $UA = \begin{pmatrix} I_n \\ (0) \end{pmatrix}$. In that case, the fact that $A\mathbf{s} = \mathbf{b}$ implies that

$$UA\mathbf{s} = \begin{pmatrix} s_1 & & \\ & \ddots & \\ & & s_n \\ & (0) & \end{pmatrix} = U\mathbf{b}$$

So we can read the values $s_1, \dots, s_n$ out of the first n entries of the vector $U\mathbf{b} \in \mathbb{Z}_q^m$. The construction of U from elementary row operations (swap two rows, multiply a row by a non-zero element, and add to one rwo a multiple of another one) is known as the *Gaussian elimination*.

Swap of rows The swap of rows of index i and j can be achieved with the following matrix multiplication:

$$\begin{pmatrix} (I_{k_1}) & & & & \\ & 0 & & 1 & \\ & & (I_{k_2}) & & \\ & 1 & & 0 & \\ & & & & (I_{k_3}) \end{pmatrix} \begin{pmatrix} a_{i,1} & a_{i,2} & \dots & a_{i,n} \\ a_{j,1} & a_{j,2} & \dots & a_{j,n} \end{pmatrix} = \begin{pmatrix} a_{j,1} & a_{j,2} & \dots & a_{j,n} \\ a_{i,1} & a_{i,2} & \dots & a_{i,n} \end{pmatrix},$$

where (I_k) denotes the k by k identity matrix. Note that the swap of two rows is always invertible (i.e. the matrix U defined above is always invertible).

Multiplication by λ The multiplication of the rown of index i by $\lambda \in \mathbb{Z}_q$ can be achieved with the following matrix multiplication:

$$\begin{pmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & 1 & & & & \\ & & & \lambda & & & \\ & & & & 1 & & \\ & & & & & \ddots & \\ & & & & & & 1 \end{pmatrix} \begin{pmatrix} a_1 & a_2 & \dots & a_n \end{pmatrix} = \begin{pmatrix} \lambda v_1 & \lambda v_2 & \dots & \lambda v_n \end{pmatrix},$$

where the λ factor in the U matrix is in the i-th row (the one we want to multiply by λ). Note that the multiplication by λ that is invertible (i.e. a class $[a] \in \mathbb{Z}_q$ such that $\gcd(a, q) = 1$) is always an invertible operation (i.e. the matrix U is invertible).

Adding a multiple of row to another one The replacement of row of index i by its addition by λ times the row of index j can be achieved with the following matrix multiplication:

$$\begin{pmatrix} (I_{k_1}) & & & & \\ & 1 & & \lambda & \\ & & (I_{k_2}) & & \\ & 0 & & 1 & \\ & & & & (I_{k_3}) \end{pmatrix} \begin{pmatrix} a_{i,1} & a_{i,2} & \dots & a_{i,n} \\ a_{j,1} & a_{j,2} & \dots & a_{j,n} \end{pmatrix} = \begin{pmatrix} a_{i,1} + \lambda a_{j,1} & a_{i,2} + \lambda a_{j,2} & \dots & a_{i,n} + \lambda a_{j,n} \\ a_{j,1} & a_{j,2} & \dots & a_{j,n} \end{pmatrix}.$$

Note that this operation is always invertible (even when λ is not invertible in $\mathbb{Z}_q$).

The Gaussian elimination process During the actual process of Gaussian elimination, we do not compute the matrice U defined above. It is clear however that if we had the sequence $U_1, \dots, U_k$ of matrices corresponding to each of the elementary operations applied to A, then the matrix performing them all in one go would be $U = U_k U_{k-1} \dots U_1$. But we are interested in $U\mathbf{b}$, not by U itself. Therefore, during the Gaussian elimination, we perform elementary operations of the kinds described above, leading to $= \begin{pmatrix} I_n \\ (0) \end{pmatrix}$. At each step of the way, we perform similar operations on $\mathbf{b}$, which ultimately leads to $U\mathbf{b}$. The next question is: in what order can be perform these operations to guarantee the result? At the beginning, we set the pivot index to $k = n$. Then for each k from n to 1, we perform the following treatment:

1. Identify an index $i \leq k$ such that $a_{i,k}$ is invertible.
2. Multiply row of index i by λ such that $\lambda \cdot a_{i,k} = 1$.

3. Swap row k and row i.
4. For all $j \neq i$, replace row j by $R_j - a_{j,k}R_k$ (where R_j denotes row j).

Sometimes, we perform these operations on the *augmented matrix* which is $(A|\mathbf{b})$. Below, we just keep track separately of UA and $U\mathbf{b}$ where U is the matrix that encodes all elementary operations up to the current step.

Example 1 *Let us illustrate this procedure for* $q = 3$, $\mathbf{b} = \begin{pmatrix}1\\1\\0\end{pmatrix}$ *and* $A = \begin{pmatrix}2&2&0\\0&1&1\\0&2&0\end{pmatrix}$. *The first step is the swap between rows* 2 *and* 3*:*

$$UA = \begin{pmatrix}2&2&0\\0&2&0\\0&1&1\end{pmatrix} \quad U\mathbf{b} = \begin{pmatrix}1\\0\\1\end{pmatrix}.$$

Now we multiply Row 2 by the inverse of 2 *(which is* 2*), thus getting*

$$UA = \begin{pmatrix}2&2&0\\0&1&0\\0&1&1\end{pmatrix} \quad U\mathbf{b} = \begin{pmatrix}1\\0\\1\end{pmatrix}.$$

Next, we perform the following operations: $R_1 \leftarrow R_1 - 2R_2$ *and* $R_3 \leftarrow R_3 - R_1$. *This yields*

$$UA = \begin{pmatrix}2&0&0\\0&1&0\\0&0&1\end{pmatrix} \quad U\mathbf{b} = \begin{pmatrix}-2\\0\\1\end{pmatrix} = \begin{pmatrix}1\\0\\1\end{pmatrix}.$$

The last operation is the multiplication of the first row by the inverse of 2 *(which is* 2*), which finally gives us*

$$UA = \begin{pmatrix}1&0&0\\0&1&0\\0&0&1\end{pmatrix} \quad U\mathbf{b} = \begin{pmatrix}2\\0\\1\end{pmatrix}.$$

This means that $\mathbf{s} = \begin{pmatrix}2\\0\\1\end{pmatrix}$ *is the solution to our problem. One can indeed verify that* $A\mathbf{s} = \mathbf{b}$.

So if solving $A\mathbf{s} = \mathbf{b}$ is easy when $m \geq n$, then what could constitute a hard problem? Reducing the number of available samples m would only create more possible solutions $\mathbf{s}$ without making the retrieval of one of them any harder (still with Gaussian elimination). Instead, the right tweak consists in only revealing approximations $b_i \approx \langle \mathbf{a}_i, \mathbf{s}\rangle$ instead of their exact values. Hence the instance of the problem is given by the m vectors $\mathbf{a}_i$, and $b_i = \langle \mathbf{a}_i, \mathbf{s}\rangle + e_i$ where e_i is a random value in $\mathbb{Z}_q$.

14.2 The rounded Gaussian distribution

In our progression towards the formal definition of the LWE problem, we need to specify the *probability distribution* of the error terms e_i defined above. Depending on the sources, the LWE problem may be defined with the e_i belonging to "some probability distribution", or this probability distribution may be explicitely defined. In the original work of Regev, the LWE problem is defined for the *rounded Gaussian distribution*. Not specifying the probability distribution might imply that we use the uniform distribution

over $\mathbb{Z}_q$ given by $\Pr(e_i = x) = \frac{1}{q}$. However, it would also leave the door open to the choice of a very bad probability distribution such as for example $\Pr(e_i = 0) = 1$, which would mean reverting back to the case of no error. Hence, it is better to specify a distribution. With certain distributions, reductions exist between the resolution of certain hard computational problems in lattices to the resolution of LWE. For other distributions, it is only conjectured that LWE is hard. In the following, we present the original probability distribution that was considered to instantiate the LWE problem.

To get to the rounded Gaussian distribution, we need to mention continuous probability distributions. This is clearly a detour because in the end, we are only trying to describe a probability distribution over the finite set $\mathbb{Z}_q$.

Definition 14.1 (Density function) *A function $f : \mathbb{R} \to \mathbb{R}_{\geq 0}$ such that $\int_{\mathbb{R}} f(x)dx = 1$ is called a probability density function. A random variable X is distributed according to f if*

$$Pr(X \geq x) = \int_x^{+\infty} f(x)dx.$$

The particular density function we are interested in is the Gaussian density function. It is a family of functions defined from $x \mapsto e^{-x^2}$. We need the following result to propaerly introduce them:

Proposition 14.2 (Gaussian integral) *The function $x \mapsto e^{-x^2}$ can be integrated over $\mathbb{R}$, and its definite integral satisfies*

$$\int_{\mathbb{R}} e^{-x^2} dx = \sqrt{\pi}.$$

Proof: We use the fact that

$$\left(\int_{\mathbb{R}} e^{-x^2} dx\right)^2 = \int_{\mathbb{R}} e^{-x^2} dx \int_{\mathbb{R}} e^{-y^2} dy = \int\int_{\mathbb{R}^2} e^{-(x^2+y^2)} dxdy.$$

Then we switch to polar coordinates. Integrating in $\mathbb{R}^2$ for x, y is the same as integrating for $(r, \theta) \in \mathbb{R}^+ \times [0, 2\pi]$:

$$\begin{aligned}\int\int_{\mathbb{R}^2} e^{-(x^2+y^2)} dxdy &= \int_0^{2\pi}\int_0^{\infty} e^{-r^2} rdrd\theta = 2\pi \int_0^{\infty} e^{-r^2} rdr = 2\pi \int_{-\infty}^{0} \frac{1}{2} e^s ds \\ &= \pi \int_{-\infty}^{0} e^s ds = \pi(e^0 - e^{-\infty}) = \pi.\end{aligned}$$

■

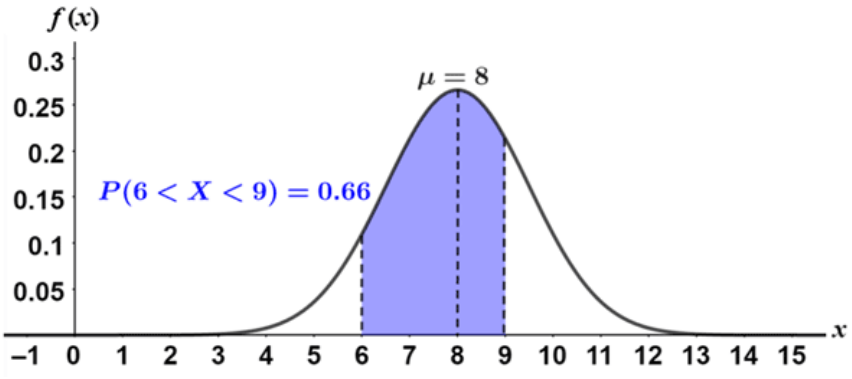
Definition 14.3 (Gaussian density function) *The Gaussian density function parametrized by μ, σ is given by*

$$g_{\mu,\sigma} : x \mapsto \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{-(x-\mu)^2}{2\sigma^2}}.$$

Proposition 14.4 *The function $g_{\mu,\sigma}$ describes the probability density function of a random variable with expected value μ and variance σ^2, i.e.*

- $\int_{\mathbb{R}} g_{\mu,\sigma}(x)dx = 1$.
- $\int_{\mathbb{R}} x g_{\mu,\sigma}(x)dx = \mu$.

Figure 14.1: Gaussian density function



- $\int_{\mathbb{R}}(x-\mu)^2 g_{\mu,\sigma}(x)dx = \sigma^2$.

To go from a continuous probability distribution over $\mathbb{R}$ to a discrete one over $\mathbb{Z}$, we can simply round. Let $X_{\mathbb{Z}}$ be the random variable with values in $\mathbb{Z}$ obtained by rounding X distributed according to $g_{\mu,\sigma}$ to the nearest integer. Then we have:

$$\Pr(X_{\mathbb{Z}} = N) = \Pr(X \in (N - \frac{1}{2}, N + \frac{1}{2}]) = \int_{N-\frac{1}{2}}^{N+\frac{1}{2}} g_{\mu,\sigma}(x)dx.$$

We are only interested in the case $\mu = 0$ (i.e. the Gaussian is centered around 0). This means that the most likely value of the e_i is 0. Now we are not done yet. The above gives a probability distribution over $\mathbb{Z}$, but our values are over $\mathbb{Z}_q$. Let $X_{\mathbb{Z}_q}$ be the random variable obtained by taking the congruence class of $X_{\mathbb{Z}}$ modulo q. We have

$$\Pr(X_{\mathbb{Z}_q} = [a]) = \Pr(X_{\mathbb{Z}} = a + kq \text{ for some } k \in \mathbb{Z}) = \sum_{k\in\mathbb{Z}} \Pr(X_{\mathbb{Z}} = a + kq) = \sum_{k\in\mathbb{Z}} \int_{a+kq-\frac{1}{2}}^{a+kq+\frac{1}{2}} g_{0,\sigma}(x)dx.$$

To "simplify" things, we might want to one integral from $a - 1/2$ to $a + 1/2$. For each k, we can perform the following change of variable $u = x - kq$:

$$\int_{a+kq-\frac{1}{2}}^{a+kq+\frac{1}{2}} \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{-x^2}{2\sigma^2}} dx = \int_{a-\frac{1}{2}}^{a+\frac{1}{2}} \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{-(u+kq)^2}{2\sigma^2}} du.$$

This immediately leads to the following description of the rounded Gaussian distribution:

Proposition 14.5 *Let X be the continuous random variable with density probability $g_{0,\sigma}$, and $X_{\mathbb{Z}}$ be the rounding of X to the nearest integer. The probability distribution of $X_{\mathbb{Z}_q}$ which is the congruence class of $X_{\mathbb{Z}}$ is given by*

$$Pr(X_{\mathbb{Z}_q} = [a]) = \int_{a-1/2}^{a+1/2} f_{\sigma,q}(x)dx,$$

where $f_{\sigma,q} : x \mapsto \sum_{k\in\mathbb{Z}} \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{-(x+kq)^2}{2\sigma^2}}$.

There is unfortunately no indefinite integral for the Gaussian function. This means that the terms that are the integral of a Gaussian between $a - 1/2$ and $a + 1/2$ need to be estimated by numerical methods. Nevertheless, the above gives us a way to compute the probability that $X_{\mathbb{Z}_q} = [a]$ for all congruence classes $[a] \in \mathbb{Z}_q$.

14.3 The definition of LWE

We call χ_σ the rounded gaussian distribution over $\mathbb{Z}_q$. When we write $e_i \leftarrow \chi_\sigma$, this means that "$e_i$ is drawn according to the distribution χ_σ. Otherwise stated, we draw e_i in $\mathbb{Z}_q$ such that

$$\Pr(e_i = [x]) = \int_{x-1/2}^{x+1/2} f_{\sigma,q}(x)dx.$$

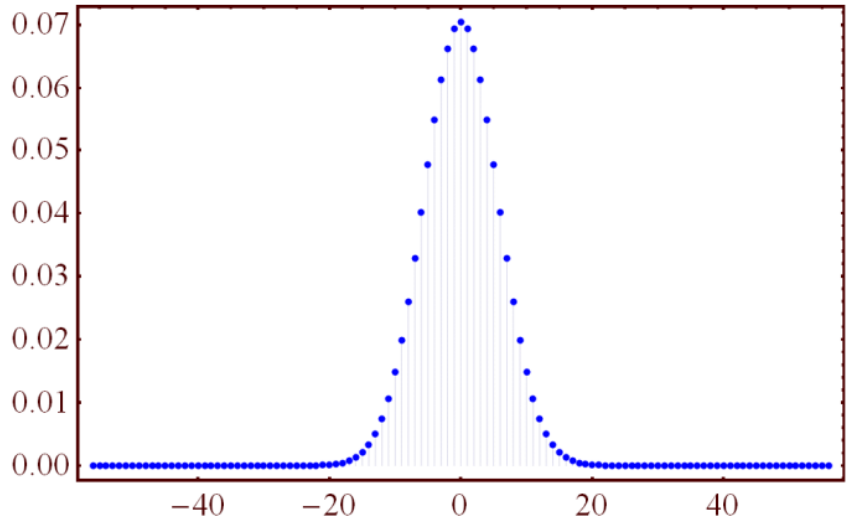
We also draw the vectors $\mathbf{a}_i$ at random, but with the uniform distribution.

Definition 14.6 (The LWE problem) *Let $q > 0$ be a modulus, $m, n > 0$ be integers, σ be a standard deviation, and χ_σ be the rounded Gaussian distribution over $\mathbb{Z}_q$ with standard deviation σ. Given m samples of the form*

$$\mathbf{a}_i, \langle \mathbf{a}_i, \mathbf{s} \rangle + e_i,$$

where $\mathbf{a}_i$ is distributed uniformly at random in $\mathbb{Z}_q^n$ and $e_i \leftarrow \chi$, find the secret $\mathbf{s} \in \mathbb{Z}_q^n$.

Figure 14.2: Error distribution for $q = 113$ and $\sigma = 0.05$ (Source: Regev 05)



Example 2 (Regev 05)

$$\begin{aligned}
14s_1 + 15s_2 + 5s_3 + 2s_4 &\approx 8 \bmod 17 \\
13s_1 + 14s_2 + 14s_3 + 6s_4 &\approx 16 \bmod 17 \\
6s_1 + 10s_2 + 13s_3 + 1s_4 &\approx 3 \bmod 17 \\
10s_1 + 4s_2 + 12s_3 + 16s_4 &\approx 12 \bmod 17 \\
9s_1 + 5s_2 + 9s_3 + 6s_4 &\approx 9 \bmod 17 \\
3s_1 + 6s_2 + 4s_3 + 5s_4 &\approx 16 \bmod 17 \\
&\vdots \\
6s_1 + 7s_2 + 16s_3 + 2s_4 &\approx 3 \bmod 17
\end{aligned}$$

In this case $\mathbf{s} = (0, 13, 9, 11)$.

14.4 A secret key encryption scheme

The standard for secret key encryption is the AES. We present this LWE secret key encryption only as a stepping stone to the full-fledged LWE public key encryption scheme.

Parameters The public parameters of the scheme are $n > 0$, $q > 0$, and the error distribution χ_σ. We require σ to be small enough with respect to q to ensure that $\Pr(|e_i| \leq q/4)$ is high. The secret key is a vector $\mathbf{s} \in \mathbb{Z}_q^n$.

Encryption We assume that we want to encrypt a message $\mu \in \{0, 1\}$ (i.e. a single bit). Longer messages have to be broken down into bits that are encrypted one by one. The steps to encryption are the following:

1. Draw $\mathbf{a} \in \mathbb{Z}_q^n$ uniformly at random.
2. Draw $e \leftarrow \chi_\sigma$.
3. Return $c = (\mathbf{a}, \langle \mathbf{a}, \mathbf{s} \rangle + e + \mu \cdot \lceil q/2 \rfloor$, where $\lceil x \rfloor$ is the rounding of $x \in \mathbb{R}$ to the nearest integer.

Decryption Assume we received $c = (\mathbf{a}, b)$. With the knowledge of $\mathbf{s}$, we can perform the operation $b - \langle \mathbf{a}, \mathbf{s} \rangle = e + \mu \cdot \lceil q/2 \rfloor$. Since we requested that $|e| \leq q/4$ with high probability, we know that

- If $\mu = 0$, $e + \mu \cdot \lceil q/2 \rfloor$ is smaller than $q/4$ with high probability.
- If $\mu = 1$, $e + \mu \cdot \lceil q/2 \rfloor$ is greater than $q/4$ with high probability.

Example 3 *For* $q = 5$*, and* $n = 3$*. Assume* $\mathbf{s} = (1, 0, 1)$*, and the message is* $\mu = 0$*.*

- ***Encryption:*** *we draw* $\mathbf{a} = (2, 4, 1)$ *and* $e = 1$*. The ciphertext is*

$$(\mathbf{a}, \langle \mathbf{a}, \mathbf{s} \rangle + e + \mu \cdot \lceil q/2 \rfloor = ((2, 4, 1), 2 \times 1 + 4 \times 0 + 1 \times 1 + 1 + 0 \times 2 \bmod 5) = ((2, 4, 1), 4)$$

- ***Decryption:*** *We receive* $c = ((2, 4, 1), 4)$*. We compute* $\langle \mathbf{a}, \mathbf{s} \rangle = 3$*, then compute* $|3 - 4| = 1$*, which is less than* $\lceil q/2 \rfloor = 2$*. Hence we conclude that* $\mu = 0$*.*

14.5 Towards public key encryption from LWE

To turn the previous scheme into a public key encryption scheme, we need to use the linearity of the ciphertexts: assume that $c_1 = (\mathbf{a}_1, \langle \mathbf{a}_1, \mathbf{s} \rangle + e_1)$ and $c_2 = (\mathbf{a}_2, \langle \mathbf{a}_2, \mathbf{s} \rangle + e_2)$ are two encryptions of $\mu = 0$, then

$$c_1 + c_2 = (\mathbf{a}_1 + \mathbf{a}_2, \langle \mathbf{a}_1, \mathbf{s} \rangle + \langle \mathbf{a}_2, \mathbf{s} \rangle + e_1 + e_2) = (\underbrace{\mathbf{a}_1 + \mathbf{a}_2}_{\mathbf{a}}, \langle \underbrace{\mathbf{a}_1 + \mathbf{a}_2}_{\mathbf{a}}, \mathbf{s} \rangle + \underbrace{e_1 + e_2}_{e}).$$

If e still satisfies that $|e| \leq q/4$ with high probability, then the decryption procedure of the secret key scheme will be correct with high probability. To turn an encryption of $\mu = 0$ into an encryption of 1, one simply add $\lceil q/2 \rfloor$ to the second coordinate.

We can therefore produce m encryptions of 0 by sampling vectors $\mathbf{a}_i \in \mathbb{Z}_q^n$ uniformly at random and error terms $e_i \langle \chi_\sigma$. Each sum of these encryption of 0 is another encryption of 0. Let A be the matrix whose rows are the $\mathbf{a}_i$, let $b_i = \langle \mathbf{a}_i, \mathbf{s} \rangle + e_i$, and $x \in \{0,1\}^m$ be the vector having a 1 in coordinate i if we want the i-th encryption of 0 to occur in the sum, then we can produce an encryption of 0 in the following way:

$$(\mathbf{x} \cdot A, \langle \mathbf{x}, \mathbf{b} \rangle) = (\mathbf{a}, \langle \mathbf{a}, \mathbf{s} \rangle + \underbrace{\langle \mathbf{x}, \mathbf{e} \rangle}_{e}).$$

If $|e| \leq q/4$, then the decryption procedure succeeds. Hence we require that σ be small enough that each e_i satisfies $|e_i| \leq q/4m$ with high probability. By triangle inequality, this guarantees that the sum of up to m terms e_i has absolute value less than $q/4$. With this encryption procedure, the public key is $A, \mathbf{b}$, and the private decryption key is $\mathbf{s}$ as in the symmetric key scheme.