# Gaussian LHL for Modules over Number Fields

Joint work with Martin R. Albrecht, Joël Felderhoff, Russell W. F. Lai, and Ivy K. Y. Woo

# Classic Leftover Hash Lemma (LHL)

$$(A, A \cdot x \mod q) \sim (A, \mathcal{U}(\mathcal{R}_q^n))$$

where

$$A \leftarrow \mathcal{U}(\mathcal{R}_q^{n \times m}), \quad x \leftarrow \chi$$

and  $\chi$  is usually a short distribution over  $\mathcal{R}^m$  or  $\mathcal{R}^m_q$ 

#### Gaussian LHL

$$(X, X \cdot \nu) \sim (X, \mathcal{D}_{\mathcal{R}^r, \sigma})$$

where

$$X \leftarrow (\mathcal{D}_{\mathcal{R}^r, \mathbf{S}_x})^m, \quad v \leftarrow \mathcal{D}_{\mathcal{R}^m, \mathbf{S}_v}$$

#### Notations

- Number field  ${\mathcal K}$  of degree d.
- Field discriminant  $\Delta_{\mathscr{H}}$ .
- The ring of integers  $\mathcal{R} = \mathcal{O}_{\mathcal{K}}$ .
- We assume the canonical embedding of  ${\mathscr R}$  has a basis  $||{f B}_{\mathscr R}||_\infty \le \delta_{\mathscr H}$ .
- We assume  $\mathcal{K}$  contains  $\mathbb{Q}(\zeta_f)$  for some  $f \geq 2$ .

#### Gaussian Linear Transform

$$(X, X \cdot v) \sim (X, \mathcal{D}_{X \cdot S_{v} \cdot S_{v}^{T} \cdot X^{T}})$$

where

$$X \leftarrow (\mathfrak{D}_{\mathbf{S}_{x}})^{m}, \quad v \leftarrow \mathfrak{D}_{\mathbf{S}_{v}}$$

#### Discrete Gaussian version

$$(\mathbf{X}, \mathbf{X} \cdot \mathbf{v}) \sim (\mathbf{X}, \mathcal{D}_{\mathbf{X} \cdot \mathcal{R}^m, \sqrt{\mathbf{X} \cdot \mathbf{S}_v \cdot \mathbf{S}_v^T \cdot \mathbf{X}^T}})$$

for 
$$\mathbf{X} \leftarrow (\mathcal{D}_{\mathcal{R}^r, \mathbf{S}_{\chi}})^m, \mathbf{v} \leftarrow \mathcal{D}_{\mathcal{R}^m, \mathbf{S}_{v}}$$

as long as 
$$S_v \ge \eta_{\varepsilon}(\Lambda^{\perp}(X))$$

#### Discrete Gaussian version

$$(\mathbf{X}, \mathbf{X} \cdot \mathbf{v}) \sim (\mathbf{X}, \mathcal{D}_{\mathbf{X} \cdot \mathcal{R}^m, \sqrt{\mathbf{X} \cdot \mathbf{S}_v \cdot \mathbf{S}_v^T \cdot \mathbf{X}^T}})$$

for 
$$\mathbf{X} \leftarrow (\mathcal{D}_{\mathcal{R}^r, \mathbf{S}_x})^m, \mathbf{v} \leftarrow \mathcal{D}_{\mathcal{R}^m, \mathbf{S}_v}$$

as long as 
$$S_v \ge \eta_{\varepsilon}(\Lambda^{\perp}(X))$$

we also consider 
$$\mathbf{X} \cdot \mathcal{R}^m = \mathcal{R}^r$$
, and  $\sqrt{\mathbf{X} \cdot \mathbf{S}_v \cdot \mathbf{S}_v^T \cdot \mathbf{X}^T} \approx \sigma \cdot I$ 

Extending [NP20] to the ring setting

• In fields Surjective = Full Rank = Non-Sigular Submatrix —> "Easy" to argue via entropy.

- In fields Surjective = Full Rank = Non-Sigular Submatrix —> "Easy" to argue via entropy.
- For  $\mathbf{X} \in \mathcal{R}^{r \times m}$  take a submatrix  $\mathbf{X}_0 \in \mathcal{R}^{r \times r}$ .
- We reduce mod every prime ideal p and prove non-zero determinant w.h.p.

- In fields Surjective = Full Rank = Non-Sigular Submatrix —> "Easy" to argue via entropy.
- For  $\mathbf{X} \in \mathcal{R}^{r \times m}$  take a submatrix  $\mathbf{X}_0 \in \mathcal{R}^{r \times r}$ .
- We reduce mod every prime ideal p and prove non-zero determinant w.h.p.
- Number of prime ideals in  $\mathcal{O}_{\mathcal{H}}$  = infinite.
- Only consider small prime ideals that could divide  $\det(X_0)$ .

- In fields Surjective = Full Rank = Non-Sigular Submatrix —> "Easy" to argue via entropy.
- For  $\mathbf{X} \in \mathcal{R}^{r \times m}$  take a submatrix  $\mathbf{X}_0 \in \mathcal{R}^{r \times r}$ .
- We reduce mod every prime ideal p and prove non-zero determinant w.h.p.
- Number of prime ideals in  $\mathcal{O}_{\mathcal{H}}$  = infinite.
- Only consider small prime ideals that could divide  $\det(X_0)$ .
- Bound the number for  $N(p) \le B$  by  $B/\log(B)$  (uses GRH).

#### The result

$$\Pr(\mathbf{X} \leftarrow (\mathcal{D}_{\mathcal{R}^r,\mathbf{S}_x})^m \text{ is surjective}) \geq 1 - 2^{-\lambda}$$

when 
$$m \ge 2r + \frac{\lambda}{\log(N_{\mathcal{R}})}$$
,  $N_{\mathcal{R}}$  - the norm of the smallest ideal

Lifting [AR16] to the ring setting

From the the kernel to the unit vector preimages

- Find short  $\mathbf{U} \in \mathcal{R}^{m \times r}$  such that  $\mathbf{X} \cdot \mathbf{U} = \mathbf{I}_r$
- Then for  $\mathbf{I}_m \mathbf{U} \cdot \mathbf{X}$  we have  $\mathbf{X} \cdot (\mathbf{I}_m \mathbf{U} \cdot \mathbf{X}) = \mathbf{X} \mathbf{I}_r \cdot \mathbf{X} = \mathbf{0}$

From the the kernel to the unit vector preimages

- Find short  $\mathbf{U} \in \mathcal{R}^{m \times r}$  such that  $\mathbf{X} \cdot \mathbf{U} = \mathbf{I}_r$
- Then for  $\mathbf{I}_m \mathbf{U} \cdot \mathbf{X}$  we have  $\mathbf{X} \cdot (\mathbf{I}_m \mathbf{U} \cdot \mathbf{X}) = \mathbf{X} \mathbf{I}_r \cdot \mathbf{X} = \mathbf{0}$

- Similarly if  $\mathbf{X} \cdot \mathbf{U} = f \cdot \mathbf{I}_r$
- Then for  $f \cdot \mathbf{I}_m \mathbf{U} \cdot \mathbf{X}$  (still short) we have  $\mathbf{X} \cdot (f \cdot \mathbf{I}_m \mathbf{U} \cdot \mathbf{X}) = \mathbf{X} \cdot f f \cdot \mathbf{I}_r \cdot \mathbf{X} = \mathbf{0}$

Integer case

For 
$$\mathbf{X} = [\mathbf{x}_1, ..., \mathbf{x}_m]$$
 with  $||\mathbf{x}_i||_{\infty} \leq B$  consider the set  $S = \sum_{i=1}^{s} \{0,1\} \cdot \mathbf{x}_i$ 

Integer case

For 
$$\mathbf{X} = [\mathbf{x}_1, ..., \mathbf{x}_m]$$
 with  $||\mathbf{x}_i||_{\infty} \leq B$  consider the set  $S = \sum_{i=1}^J \{0,1\} \cdot \mathbf{x}_i$ 

As j grows the norms in S grow slower than cardinality.

For  $\mathbf{v} \in S$ :  $||\mathbf{v}||_{\infty} \le j \cdot B$  this is a set of size  $(2jB+1)^r$  but  $|S|=2^j$ .

When 
$$2^j \ge (2jB+1)^r$$
 (or  $j \ge 2r\log(Br)$ ) we have a collision or  $\sum_{i=1}^J \{0, \pm 1\} \cdot \mathbf{x}_i = 0$ 

Sets with a shift

Consider the set 
$$S_j = \sum_{i=1}^{j} \{0, \pm 1\} \cdot \mathbf{x}_i$$

We define random variables:

- Win<sub>j</sub>: 
$$S_j \cap S_j + \mathbf{e}_1 \neq \emptyset$$

- Gain<sub>j</sub>: 
$$\mathbf{x}_{j+1} \notin S_j \land ||\mathbf{x}_{j+1}||_{\infty} \leq \sigma_x \sqrt{m} = B$$

Sets with a shift

Consider the set 
$$S_j = \sum_{i=1}^{j} \{0, \pm 1\} \cdot \mathbf{x}_i$$

We define random variables:

- Win<sub>j</sub>: 
$$S_j \cap S_j + \mathbf{e}_1 \neq \emptyset$$

- Gain<sub>j</sub>: 
$$\mathbf{x}_{j+1} \notin S_j \land ||\mathbf{x}_{j+1}||_{\infty} \leq \sigma_x \sqrt{m} = B$$

If 
$$\mathsf{Gain}_j = 1$$
 for  $j \in [1, 2r \log(Br)]$  then  $\sum_{j=1}^{2r \log(Br)} \{0, \pm 1\} \cdot \mathbf{x}_j = 0$  and  $\pm \mathbf{x}_{\mathsf{last}} = \sum \{0, \pm 1\} \cdot \mathbf{x}_j$ . Sasha Lapiha, Royal Holloway University of London

We Gain often enough

When  $\neg$  Win so  $S_j \cap S_j + \mathbf{e}_1 = \emptyset$  we have  $\rho_{\sigma_{\mathcal{X}}}(S_j) + \rho_{\sigma_{\mathcal{X}}}(S_j + \mathbf{e}_1) \leq \rho_{\sigma_{\mathcal{X}}}(\mathbb{Z}^m)$ 

And also  $\rho_{\sigma_{x}}(S_{j}) \approx_{\delta} \rho_{\sigma_{x}}(S_{j} + \mathbf{e}_{1})$ 

Hence 
$$\Pr(S_j) \le \frac{1}{2} + \delta$$
 and  $\Pr(\neg S_j) \ge \frac{1}{2} - \delta$ 

Issues with integers

For 
$$\mathbf{X} = [\mathbf{x}_1, ..., \mathbf{x}_m]$$
 with  $||\mathbf{x}_i||_{\infty} \leq B$  consider the set  $S = \sum_{i=1}^{J} \{0,1\} \cdot \mathbf{x}_i$ 

As j grows the norms in S grow slower than cardinality.

For  $\mathbf{v} \in S$ :  $||\mathbf{v}||_{\infty} \le j \cdot B$  this is a set of size  $(2jB+1)^r$  but  $|S|=2^j$ .

When 
$$2^j \ge (2jB+1)^r$$
 (or  $j \ge 2r\log(Br)$ ) we have a collision or  $\sum_{i=1}^J \{0, \pm 1\} \cdot \mathbf{x}_i = 0$ 

We change the set

Consider the set 
$$A = \mathbf{B}_{\mathcal{R}} \cdot \{0,1\}^d$$
 and  $S_j = \{\sum_{i=1}^J a_i \cdot \mathbf{x}_i \mid a_i \in A\}.$ 

Now 
$$|S_j| = 2^{dj}$$
 with norm  $\mathbf{v} \in S_j$ :  $||\mathbf{v}||_{\infty} \le jB \cdot d \cdot \delta_{\mathcal{K}}$ 

We change the set

Consider the set 
$$A = \mathbf{B}_{\mathcal{R}} \cdot \{0,1\}^d$$
 and  $S_j = \{\sum_{i=1}^J a_i \cdot \mathbf{x}_i \mid a_i \in A\}.$ 

Now  $|S_j| = 2^{dj}$  with norm  $\mathbf{v} \in S_j$ :  $||\mathbf{v}||_{\infty} \le jB \cdot d \cdot \delta_{\mathcal{K}}$ 

Pigeonhole gives  $dr \cdot \log(2jB \cdot d\delta_{\mathcal{K}} + 1) \leq dj \cdot \log(2)$  so  $m \approx r \log(rB \cdot d\delta_{\mathcal{K}})$ 

\*\* we actually take  $A = \mathbf{B}_{\mathcal{R}} \cdot \{0, \pm 1\}^d$ 

The proof doesn't work

Now for 
$$S_j = \sum_{i=1}^{j} A \cdot \mathbf{x}_i$$

We define random variables:

- Win<sub>j</sub>: 
$$S_j \cap S_j + \mathbf{e}_1 \neq \emptyset$$

- Gain<sub>j</sub>: 
$$\mathbf{x}_{j+1} \notin S_i \land ||\mathbf{x}_{j+1}||_{\infty} \leq \sigma_x \sqrt{m} = B$$

j=1

#### Change further

Consider two sets 
$$S_j = \sum_{i=1}^j A \cdot \mathbf{x}_i$$
 and  $\hat{S}_j = \{s/a \mid s \in S_j, a \in A \setminus \{0\}\}$ 

We define random variables:

- Win<sub>j</sub>: 
$$\hat{S}_j \cap \hat{S}_j + \mathbf{e}_1 \neq \emptyset$$

- 
$$\operatorname{Gain}_j: \mathbf{x}_{j+1} \notin \hat{S}_j \land ||\mathbf{x}_{j+1}||_{\infty} \leq \sigma_x \sqrt{m} = B \text{ implies } \mathbf{x}_{j+1} \notin S_j \subset \hat{S}_j$$

If 
$$\mathsf{Gain}_j = 1$$
 for  $j \in [1, \approx r \log(Br \cdot d\delta_{\mathcal{K}})]$  then 
$$\sum_{j=1}^{\approx r \log(Br d\delta_{\mathcal{K}})} a_j \cdot \mathbf{x}_j = 0 \text{ and } a_{\mathsf{last}} \cdot \mathbf{x}_{\mathsf{last}} = \sum a_j \cdot \mathbf{x}_j.$$

One more improvement

Consider two sets 
$$S_j = \sum_{i=1}^j A \cdot \mathbf{x}_i$$
 and  $\hat{S}_j = \{s/a \mid s \in S_j, a \in A \setminus \{0\}\}$ 

We define random variables:

- Win<sub>j</sub>: 
$$\hat{S}_j + \zeta^x \cdot \mathbf{e}_1 \cap \hat{S}_j + \zeta^y \cdot \mathbf{e}_1 \neq \emptyset$$
 now  $\Pr(\neg \hat{S}_j) \geq 1/f - \delta$ 

- 
$$\operatorname{Gain}_j: \mathbf{x}_{j+1} \notin \hat{S}_j \land ||\mathbf{x}_{j+1}||_{\infty} \leq \sigma_x \sqrt{m} = B \text{ implies } \mathbf{x}_{j+1} \notin S_j \subset \hat{S}_j$$

If 
$$\mathsf{Gain}_j = 1$$
 for  $j \in [1, \approx r \log(Br \cdot d\delta_{\mathcal{K}})]$  then 
$$\sum_{j=1}^{\approx r \log(Br d\delta_{\mathcal{K}})} a_j \cdot \mathbf{x}_j = 0 \text{ and } a_{\mathsf{last}} \cdot \mathbf{x}_{\mathsf{last}} = \sum a_j \cdot \mathbf{x}_j.$$

Win condition

$$\hat{S}_j + \zeta^x \cdot \mathbf{e}_1 \cap \hat{S}_j + \zeta^y \cdot \mathbf{e}_1 \neq \emptyset$$
 implies

$$s_1/a_1 - s_2/a_2 = (\zeta^y - \zeta^x) \cdot \mathbf{e}_1 = \zeta^y \cdot (1 - \zeta^{x-y}) \cdot \mathbf{e}_1$$

Win condition

$$\begin{split} \hat{S}_j + \zeta^x \cdot \mathbf{e}_1 \cap \hat{S}_j + \zeta^y \cdot \mathbf{e}_1 &\neq \varnothing \quad \text{implies} \\ s_1/a_1 - s_2/a_2 &= (\zeta^y - \zeta^x) \cdot \mathbf{e}_1 = \zeta^y \cdot (1 - \zeta^{x-y}) \cdot \mathbf{e}_1 \\ u \cdot (s_1 \cdot a_2 - s_2 \cdot a_1) &= a_1 a_2 \cdot f \cdot \mathbf{e}_1 \\ \text{where } u &= \zeta^{f-y} \cdot \frac{f}{1 - \zeta^{x-y}} \in \mathscr{R} \text{ is short} \end{split}$$

However for every  $\mathbf{e}_i$  the value  $a_1 a_2$  may be different.

We run it twice

Define  $B = \{b \text{ s.t. } ||b||_{\infty} \le R$ , coprime with  $a_1 a_2\}$ 

We prove there exists  $|B| \ge 2^d$  for  $R = O(\Delta_{\mathcal{K}}^{1/d} \cdot d^{2.5} \cdot \delta_{\mathcal{K}}^3)$ .

We get  $u' \cdot (s'_1 \cdot b_2 - s'_2 \cdot b_1) = b_1 b_2 \cdot f \cdot \mathbf{e}_1$  using small Bezout identity  $\alpha \cdot a_1 a_2 + \beta \cdot b_1 b_2 = 1$ 

We finally get  $\alpha \cdot u \cdot (s_1 \cdot a_2 - s_2 \cdot a_1) + \beta \cdot u' \cdot (s'_1 \cdot b_2 - s'_2 \cdot b_1) = f \cdot \mathbf{e}_1$ 

The result and open problems

$$\Pr\left(\eta_{\varepsilon}(\Lambda^{\perp}(\mathbf{X})) \geq O(m^{3/2} \cdot d^{12} \cdot f^2 \cdot \delta_{\mathcal{K}}^{14} \cdot \Delta_{\mathcal{K}}^{4/d} \cdot s_{\min}(\mathbf{S}_{x}))\right) \geq 1 - 2^{-\lambda}$$

when 
$$m \ge r \log(d \cdot r \cdot s_{\max}(\mathbf{S}_x))^{1+o(1)} + \frac{\lambda}{\log f}$$
.

The result and open problems

$$\Pr\left(\eta_{\varepsilon}(\Lambda^{\perp}(\mathbf{X})) \leq O(m^{3/2} \cdot d^{12} \cdot f^2 \cdot \delta_{\mathcal{K}}^{14} \cdot \Delta_{\mathcal{K}}^{4/d} \cdot s_{\min}(\mathbf{S}_{x}))\right) \geq 1 - 2^{-\lambda}$$

when 
$$m \ge r \log(d \cdot r \cdot s_{\max}(\mathbf{S}_x))^{1+o(1)} + \frac{\lambda}{\log f}$$
.

- Find a different set A that contains:
  - Short unit elements, with short inverses.
- Find a set A where all elements have many coprime short values.

First approach

Set  $\mathbf{S}_v$  as pseudo-inverse of  $\mathbf{X}$  scaled by  $\boldsymbol{\sigma}$ . Then  $\sqrt{\mathbf{X} \cdot \mathbf{S}_v \cdot \mathbf{S}_v^T \cdot \mathbf{X}^T} = \boldsymbol{\sigma} \cdot \boldsymbol{I}$ 

And 
$$\frac{\sigma}{s_{\max}(\mathbf{X})} \le s_{\min}(\mathbf{S}_v) \le s_{\max}(\mathbf{S}_v) \le \frac{\sigma}{s_{\min}(\mathbf{X})}$$

Second approach

- [AGHS13] conjectured that singular values of  $\sqrt{\mathbf{X}\cdot\mathbf{X}^T}$  are close for a large enough m
- We refine their bound and apply it to rings

Rotating a Real Gaussian Matrix [Sil85]

Assume  $\mathbf{Z} \leftarrow (\mathcal{D}_{\sigma=1})^{r\times m}$  then there exist orthogonal  $\mathbf{O}_i$  s.t.

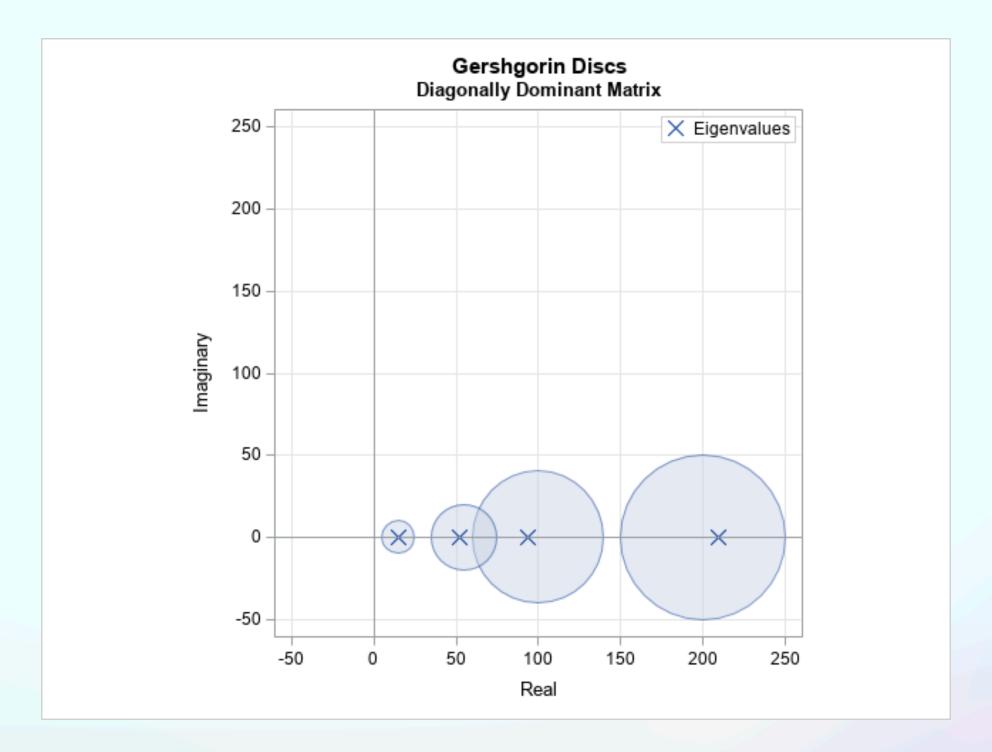
$$\mathbf{Z} = \begin{bmatrix} \mathbf{-} & \mathbf{z}_1^\mathsf{T} & \mathbf{-} \\ \mathbf{-} & \mathbf{z}_2^\mathsf{T} & \mathbf{-} \\ \vdots & \tilde{\mathbf{C}}_1 \end{bmatrix} \quad \mathbf{Z} \cdot \mathbf{O}_0 = \begin{bmatrix} X_m & 0 & \dots & 0 \\ \tilde{\mathbf{Z}}_0 & \end{bmatrix} \quad \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & \\ \vdots & \tilde{\mathbf{O}}_1 \end{bmatrix} \cdot \mathbf{Z} \cdot \mathbf{O}_0 = \begin{bmatrix} X_m & 0 & \dots & 0 \\ Y_{r-1} & \\ \vdots & \tilde{\mathbf{Z}}_1 \end{bmatrix}$$

where  $X_m$ ,  $Y_{r-1}$  are Chi random variables of corresponding dimension.

Gershgorin circle theorem

Then [Sil85] applies the Gershgorin circle theorem and Chi tail bounds

Gershgorin circle theorem



Then [Sil85] applies the Gershgorin circle theorem and Chi tail bounds

Moving to Discreet Gaussians

$$\Pr\left(1 - \frac{1}{rd} \le \lambda_{\min}/m \le \lambda_{\max}/m \le 1 + \frac{1}{rd}\right) \ge 1 - 4r \cdot \exp(-\lambda) \quad \text{for} \quad m \ge 49 \max(r, \lambda) \cdot (rd)^2$$

- We add continuos noise to embedding of X such that Z=X+Y is close to continuos
- Apply [Sil85] result
- We prove RD between this distribution and spherical Gaussian is a constant

# Application to k-SIS

Gaussian matrix with a trapdoor from [LPSS14].

```
\mathsf{HintTG}_{m,k}(\varsigma_1,\varsigma_2) \to (\mathbf{X} \in \mathcal{O}_{\mathcal{K}}^{k \times (m+k)}, \mathbf{U} \in \mathcal{O}_{\mathcal{K}}^{(m+k) \times (m+k)})
1: \mathbf{X}_1 \leftarrow \mathcal{D}^m_{\mathcal{O}^k_{\mathcal{K}}, \varsigma_1}
2: \mathbf{\Sigma} \coloneqq \sqrt{\mathbf{S} \cdot \mathbf{S}^{\mathsf{T}}} \text{ s.t. } \varsigma_2^2 \cdot \mathbf{I}_{dr} = \tilde{\Phi}(\mathbf{X}_1) \cdot \mathbf{S} \cdot \mathbf{S}^{\mathsf{T}} \cdot \tilde{\Phi}(\mathbf{X}_1)^{\mathsf{T}}
3: \forall : i \in [k] : \mathbf{r}_i \leftarrow \mathcal{D}_{\mathcal{O}_{\kappa}^m, \sqrt{\Sigma}}. \text{ Let } \mathbf{R} \coloneqq (\mathbf{r}_1, \dots, \mathbf{r}_k)
4: \mathbf{X}_2 \coloneqq \mathbf{X}_1 \cdot \mathbf{R} + \mathbf{I}_k \in \mathcal{O}_{\mathcal{K}}^{k \times k}
5: \mathbf{U} \coloneqq \begin{bmatrix} -\mathbf{R} & -\mathbf{I}_m - \mathbf{R}\mathbf{X}_1 \\ \mathbf{I}_k & \mathbf{X}_1 \end{bmatrix} \in \mathcal{O}_{\mathcal{K}}^{(m+k) \times (m+k)}
 6: return ((\mathbf{X}_1, \mathbf{X}_2), \mathbf{U})
```

# Thank you!

#### References

- [AGHS13] Shweta Agrawal, Craig Gentry, Shai Halevi, and Amit Sahai. Discrete Gaussian leftover hash lemma over infinite domains.
- [AR16] Divesh Aggarwal and Oded Regev. A note on discrete gaussian combinations of lattice vectors
- [LPSS14] San Ling, Duong Hieu Phan, Damien Stehle, and Ron Steinfeld. Hardness of k-LWE and applications in traitor tracing.
- [NP20] Hoi H. Nguyen and Elliot Paquette. Surjectivity of near-square random matrices.
- [Sil85] Jack W. Silverstein. The smallest eigenvalue of a large dimensional wishart matrix.