A Gaussian Leftover Hash Lemma for Modules over Number Fields

Martin R. Albrecht[1,2], Joël Felderhoff[1], Russell W. F. Lai[3], Oleksandra Lapiha[4], and Ivy K. Y. Woo[3]

[1] King's College London
{martin.albrecht, joel.felderhoff}@kcl.ac.uk
[2] SandboxAQ
martin.albrecht@sandboxaq.com
[3] Aalto University
{russell.lai,ivy.woo}@aalto.fi
[4] Royal Holloway, University of London
sasha.lapiha.2021@live.rhul.ac.uk

Abstract. Given a Gaussian matrix $\mathbf{X}$, a Gaussian Leftover Hash Lemma (LHL) states that $\mathbf{X} \cdot \mathbf{v}$ for a Gaussian $\mathbf{v}$ is an essentially independent Gaussian sample. It has seen numerous applications in cryptography for hiding sensitive distributions of $\mathbf{v}$. We generalise the Gaussian LHL initially stated over $\mathbb{Z}$ by Agrawal, Gentry, Halevi, and Sahai (2013) to modules over number fields. Our results have a sub-linear dependency on the degree of the number field and require only polynomial norm growth: $\|\mathbf{v}\|/\|\mathbf{X}\|$. To this end, we also prove when $\mathbf{X}$ is surjective (assuming the Generalised Riemann Hypothesis) and give bounds on the smoothing parameter of the kernel of $\mathbf{X}$. We also establish when the resulting distribution is independent of the geometry of $\mathbf{X}$ and establish the hardness of the k-SIS and k-LWE problems over modules (k-M-SIS/k-M-LWE) based on the hardness of SIS and LWE over modules (M-SIS/M-LWE) respectively, which was assumed without proof in prior works.

1 Introduction

The classic Leftover Hash Lemma (LHL) is an argument about randomness extraction. It states that if $\mathbb{H}$ is a distribution over a set of universal hash functions and X is a random variable with some guaranteed entropy then the distributions $(\mathbb{H}, \mathcal{U})$ and $(H \leftarrow \mathbb{H},\ H(X))$ are statistically close (where $\mathcal{U}$ is the uniform distribution). In lattice-based cryptography usage of the LHL appears in a context when the image $\mathbb{Z}^r$ (or $\mathcal{O}_\mathcal{K}^r$) of $H(\cdot)$ is taken modulo some q, for example in LWE or SIS based cryptography. In that case, the domain $\mathcal{O}_\mathcal{K}^r/q$ being finite allows the use of variants of the classical LHL, with some modifications based e.g. on how q splits in $\mathcal{O}_\mathcal{K}$. An overview of the LHL in the context of structured lattice based cryptography was given in [BL25]. A limitation of the classical LHL and its variants mentioned above is that it only applies when the image of $H(\cdot)$ is finite. Critically, those statements do not cover the case when $H(X)$ is close to a discrete Gaussian distribution over a lattice – an infinite domain.

In [AGHS13], Agrawal, Gentry, Halevi, and Sahai prove an LHL-type statement that holds over lattices rather than in finite groups. For $\mathbf{X} \in \mathbb{Z}^{r \times m}$ with columns sampled independently from $\mathcal{D}_{\mathbb{Z}^r,\varsigma}$ – a discrete Gaussian over $\mathbb{Z}^r$ – and $\mathbf{v}$ sampled from $\mathcal{D}_{\mathbb{Z}^m,\varsigma'}$, they upper bound the statistical distance between $(\mathbf{X}, \mathbf{X}\cdot\mathbf{v})$ and $(\mathbf{X}, \mathcal{D}_{\mathbb{Z}^r,\varsigma'\cdot\sqrt{\mathbf{X}\mathbf{X}^\intercal}})$ for any $m \geq r\ln(r\varsigma)$ (constant factors are omitted for this introduction). In follow-up works, Aggarwal and Regev [AR16] improve this result with an alternative proof. In 2020, Kirshanova, Nguyen, Stehlé and Wallet [KNSW20] further improve the result with a better lower bound $m \geq r\log(\varsigma)$. Hereafter, we refer to statements of this type as *Gaussian LHLs*.

As stated the output distribution of $\mathbf{X}\cdot\mathbf{v}$ depends on the matrix $\mathbf{X}$ and might be far from a spherical Gaussian. In order to have a spherical output, the solution is usually to sample $\mathbf{X}$ and then to sample $\mathbf{v}$ from the Gaussian distribution with parameter $\varsigma'\cdot\widetilde{\mathbf{X}}$ where $\widetilde{\mathbf{X}}$ is a pseudo-inverse of $\mathbf{X}$. However, in some applications this choice of the distribution of $\mathbf{v}$ is not available. The 'sphericity' of the distribution $\mathcal{D}_{\mathcal{O}_\mathcal{K}^r,\sqrt{\mathbf{X}\mathbf{X}^\intercal}}$ is controlled by the singular values of $\mathbf{X}$. In [AGHS13, Lemma 8], the authors argue that with high probability for a wide enough Gaussian matrix $\mathbf{X} \in \mathbb{Z}^{r\times m}$ of parameter ς, its singular values are within a constant factor of $\sqrt{m}\varsigma$. Together, this establishes when $\mathbf{X}\cdot\mathbf{v}$ follows a distribution close to independent from both of its inputs, enabling various applications discussed below.

A limitation of existing Gaussian LHLs is that they do not take advantage of the algebraic structure of the lattice, while most practical primitives based on lattices currently use module lattices: if $\mathcal{M}$ is a rank-r module over the ring of integers $\mathcal{O}_\mathcal{K}$ of a degree-d number field $\mathcal{K}$, it is a rank-$d\cdot r$ lattice when considered through, e.g. the canonical embedding of $\mathcal{K}$. In [KNSW20], the authors give a bound of $m \geq dr\ln(dr)$ independent vectors required for the Gaussian LHL result to work over $\mathcal{O}_\mathcal{K}^r$. However, a better bound for the number of independent vectors is to be expected, since any single vector $\mathbf{v} \in \mathcal{O}_\mathcal{K}^r$ gives $d-1$ other vectors (over $\mathbb{Z}$) 'for free' (think $X^i\cdot\mathbf{v}$ in the cyclotomic case). We would then expect a lower bound on m that grows linearly in r (the rank of the module) and sub-linear in d (the degree of the number field).

Contributions. In this work, we use the structure of $\mathcal{O}_\mathcal{K}$ to prove that only $m \geq r\cdot(\log(dr\varsigma))^{1+o(1)}$ vectors are necessary to get a Gaussian LHL in the context of module lattices. In §5, we prove that if $m \geq r\cdot(\log(dr\varsigma))^{1+o(1)}$, $\mathbf{X} \in \mathcal{O}_\mathcal{K}^{r\times m}$ is a Gaussian matrix and $\mathbf{v} \in \mathcal{O}_\mathcal{K}^m$ is a Gaussian vector with parameter $\sqrt{\mathbf{\Sigma}}$ then the distributions $(\mathbf{X}, \mathbf{X}\cdot\mathbf{v})$ and $(\mathbf{X}, \mathcal{D}_{\mathcal{O}_\mathcal{K}^r,\sqrt{\mathbf{X}\cdot\mathbf{\Sigma}\cdot\mathbf{X}^\intercal}})$ have negligible statistical distance. In particular, in §3 we establish when $\mathbf{X}$ generates $\mathcal{O}_\mathcal{K}^r$ and in §4 we upper bound $\eta_\varepsilon(\Lambda^\perp(\mathbf{X}))$. The composition of these two results establishes the claim, but these results may be of independent interest.

Moreover, in §6 we establish a bound when we may expect $\mathbf{X}\cdot\mathbf{v}$ to be close to spherical for any ring of integers of a number field: if $\mathbf{X} \in \mathcal{O}_\mathcal{K}^{r\times m}$ is a Gaussian matrix with $m \geq (rd)^3$ with parameter ς, then the Rényi divergence between the distribution $\mathcal{D}_{\mathcal{O}_\mathcal{K}^r,\sqrt{\mathbf{X}\mathbf{X}^\intercal}}$ and a spherical Gaussian of parameter $\approx\varsigma\sqrt{m}$ is $O(1)$.

This result permits comparing the distributions $(\mathbf{X}, \mathbf{X} \cdot \mathbf{v})$ and $(\mathbf{X}, \mathbf{v}')$, where $\mathbf{v}'$ is sampled from a spherical Gaussian depending only on r, m and ς.

Finally, as a first explicit application, in §7 we give a reduction from M-SIS to k-M-SIS that supports a linear number of hints and discuss how a similar reduction from M-LWE to k-M-LWE can be obtained.

1.1 Technical overview

Let $\mathbf{X} = [\mathbf{x}_1, \ldots, \mathbf{x}_m] \in \mathbb{Z}^{r \times m}$ be a matrix whose entries are sampled i.i.d. from a Gaussian distribution over $\mathbb{Z}$. As mentioned before, the Gaussian LHL for $\mathbb{Z}$-lattices [AGHS13,AR16,KNSW20] states that the distribution $\mathbf{X} \cdot \mathbf{v}$ when $\mathbf{v}$ is a Gaussian vector in $\mathbb{Z}^m$ is close to a Gaussian distribution, even given $\mathbf{X}$. A central step in the proof of this result is to give an upper bound for the smoothing parameter of the kernel lattice $\Lambda^\perp(\mathbf{X}) = \{\mathbf{x} \in \mathbb{Z}^m,\ \mathbf{X} \cdot \mathbf{v} = \mathbf{0}\}$ of the matrix $\mathbf{X}$. Two different techniques were introduced to do it. The authors of [AGHS13,KNSW20] consider the lattice $\Lambda_q(\mathbf{X}) \subset \mathbb{Z}^m$ generated by the rows of $\mathbf{X}$ and scaled standard unit vectors $q \cdot \mathbf{e}_i$ for a well-chosen prime q. Then they prove that with overwhelming probability, all short vectors of $\Lambda_q(\mathbf{X})$ are of the form $\mathbf{X}^\mathsf{T} \cdot \mathbf{v}$, and link the lattice minima of $\Lambda_q(\mathbf{X})$ to the shortest basis of $\Lambda^\perp(\mathbf{X})$. This is done using transference theorems in [AGHS13], and a more involved argument in [KNSW20]. The method used in [KNSW20] gives the best known parameters for Gaussian LHL over $\mathbb{Z}$ in terms of number of vectors and of ς'. However, we did not manage to leverage the algebraic structure of $\mathcal{O}_\mathcal{K}$ to improve it.

In [AR16], another approach was used. We describe it here in more detail, since we use it as our starting point. The authors first prove the existence of a short preimage $\mathbf{U} \in \mathbb{Z}^{m \times r}$ of $\mathbf{I}_r$ in $\mathbf{X} \cdot \mathbb{Z}^m$. This matrix is then used to build a short basis of $\Lambda^\perp(\mathbf{X})$, using the fact that if $\mathbf{X} \cdot \mathbf{U} = \mathbf{I}_r$ then all the columns of $\mathbf{U} \cdot \mathbf{X} - \mathbf{I}_m$ are in the kernel of $\mathbf{X}$. The existence of $\mathbf{U}$ is proven as follows. Let S be the set of formal $0, 1$ combinations of columns of $\mathbf{X}$: $S = \{\sum_{i=1}^m \{0,1\} \cdot \mathbf{x}_i\}$. The formal set S has size 2^m by construction, and the vectors inside it take values in the ball of radius $m\sqrt{r} \cdot \varsigma$ with overwhelming probability. By the pigeonhole principle, this set must contain a collision when $m \geq r \log(nm\varsigma)$ (again, constant factors are omitted for this introduction). The collision implies that some $\mathbf{x}_j$ must be a $0, \pm 1$ combination of columns of $\mathbf{X}$, and the authors then argue that this implies with very high probability that the set $S \cap (S + \mathbf{e}_1)$ is non-empty, leading to a preimage of $\mathbf{e}_1$ that has coefficients in $\{0, \pm 1, \pm 2\}$. The same argument is then repeated with $\mathbf{e}_j$ for $2 \leq j \leq n$, leading to the existence of $\mathbf{U}$ with coefficients in $\{0, \pm 1, \pm 2\}$.

We generalise this proof strategy. Assume that $\mathbf{X} = [\mathbf{x}_1, \ldots, \mathbf{x}_m]$ is a matrix in $\mathcal{O}_\mathcal{K}^{r \times m}$ whose coefficients are sampled i.i.d. from a discrete Gaussian distribution over $\mathcal{O}_\mathcal{K}$. We aim to prove that for a Gaussian vector $\mathbf{v} \in \mathcal{O}_\mathcal{K}^m$, the output distribution $\mathbf{X} \cdot \mathbf{v}$ is close to the Gaussian distribution of covariance parameter $\sqrt{\mathbf{X} \cdot \mathbf{X}^\mathsf{T}}$. Our proof, following the same blueprint as [AR16], requires two intermediary steps.

1. Proving that the matrix $\mathbf{X}$ is surjective in $\mathcal{O}_\mathcal{K}^r$ with high probability.

2. Proving an upper bound on the smoothing parameter of the kernel lattice $\Lambda^{\perp}(\mathbf{X})$ that holds with high probability.

Surjective Gaussian matrices. To our knowledge, a general result about the surjectivity of discrete Gaussian matrices over $\mathcal{O}_{\mathcal{K}}$ has not been established in prior work. Our strategy relies on methods developed within the theory of random integral matrices. In §3, we generalise the proof of [NP20], which focuses on matrices over $\mathbb{Z}$ and only gives an asymptotic estimate for the probability. The surjectivity proof over $\mathbb{Z}$ consists of looking at the random matrix modulo every prime number p and showing that with high probability $\mathbf{X}$ is non-singular modulo p. If $\mathbf{X}$ is non-singular modulo every prime p, it implies that the determinant of the lattice spanned by $\mathbf{X}$ is not divisible by any primes, and hence that $\mathbf{X}$ is surjective. In the number field case, we have to consider prime ideals instead of prime numbers: when a matrix $\mathbf{X} \in \mathcal{O}_{\mathcal{K}}^{r \times m}$ is surjective, in particular its reduction modulo any prime ideal $\mathfrak{p}$ of $\mathcal{O}_{\mathcal{K}}$ is surjective. We prove that the converse also holds: $\mathbf{X}$ is surjective over $\mathcal{O}_{\mathcal{K}}^{r}$ if and only if its reduction modulo any ideal $\mathfrak{p}$ is surjective in $\mathcal{O}_{\mathcal{K}}/\mathfrak{p}$. We then prove that a random $\widetilde{\mathbf{X}}$ is surjective with high probability if its coefficients are independent and sampled from a distribution with enough min-entropy. Finally, using a lemma from [JLWG25], we prove that this min-entropy condition holds for Gaussian matrices modulo ideals. Then since $\mathcal{O}_{\mathcal{K}}$ has an infinite number of prime ideals, we give an upper bound (depending on the Generalised Riemann Hypothesis) on the number of prime ideals that need to be considered to prove $\mathbf{X}$'s surjectivity. We do so by considering the first $r \times r$ submatrix of $\mathbf{X}$, proving that it has a non-zero determinant with high probability, and then that only the primes dividing this determinant need to be considered.

The probability that $\mathbf{X}$ is surjective highly depends on the number field considered (which was not an issue in the rational integer setting since the field is fixed to $\mathbb{Q}$). We prove that the probability that $\mathbf{X} \in (\mathcal{O}_{\mathcal{K}}/\mathfrak{p})^{r \times m}$ is surjective in $(\mathcal{O}_{\mathcal{K}}/\mathfrak{p})^{r}$ is $\geq 1 - r \cdot (1/\mathcal{N}(\mathfrak{p}))^{m-r+1}$ (Lemma 3.1). In particular, if $\mathcal{O}_{\mathcal{K}}$ has many ideals of small norm[5], then $\mathbf{X}$ has to 'be surjective in many different finite fields at the same time' in order to be surjective over $\mathcal{O}_{\mathcal{K}}^{r}$. The formula we obtain depends on the so-called Prime Zeta function of the number field $P_{\mathcal{K}}(s) = \sum_{\mathfrak{p} \subset \mathcal{O}_{\mathcal{K}}} \mathcal{N}(\mathfrak{p})^{-s}$ which 'counts' the small ideals of $\mathcal{O}_{\mathcal{K}}$. We also give a script estimating the growth of this function for cyclotomic fields in Appendix C.2. We present the overall result, both in the general form and in the case of prime-power cyclotomics in §3.5. Informally it states:

Theorem 1.1 (Assuming GRH, Informal). *If $\mathcal{K}$ is a number field of degree d and $N_{\mathcal{K}}$ is the norm of its smallest ideal then for any $r \geq 1$, $\varepsilon > 0$, ς large enough and $m \geq 2r + \log(1/\varepsilon)/\log(N_{\mathcal{K}})$, it holds that*

$$\Pr_{\mathbf{X} \leftarrow \mathcal{D}_{\mathcal{O}_{\mathcal{K}},\varsigma}^{r \times m}} \left(\mathbf{X} \textit{ is surjective} \right) \gtrsim 1 - \varepsilon.$$

[5] For example, one can construct a degree d multi-quadratic number field that has d different prime ideals of norm 2.

Short basis of $\Lambda^\perp(\mathbf{X})$ and smoothing parameter. The second step follows and generalises the proof of [AR16]. In that work, the authors first prove the existence of a short block-diagonal preimage of $\mathbf{I}_r$, and then use it to build a short basis of $\Lambda^\perp(\mathbf{X})$. To prove that a short preimage exists they define a formal set $S = \{\sum_{i=1}^{m}\{0,1\} \cdot \mathbf{x}_i\}$ (where the $\mathbf{x}_i$ are the columns of $\mathbf{X}$). This set has size 2^m, which must contain a collision when $m \geq r\log(r\varsigma)$ by the pigeonhole principle. This collision is then used to build an element in the intersection $S \cap (S + \mathbf{e}_i)$ which implies a short preimage of $\mathbf{e}_i$ for $1 \leq i \leq r$.

In order to take advantage of the structure of $\mathcal{O}_\mathcal{K}$, we now consider the set $S = \{\sum_{i=1}^{m} a_i \cdot \mathbf{x}_i\}$, where the a_i are small elements of a subset $A \subset \mathcal{O}_\mathcal{K}$ (taking $A = \{0,1\}$ recovers the proof of [AR16]). If the set A is chosen[6] to have size 2^d then the set S has size 2^{dr}, and by running a similar pigeonhole principle argument we obtain a collision for $m \geq r\log(dr\varsigma)$ (instead of $m \geq dr\log(dr\varsigma)$ when taking binary combinations).

However, changing the set A also changes the result. Instead of a preimage for $\mathbf{e}_i$ we only obtain a preimage for $a \cdot \mathbf{e}_i$ for $a \in \{a_1 \cdot a_2, a_i \in A\}$ (Lemma 4.5). Then, if this element $a \in \mathcal{O}_\mathcal{K}$ is not invertible, this does not result in an integral preimage of $\mathbf{e}_i$ we were looking for. In order to solve this problem, a solution could be to pick $A \subset \mathcal{O}_\mathcal{K}^\times$ such that all elements of A and their inverses have small ℓ_2 norm. We did not pursue this approach since the best known subset of $\mathcal{O}_\mathcal{K}^\times$ of size 2^d in the literature (to our knowledge) contains elements of norm $2^{O(\sqrt{d})}$ [CDPR16, Section 6.2]. The next option is to run the argument again to get a preimage for $a' \cdot \mathbf{e}_i$, with a new a'. Then prove that with high probability a' is coprime with a and run an effective version of Bezout's identity (Lemma 2.3) to get a preimage of $(au + a'v) \cdot \mathbf{e}_i = \mathbf{e}_i$. However, the worst-case nature of the pigeonhole principle argument forbids us from making this argument probabilistic. Thus, instead, we have chosen to construct the set $A' = \{a' \in \mathcal{O}_\mathcal{K}, a' \text{ coprime with } a,\ \|a'\| \leq R\}$ for some R, and to run the argument with this A'. Running the counting argument demands A' to be of exponential size. We prove in Lemma 4.11 that this is the case for large enough $R = \mathsf{poly}\left(d, \Delta_\mathcal{K}^{1/d}\right)$. We think that this coprime-counting lemma might have applications outside the scope of this work.

As an additional contribution, we improve the proof by considering the intersections $S + \zeta_f^j \cdot \mathbf{e}_i \cap S + \zeta_f^k \cdot \mathbf{e}_i$ when $\mathcal{K}$ is (or contains) the cyclotomic field of conductor f, allowing us to reduce the lower bound of m by a factor of $\log(f)$. This yields a matrix $\mathbf{U} \in \mathcal{O}_\mathcal{K}^{m \times r}$ whose norm is bounded by a polynomial in $d, \Delta_K^{1/d}$ and m such that $\mathbf{X} \cdot \mathbf{U} = f \cdot \mathbf{I}_r$. As in [AR16], we then argue that $f \cdot \mathbf{I}_m - \mathbf{U} \cdot \mathbf{X}$ is a "short" independent set of vectors of $\Lambda^\perp(\mathbf{X})$, which leads to a 'short' basis and a polynomial upper bound on the smoothing parameter of $\Lambda^\perp(\mathbf{X})$. This result is presented in Theorem 4.2, and can be summarised as follows:

Theorem 1.2 (Informal). *If $\mathcal{K}$ is a number field of degree d such that $\mathbb{Q}(\zeta_f) \subseteq \mathcal{K}$ then for any $r \geq 1$, $\varepsilon > 0$, ς large enough and $m \geq r \cdot (\log(dr\varsigma))^{1+o(1)} +$*

[6] The reader used to cyclotomic fields can think about $A = \left\{\sum_{i=1}^{d}\{0,1\} \cdot \zeta^i\right\}$.

$\log(1/\varepsilon)/\log(f)$, *there exists an absolute polynomial* P *such that*

$$\Pr_{\mathbf{X} \leftarrow \mathcal{D}^{r \times m}_{\mathcal{O}_\mathcal{K}, \varsigma}} \left(\eta_\varepsilon \left(\Lambda^\perp(\mathbf{X}) \right) \leq P(d, \Delta_K^{1/d}, m) \right) \ \geq \ 1 - \varepsilon.$$

We stress that the polynomial in the above theorem has a large degree in d and in $\delta_\mathcal{K}$ (the size of an $\mathcal{O}_\mathcal{K}$ basis as a $\mathbb{Z}$-module), leading to a large upper bound on the smoothing parameter of $\Lambda^\perp(\mathbf{X})$. We believe this massive size is a proof artefact coming from the worst-case nature of our existence argument, and we support this claim by computing small bases of random kernel lattices of various ranks for cyclotomic fields of various conductors in Appendix A. It should also be noted that since $\zeta_2 = -1$, the previous statement holds for $f = 2$ for any $\mathcal{K}$.

To summarise, in §5, we give an explicit statement of the Gaussian Leftover Hash Lemma for structured Gaussian matrices in Theorem 5.1.

Almost-spherical Gaussians. A limitation of the Gaussian LHL is that the distribution of $\mathbf{X} \cdot \mathbf{v}$, even if being close to a Gaussian, still depends on $\mathbf{X}$ through its centre and its covariance matrix, which is equal to $\mathbf{X}\mathbf{\Sigma}\mathbf{X}^\mathsf{T}$ if $\mathbf{v}$ is sampled from $\mathcal{D}_{\mathcal{O}_\mathcal{K}^m, \sqrt{\mathbf{\Sigma}}}$. In §6, we give two options for removing the dependency on the covariance (the centre can usually be set to $\mathbf{0}$). One option is to adapt the covariance of $\mathbf{v}$ and follows from previous work.

When the first option is not available, e.g. $\mathbf{v}$ is received from another party or algorithm, we can still extract an almost spherical Gaussian. To overcome this difficulty, we note that if $\mathbf{X}$ is of sufficient width, then it is already almost spherical with very high probability. This fact was already mentioned in [AGHS13, Lemma 8] for the $\mathbb{Z}$ setting, which we refine and generalise to the ring of integers setting. We first prove that the geometry of a discrete Gaussian matrix $\mathbf{X} \in \mathcal{O}_\mathcal{K}^{r \times m}$ is close to that of a continuous Gaussian matrix $\widetilde{\mathbf{X}} \leftarrow \mathcal{D}_\varsigma^{r \times m}$, which can be studied using tools from the field of random matrices. We give bounds on the singular values of $\widetilde{\mathbf{X}}$ following the proof of [Sil85], in which the author only studied their asymptotic behaviour. In Lemma 6.2, we prove that its singular values can be described as sums and products of independent χ^2 random variables. We then use concentration inequalities of the χ^2 distribution to show that these singular values are all close to $\sqrt{m} \cdot \varsigma$ with overwhelming probability. Finally, we show that this closeness implies that the Rényi divergence between the distribution of the covariance $\mathbf{X}\mathbf{\Sigma}\mathbf{X}^\mathsf{T}$ and $\sqrt{m} \cdot \varsigma$ is $O(1)$ with overwhelming probability. This result is presented in Theorem 6.4 and can be summarised as follows:

Theorem 1.3 (Informal). *Let* $\mathcal{K}$ *be a number field of degree* d, $\mathcal{O}_\mathcal{K}$ *its ring of integers. Let* $\mathbf{X} \in \mathcal{O}_\mathcal{K}^{r \times m}$ *be sampled from* $\mathcal{D}^m_{\mathcal{O}_\mathcal{K}^r, \varsigma}$ *with* $m \geq (rd)^3$. *Then*

$$\mathrm{RD}(\mathcal{D}_{\mathcal{O}_\mathcal{K}^r, \varsigma}; \mathcal{D}_{\mathcal{O}_\mathcal{K}^r, \sqrt{\mathbf{X} \cdot \mathbf{X}^\mathsf{T}}}) \leq 81.$$

1.2 Applications

Some direct applications of our Gaussian Leftover Hash Lemma over modules are reductions from M-SIS and M-LWE to k-M-SIS and k-M-LWE respectively that support $k = O(m)$ following the blueprint from [LPSS14].[7] In this work, we only give the reduction for M-SIS to k-M-SIS since the M-LWE to k-M-LWE reduction proceeds analogously. This allows us to lift constructions based on k-SIS and k-LWE to the module setting, such as linearly homomorphic signature schemes [BF11], traitor tracing [LPSS14] and partial GPV-style trapdoors [ALLW25]. Moreover, k-SIS may be considered as the oldest and most well-founded (due to the reduction from SIS) member of a family of "SIS with hints" assumptions that have been proposed recently. Indeed, works such as [ACL+22] prove the existence of hard instances of their newly introduced assumption by proving that these instances are as hard as k-M-SIS. However, in [ACL+22,ALLW25] the hardness of k-M-SIS is only assumed and not established.

Our generalisation of the Gaussian Leftover Hash Lemma to modules also enables the translation of various applications of [AGHS13,AR16] for rerandomising ciphertexts to achieve a form of confidentiality, such as the rerandomisation of GSW ciphertexts [BdPMW16] or the speculative suggestion in [ACL+22] to use similar techniques to make their construction zero-knowledge. Similarly, the special 1×2 case of a Gaussian LHL over rings has been established and used in [LSS14] to prove that a given distribution does not leak secret values. More generally, our result implies a generalisation of Gaussian convolution theorems such as [GMPW20, Theorem 3] to the module setting. Such tools have been used to give tighter reductions for SIS and LWE [MP13].

Moreover, the 1×2 case was generalised to more rings and used in [PS21] to establish that NTRU instances can be rerandomised. This in turn establishes a reduction to the decision NTRU problem. Our work is a building block to generalise such results to larger ranks $r > 1$. Similarly, in [SS13] the primitivity of the matrix of the special case 1×2 was considered to prove that NTRU signing keys can be efficiently sampled. Our work is a building block to generalising such reductions to a module setting.

Finally, in LWE-based encryption schemes, showing correctness often involves upper bounding the norm of the product $\mathbf{X} \cdot \mathbf{e}$ between a short matrix $\mathbf{X}$ from the decryption key and a short vector $\mathbf{e}$ from the encryption randomness. A naive upper bound is $\|\mathbf{X}\| \cdot \|\mathbf{e}\|$. This can be further upper bounded by $\sqrt{r} \cdot d \cdot m \cdot s^2$ if both $\mathbf{X}$ follow a discrete Gaussian distribution $\mathcal{D}^{r \times m}_{\mathcal{O}_\mathcal{K}, s}$ and $\mathbf{e}$ follows $\mathcal{D}^{m}_{\mathcal{O}_\mathcal{K}, s}$. A better approach is to consider the product for a fixed $\mathbf{X}$ with $\mathbf{X} \cdot \mathcal{O}^m_\mathcal{K} = \mathcal{O}^r_\mathcal{K}$ and $s \geq \eta_\epsilon(\Lambda^\perp(\mathbf{X}))$. In this case, $\mathbf{X} \cdot \mathbf{e}$ is statistically close to a sample from $\mathcal{D}^{r}_{\mathcal{O}_\mathcal{K}, s \cdot \sqrt{\mathbf{X} \cdot \mathbf{X}^\mathsf{T}}}$. Further upper bounding $\|\mathbf{X}\|$ and $\left\| \mathcal{D}_{\mathcal{O}^r_\mathcal{K}, \sqrt{\mathbf{X} \cdot \mathbf{X}^\mathsf{T}} \cdot s} \right\|$ then yields a tighter bound of $\sqrt{r \cdot d \cdot m} \cdot s^2$ at the cost of additional correctness error from the Gaussian LHL. In the literature, the second approach seems to have been applied in the setting where $\mathcal{O}_\mathcal{K} = \mathbb{Z}$ or when $r = 1$ but not in more general

[7] The reduction from SIS to k-SIS in [BF11] which supports $k = O(1)$ translates without our Gaussian LHL generalisation.

settings, probably due to the difficulty of upper bounding $\eta_\epsilon(\Lambda^\perp(\mathbf{X}))$, which we take the first steps to address in §4.

1.3 Open questions and directions

Even if our Gaussian LHL gives a polynomially bounded norm growth between the norm of the matrix $\mathbf{X}$ and the norm of the vector $\mathbf{v}$, the degree of the norm growth is large (see Theorem 5.1). This blowup comes from the fact that the upper bound on the smoothing parameter of $\Lambda^\perp(\mathbf{X})$ is very loose. We think that this is a proof artefact and provide experiments backing this claim in Appendix A for cyclotomic number fields.

The looseness of this bound comes from the fact that the proof of Lemma 4.12 is fundamentally "worst-case". A first way to improve it would be to use more probabilistic arguments instead of pigeonhole-like ones (see §4.3). Another blowup factor in the proof of Lemma 4.12 is that when constructing a set of size 2^d of elements coprime to some $a \in \mathcal{O}_\mathcal{K}$, Lemma 4.11 gives elements of size $\simeq \|a\|^2$. We think that restricting Lemma 4.11 to specific number fields (e.g. prime-power cyclotomics, where the prime ideal distribution is better known) will make our result more usable in practice. One last way of improving the bound of Lemma 4.12 would be to find a trade-off relation between the size of the smoothing parameter of $\Lambda^\perp(\mathbf{X})$ and the width of $\mathbf{X}$. Currently, our proof shows a threshold effect: if the width of $\mathbf{X}$ is less than a certain m_0, then the smoothing parameter is large, if it is above m_0, we have an upper bound on it which does not depend much on m. Finding more subtle relations between the two would improve the bound at the cost of increasing the width of $\mathbf{X}$.

2 Preliminaries

For $i, j \in \mathbb{Z}$ denote $[i, j] = \{i, \ldots, j-1\}$. We denote by $\mathbb{Z}$ the set of integers, $\mathbb{Q}$ the rationals, $\mathbb{R}$ the real numbers and $\mathbb{C}$ the complex numbers. For any positive function $f : \mathbb{R}^m \to \mathbb{R}$ and discrete set $S \subset \mathbb{R}^m$, we write $f(S) = \sum_{x \in S} f(x)$.

We use the convention that $\log = \log_2$ and $\ln = \log_e$. Vectors are denoted by bold-lower-case letters, such as $\mathbf{v}$. Matrices are denoted by bold-upper-case letters, such as $\mathbf{M}$. When $\mathbf{A}, \mathbf{B}$ are two matrices with compatible sizes, we let $[\mathbf{A} \mid \mathbf{B}]$ denote the horizontally stacked block matrix and $[\mathbf{A} \mid\mid \mathbf{B}] \coloneqq [\mathbf{A}^\mathbf{T} \mid \mathbf{B}^\mathbf{T}]^\mathbf{T}$ the vertically stacked one. When $\mathbf{A}$ is a matrix in $\mathbb{C}^{r \times m}$, we denote $\|\mathbf{A}\|$ the maximum of the ℓ_2 norm of its columns. In particular, for any $\mathbf{x} \in \mathbb{C}^m$, it holds that $\|\mathbf{A} \cdot \mathbf{x}\| \leq \sqrt{m} \cdot \|\mathbf{A}\| \cdot \|\mathbf{x}\|$.

2.1 Lattices, number theory and modules

A lattice $\Lambda \subset \mathbb{R}^m$ is a discrete additive subset of $\mathbb{R}^m$. It is generated by a (non-unique) basis $\mathbf{b}_1, \ldots, \mathbf{b}_r$ (often denoted in matrix form $\mathbf{B} = [\mathbf{b}_1, \ldots, \mathbf{b}_r] \in \mathbb{R}^{m \times r}$) of $\mathbb{R}$-linearly independent vectors. When $m = r$, the lattice is said to be full-rank. The volume of a lattice $\Lambda = \mathcal{L}(\mathbf{B})$ is $\mathrm{Vol}(\Lambda) \coloneqq \det(\mathbf{B}^\mathbf{T} \cdot \mathbf{B})$, and is independent of

the choice of basis $\mathbf{B}$. Let Λ be a lattice of rank r. We denote by $\lambda_i(\Lambda)$ its ith minima for $1 \leq i \leq r$: $\lambda_i(\Lambda) = \max(\min\{(\|\mathbf{v}_1\|, \ldots, \|\mathbf{v}_i\|), \mathbf{v}_j \in \Lambda, \mathrm{span}((\mathbf{v}_j)_j) = i\})$.

For a more in-depth introduction to number fields and number theory in general, we refer the reader to [Coh93,Neu13]. A number field is a finite-dimensional extension of $\mathbb{Q}$, its degree being the dimension of the extension. We let $\mathcal{O}_\mathcal{K}$ denote its ring of integers and $\Delta_\mathcal{K}$ its discriminant. A fractional ideal (we often omit saying "fractional" in this work) is a discrete (for the natural topology of $\mathbb{Q}$-vector space of $\mathcal{K}$) non-trivial subgroup of $\mathcal{K}$ that is stable under multiplication by any element of $\mathcal{O}_\mathcal{K}$. We say that an ideal is integral if it is included in $\mathcal{O}_\mathcal{K}$. The norm of an integral ideal $\mathfrak{a}$ is $\mathcal{N}(\mathfrak{a}) = |\mathcal{O}_\mathcal{K}/\mathfrak{a}|$. The norm of a fractional ideal I is $\mathcal{N}(I) = \mathcal{N}(N \cdot I)/N$, where N is an integer such that $N \cdot I$ is integral (such an integer always exists, and it can be shown that the norm does not depend on the choice of N). In this work, we denote integral ideals by gothic letters ($\mathfrak{a}, \mathfrak{p} \ldots$) and general ideals by uppercase letters ($I, J, \ldots$). The set of ideals has a group structure, where the product of two ideals I, J is the ideal $I \cdot J = \langle \{a \cdot b, a \in I, b \in J\} \rangle$, and the inverse of an ideal I is the ideal $I^{-1} = \{x \in \mathcal{K}, x \cdot I \subseteq \mathcal{O}_\mathcal{K}\}$. It is known that $\mathcal{O}_\mathcal{K}$ is a Dedekind domain, so its ideals have a unique prime factorisation property. In $\mathcal{O}_\mathcal{K}$ an ideal $\mathfrak{p} \subset \mathcal{O}_\mathcal{K}$ is prime iff $\mathcal{O}_\mathcal{K}/\mathfrak{p}$ is a field. Hence, for every ideal $I \subset \mathcal{K}$, there exist $\mathfrak{p}_1, \ldots, \mathfrak{p}_k$ prime ideals and $e_1, \ldots, e_k \in \mathbb{Z}$ such that $I = \prod_{i=1}^{k} \mathfrak{p}_i^{e_i}$. Up to reordering, the sequences $(\mathfrak{p}_i)_i$ and $(e_i)_i$ are unique. Some of our results rely on the Generalised Riemann Hypothesis (GRH) [BS96, Conjecture 8.7.3], which implies in particular that for any number field $\mathcal{K}$ and any $\varepsilon > 0$, the prime counting function $\pi_\mathcal{K}(x) = |\{\mathfrak{p} \text{ prime ideal}, \mathcal{N}(\mathfrak{p}) \leq x\}|$ satisfies,

$$\pi_\mathcal{K}(x) = \int_2^x 1/\ln(t)dt + O_\mathcal{K}(x^{1/2+\varepsilon}).$$

A number field $\mathcal{K}$ of degree d can be embedded in $\mathbb{C}$ via d different maps, named its canonical embedding. We let $d_\mathbb{R}$ be the number of those embeddings that only take real values, and $d_\mathbb{C}$ be the number of pairs of complex-valued embeddings. It holds that $d = d_\mathbb{R} + 2d_\mathbb{C}$. Those maps define an injective map from $\mathcal{K}$ to its completion $\mathcal{K}_\mathbb{R} \coloneqq \mathcal{K} \otimes \mathbb{R}$ called the Minkowski embedding $\Phi : \mathcal{K} \to \mathcal{K}_\mathbb{R}$. The set $\mathcal{K}_\mathbb{R}$ is isomorphic as a Hermitian space to the set $\left\{\mathbf{x} \in \mathbb{R}^{d_\mathbb{R}} \times \mathbb{C}^{2d_\mathbb{C}} \mid x_{d_\mathbb{R}+i} = \overline{x_{d_\mathbb{R}+d_\mathbb{C}+i}}\right\}$. This set is in turn isomorphic to $\mathbb{R}^d$ as an inner-product space. The Minkowski embedding allows us to endow $\mathcal{K}$ with a geometric structure: for any $x \in \mathcal{K}$, we denote by $\|x\|$ (respectively $\|x\|_\infty$) the ℓ_2 (respectively ℓ_∞) norm of its Minkowski embedding (in $\mathbb{C}^d$). Note that in particular, for any $a, b \in \mathcal{K}$, it holds that $\|a \cdot b\| \leq \|a\|_\infty \cdot \|b\|$. In this geometric setting, any ideal I of $\mathcal{K}$ (including its ring of integers) is a lattice in $\mathcal{K}_\mathbb{R}$ of volume $\mathcal{N}(I) \cdot \sqrt{\Delta_\mathcal{K}}$.

Throughout this work, we assume that we know a basis $\mathbf{B}^{\mathcal{O}_\mathcal{K}} \in \mathbb{R}^{d \times d}$ of $\Phi(\mathcal{O}_\mathcal{K})$ as a $\mathbb{Z}$-module. We denote by $\delta_\mathcal{K}$ the maximal ℓ_∞ norm of its columns (in particular $\delta_\mathcal{K} = 1$ for cyclotomic fields with the power basis). The results of this work are existential instead of constructive, so we use the 'best' basis of $\mathcal{O}_\mathcal{K}$. According to [MG02, Corollary 7.2] there always exists a basis of $\mathcal{O}_\mathcal{K}$ such that $1 \leq \delta_\mathcal{K} \leq \sqrt{d} \cdot \lambda_d(\mathcal{O}_\mathcal{K})/2$.

An $\mathcal{O}_\mathcal{K}$-module $M \subset \mathcal{K}^m$ is a discrete subset of $\mathcal{K}^m$ that is additive and stable under multiplication by any element of $\mathcal{O}_\mathcal{K}$ (fractional ideals of $\mathcal{O}_\mathcal{K}$ are rank-1

modules). Any $\mathcal{O}_\mathcal{K}$-module M can be represented as a (non-unique) pseudo basis $((I_i, \mathbf{b}_i))_{i=1,\dots,r}$, where I_i are ideals of $\mathcal{K}$ and $\mathbf{b}_i \in \mathcal{K}^m$ are $\mathcal{K}$-linearly independent. When $r = m$, we say that M is full-rank. By the Minkowski embedding, every $\mathcal{O}_\mathcal{K}$-module of rank r in $\mathcal{K}^m$ can be seen as a lattice of rank $d \cdot r$ in $\mathbb{R}^{d \cdot m}$.

Let $M \subset \mathcal{K}^r$ be a full-rank module given by a pseudo-basis $((\mathbf{b}_i, I_i))_{1 \leq i \leq r}$. Its determinantal ideal is the fractional ideal $\det(M) \coloneqq \det_\mathcal{K}(\mathbf{B}) \cdot \prod_i I_i$, and it is independent of the choice of the pseudo-basis. For any submodule $M' \subseteq M$, it holds that $\det(M') | \det(M)$ with equality if and only if $M = M'$.

Lemma 2.1 (Adapted from [MG02, Lemma 7.1]). *Let $M \subset \mathcal{K}^r$ be a full-rank module, and $\mathbf{v_1}, \dots, \mathbf{v_r} \in M$ be a $\mathcal{K}$-free set of vectors. Then there exists a matrix $\mathbf{B} \in \mathbb{R}^{dr \times dr}$ that generates $\Phi(M)$ as a $\mathbb{Z}$-lattice satisfying*

$$\|\mathbf{B}\| \leq \sqrt{rd} \cdot \delta_\mathcal{K} \cdot \max \|\mathbf{v}_i\|.$$

Lemma 2.2. *Let $M \subset \mathcal{K}^r$ be a full-rank module with basis $\mathbf{b}_1, \dots, \mathbf{b}_r \in M$, let $\mathbf{B} = [\mathbf{b}_1, \dots, \mathbf{b}_r]$ and $\mathfrak{p}$ a prime ideal of $\mathcal{O}_\mathcal{K}$ with residue field $F = \mathcal{O}_\mathcal{K}/\mathfrak{p}$. Then $\det_\mathcal{K}(\mathbf{B}) \in \det(M)$, and furthermore it holds that $\det_F(\mathbf{B} \bmod \mathfrak{p}) = \det_\mathcal{K}(\mathbf{B}) \bmod (\det(M) \cdot \mathfrak{p})$.*

Proof. That $\det_\mathcal{K}(\mathbf{B}) \in \det(M)$ comes from the fact that $\mathbf{B} \cdot \mathcal{O}_\mathcal{K}^r \subseteq M$ and that $\det_\mathcal{K}(\mathbf{B}) \cdot \mathcal{O}_\mathcal{K} = \det(\mathbf{B} \cdot \mathcal{O}_\mathcal{K}^r)$. This implies that $\det(\mathbf{B}) \bmod \det(M) \cdot \mathfrak{p}$ is well-defined. Now let us define the function $\phi(\mathbf{b}_1, \dots, \mathbf{b}_r) \coloneqq \det_\mathcal{K}(\mathbf{B}) \bmod \det(M) \cdot \mathfrak{p}$. This function is a multilinear map satisfying that for any $\mathbf{m} \in M$ and $p \in \mathfrak{p}$, $\phi(\mathbf{b}_1 + p \cdot \mathbf{m}, b_2, \dots, b_r) = \phi(\mathbf{B}) + p \cdot \det_\mathcal{K}(\mathbf{m}, \mathbf{b}_2, \dots, \mathbf{b}_r) = \phi(\mathbf{B}) \bmod \mathfrak{p} \det(M)$, so it can be lifted as a multilinear map $\phi' : (M/\mathfrak{p}M)^r \to F$. The multilinear map ϕ' is alternating and clearly satisfies $\phi'(\mathbf{I}_r) = 1$. It is thus equal to the determinant over $(M/\mathfrak{p}M)^r$. □

Lemma 2.3. *Let $a, b \in \mathcal{O}_\mathcal{K}$ be two coprime elements. Then there exists $u, v \in \mathcal{O}_\mathcal{K}$ with $\|u\|, \|v\| \leq \max(\|a\|_\infty, \|b\|_\infty) \cdot \sqrt{d} \cdot \lambda_d(\mathcal{O}_\mathcal{K})$ such that $a \cdot u + b \cdot v = 1$.*

Proof. Without loss of generality, assume that $\|a\|_\infty \leq \|b\|_\infty$. Let $b_1, \dots, b_d$ be a $\mathbb{Z}$-basis of $\mathcal{O}_\mathcal{K}$ satisfying $\|b_i\| \leq \max(1, \sqrt{i}/2) \cdot \lambda_i(\mathcal{O}_\mathcal{K})$. Let M be the rank-1 module in $\mathcal{K}^2$ generated by $(b_i \cdot (-b, a)^T)_{1 \leq i \leq d}$. The norm of the elements of its basis are bounded by $\lambda_d(\mathcal{O}_\mathcal{K}) \cdot \|a\|_\infty$, hence the covering radius of M is at most $\sqrt{d} \cdot \lambda_d(\mathcal{O}_\mathcal{K}) \cdot \|a\|_\infty$. Let $u_0, v_0 \in \mathcal{O}_\mathcal{K}$ such that $a \cdot u_0 + b \cdot v_0 = 1$ and let (u, v) be the closest element of M to (u_0, v_0). It holds that $a \cdot (u_0 - u) + b \cdot (v_0 - v) = 1$ and that $\|(u_0 - u, v_0 - v)\| \leq \sqrt{d} \cdot \lambda_d(\mathcal{O}_\mathcal{K}) \cdot \|a\|_\infty$, hence the result. □

When $\mathbf{X} \in \mathcal{K}^{r \times m}$ is a matrix, we define its kernel module $\Lambda^\perp(\mathbf{X}) \coloneqq \{\mathbf{x} \in \mathcal{O}_\mathcal{K}^m, \mathbf{X} \cdot \mathbf{x} = \mathbf{0}\}$. When $\mathbf{X}$ has $\mathcal{K}$-rank l, it is a $\mathcal{O}_\mathcal{K}$-module of rank $m - l$.

2.2 Matrices and Singular Value Decomposition

Let $\mathbf{M} \in \mathbb{R}^{n \times n}$ be a symmetric matrix. We say that $\mathbf{M}$ is positive definite if for any $\mathbf{x} \in \mathbb{R}^n \setminus \{\mathbf{0}\}$, it holds that $\mathbf{x}^\intercal \cdot \mathbf{M} \cdot \mathbf{x} > 0$. For any positive definite matrix

$\mathbf{M}$, there exists a symmetric matrix $\mathbf{S}$ such that $\mathbf{M} = \mathbf{S} \cdot \mathbf{S}^{\mathsf{T}}$. We call this matrix a square root of $\mathbf{M}$, it is not unique.

Let $\mathbf{\Sigma}_1, \mathbf{\Sigma}_2$ be two positive definite matrices. We say that $\mathbf{\Sigma}_1 \leq \mathbf{\Sigma}_2$ if $\mathbf{\Sigma}_2 - \mathbf{\Sigma}_1$ is positive semi-definite. Let $\varsigma > 0$, we abuse the notation and write $\mathbf{\Sigma} \geq \varsigma^2$ for $\mathbf{\Sigma} \geq \varsigma^2 \mathbf{I}_{dr}$. For any two full-rank matrices $\mathbf{S}, \mathbf{T}$, we write $\mathbf{S} \leq \mathbf{T}$ if $\mathbf{S} \cdot \mathbf{S}^{\mathsf{T}} \leq \mathbf{T} \cdot \mathbf{T}^{\mathsf{T}}$.

For any matrix $\mathbf{M} \in \mathbb{R}^{n \times m}$ of rank r, there exist real values $0 < s_r(\mathbf{M}) \leq \ldots \leq s_1(\mathbf{M})$, and two orthogonal matrices $\mathbf{U}_{\mathbf{M}} \in \mathbb{R}^{n \times n}, \mathbf{V}_{\mathbf{M}} \in \mathbb{R}^{m \times m}$ such that

$$\mathbf{M} = \mathbf{U}_{\mathbf{M}} \cdot \begin{pmatrix} \mathrm{diag}(\{s_i(\mathbf{M})\}_{i=1}^{r}) & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix} \cdot \mathbf{V}_{\mathbf{M}}.$$

The $s_i(\mathbf{M})$ are called the singular values of $\mathbf{M} \in \mathbb{R}^{n \times m}$. We denote $s_{\max}(\mathbf{M}) \coloneqq s_1(\mathbf{M})$ and $s_{\min}(\mathbf{M}) \coloneqq s_r(\mathbf{M})$. For any $\mathbf{x} \in \mathbb{R}^m$, it holds that $\|\mathbf{M} \cdot \mathbf{x}\| \leq s_{\max}(\mathbf{M}) \cdot \|\mathbf{x}\|$. If $\mathbf{M}$ is square and full-rank, it also holds that $s_{\min}(\mathbf{M}) \cdot \|\mathbf{x}\| \leq \|\mathbf{M} \cdot \mathbf{x}\|$.

Lemma 2.4 (Corollary of Weyl's inequality.). *Let $\mathbf{X}, \mathbf{Y} \in \mathbb{R}^{n \times m}$ with $0 < n \leq m$. If $s_{\max}(\mathbf{Y}) \leq \delta \cdot s_{\min}(\mathbf{X})$ for some $\delta < 1$ then*

$$s_{\max}(\mathbf{X} \pm \mathbf{Y}) \in [1 - \delta, 1 + \delta] \cdot s_{\max}(\mathbf{X}) \ ,$$
$$s_{\min}(\mathbf{X} \pm \mathbf{Y}) \in [1 - \delta, 1 + \delta] \cdot s_{\min}(\mathbf{X}) \ .$$

Lemma 2.5 ([LPSS14, Adapted Lemma 16 (eprint)]). *Let $1 \leq n \leq m$ be integers and $\varsigma > 0$. Let $\mathbf{M} \in \mathbb{R}^{n \times m}$ be a full-rank matrix. There is a polynomial-time algorithm in the size of its input constructing a matrix $\mathbf{S} \in \mathbb{R}^{m \times m}$ with $s_1(\mathbf{S}) \leq \varsigma / s_n(\mathbf{M})$ such that*

$$\varsigma^2 \cdot \mathbf{I}_n = \mathbf{M} \cdot \mathbf{S} \cdot \mathbf{S}^{\mathsf{T}} \cdot \mathbf{M}^{\mathsf{T}}.$$

Proof. Assume that $\mathbf{M} = \mathbf{V}_{\mathbf{M}} \cdot \mathrm{diag}(\{s_i(\mathbf{M})\}_{i=1}^{n}) \cdot \mathbf{U}_{\mathbf{M}}$ is a singular value decomposition of $\mathbf{M}$ with $\mathbf{V}_{\mathbf{M}} \in \mathbb{R}^{n \times n}$, $\mathbf{U}_{\mathbf{M}} \in \mathbb{R}^{n \times m}$ orthogonal. Set $\mathbf{S} = [\mathbf{U}_{\mathbf{M}}^{\mathsf{T}} | \mathbf{U}'] \cdot \mathrm{diag}(\{s_i\}_{i=1}^{m})$ where $\mathbf{U}'$ is an arbitrary extension of $\mathbf{U}_{\mathbf{M}}^{\mathsf{T}}$ to an orthonormal basis of $\mathbb{R}^m$ and $s_i = \varsigma / s_i(\mathbf{M})$ for $i \leq n$ and $s_i = 1$ for $n < i \leq m$. Computing the product we obtain the statement. □

2.3 Probability

Let D_1, D_2 be two (discrete or absolutely continuous w.r.t. the Lebesgue measure over $\mathbb{R}^n$ for some $n \geq 1$) probability distributions defined over the same σ-algebras. If $\Omega = \mathrm{Supp}(D_1) \cup \mathrm{Supp}(D_2)$, then the statistical distance between D_1 and D_2 is $\mathrm{SD}(D_1, D_2) \coloneqq \int_{t \in \Omega} |D_1(t) - D_2(t)| dt / 2$. If $\mathrm{Supp}(D_1) \subseteq \mathrm{Supp}(D_2)$, then order-2 Rényi divergence $\mathrm{RD}_2(D_1 || D_2) \coloneqq \int_{t \in \mathrm{Supp}(D_1)} D_1(t)^2 / D_2(t) dt$. In particular, for any event $E \subseteq \mathrm{Supp}(D_2)$, it holds that

$$D_2(E) \geq D_1(E)^2 / \mathrm{RD}(D_1 \| D_2) \tag{1}$$

Let $\mathbf{S} \in \mathbb{R}^{dr \times dr}$ be full-rank and $\mathbf{c} \in \mathcal{K}^r$ we define for any $x \in \mathcal{K}^r$,

$$\rho_{\mathbf{S},\mathbf{c}}(\mathbf{x}) \coloneqq \exp(-\pi \cdot \Phi(\mathbf{x} - \mathbf{c})^{\mathsf{T}} \cdot (\mathbf{S} \cdot \mathbf{S}^{\mathsf{T}})^{-1} \cdot \Phi(\mathbf{x} - \mathbf{c}))$$

If $\mathbf{S} = \varsigma \cdot \mathbf{I}_{dr}$ we simplify to $\rho_{\varsigma,\mathbf{c}}$, and if $\mathbf{c} = \mathbf{0}$, we omit it. For any matrix $\mathbf{S} \in \mathbb{R}^{dr \times dr}$ and centre $\mathbf{c}$, we denote by $\mathcal{D}_{\mathbf{S},\mathbf{c}}$ the continuous Gaussian distribution of covariance $\mathbf{S} \cdot \mathbf{S}^{\mathsf{T}}$ and centre $\mathbf{c}$. For any lattice $\Lambda \subset \mathbb{R}^r$, the discrete Gaussian distribution over Λ of parameter $\mathbf{S}$ and centre $\mathbf{c}$, denoted $\mathcal{D}_{\Lambda,S,\mathbf{c}}$ is the distribution

$$\mathcal{D}_{\Lambda,\mathbf{S},\mathbf{c}}(\mathbf{x}) = \frac{\rho_{\mathbf{S},\mathbf{c}}(\mathbf{x})}{\rho_{\mathbf{S},\mathbf{c}}(\Lambda)}.$$

For a positive definite matrix $\mathbf{\Sigma}$, if a statement does not depend on the particular choice of square root, we write $\sqrt{\mathbf{\Sigma}}$ to denote an arbitrary choice of the square root. In particular, we write $\mathcal{D}_{\Lambda,\sqrt{\mathbf{\Sigma}},\mathbf{c}}$ for the Gaussian distribution.

The smoothing parameter of a lattice Λ is a quantity defined for any $\varepsilon > 0$ as $\eta_\varepsilon(\Lambda) = \inf_{\varsigma>0}(\rho_{1/\varsigma}(\Lambda^\star) \leq 1 + \varepsilon)$. For any $\varepsilon > 0$ and $k \geq 1$, we define $\eta_\varepsilon^{(k)} = \sqrt{\ln(2k \cdot (1 + 1/\varepsilon))/\pi}$ and it holds that $\eta_\varepsilon(\Lambda) \leq \lambda_k(\Lambda) \cdot \eta_\varepsilon^{(k)}$ for rank-k lattices Λ ([KNSW20, Lemma 7 (eprint)]). For any full-rank matrix $\mathbf{S} \in \mathbb{R}^{n \times n}$ and any full-rank lattice $\Lambda \subset \mathbb{R}^n$, we say that $\eta_\varepsilon(\Lambda) \leq \mathbf{S}$ if $\eta_\varepsilon(\mathbf{S}^{-1}\Lambda) \leq 1$. Note that this only depends on $\mathbf{S} \cdot \mathbf{S}^{\mathsf{T}}$, and is coherent with the partial ordering of matrices we previously defined.

Lemma 2.6 (Adapted from [MP12, Lemma 2.6],[Pei08, Lemma 5.3]). *For $\Lambda \subset \mathbb{R}^n$ an n-dimensional lattice and $\varsigma, t > 0$,*

$$\Pr(\|\mathcal{D}_{\Lambda,\varsigma}\| \geq \varsigma\sqrt{n}) \leq 2^{-n},$$
$$\Pr(\|\mathcal{D}_{\Lambda,\varsigma}\|_\infty \geq \varsigma \cdot t) \leq 2en \cdot e^{-\pi t^2}.$$

Lemma 2.7. *For $\Lambda \subset \mathbb{R}^n$ an n-dimensional lattice, $t > 1/\sqrt{2\pi}$ and $\mathbf{S} \in \mathbb{R}^{n \times n}$ a non-singular matrix,*

$$\Pr(\|\mathcal{D}_{\Lambda,\mathbf{S}}\| \geq s_{\max}(\mathbf{S}) \cdot t \cdot \sqrt{n}) \leq \beta_n(t),$$

where $\beta_n(t) = (t \cdot \sqrt{2\pi e}\exp(-\pi \cdot t^2))^n$.

Proof. Since $\mathbf{S}$ is square and non-singular, it holds that $\mathbf{S}^{-1} \cdot \mathcal{D}_{\Lambda,\mathbf{S}}$ is the same distribution as $\mathcal{D}_{\mathbf{S}^{-1}\Lambda,1}$. Now, by [Ban93, Lemma 1.5], it holds that $\Pr(\left\|\mathcal{D}_{\mathbf{S}^{-1}\Lambda,1}\right\| > t \cdot \sqrt{n}) \leq (t \cdot \sqrt{2\pi e}\exp(-\pi \cdot t^2))^n$. Then $\|\mathcal{D}_{\Lambda,\mathbf{S}}\| = \left\|\mathbf{S} \cdot \mathbf{S}^{-1} \cdot \mathcal{D}_{\Lambda,\mathbf{S}}\right\| \leq s_{\max}(\mathbf{S}) \cdot \left\|\mathbf{S}^{-1} \cdot \mathcal{D}_{\Lambda,\mathbf{S}}\right\| \leq s_{\max}(\mathbf{S}) \cdot t \cdot \sqrt{n}$ with probability at least $1 - (t \cdot \sqrt{2\pi e}\exp(-\pi \cdot t^2))^n$. □

Lemma 2.8 ([GMPW20, Theorem 3.1 (eprint)]). *Let $\varepsilon \in [0, 1)$ and $1 \leq n \leq m$ be integers. Let $\Lambda \subset \mathbb{R}^m$ be an m-dimensional lattice, $\mathbf{c} \in \mathbb{R}^m$ and $\Sigma \in \mathbb{R}^{m \times m}$ be positive definite. For a matrix $\mathbf{T} \in \mathbb{R}^{n \times m}$ such that $\mathrm{span}_{\mathbb{R}}(\Lambda \cap \ker(\mathbf{T})) = \mathrm{span}_{\mathbb{R}}(\ker(\mathbf{T}))$ and $\eta_\varepsilon(\Lambda \cap \ker(\mathbf{T})) \leq \sqrt{\mathbf{\Sigma}}$, it holds that*

$$\mathrm{SD}(\mathbf{T} \cdot \mathcal{D}_{\Lambda,\sqrt{\mathbf{\Sigma}},\mathbf{c}}, \mathcal{D}_{\mathbf{T}\cdot\Lambda,\sqrt{\mathbf{\Sigma}'},\mathbf{T}\cdot\mathbf{c}}) \leq \frac{\varepsilon}{1-\varepsilon}$$

where $\mathbf{\Sigma}' = \mathbf{T} \cdot \mathbf{\Sigma} \cdot \mathbf{T}^{\mathsf{T}}$.

Lemma 2.9 ([Reg09, Claim 3.9]). *Let Λ be an n-dimensional lattice, $\varepsilon \in (0, 1/2)$, and $r, s > 0$ such that $\frac{r \cdot s}{\sqrt{r^2+s^2}} \geq \eta_\varepsilon(\Lambda)$. Then*

$$\mathrm{SD}(\mathcal{D}_{\Lambda,r} + \mathcal{D}^n_s, \mathcal{D}^n_{\sqrt{r^2+s^2}}) \leq 4\,\varepsilon.$$

Lemma 2.10 ([LM00, Lemma 1 (p. 1325)]). *Let χ^2_m be a Chi-squared random variable of degree $m > 0$, then*

$$\Pr(\chi^2_m \geq m + 2\sqrt{mk} + 2k) \leq \exp(-k),$$
$$\Pr(\chi^2_m \leq m - 2\sqrt{mk}) \leq \exp(-k).$$

Lemma 2.11. *Let $S \subset \mathcal{O}^r_\mathcal{K}$ be a set such that $S = -S$ and $\mathbf{S}$ an invertible matrix. Then for any $\mathbf{c} \in \mathcal{O}^r_\mathcal{K}$*

$$\rho_\mathbf{S}(\mathbf{c} + S) \geq \rho_\mathbf{S}(\mathbf{c}) \cdot \rho_\mathbf{S}(S)$$

Proof. For any $\mathbf{s} \in S$, it holds that $\rho_\mathbf{S}(\mathbf{c}+\mathbf{s}) + \rho_\mathbf{S}(\mathbf{c}-\mathbf{s}) = \rho_\mathbf{S}(\mathbf{s}) \cdot \rho_\mathbf{S}(\mathbf{c}) \cdot (\rho_\mathbf{S}(-2\mathbf{s} \cdot \mathbf{c}) + \rho_\mathbf{S}(2\mathbf{s} \cdot \mathbf{c})) = 2\rho_\mathbf{S}(\mathbf{s}) \cdot \rho_\mathbf{S}(\mathbf{c}) \cdot \cosh(\pi \cdot \mathbf{s} \cdot \mathbf{c}/\varsigma^2) = (\rho_\varsigma(\mathbf{s}) + \rho_\varsigma(-\mathbf{s})) \cdot \rho_\varsigma(\mathbf{c}) \cdot \cosh(\pi \cdot (\mathbf{S}^{-1} \cdot \mathbf{s}) \cdot (\mathbf{S}^{-1} \cdot \mathbf{c})) \geq (\rho_\varsigma(\mathbf{s}) + \rho_\varsigma(-\mathbf{s})) \cdot \rho_\varsigma(\mathbf{c})$ and then by adding over all pairs $-\mathbf{s}, \mathbf{s}$ in S, the result follows. □

Lemma 2.12 (Adapted [MR07], [GMPW20, Lem 2.6 eprint]). *Let Λ be an n-dimensional lattice and $\sqrt{\mathbf{\Sigma}} \geq \eta_\varepsilon(\Lambda)$. Then*

$$\rho_{\sqrt{\mathbf{\Sigma}}}(\Lambda) \in [1-\varepsilon, 1+\varepsilon] \cdot \frac{\det(\sqrt{\mathbf{\Sigma}})}{\det(\Lambda)}$$

Proof. If $\sqrt{\mathbf{\Sigma}} \geq \eta_\varepsilon(\Lambda)$ then $1 \geq \eta_\varepsilon(\sqrt{\mathbf{\Sigma}}^{-1} \cdot \Lambda)$ and the statement follows. □

Lemma 2.13. *Let $\Lambda \subset \mathbb{R}^n$ be a full-rank lattice. Let $R \geq \sqrt{n} \cdot \eta_{1/2}(\Lambda)$ be a real. Then*

$$|\Lambda \cap B(0, R)| \leq 1.5 \cdot \exp(\pi \cdot n) \cdot \left(\frac{R}{\sqrt{n}}\right)^n \cdot \frac{1}{\det(\Lambda)}.$$

Proof. Let $\varsigma = R/\sqrt{n}$. By Lemma 2.12, it holds that

$$\rho_\varsigma(\Lambda \cap B(0, R)) \leq \rho_\varsigma(\Lambda) \leq 1.5 \frac{\varsigma^n}{\det \Lambda},$$

the result follows from the fact that $\rho_\varsigma(\Lambda \bigcap B(0, R))$ is a sum of $|\Lambda \bigcap B(0, R)|$ terms, all greater than $\rho_\varsigma(R) = \exp(-\pi n)$. □

Lemma 2.14 (Simplified from [ALLW25, Theorem 9 (eprint)]). *Let g, k, m, n, q be positive integers, $a \geq 1$ be a real number, and $\mathbf{\Sigma}$ be positive semi-definite in $\mathbb{R}^{dm \times dm}$ satisfying*

- *$k, n \leq m$, $d \cdot (m - k) \geq \Omega(\lambda)$,*
- *q be an unramified prime s.t. $\langle q \rangle = \prod_{j=1}^g \mathfrak{q}_j$ with norm $\mathcal{N}(\mathfrak{q}_j) = q^{d/g}$ in $\mathcal{O}_\mathcal{K}$,*

– $ng/q^{d(m-n+1)/g} \le \mathsf{negl}(\lambda)$, *and there exists* $\epsilon \le \mathsf{negl}(\lambda)$ *such that*

$$\max\left\{\eta_A, 2\sqrt{d}\cdot(a^k q^n)^{1/(m-k)}\right\} \le s_{\min}\left(\sqrt{\mathbf{\Sigma}}\right),$$
$$s_{\max}\left(\sqrt{\mathbf{\Sigma}}\right) \le \min\left\{q^{1/g}/\sqrt{m}, a\cdot s_{\min}\left(\sqrt{\mathbf{\Sigma}}\right)\right\},$$

for some $\eta_A \ge 8\,d\sqrt{m}\cdot q^{n/m+2/(d\cdot m)}$.

Then the following distributions are statistically close in λ*:*

$$\left\{(\mathbf{A},\mathbf{X}) \;\middle|\; \begin{array}{l}\mathbf{A}\leftarrow(\mathcal{O}_\mathcal{K}/q)^{n\times m}\\ \mathbf{X}\leftarrow(\mathcal{D}_{\Lambda_q^\perp(\mathbf{A}),\sqrt{\mathbf{\Sigma}}})^k\end{array}\right\} \approx_s \left\{(\mathbf{A},\mathbf{X}) \;\middle|\; \begin{array}{l}\mathbf{X}\leftarrow(\mathcal{D}_{\mathcal{O}_\mathcal{K}^m,\sqrt{\mathbf{\Sigma}}})^k\\ \mathbf{A}\leftarrow(\mathcal{O}_\mathcal{K}/q)^{n\times m} : \mathbf{A}\cdot\mathbf{X}\equiv\mathbf{0}\end{array}\right\}.$$

Lemma 2.15 (Simplified from [ALLW25, Lemma 26 (eprint)]). *Let* g,k,m,n,q *be positive integers,* $a\ge 1$ *be a real number,* $\mathbf{A}\in(\mathcal{O}_\mathcal{K}/q)^{n\times m}$ *be primitive, and* $\mathbf{\Sigma}\in\mathbb{R}^{dm\times dm}$ *be positive semi-definite satisfying*

– $k,n\le m$, $d\cdot(m-k)\ge\Omega(\lambda)$,
– q *is an unramified prime s.t.* $\langle q\rangle=\prod_{j=1}^g\mathfrak{q}_j$ *with* $\mathcal{N}(\mathfrak{q}_j)=q^{d/g}$, *and*
– *there exists* $\epsilon\le\mathsf{negl}(\lambda)$ *such that*

$$\max\left\{\eta_\epsilon(\Lambda_q^\perp(\mathbf{A})), 2\sqrt{d}\cdot(a^k q^n)^{1/(m-k)}\right\} \le s_{\min}\left(\sqrt{\mathbf{\Sigma}_i}\right),$$
$$s_{\max}\left(\sqrt{\mathbf{\Sigma}_i}\right) \le \min\left\{q^{1/g}/\sqrt{m}, a\cdot s_{\min}\left(\sqrt{\mathbf{\Sigma}_i}\right)\right\}.$$

Then the columns of $\mathbf{X}\leftarrow(\mathcal{D}_{\Lambda_q^\perp(\mathbf{A}),\sqrt{\mathbf{\Sigma}}})^k$ *are* $(\mathcal{O}_\mathcal{K}/q)$*-linearly independent with overwhelming probability in* λ.

Lemma 2.16 ([ALLW25, Lemma 3 (eprint)]; Generalisation of [LPR13, Theorem 4.1]). *Let* n,m,q *be positive integers with* $n\le m\le\mathsf{poly}(d)$. *For* $\mathbf{A}\leftarrow\mathcal{R}_q^{n\times m}$, *with probability* $1-2^{-\Omega(dm)}$ *we have* $\eta_{2^{-\Omega(dm)}}(\Lambda_q^\perp(\mathbf{A}))\le 8d\sqrt{m}\cdot q^{n/m+2/(d\cdot m)}$.

Lemma 2.17 ([LSS14, Lemma 4.2]). *For any* m*-dimensional lattice* $\Lambda\subseteq\mathbb{R}^m$ *and rank-*m *matrix* $\sqrt{\mathbf{\Sigma}}\in\mathbb{R}^{m\times m}$, *let* $P=\mathcal{D}_{\Lambda,\sqrt{\mathbf{\Sigma}},\mathbf{w}}$ *and* $Q=\mathcal{D}_{\Lambda,\sqrt{\mathbf{\Sigma}},\mathbf{z}}$ *for some fixed* $\mathbf{w},\mathbf{z}\in\mathbb{R}^m$. *If* $\mathbf{w},\mathbf{z}\in\Lambda$, *let* $\epsilon=0$. *Otherwise, fix* $\epsilon\in(0,1)$ *and assume that* $s_{\min}(\sqrt{\mathbf{\Sigma}})\ge\eta_\epsilon(\Lambda)$. *Then* $\mathrm{RD}(P\|Q)\le(\frac{1+\epsilon}{1-\epsilon})^2\cdot\exp(2\pi\|\mathbf{w}-\mathbf{z}\|^2/s_{\min}(\sqrt{\mathbf{\Sigma}})^2)$.

Lemma 2.18. *Let* $\mathcal{K}$ *be a number field of degree* d. *Let* $n,m,k\in\mathbb{N}$. *Let* $\mathbf{A}\in\mathcal{R}_q^{n\times m}$ *be such that* $\Lambda_q^\perp(\mathbf{A})$ *is full-rank, and let* $\sqrt{\mathbf{\Sigma}}\in\mathbb{R}^{dm\times dm}$ *satisfy* $s_{\min}(\sqrt{\mathbf{\Sigma}})\ge\eta_\epsilon(\Lambda_q^\perp(\mathbf{A}))$ *for some* $0<\epsilon\le 2^{-2k}$. *Let* $\mathbf{e}_i\in\mathcal{R}^m$ *be the* i*-th unit vector for* $i\in[k]$. *The Rényi divergence between the distributions* $P:=\prod_{i\in[k]}\mathcal{D}_{\Lambda_q^\perp(\mathbf{A}),\sqrt{\mathbf{\Sigma}}}$ *and* $Q:=\prod_{i\in[k]}\mathcal{D}_{\Lambda_q^\perp(\mathbf{A}),\sqrt{\mathbf{\Sigma}},\mathbf{e}_i}$ *is upper bounded by*

$$\mathrm{RD}(P\|Q)\le 5\exp\left(\frac{2\pi\cdot dk}{s_{\min}(\sqrt{\mathbf{\Sigma}})^2}\right).$$

Proof. Denote by $1+\delta \coloneqq (1+\epsilon)/(1-\epsilon)$, equivalently $\delta = 2\epsilon/(1-\epsilon)$. Using the multiplicative property of Rényi divergence, we have

$$\begin{aligned}\mathrm{RD}(P\|Q) &\leq \prod_{i\in[k]} (1+\delta)^2 \cdot \exp\left(2\pi\|\mathbf{e}_i\|^2/s_{\min}(\sqrt{\boldsymbol{\Sigma}})^2\right) \\ &\leq (1+(2^{2k}-1)\cdot\delta)\cdot\exp\left(\frac{2\pi\cdot dk}{s_{\min}(\sqrt{\boldsymbol{\Sigma}})^2}\right) \quad \leq 5\cdot\exp\left(\frac{2\pi\cdot dk}{s_{\min}(\sqrt{\boldsymbol{\Sigma}})^2}\right),\end{aligned}$$

where the first inequality is by Lemma 2.17, the second uses $(1+x)^r \leq 1+(2^r-1)x$ for any $x\in[0,1]$ and $r\in\mathbb{R}\setminus(0,1)$, and the last is by $2^{2k}\cdot\delta = 2^{2k}\cdot 2\epsilon/(1-\epsilon) \leq 4$. □

2.4 Cryptographic Assumptions

Definition 2.19 (SIS Problem).
Let the parameters $\mathsf{params} = (\mathcal{O}_\mathcal{K}, q, n, m, \beta)$ *be parametrised by* λ*, where* $\mathcal{O}_\mathcal{K}$ *is the ring of integers of a number field* $\mathcal{K}$*,* q, n, m *are positive integers and* $\beta > 0$*. Write* $\mathcal{R}_q \coloneqq \mathcal{O}_\mathcal{K}/(q\cdot\mathcal{O}_\mathcal{K})$*. A* $\mathsf{SIS}_\mathsf{params}$ *problem asks the following: Given a uniform matrix* $\mathbf{A} \leftarrow \mathcal{R}_q^{n\times m}$*, find a vector* $\mathbf{u}\in\mathcal{O}_\mathcal{K}^m$ *such that*

$$(1)\quad \mathbf{A}\cdot\mathbf{u} = \mathbf{0} \bmod q \qquad \textit{and} \qquad (2)\quad 0 < \|\mathbf{u}\| < \beta.$$

For any PPT $\mathcal{A}$*, its advantage* $\mathsf{Adv}_\mathcal{A}^{\mathsf{SIS}_\mathsf{params}}(\lambda)$ *against this problem is equal to the probability that it solves the problem.*

Definition 2.20 (LWE Problem).
Let the parameters $\mathsf{params} = (\mathcal{O}_\mathcal{K}, q, n, m, \chi)$ *be parametrised by* λ*, where* $\mathcal{O}_\mathcal{K}$ *is the ring of integers of a number field* $\mathcal{K}$*,* q, n, m *are positive integers and* χ *is a distribution over* $\mathcal{O}_\mathcal{K}$*. Write* $\mathcal{R}_q \coloneqq \mathcal{O}_\mathcal{K}/(q\cdot\mathcal{O}_\mathcal{K})$*. An* $\mathsf{LWE}_\mathsf{params}$ *problem asks the following: Given a uniform matrix* $\mathbf{A}\leftarrow\mathcal{R}_q^{n\times m}$*, and a vector* $\mathbf{b}\in\mathcal{R}_q^m$ *which is either*

$$\begin{aligned}&(1)\quad \mathbf{b}^\mathsf{T} = \mathbf{s}^\mathsf{T}\mathbf{A} + \mathbf{e}^\mathsf{T} \bmod q \textit{ where } \mathbf{s}\leftarrow\mathcal{R}_q^n,\ \mathbf{e}\leftarrow\chi^m,\\ \textit{or}\quad &(2)\quad \mathbf{b}\leftarrow\mathcal{R}_q^m \textit{ is uniformly random},\end{aligned}$$

distinguish which one is the case. For any PPT $\mathcal{A}$*, its advantage* $\mathsf{Adv}_\mathcal{A}^{\mathsf{LWE}_\mathsf{params}}(\lambda)$ *against this problem is equal to the absolute difference in the probability of its response under these two cases.*

Definition 2.21 (k-SIS Problem).
Let the parameters $\mathsf{params} = (\mathcal{O}_\mathcal{K}, q, n, m, k, \boldsymbol{\Sigma}, \beta)$ *be parametrised by* λ*, where* $\mathcal{O}_\mathcal{K}$ *is the ring of integers of a number field* $\mathcal{K}$ *of degree* d*,* q, n, m, k *are positive integers with* $k < m$*,* $\boldsymbol{\Sigma}\in\mathbb{R}^{dm\times dm}$ *and* $\beta > 0$*. Write* $\mathcal{R}_q \coloneqq \mathcal{O}_\mathcal{K}/(q\cdot\mathcal{O}_\mathcal{K})$*. A* k-$\mathsf{SIS}_\mathsf{params}$ *problem asks the following: Given a uniform matrix* $\mathbf{A}\leftarrow\mathcal{R}_q^{n\times m}$ *and* k *vectors* $\mathbf{X} = (\mathbf{x}_1,\ldots,\mathbf{x}_k)\in\mathcal{O}_\mathcal{K}^{m\times k}$ *where* $\mathbf{x}_i\leftarrow\mathcal{D}_{\Lambda_q^\perp(\mathbf{A}),\sqrt{\boldsymbol{\Sigma}}}$ *for each* $i\in[k]$*, find a vector* $\mathbf{u}\in\mathcal{O}_\mathcal{K}^m$ *such that*

$$(1)\quad \mathbf{A}\cdot\mathbf{u} = \mathbf{0} \bmod q, \quad (2)\quad 0 < \|\mathbf{u}\| < \beta, \quad \textit{and} \quad (3)\quad \mathbf{u}\notin\mathcal{K}\text{-}\operatorname{span}(\mathbf{X}).$$

For any PPT $\mathcal{A}$, its advantage $\mathsf{Adv}^{k\text{-}\mathsf{SIS}_{\mathsf{params}}}_{\mathcal{A}}(\lambda)$ against this problem is equal to the probability that it solves the problem.

Definition 2.22 (k-LWE Problem).
Let the parameters $\mathsf{params} = (\mathcal{O}_{\mathcal{K}}, q, n, m, k, \mathbf{\Sigma}, \chi)$ be parametrised by λ, where $\mathcal{O}_{\mathcal{K}}$ is the ring of integers of a number field $\mathcal{K}$ of degree d, q, n, m, k are positive integers with $k < m$, $\mathbf{\Sigma} \in \mathbb{R}^{dm \times dm}$ and χ is a distribution over $\mathcal{O}_{\mathcal{K}}$. Write $\mathcal{R}_q \coloneqq \mathcal{O}_{\mathcal{K}}/(q \cdot \mathcal{O}_{\mathcal{K}})$. An $\mathsf{LWE}_{\mathsf{params}}$ problem asks the following: Given a uniform matrix $\mathbf{A} \leftarrow \mathcal{R}_q^{n \times m}$, k vectors $\mathbf{X} = (\mathbf{x}_1, \ldots, \mathbf{x}_k) \in \mathcal{O}_{\mathcal{K}}^{m \times k}$ where $\mathbf{x}_i \leftarrow \mathcal{D}_{\Lambda_q^{\perp}(\mathbf{A}), \sqrt{\mathbf{\Sigma}}}$ for each $i \in [k]$, and a vector $\mathbf{b} \in \mathcal{R}_q^m$ which is either

$$\begin{aligned} &(1) \quad \mathbf{b}^{\mathsf{T}} = \mathbf{s}^{\mathsf{T}}\mathbf{A} + \mathbf{e}^{\mathsf{T}} \bmod q \text{ where } \mathbf{s} \leftarrow \mathcal{R}_q^n,\ \mathbf{e} \leftarrow \chi^m, \\ \textit{or} \quad &(2) \quad \mathbf{b}^{\mathsf{T}} = \mathbf{d} + \mathbf{e}^{\mathsf{T}} \bmod q \text{ where } \mathbf{d} \leftarrow \left\{\mathbf{d} \in \mathcal{R}^m : \mathbf{d}^{\mathsf{T}}\mathbf{X} = \mathbf{0} \bmod q\right\}, \end{aligned}$$

distinguish which one is the case. For any PPT $\mathcal{A}$, its advantage $\mathsf{Adv}^{k\text{-}\mathsf{LWE}_{\mathsf{params}}}_{\mathcal{A}}(\lambda)$ against this problem is equal to the absolute difference in the probability of its response under these two cases.

3 Generating module lattices with Gaussian vectors

In this section, we study the surjectivity of Gaussian matrices. For a matrix where each column is sampled from an $\mathcal{O}_{\mathcal{K}}$-module $\mathcal{M}$, we say that it is surjective if $\mathcal{O}_{\mathcal{K}}$-linear combinations of its columns generate the module $\mathcal{M}$.

3.1 Surjective matrices over finite fields

We first establish some facts about random matrices over finite fields $\mathbb{F}$ to be used in subsequent proofs. First, extending the above notion of surjectivity to $\mathbb{F}$-vector spaces, we show that a random matrix over $\mathbb{F}$ is surjective with overwhelming probability in its dimensions if each entry of the matrix is independently distributed and has constant min-entropy.

Lemma 3.1. *Let $r, m \geq 0$ be integers. Let $\mathbb{F}$ be a finite field, and $(D_{i,j})_{1 \leq i \leq r, 1 \leq j \leq m}$ be distributions over $\mathbb{F}$ such that there exists $\alpha \in (0, 1)$ such that for any i, j, $\max_{x \in \mathbb{F}} D_{i,j}(x) \leq \alpha$. If $\mathbf{M} \in \mathbb{F}^{r \times m}$ is a random matrix such that*

- *all the entries of $\mathbf{M}$ are independent random variables, and*
- *the (i, j)-th entry of $\mathbf{M}$ is sampled from $D_{i,j}$,*

then it holds that

$$\Pr(\mathbf{M} \cdot \mathbb{F}^m = \mathbb{F}^r) \geq 1 - r \cdot \alpha^{m-r+1}.$$

Proof. Let H be a subspace of $\mathbb{F}^m$ of dimension k and $i = 1, \ldots, r$, it holds that the probability of a vector in $\mathbb{F}^m$ sampled from $D_{i,1} \times \ldots \times D_{i,m}$ belonging to this subspace is upper bounded by α^{m-k} [NP20, Generalisation of Lemma 2.3]. As $\mathbb{F}$ is a field, the matrix $\mathbf{M}$ is surjective if and only if $\mathbf{M}^T$ is injective. We now

lower-bound the probability of $\mathbf{M}^T$ being injective when $\mathbf{M}$ is built row by row. Let $\mathbf{M}_i$ be the matrix composed of the first i rows of $\mathbf{M}$. For $i = 1, \ldots, r$, the matrix $\mathbf{M}_i^T$ is injective if and only if $\mathbf{M}_{i-1}^T$ is injective and the i-th row of $\mathbf{M}$ is not in the span of the $i-1$ previous ones, which has dimension $i-1$. The probability of $\mathbf{M}^T$ being injective is then lower-bounded by

$$\prod_{i=1}^{r}(1-\alpha^{m-i+1}) = \prod_{i=m-r+1}^{m}(1-\alpha^i) \geq (1-\alpha^{m-r+1})^r \geq 1 - r\cdot\alpha^{m-r+1}. \qquad \square$$

The next lemma shows the equivalence between a matrix $\mathbf{B}$ with columns in $\mathcal{M}$ being surjective and the simpler condition of ($\mathbf{B} \bmod \mathfrak{p}$) being surjective for prime ideals $\mathfrak{p}$ of $\mathcal{O}_\mathcal{K}$.

Lemma 3.2. *For any integers $r, m \geq 1$, a matrix $\mathbf{B} \in \mathcal{M}^m$ satisfies $\mathbf{B}\cdot\mathcal{O}_\mathcal{K}^m = \mathcal{M}$ if and only if ($\mathbf{B} \bmod \mathfrak{p}$) is surjective in $\mathcal{M}/\mathfrak{p}\mathcal{M}$ for any prime ideal $\mathfrak{p}$ of $\mathcal{O}_\mathcal{K}$.*

Proof (Adaptation of the proof of [NP20, Lemma 2.1]). If $\mathbf{B}$ is surjective in $\mathcal{M}$, then it is clearly surjective modulo any prime $\mathfrak{p}$. Now we assume that $\mathbf{B} \bmod \mathfrak{p}$ is surjective modulo every $\mathfrak{p}$. Let $\mathcal{M}' = \mathbf{B}\cdot\mathcal{O}_\mathcal{K}^m$. Let $\mathfrak{p}$ be any ideal, we will show that $\mathfrak{p}$ does not divide $\det(\mathcal{M}')/\det(\mathcal{M})$. Since the matrix $\mathbf{B} \bmod \mathfrak{p}$ is surjective, there exists a square submatrix $\mathbf{B}'$ of $\mathbf{B}$ which is invertible modulo $\mathfrak{p}$, and hence full-rank over $\mathcal{K}$. Let $\mathcal{M}'' = \mathbf{B}'\cdot\mathcal{O}_\mathcal{K}^r \subset \mathcal{M}'$. Assume that the ideal $\mathfrak{p}$ divides $\det(\mathcal{M}')/\det(\mathcal{M})$, then it also divides $\det(\mathcal{M}'')/\det(\mathcal{M})$ by inclusion, which is equivalent to that $\det_\mathcal{K}(\mathbf{B}') \in \det(\mathcal{M})\cdot\mathfrak{p}$, which contradicts the fact that $\mathbf{B}'$ is invertible in $\mathcal{M}/\mathfrak{p}\mathcal{M}$ by Lemma 2.2. Since $\mathfrak{p}$ can be any ideal, it holds that $\det(\mathcal{M}') = \det(\mathcal{M})$, and hence that $\mathbf{B}$ is surjective in $\mathcal{M}$. $\square$

3.2 Surjective matrices over free modules

In this subsection, we prove that Gaussian matrices of size $r \times m$ over some full-rank modules $\mathcal{M}$ are surjective over $\mathcal{M}$ with high probability when $m \geq 2r$. The main idea of the proof is to consider the matrix modulo all the primes $\mathfrak{p}$ of the number field, showing that with high probability (using Lemmas 3.1 and 3.2), the matrix modulo $\mathfrak{p}$ is surjective over $\mathcal{O}_\mathcal{K}/\mathfrak{p}$. We first prove it in the case of free modules in Lemma 3.4, where the main contribution of this subsection lies, and then generalise it in Theorem 3.5. We first prove that the Gaussian distribution over some ideal I modulo the ideal $I\cdot\mathfrak{p}$ satisfies the distribution hypothesis of Lemma 3.1.

Lemma 3.3 (Generalised from [JLWG25, Corollary 7]). *Let $\varepsilon \in (0, 1/\sqrt{2})$, I be a (possibly fractional) ideal of $\mathcal{O}_\mathcal{K}$, $B \geq 1$, $c \in \mathcal{K}_\mathbb{R}$ be a centre and $\varsigma \geq B\cdot\Delta_\mathcal{K}^{1/d}\cdot\mathcal{N}(I)^{1/d}\cdot\eta_\varepsilon^{(d)}$. For any integral ideal $\mathfrak{a} \subset \mathcal{O}_\mathcal{K}$, we define $\mathcal{D}_{I,\varsigma,c}^{(\mathfrak{a})}$ the distribution $\mathcal{D}_{I,\varsigma,c}$ modulo $\mathfrak{a}\cdot I$. Then for any $a \in I/(I\cdot\mathfrak{a})$, it holds that*

$$\mathcal{D}_{I,\varsigma,c}^{(\mathfrak{a})}(a) \leq (1+4\varepsilon)\cdot\begin{cases}\frac{1}{\mathcal{N}(\mathfrak{a})} & \text{if } \mathcal{N}(\mathfrak{a}) \leq B^d \\ \frac{1}{B^d} & \text{else .}\end{cases}$$

Lemma 3.4 (Assuming GRH). *There exists an absolute constant $c > 1$ such that the following holds. Let $r \geq 1$ and $m \geq 2r$ be integers. Let $\mathbf{C} = [\mathbf{c}_1, \ldots, \mathbf{c}_m] \in \mathbb{R}^{rd \times rm}$ be a centre matrix and $\varsigma > 0$. Let $\mathbf{S} \geq \sqrt{2}\varsigma\mathbf{I}$ or equal to $\varsigma\mathbf{I}$. Let $\mathcal{M} \subset \mathcal{K}^r$ be a free module of rank r. Let $\varepsilon \in (0, 1/4)$. Let $B > \log(\Delta_\mathcal{K})^{c/d}$ be such that*

$$B^d \geq r \cdot d^2 \ln(dr \cdot \varsigma) \qquad \text{and} \qquad \varsigma \geq B \cdot \Delta_\mathcal{K}^{1/d} \cdot \eta_\varepsilon(\mathcal{M}).$$

Let $\mathbf{X} = [\mathbf{x}_1, \ldots, \mathbf{x}_m] \in \mathcal{K}^{r \times m}$ such that $\mathbf{x}_i \leftarrow \mathcal{D}_{\mathcal{M}, \mathbf{S}, \mathbf{c}_i}$. It holds that

$$\Pr_{\mathbf{X}}[\mathbf{X} \cdot \mathcal{O}_\mathcal{K}^m \neq \mathcal{M}] \leq \frac{m \cdot \varepsilon}{1 - \varepsilon} + r \cdot \left(\frac{B}{2}\right)^{-(r+1)d} + r \cdot 2^{-dr} + (1 + 4\varepsilon)^{m-2r+1} \cdot r \cdot P_\mathcal{K}(m - 2r + 1),$$

where $P_\mathcal{K}(s) = \sum_\mathfrak{p} \mathcal{N}(\mathfrak{p})^{-s}$ is the prime zeta function of $\mathcal{K}$.

Proof. We abuse the notation and write the distribution of $\mathbf{X}$ as $\mathcal{D}_{\mathcal{M}^m, \mathbf{S}, \mathbf{C}}$.

From free modules to $\mathcal{O}_\mathcal{K}^r$. Let $\mathbf{B} \in \mathcal{K}^{r \times r}$ generating $\mathcal{M}$. It holds that the distribution of the columns of $\mathbf{X}$ is the same as the distribution $\mathbf{B} \cdot \mathcal{D}_{\mathcal{O}_\mathcal{K}^{r \times m}, \mathbf{S}', \mathbf{C}'}$ where $\mathbf{S}' = \mathbf{B}^{-1} \cdot \mathbf{S}$ and $\mathbf{C}' = \mathbf{B}^{-1} \cdot \mathbf{C}$. Note that the condition on $\mathbf{S}$ implies that $\mathbf{S}' \geq B \cdot \Delta_\mathcal{K}^{1/d} \cdot \eta_\varepsilon(\mathcal{O}_\mathcal{K}^r)$. We can then assume without loss of generality that $\mathcal{M} = \mathcal{O}_\mathcal{K}^r$.

From elliptic to circular Gaussian. Let $\mathbf{\Sigma} = \mathbf{S}\mathbf{S}^T - \varsigma^2\mathbf{I}$, note that $\mathbf{\Sigma}$ is positive definite, let $\sqrt{\mathbf{\Sigma}}$ be a fixed square root of $\mathbf{\Sigma}$. Now, let $\mathcal{D}_{\widetilde{\mathcal{O}_\mathcal{K}^{r \times m}}, \mathbf{S}, \mathbf{C}}$ be the distribution $\mathcal{D}_{\mathcal{O}_\mathcal{K}^{r \times m}, \varsigma, \mathbf{C} - \mathbf{B}}$ where $\mathbf{B} \leftarrow \mathcal{D}_{\mathcal{O}_\mathcal{K}^{r \times m}, \sqrt{\mathbf{\Sigma}}}$. By [GMPW20, Theorem 4.5], the condition on $\mathbf{S}$ and Lemma B.2, it holds that the statistical distance between $\mathcal{D}_{\widetilde{\mathcal{O}_\mathcal{K}^{r \times m}}, \mathbf{S}, \mathbf{C}}$ and $\mathcal{D}_{\mathcal{O}_\mathcal{K}^{r \times m}, \mathbf{S}, \mathbf{C}}$ is less than $m \cdot \varepsilon/(1 - \varepsilon)$. It then holds that

$$\Pr_{\mathbf{X}}[\mathbf{X} \cdot \mathcal{O}_\mathcal{K}^m \neq \mathcal{O}_\mathcal{K}^r] \leq m \cdot \frac{\varepsilon}{1 - \varepsilon} + \sum_{\mathbf{B} \in \mathcal{O}_\mathcal{K}^{r \times m}} \mathcal{D}_{\mathcal{O}_\mathcal{K}^{r \times m}, \sqrt{\mathbf{\Sigma}'}}(\mathbf{B}) \Pr_{\mathbf{A} \leftarrow \mathcal{D}_{\mathcal{O}_\mathcal{K}^{r \times m}, \varsigma, \mathbf{C} - \mathbf{B}}} (\mathbf{A} \cdot \mathcal{O}_\mathcal{K}^m \neq \mathcal{O}_\mathcal{K}^r).$$

Bound for circular Gaussian. We now assume that $\mathbf{S} = \varsigma \cdot \mathbf{I}$ with $\varsigma \geq B \cdot \Delta_\mathcal{K}^{1/d} \cdot \eta_\varepsilon(\mathcal{O}_\mathcal{K}^r)$. Note that in that case, the coefficients of $\mathbf{X}$ are all independent and sampled from $\mathcal{D}_{\mathcal{O}_\mathcal{K}, \varsigma, c}$ for some $c \in \mathcal{K}$, so by Lemma 3.3, Lemma 3.1 holds for any submatrix of $\mathbf{X}$. Write $\mathbf{X} = [\mathbf{X}_1 | \mathbf{X}_2]$ with $\mathbf{X}_1 \in \mathcal{O}_\mathcal{K}^{r \times r}$, and let $\mathfrak{p}$ be a prime ideal of norm $\geq B^d$. Then by Lemma 3.2, $\mathbf{X}_1$ is invertible modulo $\mathfrak{p}$ with probability $\geq 1 - r(B/2)^{-(r+1)d}$ and in that case, in particular $\det_K(\mathbf{X}_1) \neq 0$, we assume that this is the case for the rest of the proof. By [Ban93, Lemma 1.5], with probability $1 - r \cdot 2^{dr}$, the norm of every column of $\mathbf{X}_1$ is bounded by $\sqrt{rd} \cdot \varsigma$, and hence its determinant is bounded by $(\sqrt{rd}\varsigma)^{rd}$. Let $\mathcal{P}$ be the set of prime ideals dividing $\det_K(\mathbf{X}_1)$, by Lemma B.1, it holds that $|\mathcal{P}| \leq rd^2 \cdot \log(\sqrt{rd}\varsigma) =: l$. We now prove that $\mathbf{X}$ is surjective modulo every prime ideal. For any prime $\mathfrak{p}$

not in $\mathcal{P}$, $\mathbf{X}_1$ is invertible modulo $\mathfrak{p}$, and so is $\mathbf{X}$. For all $\mathfrak{p} \in \mathcal{P}$, by Lemma 3.1, it holds that $\mathbf{X}_2$ is invertible mod $\mathfrak{p}$ with probability $\geq 1 - r(1-4\varepsilon)^{m-2r+1} \cdot \alpha$ for $\alpha = B^{-(m-2r+1)d}$ if $\mathcal{N}(\mathfrak{p}) \geq B^d$ and $\alpha = \mathcal{N}(\mathfrak{p})^{-(m-2r+1)}$ else.

Let us order the primes of $\mathcal{O}_\mathcal{K}$ as $\{\mathfrak{p}_1, \mathfrak{p}_2, \ldots\}$ and $\mathcal{P} = \{\mathfrak{q}_1, \ldots, \mathfrak{q}_l\}$, both in increasing norm order. It holds that $\mathcal{N}(\mathfrak{p}_i) \leq \mathcal{N}(\mathfrak{q}_i)$ for any $i = 1, \ldots, l$. We let B be a real such that the first l prime ideals of $\mathcal{O}_\mathcal{K}$ have their norm $\leq B^d$. The existence of B will be proven at the end of the proof. Under the condition that $\mathbf{X}_1$ is invertible and has bounded columns, by the union bound it holds that

$$\begin{aligned}
&\Pr(\mathbf{X} \text{ not surjective}) \leq \sum_{i=1}^{l} \Pr(\mathbf{X}_2 \text{ not surjective mod } \mathfrak{q}_i) \\
&\leq r \cdot (1+4\varepsilon)^{m-2r+1} \cdot \left(\sum_{i=1}^{a} \frac{1}{\mathcal{N}(\mathfrak{q}_i)^{m-2r+1}} + \sum_{i=a+1}^{l} \frac{1}{B^{d(m-2r+1)}} \right) \\
&\leq r \cdot (1+4\varepsilon)^{m-2r+1} \cdot \sum_{i=1}^{l} \frac{1}{\mathcal{N}(\mathfrak{p}_i)^{m-2r+1}} \leq r \cdot (1+4\varepsilon)^{m-2r+1} \cdot P_\mathcal{K}(m-2r+1).
\end{aligned}$$

Now let us prove that B satisfying our condition exists and give a lower bound on it. The condition is that B is such that the first l primes of $\mathcal{O}_\mathcal{K}$ are of norm less than B^d. By [BS96, Theorem 8.7.4], the GRH implies that for $X \geq \log(\Delta_\mathcal{K})^{O(1)}$, the number of prime ideals $\pi_\mathcal{K}(X)$ of norm less than X is less than $1.1 \cdot X / \ln(X)$. This implies that as long as $B^d \geq \log(\Delta_\mathcal{K})^{O(1)}$, we have $\pi_\mathcal{K}(B^d) \leq B^d$. Then, as long as

$$l = r \cdot d^2 \ln(dr \cdot \varsigma) \leq B^d,$$

the norm of the first l prime ideals is less than B^d. □

3.3 Generating non-free modules

From now on, we fix a full-rank (not necessarily free) module $\mathcal{M} \subset \mathcal{K}^r$ given by a pseudo-basis $\mathcal{M} = \sum_{i=1}^{r} \mathbf{b}_i \cdot \mathcal{I}_i$ where $\mathcal{I}_i$ are fixed fractional ideals of $\mathcal{O}_\mathcal{K}$ and $\mathbf{b}_i \in \mathcal{K}^r$. Our goal is to generalise Lemma 3.4 for non-free $\mathcal{M}$. For $\mathcal{M} = \mathcal{O}_\mathcal{K}^r$, we proved that if sufficiently many vectors are sampled from $\mathcal{O}_\mathcal{K}^r$, then the matrix they form is surjective over $(\mathcal{O}_\mathcal{K}/\mathfrak{p})^r$ for any ideal $\mathfrak{p}$, and then surjective over $\mathcal{O}_\mathcal{K}^r$. We will prove the same result with $\mathcal{M}/\mathfrak{p}\mathcal{M}$ for not necessarily free $\mathcal{M}$.

Let $\mathfrak{p}$ be a prime ideal with residue field $\mathbb{F} = \mathcal{O}_\mathcal{K}/\mathfrak{p}$. The set $\mathcal{M}/\mathfrak{p}\mathcal{M}$ has a natural structure of $\mathbb{F}$-vector space of dimension r. For any matrix $\mathbf{B} \in \mathcal{M}^m$, we define $(\mathbf{B} \bmod \mathfrak{p})$ the associated matrix on $(\mathcal{M}/\mathfrak{p}\mathcal{M})^m$.

Theorem 3.5 (Assuming GRH). *There exists an absolute constant $c > 1$ such that the following holds. Let $r \geq 1$ and $m \geq 2r$ be integers. Let $\mathbf{C} = [\mathbf{c}_1, \ldots, \mathbf{c}_m] \in \mathbb{R}^{rd \times rm}$ be a centre matrix and $\varsigma > 0$. Let $\mathbf{S} \geq \sqrt{2}\varsigma\mathbf{I}$ or equal to $\varsigma\mathbf{I}$. Let $\mathcal{M} \subset \mathcal{K}^r$ be a module of rank r. Let $\varepsilon \in (0, 1/4)$. For any $B > \log(\Delta_\mathcal{K})^{c/d}$, assuming that the following two inequalities hold*

$$B^d \geq r \cdot d^2 \ln(dr \cdot \varsigma) \qquad \text{and} \qquad \varsigma \geq B \cdot \Delta_\mathcal{K}^{1/d} \cdot \eta_\varepsilon(\mathcal{M}),$$

and $\mathbf{X} = [\mathbf{x}_1, \ldots, \mathbf{x}_m] \in \mathcal{K}^{r \times m}$ *such that* $\mathbf{x}_i \leftarrow \mathcal{D}_{\mathcal{M},\mathbf{S},\mathbf{c}_i}$*, then*

$$\Pr_{\mathbf{X}}[\mathbf{X} \cdot \mathcal{O}_{\mathcal{K}}^m \neq \mathcal{M}] \leq \frac{m \cdot \varepsilon}{1-\varepsilon} + r \cdot \left(\frac{B}{2}\right)^{-(r+1)d} + r \cdot 2^{-dr} + (1+4\varepsilon)^{m-2r+1} \cdot r \cdot P_{\mathcal{K}}(m-2r+1),$$

where $P_{\mathcal{K}}(s) = \sum_{\mathfrak{p}} \mathcal{N}(\mathfrak{p})^{-s}$ *is the prime zeta function of* $\mathcal{K}$*.*

Proof. By the same arguments than the proof of Lemma 3.4, we can assume without loss of generality that $\mathcal{M} = \oplus_{i=1}^{r} I_i$ with $I_1, \ldots, I_r$ ideals, and $\mathbf{S} = \varsigma \mathbf{I}_{rd}$ with $\varsigma \geq B \cdot \Delta_{\mathcal{K}}^{1/d} \cdot \eta_{\varepsilon}(\oplus_{1 \leq i \leq d} I_i)$. In particular, it holds that $\varsigma \geq B \cdot \Delta_{\mathcal{K}}^{1/d} \cdot \eta_{\varepsilon}(I_i)$ for any $1 \leq i \leq r$, in that case, Lemma 3.3 implies that Lemma 3.1 holds for any submatrix of $\mathbf{X}$, we can then run the exact same proof as the one of Lemma 3.4. □

3.4 Asymptotic behavior of $P_{\mathcal{K}}$

The upper bounds in Lemma 3.3 and Theorem 3.5 are described in terms of the prime zeta function $P_{\mathcal{K}}$. In the following, we upper bound $P_{\mathcal{K}}$ in terms of the smallest prime ideal norm $N_{\mathcal{K}}$, for which we know the explicit value when $\mathcal{K}$ is a cyclotomic field of prime-power conductor.

Lemma 3.6. *For any number field* $\mathcal{K}$ *such that there are exactly* $a_{\mathcal{K}}$ *prime ideals of norm* $N_{\mathcal{K}}$*, let* $\kappa_{\mathcal{K}} \in (0,1)$ *be the ratio between the norm of the smallest ideal and the norm of the second-smallest. Then for any* $x \geq 2$*, it holds that*

$$P_{\mathcal{K}}(x) \leq \frac{a_{\mathcal{K}}}{N_{\mathcal{K}}^x} \cdot \left(1 + \frac{2d}{a_{\mathcal{K}}} \cdot N_{\mathcal{K}} \cdot \kappa_{\mathcal{K}}^{x-1}\right) \leq \frac{3d}{N_{\mathcal{K}}^{x-1}} \ .$$

Proof. Let N be the norm of the second-smallest prime ideal of $\mathcal{O}_{\mathcal{K}}$, i.e. $\kappa_{\mathcal{K}} = N_{\mathcal{K}}/N$. For any $n \geq N$, let a_n be the number of prime ideals of norm N. For any $x \geq 1$, let $\pi_{\mathcal{K}}(x)$ be the number of prime ideals of norm $\leq x$. We use Abel's summation formula and the fact that $\pi_{\mathcal{K}}(x) \leq d \cdot x$ for any $x \geq 2$:

$$\begin{aligned} P_{\mathcal{K}}(x) - a_{\mathcal{K}} \cdot N_{\mathcal{K}}^{-x} = \sum_{n \geq N} \frac{a_n}{n^x} &= \lim_{u \to \infty} \left(\frac{\pi_{\mathcal{K}}(u)}{u^x}\right) + \int_N^{\infty} \pi_{\mathcal{K}}(u) \cdot \frac{x \cdot \mathrm{d}u}{u^{x+1}} \\ &\leq d \cdot x \cdot \int_N^{\infty} u^{-x} \mathrm{d}u = d \cdot \frac{x}{x-1} \cdot N^{-x+1} \leq 2d \cdot N^{-x+1}. \end{aligned}$$

Substituting $N = N_{\mathcal{K}}/\kappa_{\mathcal{K}}$ and rearranging gives the first inequality

$$P_{\mathcal{K}}(x) \leq \frac{a_{\mathcal{K}}}{N_{\mathcal{K}}^x} \cdot \left(1 + \frac{2d}{a_{\mathcal{K}}} \cdot N_{\mathcal{K}} \cdot \kappa_{\mathcal{K}}^{x-1}\right) \ .$$

The second inequality follows from $a_{\mathcal{K}} \leq d$ and $\kappa_{\mathcal{K}} \leq 1$. □

Lemma 3.7. *Let* $q = p^l$ *be a power of a prime and let* $\mathcal{K}_q$ *be the* q*-th cyclotomic field. Then* $N_{\mathcal{K}_q} = p$ *and the norm of the second-smallest ideal is at least* $q+1$*.*

Proof. First, it holds that p totally ramifies in $\mathcal{K}_q$, so an ideal of norm p exists in $\mathcal{K}_q$. Now let us prove that there is no smaller ideal in this field. Let $N \neq p$ be a prime and $\mathfrak{p}$ a prime of $\mathcal{K}_q$ above N. The norm of $\mathfrak{p}$ is N^l where l is the smallest integer such that $N^l = 1 \bmod q$, in particular, since $N^l \neq 1$, it holds that $N^l \geq q + 1$. □

In general, for any number field $\mathcal{K}$ there exists $b_\mathcal{K}$ such that $P_\mathcal{K}(x) \sim_{x\to\infty} b_\mathcal{K} \cdot N_\mathcal{K}^{-x}$, we provide a sagemath script computing an approximation of $b_\mathcal{K}$ for any cyclotomic number field. Note that $\kappa_\mathcal{K}$ can be arbitrarily close to 1: if $p = 2^l - 1$ is a Mersenne prime, then one can prove that $\kappa_\mathcal{K} = (p-1)/p$.

In the ideal case ($r = 1$), Theorem 3.5 can be compared to [SS11, Lemma 4.4], where the probability of generating $\mathcal{O}_\mathcal{K}$ with two elements is proven close to $1/\zeta_\mathcal{K}(2)$ for power-of-two cyclotomic number fields (note that for $m = 2, r = 1$, our upper bound is equal to ∞). A similar proof can be used to prove that for any number field, for suitable parameters, the probability of generating any ideal $\mathfrak{a}$ with m elements of $\mathfrak{a}$ is close to $1/\zeta_\mathcal{K}(m)$. One can note that when m goes to infinity, it holds that $1/\zeta_\mathcal{K}(m) \sim P_\mathcal{K}(m) \sim 1 - N_\mathcal{K}^{-m}$.

3.5 Specialised bounds

We summarise by giving bounds for Gaussian matrices being surjective for general number fields and for the special case of prime-power cyclotomic fields.

Theorem 3.8 (General Number Field, Assuming GRH). *Let $\mathcal{K}$ be a number field of degree $d \geq 2$ and of smallest ideal norm $N_\mathcal{K}$. Let $0 < \varepsilon < 2^{-6}$. Let $\mathcal{M} \subset \mathcal{K}^r$ be a full-rank module. Then for any $\mathbf{S} \in \mathbb{R}^{dr\times dr}$, $m \geq 1$ and centre $\mathbf{C} \in \mathcal{K}_\mathbb{R}^{r\times m}$ such that*

$$\begin{aligned} &\mathbf{S} \geq 9\sqrt{2} \cdot r^{\frac{1}{d}} \cdot \log\left(r \cdot \Delta_K^{\frac{1}{d}} \cdot \eta_\varepsilon(\mathcal{M})\right)^{\frac{1}{d}} \cdot \Delta_\mathcal{K}^{\frac{1}{d}} \cdot \eta_\varepsilon(\mathcal{M}) \qquad \text{and} \\ &m \geq 2r + \frac{\log(1/\varepsilon) + 1 + \log(2d)}{\log(N_\mathcal{K}/1.4)}, \end{aligned}$$

it holds that

$$\Pr_{\mathbf{X}\leftarrow\mathcal{D}^m_{\mathcal{M},\mathbf{S},\mathbf{C}}}(\mathbf{X} \cdot \mathcal{O}_\mathcal{K}^m \neq \mathcal{M}) \leq m \cdot (\varepsilon + 2^{-dr}).$$

Proof. We show below that Theorem 3.5 holds for some $B \geq 4$. The values of the parameters and Lemma 3.6 give

$$\begin{aligned} &\frac{m\cdot\varepsilon}{1-\varepsilon} + r\left(\frac{B}{2}\right)^{-(r+1)d} + r2^{-dr} \ + (1+4\varepsilon)^{m-2r+1} \cdot r \cdot P_\mathcal{K}(m-2r+1) \\ &\leq 1.2 \cdot m \cdot \varepsilon + r \cdot 2^{-d} \cdot 2^{-dr} + r2^{-dr} + r(1.4)^{m-2r+1} \cdot 3d \cdot N_\mathcal{K}^{-(m-2r)} \\ &\leq 1.2 \cdot m \cdot \varepsilon + r \cdot (2^{-d} \cdot 2^{-dr} + 2^{-dr} + 1.4 \cdot \varepsilon) \\ &\leq \varepsilon \cdot (1.2m + 1.4r) + 2r \cdot 2^{-dr} \leq m \cdot (\varepsilon + 2^{-dr}). \end{aligned}$$

Now we argue for the existence of B. Without loss of generality we assume $\mathbf{\Sigma} = \varsigma \cdot \mathbf{I}$. The hypotheses of Theorem 3.5 are satisfied if $B^d \geq r \cdot d^2 \ln(dr \cdot \varsigma)$ and $\varsigma \geq B \cdot \Delta_{\mathcal{K}}^{1/d} \cdot \eta_\varepsilon(\mathcal{M})$, that is to say if B satisfies

$$B^d \geq rd^2 \cdot \log(dr \cdot \Delta_{\mathcal{K}} \cdot \eta_\varepsilon(\mathcal{M}) \cdot B) = rd \log(B^d) + rd^2 \log(dr \cdot \Delta_{\mathcal{K}} \cdot \eta_\varepsilon(\mathcal{M})).$$

By Lemma B.5, this inequality holds for any B satisfying

$$B^d \geq 2\left(rd^2 \log(dr \cdot \Delta_{\mathcal{K}}^{1/d} \cdot \eta_\varepsilon(\mathcal{M})) + rd \log(rd)\right).$$

It holds that

$$\begin{aligned}
&\left(2\left(rd^2 \log(dr \cdot \Delta_{\mathcal{K}}^{1/d} \cdot \eta_\varepsilon(\mathcal{M})) + rd \log(rd)\right)\right)^{1/d} \\
\leq\ &\left(2r \cdot \left((d^2+d)\log(d) + d^2 \log(r\Delta_K^{1/d} \cdot \eta_\varepsilon(\mathcal{M})) + d\log(r)\right)\right)^{1/d} \\
\leq\ &(2r)^{1/d} \cdot \left(\left((d^2+d)\log(d)\right)^{1/d} + \left(d^2 \log(r\Delta_K^{1/d} \cdot \eta_\varepsilon(\mathcal{M}))\right)^{1/d} + (d\log(r))^{1/d}\right) \\
\leq\ &(2r)^{1/d} \cdot \left(2.5 + 2\log(r\Delta_K^{1/d} \cdot \eta_\varepsilon(\mathcal{M}))^{1/d} + \sqrt{2}\log(r)\right) \\
\leq\ &9 \cdot r^{1/d} \cdot \log\left(r\Delta_K^{1/d} \cdot \eta_\varepsilon(\mathcal{M})\right)^{\frac{1}{d}},
\end{aligned}$$

hence taking $B = 9 \cdot r^{1/d} \cdot \log(r\Delta_K^{1/d} \cdot \eta_\varepsilon(\mathcal{M}))^{1/d}$ gives us the desired result. □

Corollary 3.9 (Prime-Power Cyclotomics, Assuming GRH). *Let $p^k > 2$ be a prime power, and $\mathcal{K} = \mathbb{Q}(\zeta_{p^k})$ be the p^k-cyclotomic number field of degree $d = (p-1) \cdot p^{k-1}$. Let $0 < \varepsilon < 2^{-6}$. Let $\mathcal{M} \subset \mathcal{K}^r$ be a full-rank module. Then for any $\mathbf{S} \in \mathbb{R}^{dr \times dr}$, $m \geq 1$ and centre $\mathbf{C} \in \mathcal{K}_{\mathbb{R}}^{r \times m}$ such that*

$$\mathbf{S} \geq 9\sqrt{2} \cdot r^{\frac{1}{d}} \cdot \log\left(rd \cdot \eta_\varepsilon(\mathcal{M})\right)^{\frac{1}{d}} \cdot d \cdot \eta_\varepsilon(\mathcal{M}) \text{ and } m \geq 2r + \frac{\log(1/\varepsilon) + 1 + \log(2d)}{\log(p/1.4)},$$

it holds that

$$\Pr_{\mathbf{X} \leftarrow \mathcal{D}^m_{\mathcal{M},\mathbf{S},\mathbf{C}}}(\mathbf{X} \cdot \mathcal{O}_{\mathcal{K}}^m \neq \mathcal{M}) \leq m \cdot (\varepsilon + 2^{-dr}).$$

Proof. Follows from Theorem 3.8 and Lemmas 3.7 and B.3. □

4 Smoothing parameter of kernel lattices

In this section we upper bound $\eta_\varepsilon(\Lambda^\perp(\mathbf{X}))$ for $\mathbf{X} = [\mathbf{x}_1 \mid \cdots \mid \mathbf{x}_m] \in \mathcal{O}_{\mathcal{K}}^{r \times m}$, where $\mathbf{x}_i \leftarrow \mathcal{D}_{\mathcal{O}_{\mathcal{K}}^r,\varsigma}$. We emphasise that the lattice $\Lambda^\perp(\mathbf{X})$ of interest here is the rank-deficient $\mathcal{O}_{\mathcal{K}}$-kernel lattice of $\mathbf{X}$ but not the full-rank q-ary lattice $\Lambda_q^\perp(\mathbf{X}) \coloneqq \{\mathbf{x} \in \mathcal{O}_{\mathcal{K}}^m, \mathbf{X} \cdot \mathbf{x} = \mathbf{0} \bmod q\}$.

4.1 Baseline: Exponential bound for $m > r$

We first write down an approach which gives an upper bound of $\eta_\varepsilon(\Lambda^\perp(\mathbf{X}))$ which is exponential in r. It uses the short vectors in $\mathbf{X}$ to construct $m-r$ short linearly independent kernel vectors. Via the canonical embedding of $\mathcal{O}_\mathcal{K}$ we transform them into $d\cdot(m-r)$ short linearly independent kernel vectors in $\mathbb{R}^{dm}$. This way we obtain a bound on $\lambda_{d(m-r)}(\Lambda^\perp(\mathbf{X}))$ and the smoothing parameter is exponential in r. This can be seen as extending the approach of [LSS14] to $r>1, m>2$. The proof also relies on ideas from [BF11, Thm 4.2 (eprint)].

Lemma 4.1. *Let $\varsigma, t > 0$ and $0 < r < m$ be integers. Let $\mathbf{X} = [\mathbf{x}_1 \mid \cdots \mid \mathbf{x}_m] \in \mathcal{O}_\mathcal{K}^{r\times m}$, where $\mathbf{x}_i \leftarrow \mathcal{D}_{\mathcal{O}_\mathcal{K}^r,\varsigma}$. We denote $\mathbf{X}_1 \coloneqq [\mathbf{x}_1 \mid \cdots \mid \mathbf{x}_r]$, $\mathbf{X}_2 \coloneqq [\mathbf{x}_r \mid \cdots \mid \mathbf{x}_m]$ and assume $\mathbf{X}_1 \in GL_r(\mathcal{K})$. Then*

$$\Pr_{\mathbf{X}}\left(\eta_\varepsilon(\Lambda^\perp(\mathbf{X})) \le \delta_\mathcal{K}\cdot\sqrt{rd}\cdot(\sqrt{r}\cdot\varsigma\cdot t)^r\cdot\eta_\varepsilon^{(d\cdot(m-r))}\right) \ge 1 - 2me\cdot rd\cdot e^{-\pi t^2}.$$

Proof. The inverse of $\mathbf{X}_1$ can be written as $\mathbf{X}_1^{-1} = \frac{1}{\det(\mathbf{X}_1)}\cdot\mathrm{adj}(\mathbf{X}_1)$ where $\mathrm{adj}(\mathbf{X}_1)$ has entries in $\mathcal{O}_\mathcal{K}$. Then $\det(\mathbf{X}_1)\cdot\mathbf{X}_1^{-1}$ is a matrix over the ring and

$$[\mathbf{X}_1 \mid \mathbf{X}_2]\cdot\begin{bmatrix}\det(\mathbf{X}_1)\cdot\mathbf{X}_1^{-1} & \mathbf{0}\\ \mathbf{0} & \mathbf{I}_{m-r}\end{bmatrix} = [\det(\mathbf{X}_1)\cdot\mathbf{I}_r \mid \mathbf{X}_2].$$

The kernel of $[\det(\mathbf{X}_1)\cdot\mathbf{I}_r \mid \mathbf{X}_2]$ contains $\mathcal{K}$ linearly independent vectors $(\mathbf{x}_{r+i} \,||\, \det(\mathbf{X}_1)\cdot\mathbf{e}_i)$ for $i\in[m-r]$. Hence, vectors $\det(\mathbf{X}_1)\cdot(\mathbf{X}_1^{-1}\mathbf{x}_{r+i} \,||\, \mathbf{e}_i) \in \mathcal{O}_\mathcal{K}^m$ for $i\in[m-r]$ are also linearly independent and in the kernel of $\mathbf{X}$. It remains to bound the norm of these vectors. Denote $\mathbf{y}_i \coloneqq \det(\mathbf{X}_1)\cdot\mathbf{X}_1^{-1}\mathbf{x}_{r+i} \in \mathcal{O}_\mathcal{K}^r$ and let y_{ij} be the j-th coefficient of $\mathbf{y}_i$. Then $\mathbf{X}_1\cdot\mathbf{y}_i = \det(\mathbf{X}_1)\cdot\mathbf{x}_{r+i}$. By Cramer's rule the only solution to this linear equation over $\mathcal{K}$ can be expressed as $y_{ij} = \det(\tilde{\mathbf{X}}_j)/\det(\mathbf{X}_1)$. Here $\tilde{\mathbf{X}}_j$ stands for the matrix $\mathbf{X}_1$ with its j-th column replaced by $\det(\mathbf{X}_1)\cdot\mathbf{x}_{r+i}$. Cancelling the $\det(\mathbf{X}_1)$ factor we get $y_{ij} = \det\left[\mathbf{x}_1 \mid \ldots \mathbf{x}_{r+i} \ldots \mid \mathbf{x}_r\right]$. Hence,

$$\begin{aligned}
\|\Phi(\mathbf{y}_i)\|_\infty &= \max_{j\in[r]}\left(\left\|\Phi(\det\left[\mathbf{x}_1 \mid \ldots \mathbf{x}_{r+i} \ldots \mid \mathbf{x}_r\right])\right\|_\infty\right)\\
&= \max_{j\in[r]}\max_{k\in[d]}\det\left[\sigma_k(\mathbf{x}_1) \mid \ldots \sigma_k(\mathbf{x}_{r+i}) \ldots \mid \sigma_k(\mathbf{x}_r)\right]\\
&\le \max_{j\in[r]}\max_{k\in[d]}\|\sigma_k(\mathbf{x}_{r+i})\|\cdot\prod_{\ell\in[r]\setminus\{j\}}\|\sigma_k(\mathbf{x}_\ell)\|\\
&\le \max_{j\in[r]}\left(\sqrt{r}\cdot\|\mathbf{x}_{r+i}\|_\infty\cdot\prod_{\ell\in[r]\setminus\{j\}}\sqrt{r}\cdot\|\mathbf{x}_\ell\|_\infty\right)\\
&\le (\sqrt{r}\cdot\varsigma\cdot t)^r
\end{aligned}$$

with probability at least $1 - r\cdot 2edr\cdot e^{-\pi t^2}$. Here we apply Hadamard's inequality for the first upper bound, the norm inequalities for the second and the tail bound from Lemma 2.6 for the last transition. Then for the Euclidean norm

we get $\|\mathbf{y}_i\| \le \sqrt{dr} \cdot (\sqrt{r} \cdot \varsigma \cdot t)^r$ for all $i \in [m-r]$. This holds with probability $1 - m \cdot 2edr \cdot e^{-\pi t^2}$, since the tail bound must be true for all $\mathbf{x}_\ell$, $\ell \in [m]$.

As a result, we obtain $m - r$ vectors of bounded norm that are linearly independent over $\mathcal{K}$. Multiplying by the ring basis $\mathbf{B}_{\mathcal{O}_\mathcal{K}}$ we extend this set to $d \cdot (m-r)$ linearly independent vectors in $\Phi(\Lambda^\perp(\mathbf{X})) \subset \mathbb{R}^{dm}$ with norm bounded by $\delta_\mathcal{K} \cdot \sqrt{rd} \cdot (\sqrt{r} \cdot \varsigma \cdot t)^r$. Applying a classic bound on smoothing from §2.3, we get the statement. □

4.2 Polynomial bound for $m \ge \Omega(\log(\epsilon^{-1})/\log(f))$ in extensions of f-th cyclotomic field

When $r \ge \omega(1)$, which is typically the case when the ring degree d is small, the strategy in §4.1 leads to a superpolynomial bound on $\lambda_{d(m-r)}(\Lambda^\perp(\mathbf{X}))$ which is often not usable in applications. For a more useful bound we turn to the approach of [AR16] which focuses on the setting $\mathcal{K} = \mathbb{Q}$. We extend their strategy to any number field $\mathcal{K} \supseteq \mathbb{Q}(\zeta_f)$ containing the f-th cyclotomic field where the quality of the obtained bound depends on f. We recall that the cyclotomic fields with conductor 2 is $\mathbb{Q}$, and that all lemmas stated in this section hold in particular for any $\mathcal{K}$ with $f = 2$. For a high level description of our proof strategy, see §1.1. For readability, we choose to name all the implied absolute constants of the theorems and lemmas of this section C. We highlight that those refer to possibly different constants.

Theorem 4.2. *There exists an absolute constant $C > 2$ such that the following holds. Let $\mathcal{K}$ be a number field of degree d containing the cyclotomic field of conductor f. Let $\mathbf{S} \ge \max(\eta_{1/2}(\mathcal{O}_\mathcal{K}^r), \sqrt{2\pi d})$, and $\mathbf{X} \in \mathcal{O}_\mathcal{K}^{r \times m}$ be a matrix with columns sampled from $\mathcal{D}_{\mathcal{O}_\mathcal{K}^r, \mathbf{S}}$ for some $m \ge 1$. Let $\varepsilon > 0$, and $T_{\mathcal{K},r,\mathbf{S}} = r \cdot \log(dr \cdot \Delta_\mathcal{K}^{1/(2d)} \cdot s_{\min}(\mathbf{S}))$. For*

$$m \ge C \cdot T_{\mathcal{K},r,\mathbf{S}} \cdot \log(T_{\mathcal{K},r,\mathbf{S}}) + \frac{2\log(1/\varepsilon)}{\log(f/1.8)},$$

we have with probability $\ge 1 - \varepsilon$ that

$$\eta_\varepsilon(\Lambda^\perp(\mathbf{X})) = O\left(\eta_\varepsilon^{((m-r)d)} \cdot m^{1.5} \cdot d^{12} \cdot \delta_\mathcal{K}^{14} \cdot f^2 \cdot \Delta_\mathcal{K}^{4/d} \cdot s_{\min}(\mathbf{S})\right).$$

Corollary 4.3 (Cyclotomics version of Theorem 4.2). *There exists an absolute constant $C > 2$ such that the following holds. Let $f > 2$ and $\mathcal{K} = \mathbb{Q}(\zeta_{p^k})$ be the cyclotomic number field of conductor f and degree d. Let $1 \le r \le m$, $\varepsilon > 0$ and $T = r \cdot \log(r \cdot d^{3/2} \cdot s_{\min}(\mathbf{S}_X))$. Let $\mathbf{S}_X \in \mathbb{R}^{dr \times dr}$ be positive definite, $\mathbf{X} \in \mathcal{O}_\mathcal{K}^{r \times m}$ be a matrix with columns sampled from $\mathcal{D}_{\mathcal{O}_\mathcal{K}^r, \mathbf{S}_X}$ and $\mathbf{c} \in \mathcal{O}_\mathcal{K}^m$. Then for*

$$\mathbf{S}_X \ge \max(\eta_{1/2}(\mathcal{O}_\mathcal{K}^r), \sqrt{2\pi d}), \qquad m \ge C \cdot T \cdot \log(T) + \frac{2\log(1/\varepsilon)}{\log(p^k/1.8)},$$

with probability $\ge 1 - \varepsilon$ we have

$$\eta_\varepsilon(\Lambda^\perp(\mathbf{X})) = O\left(\eta_\varepsilon^{((m-r)d)} \cdot m^{1.5} \cdot f^2 \cdot d^{16} \cdot s_{\min}(\mathbf{S}_X)\right).$$

Note that the lower bound on m in Theorem 4.2 has two summands, with only the second one ($2\log(1/\varepsilon)/\log(f/1.8)$) related to the failure probability of the theorem. This is because the theorem has a sharp threshold in m, due to the "pigeonhole" nature of Lemma 4.7. That is, we show that to prove the existence of a short vector in $\Lambda(\mathbf{X})$, m must be larger than $C \cdot T_{\mathcal{K},r,\mathbf{S}} \cdot \log(T_{\mathcal{K},r,\mathbf{S}})$. Once this threshold is reached, we show in Lemma 4.9 that we can increase the width of $\mathbf{X}$ to lower the failure probability. Indeed, we show that adding a column to $\mathbf{X}$ multiplies the probability of failure of the theorem by $1 - 1/f$, where $\zeta_f \in \mathcal{K}$ is a root of unity. This leads to the $\log(1/\varepsilon)/\log(f)$ factor in the lower bound on m.

Note that this improves on [AR16] lifted to modules in two ways. The second summand is super-linear in r and logarithmic in d in contrast to lifting [AR16] directly, where the lower bound would be super-linear in $r \cdot d$. The first summand is also smaller by a factor of $\log f$.[8]

To prove Theorem 4.2, we state in Lemma 4.5 that for a large enough set $A \subseteq \mathcal{O}_\mathcal{K}$ there exists a preimage of a small multiple of the i-th unit vector $\mathbf{e}_i \in \mathcal{O}_\mathcal{K}^r$ for any i. This almost generalises Lemma 4.2 in [AR16] to the ring setting except that instead of obtaining a preimage of $\mathbf{e}_i$ we only obtain a preimage of a small multiple of $f \cdot \mathbf{e}_i$.

Definition 4.4. *We say that a set $A \subset \mathcal{O}_\mathcal{K}$ is B-admissible if it is symmetric ($-A = A$); it contains 0 and 1; its size is at least 2^d; all of $x \in A$ satisfy $\|x\|_\infty \leq B$. If a set A is B-admissible, we define*

$$A_{\Pi} = \{a \cdot b \mid a, b \in A\}, \quad A_{\Sigma\Pi} = \{a + b \mid a, b \in A_{\Pi}\}.$$

Lemma 4.5. *There exists an absolute constant $C > 2$ such that the following holds. Let $\mathcal{K} \supseteq \mathbb{Q}(\zeta_f)$ be a number field of degree d containing the cyclotomic field of conductor f. Let $\mathbf{S} \geq \max(\eta_{1/2}(\mathcal{O}_\mathcal{K}^r), \sqrt{2\pi d})$ and let $\mathbf{X} \in \mathcal{O}_\mathcal{K}^{r \times m}$ be a matrix with columns sampled from $\mathcal{D}_{\mathcal{O}_\mathcal{K}^r,\mathbf{S}}$ for some $m \geq 1$.*

Let $\varepsilon > 0$, $1 \leq i \leq r$, and $A \subset \mathcal{O}_\mathcal{K}$ be B-admissible set for some $B > 0$. Let $T = T_{\mathcal{K},r,B,\mathbf{S}} = r \cdot \log(r \cdot B \cdot s_{\min}(\mathbf{S})/\Delta_K^{1/(2d)})$, then for any

$$m \geq C \cdot T \cdot \log(T) + \frac{2\log(1/\varepsilon)}{\log(f/1.8)},$$

there exists $\mathbf{c} \in A_{\Sigma\Pi}^m, u \in \mathcal{O}_\mathcal{K}$ with $\|u\|_\infty \leq f^2/2$ and $a' \in A_{\Pi}$ such that $\mathbf{X} \cdot u \cdot \mathbf{c} = f \cdot a' \cdot \mathbf{e}_i$ with probability $\geq 1 - \varepsilon$ over the randomness of $\mathbf{X}$.

4.3 Proof of Lemma 4.5

In this subsection, we fix a number field $\mathcal{K} \supseteq \mathbb{Q}(\zeta_f)$ of degree d containing the cyclotomic field of conductor f, $\mathbf{S} \geq \max(\sqrt{2\pi d}, \eta_{1/2}(\mathcal{O}_\mathcal{K}^r))$, and $\mathbf{X} \in \mathcal{O}_\mathcal{K}^{r \times m}$ a matrix with columns sampled from $\mathcal{D}_{\mathcal{O}_\mathcal{K}^r,\mathbf{S}}$ for some $m \geq 1$.

[8] Reducing this part of the bound by a factor of d seems unlikely due to the rich structure of $\mathcal{K}$, i.e. the d $\mathbb{Z}$-vectors corresponding to an $\mathcal{O}_\mathcal{K}$-vector are highly correlated.

It remains to prove Lemma 4.5. Without loss of generality we take $i = 1$. We fix a B-admissible set A and define two set sequences, for $1 \le j \le m$:

$$S_j = \left\{ \sum_{i=1}^{j} a_i \cdot \mathbf{x}_i,\ a_i \in A \right\}, \quad \hat{S}_j = \{\mathbf{s}/a,\ \mathbf{s} \in S_j, a \in A \setminus \{0\}\}$$

For $1 \le j \le m$ we define the following events:

$\mathtt{Win}_j$: Two sets in $\left\{\hat{S}_j, \hat{S}_j + \zeta_f^x \cdot \mathbf{e_1} \mid x \in [f]\right\}$ have non-empty intersection,

$\mathtt{Gain}_j$: $\mathbf{x}_{j+1} \notin \hat{S}_j$ and $\|\mathbf{x}_{j+1}\| \le \sqrt{dr} \cdot s_{\min}(\mathbf{S})$.

The following Lemma 4.6 shows that the event $\mathtt{Win}_j$ implies the existence of a short preimage of a multiple of $\mathbf{e}_1$.

Lemma 4.6. *Let $j \ge 1$. If $\mathtt{Win}_j$ is true, then there exists $\mathbf{c} \in (A_{\Sigma\Pi})^j$, $u \in \mathcal{O}_\mathcal{K}$ with $\|u\|_\infty \le f^2/2$ and $a' \in A_\Pi$ such that $\mathbf{X}_{[1:j]} \cdot u \cdot \mathbf{c} = f \cdot a' \cdot \mathbf{e}_1$, where A_Π and $A_{\Sigma\Pi}$ are defined as in Lemma 4.5.*

Proof. Since $\mathtt{Win}_j$ is true, two cases can happen: (1) $(\hat{S}_j + \zeta_f^x \cdot \mathbf{e_1}) \cap (\hat{S}_j + \zeta_f^y \cdot \mathbf{e_1}) \neq \emptyset$ for some $x, y \in [f]$; (2) $\hat{S}_j \cap (\hat{S}_j + \zeta_f^y \cdot \mathbf{e_1}) \neq \emptyset$ for some $y \in [f]$. We focus on Case (1) and note that Case (2) can be treated similarly. In Case (1), there exist $\mathbf{a}_1, \mathbf{a}_2 \in A^j$ and $b_1, b_2 \in A$ such that $\mathbf{X} \cdot (\mathbf{a}_1/b_1 - \mathbf{a}_2/b_2) = (\zeta_f^x - \zeta_f^y) \cdot \mathbf{e}_1$. Multiplying the previous equality by $b_1 \cdot b_2 \cdot f/(\zeta_f^x - \zeta_f^y)$ and applying Lemma B.4 gives the result. □

Lemma 4.7 ([AR16, Clm. 4.1 adapted]). *There exists an absolute constant $c_1 > 1$ such that, for any sequence $1 \le i_1 < \ldots < i_k$ of integers of length $k \ge c_1 \cdot r \log(r \cdot B \cdot s_{\min}(\mathbf{S})/\Delta_K^{1/(2d)})$, if $\mathtt{Gain}_{i_l}$ is true for $l \in [1, k+1]$ then $\mathtt{Win}_{i_k}$ is true.*

Proof. Assume that $\mathtt{Win}_{i_k}$ is not true, we prove the statement by contradiction. Without loss of generality, we also remove all $\mathbf{x}_j$ for which $\mathtt{Gain}_j$ is false, and consider $i_l = l$ for $l \in [1, k+1]$. For all $l \in [1, k+1]$ and $a \in A$, we have that $a \cdot \mathbf{x}_{l+1} \notin S_l$, since else $\mathbf{x}_{l+1} \in \hat{S}_l$ which is forbidden by $\mathtt{Gain}_l$. This implies that the size of S_l is at least $|A|^l \ge 2^{dl}$. Then, since the sets $\hat{S}_j$ are non-intersecting, we have $|S_l| + \sum_a \left|S_l + \zeta_f^a \cdot \mathbf{e}_1\right| \ge (f+1)2^{dl}$.

Furthermore, note that every element of S_l and $S_l + \zeta_f^a \cdot \mathbf{e}_1$ has norm bounded by $2 \cdot \sqrt{rd} \cdot B \cdot s_{\min}(\mathbf{S}) \cdot l$. By Lemma 2.13 and the condition on $\mathbf{S}$, the set $\mathcal{O}_\mathcal{K}^r \cap B(0, C)$ is of size at most $1.5 \cdot (2\exp(\pi) B \cdot l \cdot s_{\min}(\mathbf{S})/\Delta_K^{1/(2d)})^{rd} \le C^{rd}$ for $C = O(B \cdot l \cdot s_{\min}(\mathbf{S})/\Delta_K^{1/(2d)})$. This implies that for k large enough we have $(f+1)2^{dk} > C^{rd}$, and a pair of sets in $\left\{S_j, S_j + \zeta_f^a \cdot \mathbf{e_1} \mid a \in [f]\right\}$ must have a non-empty intersection by the pigeonhole principle. This means that $\mathtt{Win}_k$ is true, and we have a contradiction. Using Lemma B.5, it can be computed that there exists an absolute $c' > 1$ such that $(f+1) \cdot 2^{dk} > C^{rd}$ holds for any $k \ge c' \cdot r \log(r \cdot B \cdot s_{\min}(\mathbf{S})/\Delta_K^{1/(2d)})$. □

We now show that as long as $\texttt{Win}_j$ is false, the event $\texttt{Gain}_j$ happens with high probability.

Lemma 4.8. *Let* $\mathbf{S} \geq \sqrt{2\pi d}$ *and* $j \geq 1$. *It holds that* $\Pr(\texttt{Gain}_j \mid \neg\texttt{Win}_j) \geq 1 - 1.8/f$.

Proof. For clarity, we omit the "$\mid \neg\texttt{Win}_j$" from the probabilities in this proof. Note that

$$\Pr(\neg\texttt{Gain}_j) = \frac{\rho_{\mathbf{S}}(\hat{S}_j \cup \mathcal{O}^r_{\mathcal{K}} \setminus B(0,R))}{\rho_{\mathbf{S}}(\mathcal{O}^r_{\mathcal{K}})} \leq \frac{\rho_{\mathbf{S}}(\hat{S}_j)}{\rho_{\mathbf{S}}(\mathcal{O}^r_{\mathcal{K}})} + \beta_{dr}(\sqrt{2\pi d}),$$

where $R = \sqrt{dr}s_1(\mathbf{S})$ and the inequality follows from Lemma 2.7. One can prove that it holds that $\beta_{rd}(\sqrt{2\pi d}) \leq 1/(10d)$ for any $d \geq 2$, $r \geq 1$. For any $i \in [1, f]$, by Lemma 2.11 and the bound on $\mathbf{S}$, we have that $\rho_{\mathbf{S}}(\hat{S}_j + \zeta^i_f \mathbf{e}_1) \geq \rho_{\mathbf{S}}(\hat{S}_j)/\sqrt{e}$. Now since $\texttt{Win}_j$ is false, it holds that $\rho_{\mathbf{S}}(\hat{S}_j) + \sum_i \rho_{\mathbf{S}}(\hat{S}_j + \zeta^i_f \mathbf{e}_1) \leq \rho_{\mathbf{S}}(\mathcal{O}^r_{\mathcal{K}})$, and then $\rho_{\mathbf{S}}(\hat{S}_j)/\rho_{\mathbf{S}}(\mathcal{O}^r_{\mathcal{K}}) \leq 1/(1 + f/\sqrt{e}) \leq \sqrt{e}/f$. Finally, $\Pr(\neg\texttt{Gain}_j) \leq 1.8/f$ □

Lemma 4.9. *There exists an absolute constant* $c_2 > 2$ *such that the following holds. Let* $Y_1, Y_2, \ldots$ *be a sequence of not necessarily independent Boolean random variables such that for any* $i \geq 1$ *and any* $\{y_1, \ldots y_{i-1}\} \in \{0,1\}^{i-1}$ *it holds that*

$$\Pr(Y_i = 1 \mid Y_{i-1} = y_{i-1}, \ldots Y_1 = y_1) \geq 1 - \alpha \tag{2}$$

for some $\alpha \in (0, 2/3)$. *Let* $T \geq 3$ *be a target value. Then for any* $\varepsilon \in (0, 1/2)$ *and any*

$$m \geq c_2 \cdot T\log(T) + 2 \cdot \frac{\log(1/\varepsilon)}{\log(1/\alpha)},$$

it holds that

$$\Pr(\sum_{i=1}^{m} Y_i \leq T) \leq \varepsilon.$$

Furthermore, if we assume $T \geq 13$ *and* $\alpha \leq 0.1$, *one can choose* $c_2 = 4$.

Lemma 4.10. *Let* $Y_1, Y_2, \ldots$ *be a sequence of not necessarily independent Boolean random variables such that for any* $i \geq 1$ *and any* $\{y_1, \ldots y_{i-1}\} \in \{0,1\}^{i-1}$ *it holds that*

$$\Pr(Y_i = 1 \mid Y_{i-1} = y_{i-1}, \ldots Y_1 = y_1) \geq 1 - \alpha$$

for some $\alpha \in (0,1)$. *Let* $\tau \in \{0,1\}^m$ *and* $|\tau|$ *its Hamming weight. It holds that*

$$\Pr((Y_1, \ldots, Y_m) = \tau) \leq \alpha^{m-|\tau|}.$$

Proof. It holds that

$$\Pr((Y_1, \ldots, Y_m) = \tau) = \prod_{i=1}^{m} \Pr(Y_i = \tau_i | (Y_1, \ldots, Y_{i-1}) = \tau_{[1:i-1]}),$$

we upper bound each $\Pr(Y_i = \tau_i | (Y_1, \ldots, Y_{i-1}) = \tau_{[1:i-1]})$ by 1 if $\tau_i = 1$, and by α if $\tau_i = 0$, giving the result. □

Proof (Of Lemma 4.9). By Lemma 4.10, it holds that

$$\begin{aligned}\Pr(\sum_{i=1}^{m} Y_i \leq T) &= \sum_{\substack{\tau\in\{0,1\}^m \\ |\tau|\leq T}} \Pr((Y_1,\ldots,Y_m)=\tau) \\ &\leq \sum_{k=0}^{T} \binom{m}{k}\alpha^{m-k} \leq \alpha^{m-T}\cdot\binom{m}{T}\cdot(T+1) \\ &\leq \alpha^{m-T}\cdot\frac{e^T\cdot(T+1)\cdot m^T}{T^T} \\ &\leq 7\cdot\alpha^{m-T}\cdot m^T,\end{aligned}$$

where we used the fact that $T \leq m/2$, by assumption. Now, note that

$$7\cdot\alpha^{m-T}\cdot m^T \leq \varepsilon \Leftrightarrow m \geq T + \frac{\log(1/\varepsilon)+\log(7)}{\log(1/\alpha)} + \frac{T}{\log(1/\alpha)}\cdot\log(m),$$

which, by Lemma B.5 holds for any

$$m \geq 2\cdot\left(T + \frac{\log(1/\varepsilon)+\log(7)}{\log(1/\alpha)} + \frac{T}{\log(1/\alpha)}\cdot\log\left(\frac{T}{\log(1/\alpha)}\right)\right),$$

a proper analysis proves that for the range of parameters $T \geq 13$ and $\alpha \leq 0.1$, the above bound is less than

$$4T\log(T) + \frac{2\log(1/\varepsilon)}{\log(1/\alpha)},$$

which concludes the proof. □

With all of these results, we can now prove Lemma 4.5.

Proof (of Lemma 4.5). Combining Lemmas 4.6, 4.7 and 4.9 with $\alpha = 1.8/f$ ($f \geq 2$ since $\zeta_2 = -1 \in \mathcal{K}$) by Lemma 4.8 proves Lemma 4.5. □

4.4 From short preimages of $f \cdot a \cdot \mathbf{e}_i$ to short preimage of $f \cdot \mathbf{e}_i$

As before, we fix a number field $\mathcal{K} \supseteq \mathbb{Q}(\zeta_f)$ of degree d containing the cyclotomic field of conductor f, $\mathbf{S} \geq \max(\sqrt{2\pi d}, \eta_{1/2}(\mathcal{O}_{\mathcal{K}}^r))$, and $\mathbf{X} \in \mathcal{O}_{\mathcal{K}}^{r\times m}$ a matrix with columns sampled from $\mathcal{D}_{\mathcal{O}_{\mathcal{K}}^r,\mathbf{S}}$ for some $m \geq 1$.
Lemma 4.5 gives a probabilistic guarantee for the existence of a short preimage of $f\cdot a\cdot\mathbf{e}_i$ for some short $a \in A_{\Pi}$ defined relative to a set A. To obtain preimages of $f\cdot\mathbf{e}_i$, our strategy is to invoke Lemma 4.5 again on a different B'-admissible set A' whose elements are coprime with a. This gives a short preimage of $f\cdot a'\cdot\mathbf{e}_i$ for a short $a' \in A'_{\Pi}$. We then use the co-primality of a and a' and an effective version of Bezout's identity (Lemma 2.3) to create a preimage of $f\cdot\mathbf{e}_i$. We discuss why preimages of $f\cdot\mathbf{I}_r$ are sufficient in the statement on the smoothing parameter in Lemma 4.14. We did not find a formula for the size of the set A' in the literature, so we provide one in Lemma 4.11, which we think might have applications beyond this work.

Lemma 4.11. *Let $\mathcal{K}$ be a number field of degree d with ring of integers $\mathcal{O}_\mathcal{K}$. Let $\mathfrak{a}$ be an integral ideal of $\mathcal{O}_\mathcal{K}$ with prime factorisation $\mathfrak{a} = \prod_{i=1}^{l} \mathfrak{p}_i^{e_i}$ and $\varepsilon = \min(2^{-d}, 3^{-l}/10)$. Then for*

$$R \geq \frac{\mathcal{N}(\mathfrak{a})^{1/d} \cdot \Delta_\mathcal{K}^{1/d} \cdot \ln(1/\varepsilon^{1/d})}{2\sqrt{d}},$$

we have

$$\left|\left\{x \in \mathcal{O}_\mathcal{K},\ \|x\| \leq R,\ x \text{ coprime with } \mathfrak{a}\right\}\right| \geq \frac{R^d}{4\sqrt{\Delta_\mathcal{K}} \cdot d^{d/2} \cdot 2^{d+l}}.$$

Proof. Let D denote the Gaussian distribution over $\mathcal{O}_\mathcal{K}$ with parameter $\varsigma \geq \mathcal{N}(a)^{1/d} \cdot \Delta_\mathcal{K}^{1/d} \cdot \ln(1/\varepsilon)/d$ tail-cut for $\|x\| \geq R \coloneqq 2 \cdot \sqrt{d} \cdot \varsigma$. By [Ban93, Lemma 1.5] and [MR07, Lemmas 3.3 and 4.4] and some computations [FPS22, Lemma 2.7], it holds that for any ideal $\mathfrak{a}$ such that $\varsigma \geq \eta_\varepsilon(\mathfrak{a})$, $D(\mathfrak{a}) \in [1-5\varepsilon, 1+5\varepsilon] \cdot 1/\mathcal{N}(\mathfrak{a})$.

We write $\mathfrak{a} = \prod_{i=1}^{l} \mathfrak{p}_i^{e_i}$. An element $x \in \mathcal{O}_\mathcal{K}$ is coprime with $\mathfrak{a}$ if and only if it does not belong to any of the $\mathfrak{p}_i$. We now compute the probability of sampling an element coprime to $\mathfrak{a}$ from D. By the inclusion-exclusion principle and the fact that for any prime ideals $\mathfrak{p}, \mathfrak{q}$, $\mathfrak{p} \cap \mathfrak{q} = \mathfrak{p}\mathfrak{q}$, it holds that

$$\begin{aligned}
D(\mathcal{O}_\mathcal{K} \setminus \bigcup_{i=1}^{l} \mathfrak{p}_i) &= \sum_{k=0}^{l} (-1)^k \sum_{|I|=k} D(\prod_{i \in I} \mathfrak{p}_i) \geq \sum_{k=0}^{l} (-1)^k \sum_{|I|=k} \frac{1-(-1)^k \cdot 5\varepsilon}{\mathcal{N}(\prod_{i \in I} \mathfrak{p}_i)} \\
&= \sum_{k=0}^{l} (-1)^k \sum_{|I|=k} \frac{1}{\mathcal{N}(\prod_{i \in I} \mathfrak{p}_i)} - 5\varepsilon \sum_{k=0}^{l} \sum_{|I|=k} \frac{1}{\mathcal{N}(\prod_{i \in I} \mathfrak{p}_i)} \\
&= \prod_{i=1}^{l} \left(1 - \frac{1}{\mathcal{N}(\mathfrak{p}_i)}\right) - 5\varepsilon \cdot \prod_{i=1}^{l} \left(1 + \frac{1}{\mathcal{N}(\mathfrak{p}_i)}\right)
\end{aligned}$$

Setting $\varepsilon = \min(2^{-d}, 0.1 \cdot \prod_i (\mathcal{N}(\mathfrak{p}_i) - 1)/(\mathcal{N}(\mathfrak{p}_i) + 1))$, we get that $D(\mathcal{O}_\mathcal{K} \setminus \bigcup_{i=1}^{l} \mathfrak{p}_i) \geq 0.5 \cdot \prod_i (1 - 1/\mathcal{N}(\mathfrak{p}_i))$. To make this into a counting argument it suffices to say that the maximal value of $D(x)$ is equal to $D(0) \leq (1+5\varepsilon) \cdot \sqrt{\Delta_\mathcal{K}}/\varsigma^d \leq 2\sqrt{\Delta_\mathcal{K}}/\varsigma^d$. Then the number of ring elements of ℓ_2 norm less than R coprime with $\mathfrak{a}$ is greater than

$$\frac{D(\mathcal{O}_\mathcal{K} \setminus \bigcup_{i=1}^{l} \mathfrak{p}_i)}{D(0)} \geq \frac{\prod_i (1 - 1/\mathcal{N}(\mathfrak{p}_i)) \cdot \varsigma^d}{4\sqrt{\Delta_\mathcal{K}}} = \frac{\prod_i (1 - 1/\mathcal{N}(\mathfrak{p}_i)) \cdot R^d}{4\sqrt{\Delta_\mathcal{K}} \cdot d^{d/2} \cdot 2^d}.$$

Note that since $\mathcal{N}(\mathfrak{p}_i) \geq 2$ for all i, then $\prod_i (1 - 1/\mathcal{N}(\mathfrak{p}_i)) \geq 2^{-l}$ and $\prod_i (\mathcal{N}(\mathfrak{p}_i) - 1)/(\mathcal{N}(\mathfrak{p}_i) + 1) \geq 3^{-l}$, which concludes the proof. □

Lemma 4.12. *There exists an absolute constant $C > 2$ such that the following holds. Let $\mathcal{K} \supseteq \mathbb{Q}(\zeta_f)$ be a number field of degree d containing the cyclotomic field of conductor f. Let $\mathbf{S} \geq \max(\eta_{1/2}(\mathcal{O}_\mathcal{K}^r), \sqrt{2\pi d})$ and $\mathbf{X} \in \mathcal{O}_\mathcal{K}^{r \times m}$ be a matrix with columns sampled from $\mathcal{D}_{\mathcal{O}_\mathcal{K}^r, \mathbf{S}}$ for some $m \geq 1$.*

Let $\varepsilon > 0$, and $T_{\mathcal{K},r,\mathbf{S}} = r \cdot \log(dr \cdot \delta_K \cdot \Delta_{\mathcal{K}}^{1/(2d)} \cdot s_{\min}(\mathbf{S}))$. For

$$m \geq C \cdot T_{\mathcal{K},r,\mathbf{S}} \cdot \log(T_{\mathcal{K},r,\mathbf{S}}) + \frac{2\log(1/\varepsilon)}{\log(f/1.8)},$$

there exists a matrix $\mathbf{U} \in \mathcal{O}_{\mathcal{K}}^{m \times r}$ with $\|\mathbf{U}\| \leq O(\sqrt{md} \cdot f^2 \cdot \Delta_{\mathcal{K}}^{4/d} \cdot d^{10.5} \cdot \delta_{\mathcal{K}}^{13})$ and $\mathbf{X} \cdot \mathbf{U} = f \cdot \mathbf{I}_r$, with probability $\geq 1 - \varepsilon$ over the randomness of $\mathbf{X}$.

Proof. Let $A = \mathbf{B}_{\mathcal{O}_\mathcal{K}} \cdot \{\pm 1, 0\}^d \subset \mathcal{O}_\mathcal{K}$, which is bounded in the ℓ_∞-norm by $d \cdot \delta_\mathcal{K}$, symmetric and contains 3^d elements. As long as m is large enough by Lemma 4.5, there exists a short preimage of a short multiple of $f \cdot \mathbf{e}_1$ with probability $1 - \varepsilon/(2r)$. Namely, there exists $\mathbf{c} \in A_{\Sigma\Pi}^m, u \in \mathcal{O}_\mathcal{K}$ with $\|u\|_\infty \leq f^2/2$, $\|\mathbf{c}\|_\infty \leq 2d^2\delta_\mathcal{K}^2$ and $a' \in A_\Pi$ such that $\mathbf{X} \cdot u \cdot \mathbf{c} = f \cdot a' \cdot \mathbf{e}_1$. Note that by construction $\|a'\|_\infty \leq d^2\delta_\mathcal{K}^2$ and that a' belongs to a maximum of $l = \log_2(\mathcal{N}(a')) \leq d\log(d^2\delta_\mathcal{K}^2)$ different prime ideals.

Let $\widetilde{A}$ be the set of elements of $\mathcal{O}_\mathcal{K}$ which are coprime with a' of norm bounded by $R = 4 \cdot \Delta_\mathcal{K}^{1/d} \cdot d^{2.5} \cdot \delta_\mathcal{K}^3$. One can check that R follows the hypothesis of Lemma 4.11, and that in that case $\left|\widetilde{A}\right| \geq 2^d$. The set $\widetilde{A}$ is then R-admissible. Then by Lemma 4.5 there exists $\widetilde{\mathbf{c}} \in \widetilde{A_{\Sigma\Pi}}^m, \widetilde{u} \in \mathcal{O}_\mathcal{K}$ with $\|\widetilde{u}\|_\infty \leq f^2/2$, $\|\widetilde{\mathbf{c}}\|_\infty \leq 2R^2$ and $\widetilde{a'} \in \widetilde{A_\Pi}$, $||\widetilde{a'}||_\infty \leq R^2$ such that $\mathbf{X} \cdot \widetilde{u}\,\widetilde{\mathbf{c}} = f\widetilde{a'} \cdot \mathbf{e}_1$ with probability $1 - \varepsilon/(2r)$.

By construction, a' and $\widetilde{a'}$ are coprime, then by Lemma 2.3 there exist bounded norm $\alpha, \beta \in \mathcal{O}_\mathcal{K}$ with $\|\alpha\|_\infty, \|\beta\|_\infty \leq R^2 \cdot \sqrt{d} \cdot \lambda_d(\mathcal{O}_\mathcal{K})$ such that $a'\alpha + \widetilde{a'}\beta = 1$, and finally it holds that

$$\mathbf{X} \cdot (\alpha \cdot u \cdot \mathbf{c} + \beta \cdot \widetilde{u} \cdot \widetilde{\mathbf{c}}) = f \cdot \mathbf{e}_1.$$

We name $\mathbf{u}_1 \coloneqq \alpha \cdot u \cdot \mathbf{c} + \beta \cdot \widetilde{u} \cdot \widetilde{\mathbf{c}}$. It holds that $\|\mathbf{u}_1\|_\infty \leq 2\sqrt{d} \cdot f^2 \cdot \lambda_d(\mathcal{O}_\mathcal{K}) \cdot R^4$. Running the same argument r times for all the $\mathbf{e}_j$ and using the union bound, we get that with probability $1 - \varepsilon$, we can construct $\mathbf{U} = [\mathbf{u}_1, \ldots, \mathbf{u}_r]$ such that $\mathbf{X} \cdot \mathbf{U} = f \cdot \mathbf{I}_r$ and $\|\mathbf{U}\| \leq O(\sqrt{md} \cdot f^2 \cdot \Delta_\mathcal{K}^{4/d} \cdot d^{10.5} \cdot \delta_\mathcal{K}^{13})$. The choice of lower bound on m follows from the parameters of Lemma 4.5 when we used it during the proof. □

Corollary 4.13 (Cyclotomics version of Theorem 4.2). *There exists an absolute constant $C > 2$ such that the following holds. Let $f > 2$ and $\mathcal{K} = \mathbb{Q}(\zeta_{p^k})$ be the cyclotomic number field of conductor f and degree d. Let $1 \leq r \leq m$, $\varepsilon > 0$ and $T = r \cdot \log(r \cdot d^{3/2} \cdot s_{\min}(\mathbf{S}_X))$. Let $\mathbf{S}_X \in \mathbb{R}^{dr \times dr}$ be positive definite, $\mathbf{X} \in \mathcal{O}_\mathcal{K}^{r \times m}$ be a matrix with columns sampled from $\mathcal{D}_{\mathcal{O}_\mathcal{K}^r, \mathbf{S}_X}$ and $\mathbf{c} \in \mathcal{O}_\mathcal{K}^m$. Then for*

$$\mathbf{S}_X \geq \max(\eta_{1/2}(\mathcal{O}_\mathcal{K}^r), \sqrt{2\pi d}), \qquad m \geq C \cdot T \cdot \log(T) + \frac{2\log(1/\varepsilon)}{\log(p^k/1.8)},$$

with probability $\geq 1 - \varepsilon$ we have

$$\eta_\varepsilon(\Lambda^\perp(\mathbf{X})) = O\left(\eta_\varepsilon^{((m-r)d)} \cdot m^{1.5} \cdot f^2 \cdot d^{16} \cdot s_{\min}(\mathbf{S}_X)\right).$$

We now use the short preimage computed in Lemma 4.12 to compute a short basis of $\Lambda^\perp(\mathbf{X})$, and then an upper bound of the smoothing parameter.

Lemma 4.14. *Let $\mathbf{U} \in \mathcal{O}_\mathcal{K}^{m\times r}$ and $\mathbf{X} \in \mathcal{O}_\mathcal{K}^{r\times m}$ such that $\mathbf{X}\cdot\mathbf{U} = f\cdot\mathbf{I}_r$, then there exists a basis $\mathbf{B} \in \mathbb{R}^{dm\times d(m-r)}$ of $\Phi(\Lambda^\perp(\mathbf{X}))$ such that*

$$\|\mathbf{B}\| \le \sqrt{(m-r)\,d}\cdot\delta_\mathcal{K}\cdot(\|\mathbf{U}\|\cdot\|\mathbf{X}\| + f)\,,$$

in particular, for any $\varepsilon > 0$, it holds that

$$\eta_\varepsilon(\Lambda^\perp) \le \eta_\varepsilon^{((m-r)\,d)}\cdot\sqrt{(m-r)\,d}\cdot\delta_\mathcal{K}\cdot(\|\mathbf{U}\|\cdot\|\mathbf{X}\| + f)\,.$$

Proof. One can check that the column vectors of the matrix $\mathbf{Y} = f\cdot\mathbf{I}_m - \mathbf{U}\cdot\mathbf{X} \in \mathcal{O}_\mathcal{K}^{m\times m}$ belong to $\Lambda^\perp(\mathbf{X})$ and that their norm is bounded by $\|\mathbf{U}\|\cdot\|\mathbf{X}\| + f$. Since $\mathbf{U}\cdot\mathbf{X}$ is of rank at most r (over $\mathcal{K}$), the rank of $\mathbf{Y}$ is at least $m-r$. Hence, $\mathbf{Y}$ contains a free subset of size $m-r$ generating $\mathrm{span}_\mathcal{K}(\Lambda^\perp(\mathbf{X}))$. Then Lemma 2.1 implies the existence of a basis $\mathbf{B}$ with the corresponding norm bound. Lastly, we get a bound on the smoothing parameter by noting that $\lambda_{(m-r)d}(\Lambda^\perp(\mathbf{X})) \le \|\mathbf{B}\|$. □

Proof (of Theorem 4.2). The result follows immediately from Lemmas 4.12 and 4.14. □

5 Gaussian LHL over rings

We combine the arguments above to obtain a Gaussian Leftover Hash Lemma.

Theorem 5.1 (Gaussian LHL, Assuming GRH). *There exists an absolute constant $C > 2$ such that the following holds. Let $\mathcal{K}$ be a number field of degree $d \ge 2$ containing the cyclotomic field of conductor f and of smallest ideal norm $N_\mathcal{K}$. Let $\mathbf{c} \in \mathcal{O}_\mathcal{K}^m$, $\mathbf{S}_X \in \mathbb{R}^{dr\times dr}$ positive definite, $1 \le r \le m$, $0 < \varepsilon < 2^{-6}$, and $T = T_{\mathcal{K},r,\mathbf{S}_X} = r\cdot\log(dr\cdot\Delta_\mathcal{K}^{1/(2d)}\cdot s_{\min}(\mathbf{S}_X))$. For*

$$\mathbf{S}_X \ge \max\Big(\underbrace{\sqrt{2\pi d}}_{\text{for full image}}, \underbrace{9\cdot r^{\frac{1}{d}}\cdot\log\left(r\cdot\Delta_K^{\frac{1}{d}}\cdot\eta_\varepsilon(\mathcal{O}_\mathcal{K}^r)\right)^{\frac{1}{d}}\cdot\Delta_\mathcal{K}^{\frac{1}{d}}\cdot\eta_\varepsilon(\mathcal{O}_\mathcal{K}^r)}_{\text{for the smoothing bound}}\Big),$$

$$m \ge \max\left(\underbrace{2r + \frac{\log(1/\varepsilon)+1+\log(2d)}{\log(N_\mathcal{K}/1.4)}}_{\text{for full image}}, \underbrace{C\cdot T\cdot\log(T) + \frac{2\log(1/\varepsilon)}{\log(f/1.8)}}_{\text{for the smoothing bound}}\right),$$

$$\sqrt{\mathbf{\Sigma}} \in \omega\left(\eta_\varepsilon^{((m-r)d)}\cdot m^{1.5}\cdot d^{12}\cdot\delta_\mathcal{K}^{14}\cdot f^2\cdot\Delta_\mathcal{K}^{4/d}\cdot s_{\min}(\mathbf{S}_X)\right),$$

it holds that

$$\mathrm{SD}((\mathbf{X}\leftarrow\left(\mathcal{D}_{\mathcal{O}_\mathcal{K}^r,\mathbf{S}_X}\right)^m, \mathbf{X}\cdot\mathcal{D}_{\mathcal{O}_\mathcal{K}^m,\sqrt{\mathbf{\Sigma}},\mathbf{c}}), (\mathbf{X}\leftarrow\left(\mathcal{D}_{\mathcal{O}_\mathcal{K}^r,\mathbf{S}_X}\right)^m, \mathcal{D}_{\mathcal{O}_\mathcal{K}^r,\sqrt{\mathbf{\Sigma}'},\mathbf{X}\cdot\mathbf{c}})) \le \varepsilon'$$

where $\mathbf{\Sigma}' = \tilde{\Phi}(\mathbf{X})\cdot\mathbf{\Sigma}\cdot\tilde{\Phi}(\mathbf{X})^\mathsf{T}$, $\varepsilon' = \frac{\varepsilon}{1-\varepsilon} + m\cdot(\varepsilon + 2^{-dr}) + \varepsilon$.

Proof. Denote $\mathcal{S} \subset \mathcal{O}_{\mathcal{K}}^{r \times m}$ the set of matrices $\mathbf{X}$ satisfying $\sqrt{\mathbf{\Sigma}} \geq \eta_\varepsilon(\Lambda^\perp(\mathbf{X}))$ and $\mathbf{X} \cdot \mathcal{O}_{\mathcal{K}}^m = \mathcal{O}_{\mathcal{K}}^r$. Consider a fixed element $\mathbf{X} \in \mathcal{S}$. We verify conditions of Lemma 2.8 for the real lattice embedding of $\mathcal{O}_{\mathcal{K}}^m$ and the $\tilde{\Phi}(\cdot)$ embedding of $\mathbf{X}$. For matrix $\tilde{\Phi}(\mathbf{X})$ as a linear map over $\mathbb{R}$ the set $\ker(\tilde{\Phi}(\mathbf{X}))$ is indeed equal to $\mathrm{span}_{\mathbb{R}}(\Phi(\Lambda^\perp(\mathbf{X})))$ and $\ker(\tilde{\Phi}(\mathbf{X})) \cap \Phi(\mathcal{O}_{\mathcal{K}}^m) = \Phi(\Lambda^\perp(\mathbf{X}))$. Hence,

$$\mathrm{SD}(\mathbf{X} \cdot \mathcal{D}_{\mathcal{O}_{\mathcal{K}}^m, \sqrt{\mathbf{\Sigma}}, \mathbf{c}}, \mathcal{D}_{\mathbf{X} \cdot \mathcal{O}_{\mathcal{K}}^m, \sqrt{\mathbf{\Sigma}'}, \mathbf{X} \cdot \mathbf{c}}) \leq \frac{\varepsilon}{1-\varepsilon}.$$

Lastly, we apply $\mathbf{X} \cdot \mathcal{O}_{\mathcal{K}}^m = \mathcal{O}_{\mathcal{K}}^r$ for the support. Now for the statistical distance between tuples $E \coloneqq \mathrm{SD}((\mathbf{X}, \mathbf{X} \cdot \mathcal{D}_{\mathcal{O}_{\mathcal{K}}^m, \sqrt{\mathbf{\Sigma}}, \mathbf{c}}), (\mathbf{X}, \mathcal{D}_{\mathcal{O}_{\mathcal{K}}^r, \sqrt{\mathbf{\Sigma}'}, \mathbf{X} \cdot \mathbf{c}}))$ we have

$$\begin{aligned}
E &= \mathbb{E}_{\mathbf{X}}(\mathbf{X} \cdot \mathcal{D}_{\mathcal{O}_{\mathcal{K}}^m, \sqrt{\mathbf{\Sigma}}, \mathbf{c}}, \mathcal{D}_{\mathbf{X} \cdot \mathcal{O}_{\mathcal{K}}^m, \sqrt{\mathbf{\Sigma}'}, \mathbf{X} \cdot \mathbf{c}}) \\
&\leq \Pr(\mathbf{X} \in \mathcal{S}) \cdot \max_{\mathbf{X} \in \mathcal{S}}(\mathrm{SD}(\mathbf{X} \cdot \mathcal{D}_{\mathcal{O}_{\mathcal{K}}^m, \sqrt{\mathbf{\Sigma}}, \mathbf{c}}, \mathcal{D}_{\mathcal{O}_{\mathcal{K}}^r, \sqrt{\mathbf{\Sigma}'}, \mathbf{X} \cdot \mathbf{c}})) \\
&+ \Pr(\mathbf{X} \notin \mathcal{S}) \cdot \max_{\mathbf{X} \notin \mathcal{S}}(\mathrm{SD}(\mathbf{X} \cdot \mathcal{D}_{\mathcal{O}_{\mathcal{K}}^m, \sqrt{\mathbf{\Sigma}}, \mathbf{c}}, \mathcal{D}_{\mathcal{O}_{\mathcal{K}}^r, \sqrt{\mathbf{\Sigma}'}, \mathbf{X} \cdot \mathbf{c}})) \\
&\leq 1 \cdot \frac{\varepsilon}{1-\varepsilon} + (m \cdot (\varepsilon + 2^{-dr}) + \varepsilon) \cdot 1
\end{aligned}$$

where the first equality holds by the definition of statistical distance.

Corollary 5.2 (Prime-Power Cyclotomics version of Theorem 5.1). *There exists an absolute constant $C > 2$ such that the following holds. Let $p^k > 2$ a be prime power and $\mathcal{K} = \mathbb{Q}(\zeta_{p^k})$ be the p^k-cyclotomic number field of degree $d = (p-1) \cdot p^{k-1}$. Let $1 \leq r \leq m$, $0 < \varepsilon < 2^{-6}$, and $T = r \cdot \log(r \cdot d^{3/2} \cdot s_{\min}(\mathbf{S}_X))$. Let $\mathbf{S}_X \in \mathbb{R}^{dr \times dr}$ be positive definite, and $\mathbf{X} \in \mathcal{O}_{\mathcal{K}}^{r \times m}$ be a matrix with columns sampled from $\mathcal{D}_{\mathcal{O}_{\mathcal{K}}^r, \mathbf{S}_X}$ and $\mathbf{c} \in \mathcal{O}_{\mathcal{K}}^m$. Then for*

$$\begin{aligned}
\mathbf{S}_X &\geq 9 \cdot r^{1+\frac{1}{d}} \cdot d^{3/2} \cdot \eta_\varepsilon^{(rd)} \cdot \log^{1/d}\left((rd)^{3/2} \cdot \eta_\varepsilon^{(rd)}\right), \\
m &\geq \max\left(2r + \frac{\log(1/\varepsilon) + 1 + \log(2d)}{\log(p/1.4)}, C \cdot T \cdot \log(T) + \frac{2\log(1/\varepsilon)}{\log(p^k/1.8)}\right),
\end{aligned}$$

and for $\sqrt{\mathbf{\Sigma}} \in \omega\left(\eta_\varepsilon^{((m-r)d)} \cdot m^{1.5} \cdot p^{18k} \cdot s_{\min}(\mathbf{S}_X)\right)$, $\mathbf{\Sigma}' = \tilde{\Phi}(\mathbf{X}) \cdot \mathbf{\Sigma} \cdot \tilde{\Phi}(\mathbf{X})^{\mathsf{T}}$ it holds

$$\mathrm{SD}((\mathbf{X}, \mathbf{X} \cdot \mathcal{D}_{\mathcal{O}_{\mathcal{K}}^m, \sqrt{\mathbf{\Sigma}}, \mathbf{c}}), (\mathbf{X}, \mathcal{D}_{\mathcal{O}_{\mathcal{K}}^r, \sqrt{\mathbf{\Sigma}'}, \mathbf{X} \cdot \mathbf{c}})) \leq \frac{\varepsilon}{1-\varepsilon} + m \cdot (\varepsilon + 2^{-dr}) + \varepsilon$$

6 Spherical covariance

One way to make the distribution close to a spherical Discrete Gaussian is to adapt the covariance parameter of the input vector, as in the corollary below. Besides this, we set the centre of the distribution to $\mathbf{c} = 0$ to remove the dependence on $\mathbf{X}$ in the second distribution completely. Our proof gives similar (slightly better and general) parameters and results than [GMPW20, Theorem 6.1], with a different technique.

Corollary 6.1 (of Theorem 5.1). *Let $1 \leq r \leq m$, and $\varsigma > 0$. Following Lemma 2.5 and using that $\mathbf{M} = \tilde{\Phi}(\mathbf{X})$ is full-rank we can efficiently construct a matrix $\mathbf{S} \in \mathbb{R}^{dm \times dm}$ s.t. $\varsigma^2 \cdot \mathbf{I}_{dr} = \tilde{\Phi}(\mathbf{X}) \cdot \mathbf{S} \cdot \mathbf{S}^\intercal \cdot \tilde{\Phi}(\mathbf{X})^\intercal$. Letting $\mathbf{\Sigma} = \mathbf{S} \cdot \mathbf{S}^\intercal$ and $\mathbf{c} = \mathbf{0}$ we obtain*

$$\mathrm{SD}((\mathbf{X}, \mathbf{X} \cdot \mathcal{D}_{\mathcal{O}^m_{\mathcal{K}}, \sqrt{\mathbf{\Sigma}}}), (\mathbf{X}, \mathcal{D}_{\mathcal{O}^r_{\mathcal{K}}, \varsigma})) \leq \frac{\varepsilon}{1-\varepsilon} + m \cdot (\varepsilon + 2^{-dr}) + \varepsilon$$

Furthermore, $\varsigma / s_{\max}(\tilde{\Phi}(\mathbf{X})) \leq s_{\min}(\sqrt{\mathbf{\Sigma}}) \leq s_{\max}(\sqrt{\mathbf{\Sigma}}) \leq \varsigma / s_{\min}(\tilde{\Phi}(\mathbf{X}))$.

We also consider a different approach for removing the dependence on $\mathbf{X}$ in the case where the covariance of the input cannot be adjusted. We prove that, when the number of columns m is large enough, the singular values of $\mathbf{X} \cdot \mathbf{X}^\intercal$ are so close to each other that the distribution has a constant Rényi Divergence from a spherical Discrete Gaussian. The ideas are inspired by the arguments in [AGHS13, Lemma 8]. First we detail an adaptation of the proof of [Sil85] for continuous Gaussian matrices. To adapt it we tighten the bounds on the convergence speed of the singular values.

Lemma 6.2 (Continuous Gaussians. Adapted from [Sil85]). *Let $\mathbf{Z} \leftarrow \mathcal{D}^{r \times m}_{\varsigma}$ for $1 \leq r \leq m$, $\varsigma > 0$ and parameters $d, k \geq 1$ s.t. $rd \geq 4$. If $m \geq 49 \cdot \max(r, k) \cdot (rd)^2$, then with probability at least $1 - 4r \cdot \exp(-k)$ we have*

$$m \cdot \left(1 - \frac{1}{rd}\right) \leq s^2_{\min}(\mathbf{Z}) \leq s^2_{\max}(\mathbf{Z}) \leq m \cdot \left(1 + \frac{1}{rd}\right).$$

Proof. Denote the first row of $\mathbf{Z}$ by $\mathbf{z}_0 \leftarrow \mathcal{D}^m_{\varsigma}$. Define $\mathbf{O}_0 \in \mathbb{R}^{m \times m}$ a matrix where the first column is equal to $\mathbf{z}_0 / \|\mathbf{z}_0\|$ and the remaining columns are a non-random completion to an orthogonal matrix (e.g. taking the unit vectors and running the Gram-Schmidt orthogonalisation starting from $\mathbf{z}_0 / \|\mathbf{z}_0\|$). Then

$$\mathbf{Z} \cdot \mathbf{O}_0 = \begin{bmatrix} X_m & 0 \ldots 0 \\ & \tilde{\mathbf{Z}}_0 \end{bmatrix}$$

where $X_m \coloneqq \|\mathbf{z}_0\| \leftarrow \chi_m$. The distribution of the rows in $\tilde{\mathbf{Z}}_0$ remains independent Gaussian. This is because $\mathbf{O}_0$ is orthogonal and the Gaussian distribution is rotationally invariant. Then we define a rotation $\mathbf{O}_1 \in \mathbb{R}^{r \times r}$ on the right in a similar way such that

$$\begin{bmatrix} 1 & 0 \ldots 0 \\ 0 & \\ \vdots & \tilde{\mathbf{O}}_1 \\ 0 & \end{bmatrix} \cdot \mathbf{Z} \cdot \mathbf{O}_0 = \begin{bmatrix} X_m & 0 \ldots 0 \\ Y_{r-1} & \\ \vdots & \tilde{\mathbf{Z}}_1 \\ 0 & \end{bmatrix}$$

for $Y_{r-1} \leftarrow \chi_{r-1}$. Here the first row of $\tilde{\mathbf{O}}_1$ is a normalisation of the first column in $\tilde{\mathbf{Z}}_0$ and the rest is completed to an orthogonal matrix. We repeat the algorithm

until

$$\prod_{j=0}^{r-2} \mathbf{O}_{2j+1} \cdot \mathbf{Z} \cdot \prod_{i=0}^{r-1} \mathbf{O}_{2i} = \begin{bmatrix} X_m & 0 & 0 \dots 0 \\ Y_{r-1} & X_{m-1} & 0 \dots 0 \\ \vdots & Y_{r-2} & \\ 0 & \vdots & \dots \end{bmatrix}$$

As a consequence, there exists an orthogonal transformation $\mathbf{O} \in \mathbb{R}^{r \times r}$ s.t. $\mathbf{O} \cdot \mathbf{Z} \cdot \mathbf{Z}^{\mathsf{T}} \cdot \mathbf{O}^{\mathsf{T}} =$

$$\begin{bmatrix} X_m^2 & X_m Y_{r-1} & 0 & & & 0 \\ X_m Y_{r-1} & X_{m-1}^2 + Y_{r-1}^2 & X_{m-1} Y_{r-2} & 0 & & 0 \\ & & \dots & \dots & & \\ 0 & & & 0 & X_{m-r+2} Y_1 & X_{m-r+1}^2 + Y_1^2 \end{bmatrix}$$

where $X_i \leftarrow \chi_i, Y_i \leftarrow \chi_i$, and the variables denoted by different letters or indices are pairwise independent. By the concentration inequalities of Lemma 2.10 and the union bound, $\forall i \in [m-r+1, m+1]$, $j \in [1, r]$ it holds that $\sqrt{i - 2\sqrt{ik}} \leq X_i \leq \sqrt{i + 2\sqrt{ik} + 2k}$ and $\sqrt{j - 2\sqrt{jk}} \leq Y_j \leq \sqrt{j + 2\sqrt{jk} + 2k}$ with overall probability $\geq 1 - 4r \cdot \exp(-k)$. Consider $r > 1$ and denote $Y_r = Y_0 = 0$. Then by the Gershgorin circle theorem the eigenvalues of $\mathbf{O} \cdot \mathbf{Z} \cdot \mathbf{Z}^{\mathsf{T}} \cdot \mathbf{O}^{\mathsf{T}}$ (and hence the eigenvalues of $\mathbf{Z} \cdot \mathbf{Z}^{\mathsf{T}}$) are bounded as

$$\lambda_{\min}(\mathbf{Z} \cdot \mathbf{Z}^{\mathsf{T}}) \geq \min_{i \in [0,r]} (X_{m-i}^2 + Y_{r-i}^2 - X_{m-i+1} Y_{r-i} - X_{m-i} Y_{r-i-1})$$
$$\lambda_{\max}(\mathbf{Z} \cdot \mathbf{Z}^{\mathsf{T}}) \leq \max_{i \in [0,r]} (X_{m-i}^2 + Y_{r-i}^2 + X_{m-i+1} Y_{r-i} + X_{m-i} Y_{r-i-1})$$

Substituting in the concentration inequalities and simplifying the expressions we get that for $m \geq c^2 \cdot \max(r, k) \cdot x^2$

$$\frac{\lambda_{\min}(\mathbf{Z} \cdot \mathbf{Z}^{\mathsf{T}})}{m} \geq 1 - \frac{2}{c \cdot x} - \frac{2}{c^2 \cdot x^2} - 2\sqrt{5} \cdot \frac{1}{c \cdot x} \cdot \sqrt{1 + \frac{2}{c \cdot x} + \frac{2}{c^2 \cdot x^2}}, \quad (3)$$

$$\frac{\lambda_{\max}(\mathbf{Z} \cdot \mathbf{Z}^{\mathsf{T}})}{m} \leq 1 + \frac{2}{c \cdot x} + \frac{7}{c^2 \cdot x^2} + 2\sqrt{5} \cdot \frac{1}{c \cdot x} \cdot \sqrt{1 + \frac{2}{c \cdot x} + \frac{2}{c^2 \cdot x^2}}. \quad (4)$$

Then for $c = 7, x = rd$ for $rd \geq 4$ we have[9]

$$\Pr(1 - \frac{1}{rd} \leq \lambda_{\min}/m \leq \lambda_{\max}/m \leq 1 + \frac{1}{rd}) \geq 1 - 4r \cdot \exp(-k).$$

Lastly, we consider the special case when $r = 1$. Then using $\mathbf{z}^{\mathsf{T}} \cdot \mathbf{z} \sim \chi_m^2$, $k/m \leq 1/(7\,rd)^2$ and Lemma 2.10 we get an even tighter bound

$$1 - \frac{1}{7\,rd} \leq 1 - \sqrt{\frac{k}{m}} \leq s_{\min}^2(\mathbf{z})/m = s_{\max}^2(\mathbf{z})/m \leq 1 + \sqrt{\frac{k}{m}} + 2 \cdot \frac{k}{m} \leq 1 + \frac{1}{7\,rd}$$

with probability $\geq 1 - 2\exp(-k)$. □

[9] See the included SageMath script or Appendix C.3 for the corresponding calculations.

Now we provide an adaptation of [PS21, Lemma 2.3] on the Rényi Divergence between discrete Gaussian distributions with different covariance parameters. We provide the proof for completeness.

Lemma 6.3. *Let Λ be an n-dimensional lattice $\varepsilon \in (0,1)$ and $\mathbf{\Sigma}_0, \mathbf{\Sigma}_1 \in \mathbb{R}^{n\times n}$ be positive definite matrices s.t. $\sqrt{\mathbf{\Sigma}}_i \geq \eta_\varepsilon(\Lambda), i = 0,1$ and $s_{\max}(\sqrt{\mathbf{\Sigma}}_0) \leq s_{\min}(\sqrt{\mathbf{\Sigma}}_1)$. Then*

$$\mathrm{RD}(\mathcal{D}_{\Lambda,\sqrt{\mathbf{\Sigma}}_0}; \mathcal{D}_{\Lambda,\sqrt{\mathbf{\Sigma}}_1}) \leq \frac{1+\varepsilon}{1-\varepsilon} \cdot \frac{\det(\sqrt{\mathbf{\Sigma}}_1)}{\det(\sqrt{\mathbf{\Sigma}}_0)}.$$

Proof. By definition

$$\begin{aligned}
&\mathrm{RD}(\mathcal{D}_{\Lambda,\sqrt{\mathbf{\Sigma}}_0}; \mathcal{D}_{\Lambda,\sqrt{\mathbf{\Sigma}}_1}) = \sum_{\mathbf{s}\in\Lambda} \Pr(\mathbf{s} \leftarrow \mathcal{D}_{\Lambda,\sqrt{\mathbf{\Sigma}}_0}) \cdot \frac{\Pr(\mathbf{s} \leftarrow \mathcal{D}_{\Lambda,\sqrt{\mathbf{\Sigma}}_0})}{\Pr(\mathbf{s} \leftarrow \mathcal{D}_{\Lambda,\sqrt{\mathbf{\Sigma}}_1})} \\
&= \mathbb{E}_{\mathbf{s}\leftarrow\mathcal{D}_{\Lambda,\sqrt{\mathbf{\Sigma}}_0}} \left(\frac{\rho_{\sqrt{\mathbf{\Sigma}}_0}(\mathbf{s})}{\rho_{\sqrt{\mathbf{\Sigma}}_0}(\Lambda)} \cdot \frac{\rho_{\sqrt{\mathbf{\Sigma}}_1}(\Lambda)}{\rho_{\sqrt{\mathbf{\Sigma}}_1}(\mathbf{s})} \right) = \frac{\rho_{\sqrt{\mathbf{\Sigma}}_1}(\Lambda)}{\rho_{\sqrt{\mathbf{\Sigma}}_0}(\Lambda)} \cdot \mathbb{E}_{\mathbf{s}\leftarrow\mathcal{D}_{\Lambda,\sqrt{\mathbf{\Sigma}}_0}} \left(\frac{\rho_{\sqrt{\mathbf{\Sigma}}_0}(\mathbf{s})}{\rho_{\sqrt{\mathbf{\Sigma}}_1}(\mathbf{s})} \right) \\
&\leq \frac{1+\varepsilon}{1-\varepsilon} \cdot \frac{\det(\sqrt{\mathbf{\Sigma}}_1)}{\det(\sqrt{\mathbf{\Sigma}}_0)} \cdot \mathbb{E}_{\mathbf{s}\leftarrow\mathcal{D}_{\Lambda,\sqrt{\mathbf{\Sigma}}_0}} \exp\left(\pi \cdot \left(\left\| \sqrt{\mathbf{\Sigma}}_1^{-1} \cdot \mathbf{s} \right\|^2 - \left\| \sqrt{\mathbf{\Sigma}}_0^{-1} \cdot \mathbf{s} \right\|^2 \right) \right)
\end{aligned}$$

where the last transition follows from Lemma 2.12. Denote $\tilde{\mathbf{s}} \coloneqq \sqrt{\mathbf{\Sigma}}_0^{-1} \cdot \mathbf{s}$, then

$$\begin{aligned}
\left\| \sqrt{\mathbf{\Sigma}}_1^{-1} \cdot \mathbf{s} \right\|^2 - \left\| \sqrt{\mathbf{\Sigma}}_0^{-1} \cdot \mathbf{s} \right\|^2 &= \left\| \sqrt{\mathbf{\Sigma}}_1^{-1} \cdot \sqrt{\mathbf{\Sigma}}_0 \cdot \tilde{\mathbf{s}} \right\|^2 - \|\tilde{\mathbf{s}}\|^2 \\
&\leq \left(\frac{s_{\max}^2(\sqrt{\mathbf{\Sigma}}_0)}{s_{\min}^2(\sqrt{\mathbf{\Sigma}}_1)} - 1 \right) \cdot \|\tilde{\mathbf{s}}\|^2 \leq 0
\end{aligned}$$

since $s_{\max}(\sqrt{\mathbf{\Sigma}}_0) \leq s_{\min}(\sqrt{\mathbf{\Sigma}}_1)$. Combining the above we get the statement. □

Finally, we prove the target statement. At a high level, we first add continuous noise to the discrete Gaussian matrix $\tilde{\Phi}(\mathbf{X})$ to make it approach a continuous Gaussian sample. Then we apply the bounds from Lemma 6.2 to the continuous matrices. Lastly, we obtain bounds on the singular values of their difference equal to $\Phi(\mathbf{X})$ and use Lemma 6.3.

Theorem 6.4. *Let $1 < r \leq m$ be integers and $k \geq 1$ s.t. $m \geq 49 \cdot \max(r,k) \cdot (rd)^2$, $rd \geq 4$. Let $\varsigma_x \geq 2rd \cdot \eta_\varepsilon(\mathcal{O}_\mathcal{K})$ for $\varepsilon \in (0, 1/2)$. Let $\mathbf{X} \leftarrow \left(\mathcal{D}_{\mathcal{O}_\mathcal{K}^r,\varsigma_x}\right)^m$. For $\varsigma = \sqrt{m \cdot (\varsigma_x^2 + 2 \cdot \eta_\varepsilon^2(\mathcal{O}_\mathcal{K}))} \cdot \left(1 - \frac{1}{rd}\right)^{3/2}$, we have*

$$\mathrm{RD}(\mathcal{D}_{\mathcal{O}_\mathcal{K}^r,\varsigma}; \mathcal{D}_{\mathcal{O}_\mathcal{K}^r,\sqrt{\tilde{\Phi}(\mathbf{X})\cdot\tilde{\Phi}(\mathbf{X})^\intercal}}) \leq 4 \cdot \exp(3)$$

with probability at least $1 - 4rm \cdot \varepsilon - 4rd \cdot \exp(-k)$ over the randomness of $\mathbf{X}$.

Proof. In this proof we work with diagonal-by-block matrices[10] of dimension $dr \times dm$ with blocks of dimension $d \times d$. For a matrix $\mathbf{M}$ we denote by $\mathbf{M}_{ij}$,

[10] block matrices with diagonal blocks

$i \in [r], j \in [m]$ the corresponding $d \times d$ block. Consider $\mathbf{Z} = \tilde{\Phi}(\mathbf{X}) + \mathbf{Y}$ where $\mathbf{X} \leftarrow \left(\mathcal{D}_{\mathcal{O}_{\mathcal{K}}^r, \varsigma_x}\right)$ and $\mathbf{Y}$ is diagonal-by-block with $\mathbf{Y}_{ij} \leftarrow \operatorname{diag}\left(\mathcal{D}_{\varsigma_y}^d\right)$, $i \in [r], j \in [m]$. By Lemma 2.9 for $\mathbf{Z}$ we have $\operatorname{SD}(\mathbf{Z}_{ij}, \operatorname{diag}(\mathcal{D}^d_{\sqrt{\varsigma_x^2+\varsigma_y^2}})) \leq 4\varepsilon$ for $\varsigma_y \coloneqq \sqrt{2} \cdot \eta_\varepsilon(\mathcal{O}_{\mathcal{K}})$.

We analyse the singular values of $\mathbf{Z}$. First we multiply by unitary matrices on the right and on the left to reorder the rows and columns and obtain a **block-diagonal** matrix. Now on the diagonal we have matrices we denote $\mathbf{Z}_i$ sampled independently from $\mathbf{Z}_i \leftarrow \mathcal{D}^{r \times m}_{\sqrt{\varsigma_x^2+\varsigma_y^2}}$ and similarly $\mathbf{Y}_i \leftarrow \mathcal{D}^{r \times m}_{\varsigma_y}$ for $i \in [d]$. By Lemma 6.2 for $i \in [d]$

$$
\begin{aligned}
m \cdot (\varsigma_x^2 + \varsigma_y^2) \cdot \left(1 - \frac{1}{rd}\right) \leq s^2_{\min}(\mathbf{Z}_i) \leq s^2_{\max}(\mathbf{Z}_i) \leq m \cdot (\varsigma_x^2 + \varsigma_y^2) \cdot \left(1 + \frac{1}{rd}\right) \\
m \cdot \varsigma_y^2 \cdot \left(1 - \frac{1}{rd}\right) \leq s^2_{\min}(\mathbf{Y}_i) \leq s^2_{\max}(\mathbf{Y}_i) \leq m \cdot \varsigma_y^2 \cdot \left(1 + \frac{1}{rd}\right)
\end{aligned}
$$

with probability at least $1 - 8dr \cdot \exp(-k)$. Note that $\mathbf{X}_i = \mathbf{Z}_i - \mathbf{Y}_i$ and one can verify that $s_{\max}(\mathbf{Y}_i) \leq \delta \cdot s_{\min}(\mathbf{Z}_i)$ for $\varsigma_x \geq \varsigma_y \cdot \sqrt{2} \cdot rd$, $\delta = 1/rd$ and $rd \geq 3$, as shown below.

$$
\begin{aligned}
m \cdot \varsigma_y^2 \cdot \left(1 + \frac{1}{rd}\right) &\leq \delta^2 \cdot m \cdot (\varsigma_x^2 + \varsigma_y^2) \cdot \left(1 - \frac{1}{rd}\right) \\
\frac{\varsigma_x^2}{2 \cdot (rd)^2} \cdot \left(1 + \frac{1}{rd}\right) &\leq \delta^2 \cdot \varsigma_x^2 \cdot \left(1 - \frac{1}{rd}\right) \\
\frac{1}{2 \cdot (rd)^2} \cdot \left(1 + \frac{1}{rd}\right) &\leq \frac{1}{(rd)^2} \cdot \left(1 - \frac{1}{rd}\right) \\
1 + \frac{1}{rd} &\leq 2 - \frac{2}{rd} \\
\frac{3}{rd} &\leq 1
\end{aligned}
$$

Then using Lemma 2.4 we get

$$
\begin{aligned}
m \cdot (\varsigma_x^2 + \varsigma_y^2) \cdot \left(1 - \frac{1}{rd}\right) \cdot (1 - \delta)^2 &\leq s^2_{\min}(\mathbf{X}_i) \\
m \cdot (\varsigma_x^2 + 2 \cdot \eta_\varepsilon^2(\mathcal{O}_{\mathcal{K}})) \cdot \left(1 - \frac{1}{rd}\right)^3 &\leq s^2_{\min}(\mathbf{X}_i)
\end{aligned}
$$

and

$$
\begin{aligned}
s^2_{\max}(\mathbf{X}_i) &\leq m \cdot (\varsigma_x^2 + \varsigma_y^2) \cdot \left(1 + \frac{1}{rd}\right) \cdot (1 + \delta)^2 \\
s^2_{\max}(\mathbf{X}_i) &\leq m \cdot (\varsigma_x^2 + 2 \cdot \eta_\varepsilon^2(\mathcal{O}_{\mathcal{K}})) \cdot \left(1 + \frac{1}{rd}\right)^3
\end{aligned}
$$

Then, for $\varsigma = \sqrt{m \cdot (\varsigma_x^2 + 2 \cdot \eta_{\bar{\varepsilon}}^2(\mathcal{O}_\mathcal{K}))} \cdot \left(1 - \frac{1}{rd}\right)^{3/2}$, by Lemma 6.3 we have

$$\begin{aligned}\mathrm{RD}(\mathcal{D}_{\mathcal{O}_\mathcal{K}^r,\varsigma}; \mathcal{D}_{\mathcal{O}_\mathcal{K}^r,\sqrt{\tilde{\Phi}(\mathbf{X})\cdot\tilde{\Phi}(\mathbf{X})^\intercal}}) &\leq \frac{1+\varepsilon}{1-\varepsilon} \cdot \frac{\left(1+\frac{1}{rd}\right)^{3rd/2}}{\left(1-\frac{1}{rd}\right)^{3rd/2}} \\ &\leq \frac{1+\varepsilon}{1-\varepsilon} \cdot \frac{\exp(3/2)}{\exp(-3/2)\cdot\left(1-\frac{1}{rd}\right)} \leq 4\exp(3)\end{aligned}$$

with probability at least $1 - 4rm \cdot \varepsilon - 4rd \cdot \exp(-k)$ over the randomness of $\mathbf{X}$. The last inequality uses $\exp(x) \cdot (1 - \frac{x^2}{n}) \leq (1 + \frac{x}{n})^n \leq \exp(x)$ for $n \geq 1$, $|x| \leq n$, $\varepsilon \leq 1/2$ and $rd \geq 4$.

Remark 6.5. We expect an equivalent statement for $\mathsf{negl}(\lambda)$ statistical distance requires a number of columns exponential in rd. We leave this as an open problem.

As a side-result, we give a statement of [AGHS13, Lemma 8] with explicit constants, in the general case of ring of integers.

Lemma 6.6 (Discrete Gaussians). *Let $1 < r \leq m$ be integers and $k \geq 1$ s.t. $m \geq 64 \cdot \max(r, k)$. Let $\varsigma_x \geq 10 \cdot \eta_\varepsilon(\mathcal{O}_\mathcal{K})$ for $\varepsilon \in (0, 1/2)$. Let $\mathbf{X} \leftarrow \left(\mathcal{D}_{\mathcal{O}_\mathcal{K}^r,\varsigma_x}\right)^m$. It holds that*

$$\frac{\sqrt{m} \cdot \varsigma_x}{7} \leq s_{\min}(\tilde{\Phi}(\mathbf{X})) \leq s_{\max}(\tilde{\Phi}(\mathbf{X})) \leq 2.2 \cdot \sqrt{m} \cdot \varsigma_x,$$

each inequality holding with probability at least $1 - 4rm \cdot \varepsilon - 4rd \cdot \exp(-k)$ over the randomness of $\mathbf{X}$.

Proof. The proof is similar to that of Theorem 6.4. We work with diagonal-by-block matrices[11] of dimension $dr \times dm$ with blocks of dimension $d \times d$. For matrix $\mathbf{M}$ we denote $\mathbf{M}_{ij}$, $i \in [r], j \in [m]$ the corresponding $d \times d$ block. Let $\varsigma_y \coloneqq \varsigma_x/\sqrt{96}$, and consider $\mathbf{Z} = \tilde{\Phi}(\mathbf{X}) + \mathbf{Y}$ where $\mathbf{X} \leftarrow \left(\mathcal{D}_{\mathcal{O}_\mathcal{K}^r,\varsigma_x}\right)$ and $\mathbf{Y}$ is diagonal-by-block with $\mathbf{Y}_{ij} \leftarrow \mathrm{diag}\left(\mathcal{D}^d_{\varsigma_y}\right)$, $i \in [r], j \in [m]$. One can compute that $\varsigma_x \cdot \varsigma_y / \sqrt{\varsigma_x^2 + \varsigma_y^2} \geq \varsigma_x/10$, and then by the condition on ς_x we can use Lemma 2.9 for $\mathbf{Z}$. We then have $\mathrm{SD}(\mathbf{Z}_{ij}, \mathrm{diag}(\mathcal{D}^d_{\sqrt{\varsigma_x^2+\varsigma_y^2}})) \leq 4\varepsilon$.

We analyse the singular values of $\mathbf{Z}$. First we multiply by unitary matrices on the right and on the left to reorder the rows and columns and obtain a **block-diagonal** matrix. Now on the diagonal we have matrices we denote $\mathbf{Z}_i$ sampled independently from $\mathbf{Z}_i \leftarrow \mathcal{D}^{r\times m}_{\sqrt{\varsigma_x^2+\varsigma_y^2}}$ and similarly $\mathbf{Y}_i \leftarrow \mathcal{D}^{r\times m}_{\varsigma_y}$ for $i \in [d]$. We can use Eqs. (3) and (4) of the proof of Lemma 6.2 with $c \cdot x = 8$ (which matches the condition on m) to obtain, for any $i \in [d]$:

$$\begin{aligned} s_{\max}(\mathbf{Y}_i)^2 &\leq 2 \cdot m \cdot \varsigma_y^2, \\ s_{\min}(\mathbf{Z}_i)^2 &\geq \frac{m}{12} \cdot (\varsigma_x^2 + \varsigma_y^2), \\ s_{\max}(\mathbf{Z}_i)^2 &\leq 2 \cdot m \cdot (\varsigma_x^2 + \varsigma_y^2) \end{aligned}$$

[11] block matrices with diagonal blocks

$\mathsf{HintTG}_{m,k}(\varsigma_1, \varsigma_2) \to (\mathbf{X} \in \mathcal{O}_{\mathcal{K}}^{k\times(m+k)}, \mathbf{U} \in \mathcal{O}_{\mathcal{K}}^{(m+k)\times(m+k)})$

1 : $\mathbf{X}_1 \leftarrow \mathcal{D}^m_{\mathcal{O}^k_{\mathcal{K}},\varsigma_1}$

2 : $\mathbf{\Sigma} \coloneqq \sqrt{\mathbf{S}\cdot\mathbf{S}^{\mathsf{T}}}$ s.t. $\varsigma_2^2 \cdot \mathbf{I}_{dr} = \tilde{\Phi}(\mathbf{X}_1)\cdot\mathbf{S}\cdot\mathbf{S}^{\mathsf{T}}\cdot\tilde{\Phi}(\mathbf{X}_1)^{\mathsf{T}}$ from Cor. 6.1

3 : $\forall : i \in [k] : \mathbf{r}_i \leftarrow \mathcal{D}_{\mathcal{O}^m_{\mathcal{K}},\sqrt{\mathbf{\Sigma}}}$. Let $\mathbf{R} \coloneqq (\mathbf{r}_1, \ldots, \mathbf{r}_k)$

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** $((\mathbf{X}_1, \mathbf{X}_2), \mathbf{U})$

Fig. 1. HintTG procedure.

with probability $\geq 1 - 4rd\exp(-k)$. Now, one can compute that with our choice of ς_y, it holds that $s_{\max}(\mathbf{Y}_i)/s_{\min}(\mathbf{Z}_i) \leq 1/2$. Then using Lemma 2.4 and the fact that $\tilde{\Phi}(\mathbf{X}) = \mathbf{Z} - \mathbf{Y}$, we get that for any $i \in [d]$:

$$s_{\min}(\mathbf{X}_i) \geq \frac{1}{2}\cdot s_{\min}(\mathbf{Z}_i) \geq \frac{\sqrt{m}}{2\sqrt{12}}\cdot\sqrt{\varsigma_x^2+\varsigma_y^2} \geq \frac{\sqrt{m}\cdot\varsigma_x}{7},$$
$$s_{\max}(\mathbf{X}_i) \leq \frac{3}{2}\cdot s_{\max}(\mathbf{Z}_i) \leq 2.2\cdot\sqrt{m}\cdot\varsigma_x,$$

for any $i \in [d]$ with probability $\geq 1 - 4rm\cdot\varepsilon - 4rd\exp(-k)$. The union bound gives the desired result. □

7 Hardness of k-SIS and k-LWE problems

In this section we formalise a major application of our Gaussian LHL result: a reduction from the SIS problem to the k-SIS problem in the module setting. In order to be coherent with prior work's notation, in this section, let $\mathcal{K}$ be a cyclotomic number field of degree d and conductor f and we denote $\mathcal{R} = \mathcal{O}_{\mathcal{K}}$ and $\mathcal{R}_q = \mathcal{R}/(q\cdot\mathcal{R})$ for any $q \geq 2$. Based on our new Gaussian LHL (Theorem 5.1), we obtain the following generalisation and adaption of [LPSS14, Theorem 17 (eprint)], which is the core of our SIS-to-k-SIS reduction over modules.

Lemma 7.1. *Let $\mathcal{K}$ be a number field of degree $d \geq 2$ containing the cyclotomic field of conductor f. Let $k = r \geq 1$, $m \geq 64\max(\lambda/\log(dk), k)$, $0 < \varepsilon \leq \mathsf{negl}(\lambda)$, $d\cdot k \geq \lambda$ and $\varsigma_1 \cdot \mathbf{I} = \mathbf{S}_X \in \mathbb{R}^{dk\times dk}$ follow constraints of Theorem 5.1 and satisfy $\varsigma_1 \geq 10\cdot\eta_\epsilon(\mathcal{O}_{\mathcal{K}})$. Additionally, let $\varsigma_2 \geq \varsigma_1 d\sqrt{mk}\cdot\omega\left(\eta_\varepsilon^{((m-r)d)}\cdot m^{1.5}\cdot d^{12}\cdot\delta_{\mathcal{K}}^{14}\cdot f^2\cdot\Delta_{\mathcal{K}}^{4/d}\cdot\varsigma_1\right)$. Denote by $\mathbf{e}_i \in \mathcal{O}_{\mathcal{K}}^k$ the i-th unit vector. Then there exists a PPT algorithm $\mathsf{HintTG}_{m,k}$ (Fig. 1) which, on input parameters $(\varsigma_1, \varsigma_2)$, outputs $((\mathbf{X}_1, \mathbf{X}_2), \mathbf{U}) \in \mathcal{O}_{\mathcal{K}}^{k\times(m+k)} \times \mathcal{O}_{\mathcal{K}}^{(m+k)\times(m+k)}$ satisfying the following:*

1. $\mathrm{SD}\left((\mathbf{X}_1, \mathbf{X}_2),\ ((\mathcal{D}_{\mathcal{O}^k_{\mathcal{K}},\varsigma_1})^m, \prod_{i\in[k]}\mathcal{D}_{\mathcal{O}^k_{\mathcal{K}},\varsigma_2,\mathbf{e}_i})\right) \leq \mathsf{negl}(\lambda)$.

2. *We have* $\det(\mathbf{U}) = 1$*, and letting* $\bar{\mathbf{U}} \in \mathcal{O}_{\mathcal{K}}^{(m+k)\times m}$ *be the last* m *columns of* $\mathbf{U}$*, then* $(\mathbf{X}_1, \mathbf{X}_2) \cdot \bar{\mathbf{U}} = \mathbf{0}_{k\times m}$*.*
3. *It holds that* $\left\|\bar{\mathbf{U}}\right\| = \Omega(\varsigma_2 \cdot d \cdot \sqrt{m})$ *with overwhelming probability in* λ*.*

Proof. We generate required values using $\mathsf{HintTG}_{m,k}$ described in Fig. 1. We show that the output from HintTG satisfies all Items 1 to 3 in the lemma.

For Item 1, applying Theorem 5.1 and Cor. 6.1 k times over all $\mathbf{X}_1 \cdot \mathbf{r}_i$ together with the union bound, we have

$$\mathrm{SD}\left((\mathbf{X}_1, \mathbf{X}_1 \cdot \mathbf{R}),\ ((\mathcal{D}_{\mathcal{O}_{\mathcal{K}}^k,\varsigma_1})^m, (\mathcal{D}_{\mathcal{O}_{\mathcal{K}}^k,\varsigma_2})^k)\right) \leq \mathsf{negl}(\lambda).$$

The only difference between $\mathbf{X}_1 \cdot \mathbf{R}$ and the output $\mathbf{X}_2 = (\mathbf{x}_{2,1}, \ldots, \mathbf{x}_{2,k})$ from HintTG is that each i-th column $\mathbf{x}_{2,i}$ of $\mathbf{X}_2$ is offset by $\mathbf{e}_i \in \mathcal{O}_{\mathcal{K}}^k$, the i-th unit vector, respectively. The claim follows by noting that for any $\mathbf{y} \in \mathcal{O}_{\mathcal{K}}^k$ following $\mathcal{D}_{\mathcal{O}_{\mathcal{K}},\varsigma_2}^k$ and any fixed $\mathbf{e}_i \in \mathcal{O}_{\mathcal{K}}^k$, it holds that $\mathbf{y} + \mathbf{e}_i$ follows $\mathcal{D}_{\mathcal{O}_{\mathcal{K}}^k,\varsigma_2,\mathbf{e}_i}$. For Item 2, observe that

$$\mathbf{U} = \begin{pmatrix} \mathbf{I}_m & -\mathbf{R} \\ & \mathbf{I}_k \end{pmatrix} \cdot \begin{pmatrix} -\mathbf{I}_m & \\ \mathbf{I}_k & \mathbf{X}_1 \end{pmatrix}.$$

Noting that both matrices are triangular with diagonal entries ± 1 implies $\det(\mathbf{U}) = 1$. For the kernel condition we note

$$(\mathbf{X}_1, \mathbf{X}_2) = (\mathbf{X}_1, \mathbf{X}_1\mathbf{R} + \mathbf{I}_k) = (\mathbf{X}_1, \mathbf{I}_k) \cdot \begin{pmatrix} \mathbf{I}_m & \mathbf{R} \\ & \mathbf{I}_k \end{pmatrix},$$

thus

$$(\mathbf{X}_1, \mathbf{X}_2) \cdot \mathbf{U} = (\mathbf{X}_1, \mathbf{I}_k) \cdot \begin{pmatrix} & -\mathbf{I}_m \\ \mathbf{I}_k & \mathbf{X}_1 \end{pmatrix} = (\mathbf{I}_k, \mathbf{0}_{k\times m}),$$
$$\text{implying} \quad (\mathbf{X}_1, \mathbf{X}_2) \cdot \bar{\mathbf{U}} = \mathbf{0}_{k\times m}.$$

Finally, for Item 3, we bound the norm of $\bar{\mathbf{U}}$, obtaining

$$\begin{aligned}
\left\|\bar{\mathbf{U}}\right\| &= \left\| \begin{pmatrix} -\mathbf{I}_m - \mathbf{R}\mathbf{X}_1 \\ -\mathbf{X}_1 \end{pmatrix} \right\| \\
&\leq (1 + \|\mathbf{R}\|^2 \cdot \|\mathbf{X}_1\|^2 + \|\mathbf{X}_1\|^2)^{1/2} \\
&\leq 1 + (\|\mathbf{R}\| + 1) \cdot \|\mathbf{X}_1\| \\
&\leq 1 + (\varsigma_2 \cdot \sqrt{d \cdot m}/s_{\min}(\tilde{\Phi}(\mathbf{X}_1)) + 1) \cdot \varsigma_1 \cdot \sqrt{d \cdot m} \\
&\leq 1 + (\varsigma_2 \cdot \sqrt{d \cdot m} \cdot 7/(\sqrt{m} \cdot \varsigma_1) + 1) \cdot \varsigma_1 \cdot \sqrt{d \cdot m} \\
&= \Omega(\varsigma_2 \cdot d \cdot \sqrt{m})
\end{aligned}$$

with probability overwhelming in $d \cdot k \geq \lambda$. In the above, the third $\leq$ is by Gaussian tail bound together with the bound on $s_{\max}(\sqrt{\mathbf{\Sigma}})$ from Cor. 6.1, the fourth $\leq$ by Lemma 6.6, and the last $=$ because $\varsigma_2 \cdot d$ dominates $\varsigma_1 \cdot \sqrt{d}$. □

In the rest, we focus on the case where $\mathcal{K}$ is a cyclotomic field. From Lemma 7.1, we have the following generalisation of [LPSS14, Lemma 19 (eprint)] without the restriction of $\mathcal{O}_{\mathcal{K}} = \mathbb{Z}$.

The original proof relies on [BF11, Lemma 4.5 and Theorem 4.3] which we replace with its generalisation to cyclotomic rings in Lemmas 2.14 and 2.15.

Lemma 7.2 (Generalisation of [LPSS14, Lemma 19 (eprint)]). *Let $\mathcal{K}$ be a cyclotomic field of degree $d \geq 2$ with conductor f. Let $k = r \geq 1$, $m \geq 64 \max(\lambda/\log(dk), k)$, $0 < \varepsilon \leq \mathsf{negl}(\lambda)$, and $\varsigma_1 \cdot \mathbf{I} = \mathbf{S}_X$ follow constraints of Theorem 5.1 and satisfy $\varsigma_1 \geq 10 \cdot \eta_\epsilon(\mathcal{O}_{\mathcal{K}})$. Additionally, let $\varsigma_2 \geq \varsigma_1 d\sqrt{mk} \cdot \omega\left(\eta_\varepsilon^{((m-r)d)} \cdot m^{1.5} \cdot d^{12} \cdot \delta_{\mathcal{K}}^{14} \cdot f^2 \cdot \Delta_{\mathcal{K}}^{4/d} \cdot \varsigma_1\right)$. Denote $\sqrt{\mathbf{\Sigma}} \coloneqq \mathrm{diag}(\varsigma_1 \mathbf{I}_m, \varsigma_2 \mathbf{I}_k)$, and suppose the following hold* [12]*:*

- *q is an unramified prime that splits into g ideals in $\mathcal{R}$;*
- *$1 \leq n \leq m$, $d \cdot (m-k) \geq \Omega(\lambda)$, and $ng/q^{d(m-n+1)/g} \leq \mathsf{negl}(\lambda)$;*
- *letting $\eta_A \geq 8\, d\sqrt{m} \cdot q^{n/m+2/(d \cdot m)}$, then there exists $\epsilon \leq \mathsf{negl}(\lambda), a \geq 1$, such that $\max\left\{\eta_A, 2\sqrt{d} \cdot (a^k q^n)^{1/(m-k)}\right\} \leq \varsigma_1$, $\varsigma_2 \leq \min\left\{q^{1/g}/\sqrt{m}, a \cdot \varsigma_1\right\}$.*

Denote by $\mathbf{e}_i \in \mathcal{R}^{(m+k)}$ the $(m+i)$-th unit vector. Let $\mathsf{HintTG}_{m,k}$ be the PPT algorithm in Fig. 1. Then, the following distributions are statistically close in λ:

$$\left\{ (\mathbf{B}, \mathbf{X}^{\mathsf{T}}) \;\middle|\; \begin{array}{l} \mathbf{A} \leftarrow \mathcal{R}_q^{n \times m} \\ (\mathbf{X}, \mathbf{U}) \leftarrow \mathsf{HintTG}_{m,k}(\varsigma_1, \varsigma_2) \\ \bar{\mathbf{U}}^{\mathsf{T}} \coloneqq \mathbf{U} \cdot (\mathbf{0}_{(m+k)\times k}, \mathbf{I}_m) \\ \mathbf{B} \coloneqq \mathbf{A} \cdot \bar{\mathbf{U}} \bmod q \end{array} \right\} \approx \left\{ (\mathbf{B}, \mathbf{X}) \;\middle|\; \begin{array}{l} \mathbf{B} \leftarrow \mathcal{R}_q^{n \times (m+k)} \\ \mathbf{X} \leftarrow \prod\limits_{i \in [k]} \mathcal{D}_{\Lambda_q^\perp(\mathbf{B}), \sqrt{\mathbf{\Sigma}}, \mathbf{e}_i} \end{array} \right\}.$$

Proof. We consider a sequence of hybrid distributions.

- H_0: The RHS distribution in the statement above.
- H_1: We first sample $\mathbf{X} \leftarrow \prod_{i \in [k]} \mathcal{D}_{\mathcal{R}^{m+k}, \sqrt{\mathbf{\Sigma}}, \mathbf{e}_i}$, then sample $\mathbf{B} \leftarrow \mathcal{R}_q^{n \times (m+k)}$ subject to $\mathbf{B} \cdot \mathbf{X} = \mathbf{0} \bmod q$.
- H_2: Same as H_1, except we sample $\mathbf{X}$ via $(\mathbf{X}^{\mathsf{T}}, \mathbf{U}) \leftarrow \mathsf{HintTG}_{m,k}(\varsigma_1, \varsigma_2)$.
- H_3: Same as H_2, except that we sample $\mathbf{A} \leftarrow \mathcal{R}_q^{n \times m}$ and set $\mathbf{B} \coloneqq \mathbf{A} \cdot \bar{\mathbf{U}} \bmod q$.

It holds that RHS $\approx_s \mathsf{H}_1$ by Lemma 2.14. Then it holds that $\mathsf{H}_1 \approx_s \mathsf{H}_2$ by Lemma 7.1 item 1. In the rest, we show $\mathsf{H}_2 \approx_s \mathsf{H}_3$, which completes the proof since $\mathsf{H}_3 \equiv$ LHS.

By Lemma 2.15, the columns of $\mathbf{X}$ sampled in H_0 are $\mathcal{R}_q$-linearly independent with overwhelming probability in λ. In both H_2 and H_3, the sampling of $(\mathbf{X}^{\mathsf{T}} = (\mathbf{X}_1, \mathbf{X}_2), \mathbf{U})$ is identical. Since we have shown RHS $\approx_s \mathsf{H}_2$, we have that columns of $\mathbf{X}$ from H_2 and H_3 are also $\mathcal{R}_q$-linearly independent with overwhelming probability in λ. Conditioned on the latter, we argue that the distribution of $\mathbf{B}$ in H_2 and H_3 are identical.

[12] The additional constraints are such that we can apply [ALLW25, Lemma 29 and Theorem 9 (eprint)] in the proof.

Consider an arbitrary sample $(\mathbf{X}, \mathbf{U})$ where $\mathbf{X}$ is $\mathcal{R}_q$-injective. In H_2, the matrix $\mathbf{B}$ is sampled uniformly from $S \coloneqq \left\{\mathbf{M} \in \mathcal{R}_q^{n\times(m+k)} : \mathbf{M}\cdot\mathbf{X} = \mathbf{0}_{n\times k} \bmod q\right\}$. Since $\mathbf{X}$ is injective, we have $|S| = q^{dn(m+k)}/q^{dnk} = q^{dnm}$. Next, observe that for $\bar{\mathbf{U}}$ in H_3 we have $\bar{\mathbf{U}}\cdot\mathbf{X} = \mathbf{0}_{m\times k}$ by Lemma 7.1 Item 2. Therefore, for any $\mathbf{A} \in \mathcal{R}_q^{n\times m}$ it holds that $\mathbf{A}\cdot\bar{\mathbf{U}}\cdot\mathbf{X} = \mathbf{0}_{m\times k} \bmod q$, implying that $\mathcal{R}_q^{n\times m}\cdot\bar{\mathbf{U}} \subseteq S$. Further, by Lemma 7.1 Item 2, $\det(\mathbf{U}) = 1$, thus (right-multiplication by) $\mathbf{U}$ is a bijective map, and (right-multiplication by) $\bar{\mathbf{U}}$ is injective. Thus, $|\mathcal{R}_q^{n\times m}\cdot\bar{\mathbf{U}}| = |\mathcal{R}_q^{n\times m}| = q^{dnm}$. Therefore, $\mathcal{R}_q^{n\times m}\cdot\bar{\mathbf{U}} = S$, and the two ways of sampling $\mathbf{B}$ are the same. This implies $\mathsf{H}_2 \approx_s \mathsf{H}_3$, concluding the proof. □

For a random matrix $\mathbf{A}$ and k Gaussian samples $(\mathbf{x}_i)_{i\in[k]}$ from $\Lambda_q^\perp(\mathbf{A})$, the k-SIS problem asks to find a short non-zero $\mathbf{u}$ which is not in the $\mathcal{K}$-span of $(\mathbf{x}_i)_{i\in[k]}$ and s.t. $\mathbf{A}\cdot\mathbf{u} = \mathbf{0} \bmod q$. See Definition 2.21 for a formal definition. The following theorem provides a reduction from M-SIS to k-M-SIS.

Theorem 7.3 (SIS-to-k-SIS reduction over cyclotomic rings). *Let $\mathcal{K}$ be a cyclotomic field of degree $d \geq 2$ with conductor f. Take the parameters $k, n, m, q \in \mathbb{N}$, $0 < \varepsilon \leq \mathsf{negl}(\lambda)$, $\varsigma_1 > 0$, $\sqrt{\boldsymbol{\Sigma}} \in \mathbb{R}^{d(m+k)\times d(m+k)}$ specified by Lemma 7.2. Further suppose that $\varsigma_2, \varsigma_1 \geq 8d\sqrt{m}\cdot q^{n/m+2/(d\cdot m)}$.*

Let $\beta_1 > 0$ and $\beta_0 \geq \Omega(\varsigma_2\cdot d\cdot\sqrt{m})\cdot\beta_1$. Denote $\mathsf{params}_0 = (\mathcal{O}_\mathcal{K}, q, n, m, \beta_0)$, $\mathsf{params}_1 = (\mathcal{O}_\mathcal{K}, q, n, m+k, k, \sqrt{\boldsymbol{\Sigma}}, \beta_1)$. For any PPT adversary $\mathcal{B}$, there is a PPT adversary $\mathcal{A}$ such that

$$\mathsf{Adv}^{\mathsf{SIS}_{\mathsf{params}_0}}_{\mathcal{A}}(\lambda) \geq (\mathsf{Adv}^{k\text{-}\mathsf{SIS}_{\mathsf{params}_1}}_{\mathcal{B}}(\lambda))^2/O(1) - \mathsf{negl}(\lambda).$$

Remark 7.4. Firstly, here we consider the hints with simple covariance matrices, but this approach extends directly to $\sqrt{\boldsymbol{\Sigma}} = \mathrm{diag}(\mathbf{S}_{X1}, \mathbf{S}_{X2})$. Secondly, the constant 64 in $m \geq 64\max(\lambda/\log(dk), k)$ can be reduced depending on the parametrisation of the problem. For example, we can get a suitable singular value lower bound for $\lambda = 128$, $k = 4\lambda$, $m = 10\max(\lambda/\log(dk), k)$. We leave more specific parameter choices to the applications.

Proof. Throughout the proof we denote by $\mathcal{R} = \mathcal{O}_\mathcal{K}$ and $\mathcal{R}_q = \mathcal{R}/q\mathcal{R}$. We consider an intermediate search problem, which asks a PPT algorithm $\mathcal{B}$ to solve Shift-k-SIS, defined below:

$\underline{\mathsf{Shift}\text{-}k\text{-}\mathsf{SIS}_\mathcal{B}(1^\lambda)}$

$\mathbf{B} \leftarrow \mathcal{R}_q^{n\times(m+k)};\ \mathbf{X} \leftarrow (\prod_{i\in[k]} \mathcal{D}_{\Lambda_q^\perp(\mathbf{B}), \sqrt{\boldsymbol{\Sigma}}, \mathbf{e}_i})$

$\mathbf{u} \leftarrow \mathcal{B}(\mathbf{B}, \mathbf{X})$

return $(\mathbf{B}\cdot\mathbf{u} = \mathbf{0} \bmod q) \wedge (\|\mathbf{u}\| \leq \beta_1) \wedge (\mathbf{u} \notin \mathcal{K}\text{-}\mathrm{span}(\mathbf{X}))$

The only difference between Shift-k-SIS and k-$\mathsf{SIS}_{\mathsf{params}_1}$ is the Gaussian parameters of $\mathbf{X}$, where in Shift-k-SIS the i-th preimage is offset by the i-th unit vector $\mathbf{e}_i$. The theorem follows from combining Lemmas 7.5 and 7.6 below.

Lemma 7.5. *Given the parameter constraints of Theorem 7.3, for any PPT algorithm $\mathcal{B}$, it holds that* $\mathsf{Adv}^{\mathsf{Shift}\text{-}k\text{-}\mathsf{SIS}}_{\mathcal{B}}(\lambda) \geq (\mathsf{Adv}^{k\text{-}\mathsf{SIS}_{\mathsf{params}_1}}_{\mathcal{B}}(\lambda))^2/O(1) - \mathsf{negl}(\lambda)$.

Proof. We define a sequence of hybrids:

- H_0: Identical to the k-$\mathsf{SIS}_{\mathsf{params}_1}$ experiment. We denote the problem instance as $(\mathbf{A}, \mathbf{X}) \in \mathcal{R}_q^{n\times(m+k)} \times \mathcal{R}^{(m+k)\times k}$.
- H_1: Same as H_0, except that we sample $\mathbf{A} \leftarrow \mathcal{R}_q^{n\times(m+k)}$ subject to $\eta_\epsilon(\Lambda_q^\perp(\mathbf{A})) \le 8d\sqrt{m} \cdot q^{n/m+2/(d\cdot m)}$ for some $\epsilon \le 2^{-2k}$.
- H_2: Same as H_1, except we shift the distribution $\mathbf{X} \leftarrow \prod_{i\in[k]} \mathcal{D}_{\Lambda_q^\perp(\mathbf{A}),\sqrt{\mathbf{\Sigma}}}$ to $\mathbf{X} \leftarrow \prod_{i\in[k]} \mathcal{D}_{\Lambda_q^\perp(\mathbf{A}),\sqrt{\mathbf{\Sigma}},\mathbf{e}_i}$.
- H_3: Same as H_2, except that we sample $\mathbf{A} \leftarrow \mathcal{R}_q^{n\times(m+k)}$ without the constraint on $\eta_\epsilon(\Lambda_q^\perp(\mathbf{A}))$, i.e. we undo the change introduced in H_1. Note that H_3 is identical to Shift-k-SIS.

Let $\mathcal{B}$ be a PPT solver against $\mathsf{H}_0 = k$-$\mathsf{SIS}_{\mathsf{params}_1}$ with advantage $\mathsf{Adv}_{\mathcal{B}}^{k\text{-}\mathsf{SIS}}(\lambda) \in \mathsf{poly}(\lambda)$. We claim:

1. $\mathcal{B}$ solves H_1 with advantage $\delta_1 \ge \mathsf{Adv}_{\mathcal{B}}^{\mathsf{SIS}}(\lambda) - 2^{-\lambda}$.
2. $\mathcal{B}$ solves H_2 with advantage $\delta_2 \ge \delta_1^2/O(1)$.
3. $\mathcal{B}$ solves $\mathsf{H}_3 = \mathsf{Shift}$-$k$-$\mathsf{SIS}$ with advantage $\delta_3 \ge \delta_2 - 2^{-\lambda}$.

Items 1 and 3 hold by Lemma 2.16, which says that the probability of $\eta_\epsilon(\Lambda_q^\perp(\mathbf{A})) \le 8d\sqrt{m} \cdot q^{n/m+2/(d\cdot m)}$ for any $\epsilon \le 2^{-2k}$ is at least $1 - 2^{-\Omega(dm)} \ge 1 - 2^{-\lambda}$. In the rest we argue that Item 2 holds, which concludes the proof.

We notice that the Rényi divergence between the distributions $P := k$-$\mathsf{SIS}_{\mathsf{params}_1}$ and $Q := \mathsf{Shift}$-k-SIS is identical to the one between the distributions of $\mathbf{X}$ in H_1 and H_2. Thus we use Lemma 2.18 to obtain $R(P\|Q) \le 5\exp\left(\frac{2\pi\cdot dk}{s_{\min}(\sqrt{\mathbf{\Sigma}})^2}\right) \in O(1)$. Note that the conditions required by Lemma 2.18 are satisfied since, for any $\epsilon \le 2^{-2k}$, we have $s_{\min}(\sqrt{\mathbf{\Sigma}}) \ge 8d\sqrt{m} \cdot q^{n/m+2/(d\cdot m)} \ge \eta_\epsilon(\Lambda_q^\perp(\mathbf{A}))$ in H_2, where the first inequality is by design of H_1. From here, Item 2 follows from Eq. (1). □

Lemma 7.6. *Following the parameter constraints of Theorem 7.3, for any PPT adversary $\mathcal{B}$, there is a PPT adversary $\mathcal{A}$ such that* $\mathsf{Adv}_{\mathcal{A}}^{\mathsf{SIS}_{\mathsf{params}_0}}(\lambda) \ge \mathsf{Adv}_{\mathcal{B}}^{\mathsf{Shift}\text{-}k\text{-}\mathsf{SIS}}(\lambda) - \mathsf{negl}(\lambda)$.

Proof. We construct the PPT algorithm $\mathcal{A}$. On input a $\mathsf{SIS}_{\mathsf{params}_0}$ instance $\mathbf{A}$, let $\mathcal{A}$ proceed as follows:

> Sample $(\mathbf{X}^\mathsf{T}, \mathbf{U}) \leftarrow \mathsf{HintTG}_{m,k}(\varsigma_1, \varsigma_2)$ using $\mathsf{HintTG}_{m,k}$ from Fig. 1.
> Let $\bar{\mathbf{U}}^\mathsf{T} \in \mathcal{R}^{(m+k)\times m}$ be the last m columns of $\mathbf{U}$ and $\mathbf{B} := \mathbf{A} \cdot \bar{\mathbf{U}} \bmod q$.
> Send $(\mathbf{B}, \mathbf{X})$ to $\mathcal{B}$, and receive a vector $\mathbf{u}$ from $\mathcal{B}$.
> Return $\mathbf{u}^* := \bar{\mathbf{U}} \cdot \mathbf{u}$.

By Lemma 7.2, the distribution of $(\mathbf{B}, \mathbf{X})$ is statistically close in λ to that from Shift-k-SIS. It remains to show that $\mathbf{u}^*$ returned by $\mathcal{A}$ is a SIS solution for $\mathbf{A}$. Suppose that $\mathbf{u}$ returned by $\mathcal{B}$ is a valid solution for Shift-k-SIS, meaning that $\mathbf{B} \cdot \mathbf{u} = \mathbf{0} \bmod q$, $0 < \|\mathbf{u}\| \le \beta_1$, and $\mathbf{u} \notin \mathcal{K}\text{-}\mathrm{span}(\mathbf{X})$. Then, we have

$$\mathbf{A} \cdot \mathbf{u}^* = \mathbf{A} \cdot \bar{\mathbf{U}} \cdot \mathbf{u} = \mathbf{B} \cdot \mathbf{u} = \mathbf{0} \bmod q.$$

Moreover, $\bar{\mathbf{U}} \cdot \mathbf{X} = \mathbf{0}$ by Lemma 7.1 Item 2, and $\mathbf{X}$ is $\mathcal{R}_q$-linearly independent with overwhelming probability by Lemma 2.15. Conditioned on the latter, $\mathbf{X}$ is an $\mathcal{R}_q$-basis of the (right-)kernel of $\bar{\mathbf{U}}$, and thus $\bar{\mathbf{U}} \cdot \mathbf{u} = \mathbf{0}$ would imply $\mathbf{u} \in \mathcal{R}_q\text{-span}(\mathbf{X})$, therefore also $\mathbf{u} \in \mathcal{K}\text{-span}(\mathbf{X})$, a contradiction. We thus conclude $\mathbf{u}^* = \bar{\mathbf{U}} \cdot \mathbf{u} \neq \mathbf{0}$ with overwhelming probability. Finally, the norm of $\mathbf{u}^*$ is bounded by

$$\|\mathbf{u}^*\| \leq \left\|\bar{\mathbf{U}}\right\| \cdot \|\mathbf{u}\| \leq \Omega(\varsigma_2 \cdot d \cdot \sqrt{m}) \cdot \beta_1 \leq \beta_0$$

with overwhelming probability, where the second inequality follows from Lemma 7.1 Item 3. Putting everything together, we conclude that $\mathbf{u}^*$ returned by $\mathcal{A}$ is a SIS solution for $\mathsf{SIS}_{\mathsf{params}_0}$ with overwhelming probability. □

□

LWE to k-LWE reduction over cyclotomic rings. Analogous to the above, using Theorem 5.1, we obtain a reduction from the LWE problem to the k-LWE problem over general cyclotomic rings. The reduction is similar to the above, we provide a sketch and refer to [LPSS14] for the details in the special case of $\mathcal{O}_\mathcal{K} = \mathbb{Z}$.

Recall that in the k-LWE problem (formal definition given in Definition 2.22) an adversary is given as challenge instance $(\mathbf{B}, \mathbf{X}, \mathbf{c})$, where $\mathbf{B} \leftarrow \mathcal{R}_q^{n \times (m+k)}$ is uniform and $\mathbf{X} \in \mathcal{R}^{(m+k) \times k}$ are Gaussian preimages satisfying $\mathbf{B}\mathbf{X} = \mathbf{0} \bmod q$, and is asked to distinguish whether $\mathbf{c} \in \mathcal{R}^{m+k}$ is an LWE sample, i.e. $\mathbf{c}^\intercal = \mathbf{s}^\intercal\mathbf{B} + \mathbf{e}^\intercal \bmod q$ for some secret $\mathbf{s}$ and error $\mathbf{e}$, or $\mathbf{c} = \mathbf{d} + \mathbf{e} \bmod q$ where $\mathbf{d}$ is a random sample over $\{\mathbf{d} : \mathbf{d}^\intercal\mathbf{X} = \mathbf{0} \bmod q\}$. In the LWE to k-LWE reduction, given an LWE instance $(\mathbf{A}, \hat{\mathbf{c}})$, we simulate $\mathbf{B}, \mathbf{X}$ using Lemma 7.2, in the same way as in the reduction in Theorem 7.3. It remains to simulate the LWE sample $\mathbf{c}$, for which we let $\mathbf{c}^\intercal := \hat{\mathbf{c}}^\intercal \cdot \bar{\mathbf{U}} + \bar{\mathbf{e}} \bmod q$ for some freshly sampled error $\bar{\mathbf{e}} \in \mathcal{R}^{m+k}$. By picking the Gaussian parameters of $\bar{\mathbf{e}}$ appropriately and appealing to standard Gaussian lemmas, we conclude that, if $\hat{\mathbf{c}}^\intercal = \mathbf{s}^\intercal\mathbf{A} + \hat{\mathbf{e}}^\intercal \bmod q$ then $\mathbf{c}^\intercal = (\mathbf{s}^\intercal\mathbf{A} + \hat{\mathbf{e}}^\intercal) \cdot \bar{\mathbf{U}} + \bar{\mathbf{e}} \approx \mathbf{s}^\intercal\mathbf{B} + \mathbf{e} \bmod q$ for $\mathbf{e}$ of the desired shape (independent of $\bar{\mathbf{U}}$), else if $\hat{\mathbf{c}}$ is uniformly random then $\mathbf{c}^\intercal \approx \mathbf{d} + \mathbf{e} \bmod q$.

Acknowledgements

We would like to thank Jonathan Husson for useful discussions on the Random Matrix Theory. Martin Albrecht's and Joël Felderhoff's work is supported by UKRI grant EP/Y02432X/1. Russell W. F. Lai and Ivy K. Y. Woo are supported by the Research Council of Finland projects No. 358951 and 358950 respectively. Oleksandra Lapiha was supported by the EPSRC and the UK Government as part of the Centre for Doctoral Training in Cyber Security for the Everyday at Royal Holloway, University of London (EP/S021817/1).

References

ACL+22. Martin R. Albrecht, Valerio Cini, Russell W. F. Lai, Giulio Malavolta, and Sri Aravinda Krishnan Thyagarajan. Lattice-based SNARKs: Publicly

verifiable, preprocessing, and recursively composable - (extended abstract). In Yevgeniy Dodis and Thomas Shrimpton, editors, *CRYPTO 2022, Part II*, volume 13508 of *LNCS*, pages 102–132. Springer, Cham, August 2022. doi:10.1007/978-3-031-15979-4_4.

AGHS13. Shweta Agrawal, Craig Gentry, Shai Halevi, and Amit Sahai. Discrete Gaussian leftover hash lemma over infinite domains. In Kazue Sako and Palash Sarkar, editors, *ASIACRYPT 2013, Part I*, volume 8269 of *LNCS*, pages 97–116. Springer, Berlin, Heidelberg, December 2013. doi:10.1007/978-3-642-42033-7_6.

ALLW25. Martin R. Albrecht, Russell W. F. Lai, Oleksandra Lapiha, and Ivy K. Y. Woo. Partial lattice trapdoors: How to split lattice trapdoors, literally. In Goichiro Hanaoka and Bo-Yin Yang, editors, *ASIACRYPT 2025, Part III*, volume 16247 of *LNCS*, pages 265–296. Springer, Singapore, December 2025. doi:10.1007/978-981-95-5099-9_9.

AR16. Divesh Aggarwal and Oded Regev. A note on discrete gaussian combinations of lattice vectors. *Chicago Journal of Theoretical Computer Science*, 2016.

Ban93. Wojciech Banaszczyk. New bounds in some transference theorems in the geometry of numbers. *Mathematische Annalen*, 296:625–635, 1993.

BdPMW16. Florian Bourse, Rafaël del Pino, Michele Minelli, and Hoeteck Wee. FHE circuit privacy almost for free. In Matthew Robshaw and Jonathan Katz, editors, *CRYPTO 2016, Part II*, volume 9815 of *LNCS*, pages 62–89. Springer, Berlin, Heidelberg, August 2016. doi:10.1007/978-3-662-53008-5_3.

BF11. Dan Boneh and David Mandell Freeman. Linearly homomorphic signatures over binary fields and new tools for lattice-based signatures. In Dario Catalano, Nelly Fazio, Rosario Gennaro, and Antonio Nicolosi, editors, *PKC 2011*, volume 6571 of *LNCS*, pages 1–16. Springer, Berlin, Heidelberg, March 2011. doi:10.1007/978-3-642-19379-8_1.

BL25. Katharina Boudgoust and Oleksandra Lapiha. Leftover hash lemma(s) over cyclotomic rings. Cryptology ePrint Archive, Paper 2025/1080, 2025. URL: https://eprint.iacr.org/2025/1080.

BS96. E. Bach and J. O. Shallit. *Algorithmic Number Theory: Efficient Algorithms*. MIT Press, 1996.

CDPR16. Ronald Cramer, Léo Ducas, Chris Peikert, and Oded Regev. Recovering short generators of principal ideals in cyclotomic rings. In Marc Fischlin and Jean-Sébastien Coron, editors, *EUROCRYPT 2016, Part II*, volume 9666 of *LNCS*, pages 559–585. Springer, Berlin, Heidelberg, May 2016. doi:10.1007/978-3-662-49896-5_20.

Coh93. Henri Cohen. Algorithms for algebraic number theory. In *A Course in Computational Algebraic Number Theory*. Springer, 1993.

FPS22. Joël Felderhoff, Alice Pellet-Mary, and Damien Stehlé. On module unique-SVP and NTRU. In Shweta Agrawal and Dongdai Lin, editors, *ASIACRYPT 2022, Part III*, volume 13793 of *LNCS*, pages 709–740. Springer, Cham, December 2022. doi:10.1007/978-3-031-22969-5_24.

GMPW20. Nicholas Genise, Daniele Micciancio, Chris Peikert, and Michael Walter. Improved discrete gaussian and subgaussian analysis for lattice cryptography. In Aggelos Kiayias, Markulf Kohlweiss, Petros Wallden, and Vassilis Zikas, editors, *PKC 2020, Part I*, volume 12110 of *LNCS*, pages 623–651. Springer, Cham, May 2020. doi:10.1007/978-3-030-45374-9_21.

JLWG25. Haoxiang Jin, Feng-Hao Liu, Zhedong Wang, and Dawu Gu. Discrete gaussians modulo sub-lattices: New leftover hash lemmas for discrete gaussians. In Tibor Jager and Jiaxin Pan, editors, *PKC 2025, Part II*, volume 15675 of *LNCS*, pages 301–330. Springer, Cham, May 2025. doi:10.1007/978-3-031-91823-0_10.

KLNO24. Michael Klooß, Russell W. F. Lai, Ngoc Khanh Nguyen, and Michal Osadnik. RoK, paper, SISsors toolkit for lattice-based succinct arguments - (extended abstract). In Kai-Min Chung and Yu Sasaki, editors, *ASIACRYPT 2024, Part V*, volume 15488 of *LNCS*, pages 203–235. Springer, Singapore, December 2024. doi:10.1007/978-981-96-0935-2_7.

KNSW20. Elena Kirshanova, Huyen Nguyen, Damien Stehlé, and Alexandre Wallet. On the smoothing parameter and last minimum of random orthogonal lattices. *DCC*, 88(5):931–950, 2020. doi:10.1007/s10623-020-00719-w.

LM00. B. Laurent and P. Massart. Adaptive estimation of a quadratic functional by model selection. *The Annals of Statistics*, 28(5):1302 – 1338, 2000. doi:10.1214/aos/1015957395.

LPR13. Vadim Lyubashevsky, Chris Peikert, and Oded Regev. A toolkit for ring-LWE cryptography. In Thomas Johansson and Phong Q. Nguyen, editors, *EUROCRYPT 2013*, volume 7881 of *LNCS*, pages 35–54. Springer, Berlin, Heidelberg, May 2013. doi:10.1007/978-3-642-38348-9_3.

LPSS14. San Ling, Duong Hieu Phan, Damien Stehlé, and Ron Steinfeld. Hardness of k-LWE and applications in traitor tracing. In Juan A. Garay and Rosario Gennaro, editors, *CRYPTO 2014, Part I*, volume 8616 of *LNCS*, pages 315–334. Springer, Berlin, Heidelberg, August 2014. doi:10.1007/978-3-662-44371-2_18.

LSS14. Adeline Langlois, Damien Stehlé, and Ron Steinfeld. GGHLite: More efficient multilinear maps from ideal lattices. In Phong Q. Nguyen and Elisabeth Oswald, editors, *EUROCRYPT 2014*, volume 8441 of *LNCS*, pages 239–256. Springer, Berlin, Heidelberg, May 2014. doi:10.1007/978-3-642-55220-5_14.

MG02. Daniele Micciancio and Shafi Goldwasser. *Complexity of lattice problems: a cryptographic perspective*, volume 671. Springer Science & Business Media, 2002.

MP12. Daniele Micciancio and Chris Peikert. Trapdoors for lattices: Simpler, tighter, faster, smaller. In David Pointcheval and Thomas Johansson, editors, *EUROCRYPT 2012*, volume 7237 of *LNCS*, pages 700–718. Springer, Berlin, Heidelberg, April 2012. doi:10.1007/978-3-642-29011-4_41.

MP13. Daniele Micciancio and Chris Peikert. Hardness of SIS and LWE with small parameters. In Ran Canetti and Juan A. Garay, editors, *CRYPTO 2013, Part I*, volume 8042 of *LNCS*, pages 21–39. Springer, Berlin, Heidelberg, August 2013. doi:10.1007/978-3-642-40041-4_2.

MR07. Daniele Micciancio and Oded Regev. Worst-case to average-case reductions based on gaussian measures. *SIAM journal on computing*, 37(1):267–302, 2007.

Neu13. Jürgen Neukirch. *Algebraic number theory*, volume 322. Springer Science & Business Media, 2013.

NP20. Hoi H. Nguyen and Elliot Paquette. Surjectivity of near-square random matrices. *Combinatorics, Probability and Computing*, 29(2):267–292, March 2020. doi:10.1017/S0963548319000348.

Pei08. Chris Peikert. Limits on the hardness of lattice problems in lp norms. *computational complexity*, 17(2):300–351, 2008.

PS21. Alice Pellet-Mary and Damien Stehlé. On the hardness of the NTRU problem. In Mehdi Tibouchi and Huaxiong Wang, editors, *ASIACRYPT 2021, Part I*, volume 13090 of *LNCS*, pages 3–35. Springer, Cham, December 2021. doi:10.1007/978-3-030-92062-3_1.

Reg09. Oded Regev. On lattices, learning with errors, random linear codes, and cryptography. *J. ACM*, 56(6), September 2009. doi:10.1145/1568318.1568324.

Sil85. Jack W. Silverstein. The smallest eigenvalue of a large dimensional wishart matrix. *The Annals of Probability*, 13(4), 1985. doi:10.1214/aop/1176992819.

SS11. Damien Stehlé and Ron Steinfeld. Making NTRU as secure as worst-case problems over ideal lattices. In Kenneth G. Paterson, editor, *EUROCRYPT 2011*, volume 6632 of *LNCS*, pages 27–47. Springer, Berlin, Heidelberg, May 2011. doi:10.1007/978-3-642-20465-4_4.

SS13. Damien Stehlé and Ron Steinfeld. Making NTRUEncrypt and NTRUSign as secure as standard worst-case problems over ideal lattices. Cryptology ePrint Archive, Report 2013/004, 2013. URL: https://eprint.iacr.org/2013/004.

A Numerical experiments on Gaussian matrix's kernels

In order to support our claim of the introduction that the bounds of Theorem 4.2 are very loose, we propose a sagemath program in Appendix C.2 (and as an attachment to this pdf). This program takes as input:

- f the conductor of a cyclotomic field;
- r a rank;
- m a width;
- β a block factor;
- ς a Gaussian parameter;
- $\mathsf{num_samples}$ an integer.

It then do $\mathsf{num_samples}$ times the following:

1. Sample $\mathbf{X} \leftarrow \mathcal{D}^{r\times m}_{\mathcal{O}_\mathcal{K},\varsigma}$;
2. Compute $\mathbf{X}' \in \mathbb{Z}^{rd\times md}$ the coefficient embedding of $\mathbf{X}$ with the power-basis;
3. Compute $\mathbf{B} \in \mathbb{Z}^{(m-r)d\times rd}$ a basis of $\Lambda^\perp(\mathbf{X}')$;
4. Compute $\mathbf{B}' \in \mathbb{Z}^{(m-r)d\times rd}$ the BKZ reduction of $\mathbf{B}$ with block-size $\beta\cdot(m-r)\cdot d$;
5. Compute the canonical embedding of each column of $\mathbf{B}'$ and return the maximal infinite norm.

The algorithm then output the quantils of the $\mathsf{num_samples}$ experiments (in Figs. 2 and 3 we give only the maximum). We could only test it for small values of m, r and f, but we give some plot in Figs. 2 and 3, where we take $\beta = 0.1$, $\varsigma = 2\sqrt{r\cdot d}$ and $\mathsf{num_samples} = 500$. We do not claim that those figures indicate anything for values of f, m and r relevant for cryptographic purposes. We included them as an indicator that more work is required to refine the bound of Lemma 4.12.

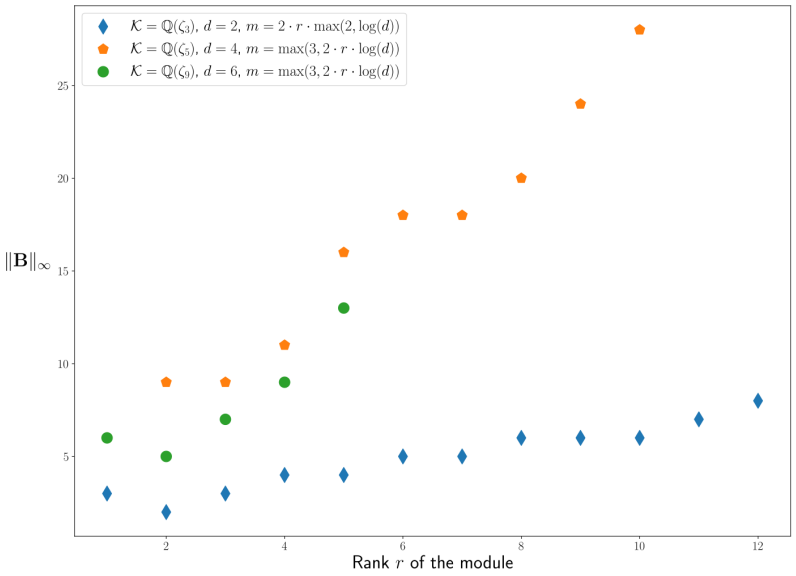


Fig. 2. Output of Appendix C.2 for constant conductor and varying rank.

B Miscellaneous results

Lemma B.1. *For any integer $N \geq 2$, the number of prime ideals dividing N is at most $d \cdot \log_2(N)$.*

Proof. If $\mathfrak{p}$ is a prime ideal dividing N, then it must be above a integral prime dividing N. Since prime numbers split in at most d prime ideals, and that the number of prime integer dividing a number N is always less than $\log_2(N)$, the result follows. □

Lemma B.2. *Let $\mathbf{\Sigma} = \mathbf{S}\mathbf{S}^\intercal - \varsigma^2\mathbf{I}$ with $\mathbf{S} \in \mathbb{R}^{dr \times dr} \geq x \cdot \varsigma\mathbf{I}$ for some $x \geq 1$. Let $(\mathbf{\Sigma}')^{-1} = (\mathbf{\Sigma}^{-1} + \varsigma^{-2}\mathbf{I})$. Then $\mathbf{\Sigma}' \geq \varsigma^2 \cdot \left(1 - 1/x^2\right)\mathbf{I}$.*

Proof. By definition, $\mathbf{\Sigma}$ is diagonalisable with eigenvalues $\geq (x^2 - 1) \cdot \varsigma^2$. The eigenvalues of $\mathbf{\Sigma}'$ are exactly the $(s^{-1} + \varsigma^{-2})^{-1}$ where the s are the eigenvalues of $\mathbf{\Sigma}$. We then have, for every $s \geq (x^2 - 1) \cdot \varsigma^2$:

$$\frac{1}{\frac{1}{s} + \frac{1}{\varsigma^2}} \geq \frac{\varsigma^2}{\frac{1}{x^2 - 1} + 1} = \varsigma^2 \cdot \left(1 - \frac{1}{x^2}\right).$$

□

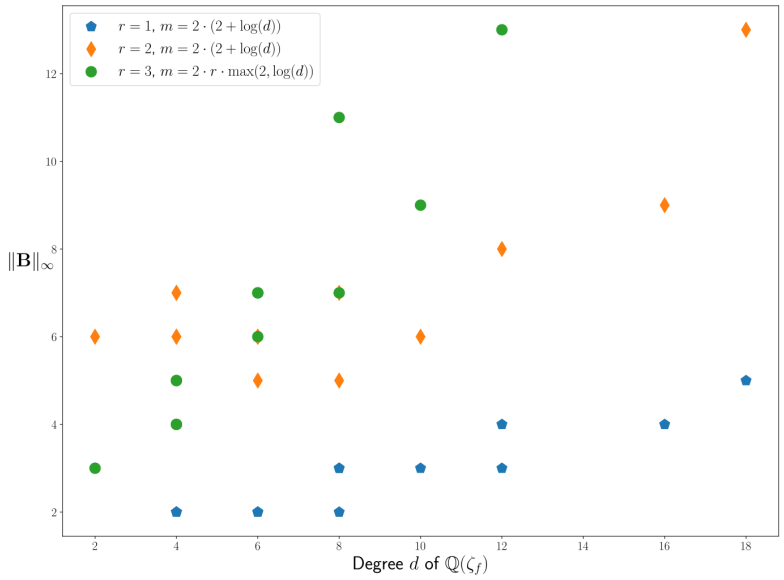


Fig. 3. Output of Appendix C.2 for constant rank and varying conductor.

Lemma B.3. *Let $\mathcal{K} = \mathbb{Q}(f)$ be the cyclotomic number field of conductor f, degree d and discriminant $\Delta_\mathcal{K}$. It holds that*

$$|\Delta_\mathcal{K}|^{1/d} \leq d.$$

Proof. Since for any $f = f_1 \cdot f_2$ with $\gcd(f_1, f_2) = 1$ it holds that

$$\left|\Delta_{\mathbb{Q}(\zeta_f)}\right|^{1/\phi(f)} = \left|\Delta_{\mathbb{Q}(\zeta_{f_1})}\right|^{1/\phi(f_1)} \cdot \left|\Delta_{\mathbb{Q}(\zeta_{f_2})}\right|^{1/\phi(f_2)},$$

and $\phi(f) = \phi(f_1) \cdot \phi(f_2)$, it suffices to prove the lemma for $f = p^k$. In that case, it holds that $d = p^{k-1}(p-1)$ and that $|\Delta_\mathcal{K}| = p^{p^{k-1}(pk-k-1)} = p^{d \cdot (pk-k-1)/(p-1)}$. Now we compute

$$\begin{aligned}\log_p(|\Delta_\mathcal{K}|^{1/d}/d) &= \frac{pk-k-1}{p-1} - \big(k - 1 + \log_p(p-1)\big) \\ &= \frac{p-2}{p-1} - \log_p(p-1) = 1 - \frac{1}{p-1} - \log_p(p-1).\end{aligned}$$

It can be computed that the function $f(x) = 1 - 1/(x-1) - \log(x-1)/\log(x)$ is ≤ 0 for $x \geq 2$, implying the result. □

Lemma B.4 ([KLNO24, Lemma 2]). *For any $f \geq 4$ and number field $\mathcal{K}$ containing a fth root of unity ζ_f, and $a \in \mathbb{Z}/f\mathbb{Z} \setminus \{0\}$, it holds that $(1-\zeta_f^a) \mid f$ and $\left\| f/(1-\zeta_f^a) \right\|_\infty \leq f^2/2$.*

Lemma B.5. *Let $A, B \geq 0$, then for any $x \geq 2(A + B\log(B))$, it holds that $x \geq A + B \cdot \log(x)$.*

C Source code

The Python source code files are available as attachments to this PDF.

C.1 Approximate the ζ function

```
import numpy as np

from sage.all import primes, mod, exp, RR, euler_phi, ln


def find_splitting(p, n):
    """
    Return (N, g), g being the number of prime above p in the nth
    cyclotomic, and N their norm https://math.stackexchange.com/a/1666456
    (Proposition 10.3 in Neukirch - Algebraic Number Theory)
    """
    k = 0
    d = n

    while d % p == 0:
        k += 1
        d /= p
    if d == 1:
        return (p, 1)

    r = mod(p, d).multiplicative_order()
    return (p**r, euler_phi(n) / (euler_phi(p**k) * r))


def find_approx_zeta_prime_cyclo(n, eps):
    """
    Return `(a, b)` such that the Prime zeta function of the cyclotomic
    field of conductor `n` is approximately equal to `b·a^x`.

    :param n:
    :param eps: a precision parameter.

    EXAMPLE::

        sage: n = 2**16 + 1
        sage: a, b = find_approx_zeta_prime_cyclo(n, 1e-100)
        sage: n, euler_phi(n), a, b
        (65537, 65536, 0.0000152585561077570, 1.00000011433444)
    """
    L = [find_splitting(p, n) for p in primes(1, n + 1)]
    L = [(N, a) for (N, a) in L if N < 1 / eps]

    def approx_zeta_prime(x):
        return sum([RR(a / N**x) for (N, a) in L])

    X = list(range(2, 20))  # No particular reason for the choice of 20
    Y1 = [approx_zeta_prime(x) for x in X]
```

```
    Y = [float(ln(y)) for y in Y1 if y != 0]
    if len(Y) == 0:
        return (None, None)

    X = X[: len(Y)]
    a, b = np.polyfit(X, Y, 1)
    return RR(exp(a)), RR(exp(b))
```

C.2 Compute the basis size of the kernel of random Gaussian matrices

```
from sys import argv
from sage.all import vector, RDF, matrix, ZZ, CyclotomicField
from sage.interfaces.r import r as r_stats

from sage.stats.distributions.discrete_gaussian_lattice import (
    DiscreteGaussianDistributionLatticeSampler as DGSL,
)

import multiprocessing
import argparse


def nf_element_to_rd_vector(x, places):
    """
    Compute the canonical embedding of x.

    (separating real and imaginary part for complex embeddings)

    :param x:
    :param places:

    """
    cc_elts = [f(x) for f in places]
    rr_elts = []
    for z in cc_elts:
        rr_elts.append(z.real())
        rr_elts.append(z.imag())
    return vector(rr_elts)


def sample_gaussian_element(DGaussian, basis_canonical_emb, basis_roi):
    """
    TODO describe function

    :param DGaussian:
    :param basis_canonical_emb:
    :param basis_roi:
    :returns:

    """
    d = len(basis_roi)
    coords = vector(RDF, DGaussian() * basis_canonical_emb.inverse())
    coords = vector([x.round() for x in coords])
    return sum([coords[i] * basis_roi[i] for i in range(d)])


def rotation_matrix(x, OK, basis_roi):
    """
    Return the multiplication matrix of x in coordinate embedding.

    :param x:
    :param OK:
    :param basis_roi:

    """
    return matrix([list(OK.coordinates(b * x)) for b in basis_roi])
```

```
def ok_matrix_to_rotation_matrix(M, OK, basis_roi):
    """
    Return a ZZ-rotation matrix from an OK matrix.

    :param M:
    :param OK:
    :param basis_roi:

    """
    L = [[rotation_matrix(x, OK, basis_roi) for x in r] for r in M.rows()]
    return matrix.block(ZZ, L)


def infinity_norm_elt(x, places):
    """
    Return the infinite norm in canonical embedding of x

    :param x:
    :param places:

    """
    return max([abs(f(x)) for f in places])


def infinity_norm_matrix(M, places):
    """
    Return the maximum infinite norm (in canonical embedding) of all
        ↪ coefficients of M.

    :param M:
    :param places:

    """
    return max([infinity_norm_elt(x, places) for x in M.coefficients()])


def print_result(f, d, m, r, beta, out):
    """
    Compute and print the quartiles of the output of run_batch

    :param f:
    :param d:
    :param m:
    :param r:
    :param beta:
    :param out:

    """
    stats = dict(
        zip(
            (Q := [0, 1 / 4, 1 / 2, 3 / 4, 1]),
            r_stats.quantile(out, Q)._sage_()["DATA"],
        )
    )
    print(
        f"f={f} d={d} r={r} m={m} (dim {(m-r)*d} kernel, block-size {int((m-r
            ↪ )*d*beta)})"
        "\n"
        f"{stats[0]}, {stats[1/4]}, {stats[1/2]}, {stats[3/4]}, {stats[1]}"
        "\n"
    )


class Worker:
    """
    This worker is used for parallel computing.
```

```
    It samples a Gaussian matrix of size r*m with parameter sigma in OK and
    compute a beta-BKZ-reduced basis of its kernel and returns the infinity
        ↪ norm
    of this basis.
    """

    def __init__(
        self, r, m, sigma, OK, basis_canonical_emb, basis_roi, places,
            ↪ block_size
    ):
        """
        TODO describe function

        :param r:
        :param m:
        :param sigma:
        :param OK:
        :param basis_canonical_emb:
        :param basis_roi:
        :param places:
        :param block_size:
        :returns:

        """
        self.r = r
        self.m = m
        self.basis_canonical_emb = basis_canonical_emb
        self.basis_roi = basis_roi
        self.block_size = block_size
        self.places = places
        self.OK = OK
        self.sigma = sigma

    def __call__(self, _):
        """
        TODO describe function

        """
        DGaussian = DGSL(self.basis_canonical_emb, self.sigma)

        # Sample a r*m matrix from the Gaussian distribution
        X = matrix(
            [
                [
                    sample_gaussian_element(
                        DGaussian, self.basis_canonical_emb, self.basis_roi
                    )
                    for _ in range(self.r)
                ]
                for _ in range(self.m)
            ]
        )
        # Make it into a ZZ-matrix to run BKZ
        coef_X = ok_matrix_to_rotation_matrix(X, self.OK, self.basis_roi)

        basis_kernel = coef_X.kernel().matrix()
        # If there is no kernel, return 0 (this means that m should be chosen
            ↪  larger)
        if basis_kernel.nrows() == 0:
            return 0

        # Run BKZ on the kernel basis
        small_basis_kernel = basis_kernel.BKZ(block_size=self.block_size)
        return infinity_norm_matrix(small_basis_kernel, self.places)


def run_batch(
```

```
    r,
    m,
    sigma,
    beta,
    OK,
    basis_canonical_emb,
    basis_roi,
    places,
    num_samples,
    num_workers,
):
    """
    Runs num_samples Workers in parallel using num_workers threads

    :param r:
    :param m:
    :param sigma:
    :param beta:
    :param OK:
    :param basis_canonical_emb:
    :param basis_roi:
    :param places:
    :param num_samples:
    :param num_workers:

    """
    block_size = max(1, int((m - r) * OK.degree() * beta))
    worker = Worker(r, m, sigma, OK, basis_canonical_emb, basis_roi, places,
        ↪ block_size)
    with multiprocessing.Pool(processes=num_workers) as pool:
        results = pool.map(worker, range(1, num_samples))
    return sorted(results)


def fixed_mode(f, r, m, sigma, beta, num_samples, num_workers):
    """
    Compute stats on the size of the kernel of Gaussian matrices with fixed
    rank r and number field QQ(zeta_f).

    :param f:
    :param r:
    :param m:
    :param sigma:
    :param beta:
    :param num_samples:
    :param num_workers:
    """
    K, zeta = CyclotomicField(f, "zeta").objgen()
    d = K.degree()
    OK = K.ring_of_integers()
    basis_roi = OK.basis()

    print(
        f"Fixed mode, parameters: f={f}, d={d}, m={m}, "
        f"sigma={sigma}, beta={beta}, num_samples={num_samples}"
    )

    m = int(eval(m))
    sigma = float(eval(sigma))
    places = K.places()

    basis_canonical_emb = matrix(
        RDF, [list(nf_element_to_rd_vector(b, places)) for b in basis_roi]
    )
    out = run_batch(
        r,
        m,
        sigma,
```

```
        beta,
        OK,
        basis_canonical_emb,
        basis_roi,
        places,
        num_samples,
        num_workers,
    )
    print_result(f, d, m, r, beta, out)


def fixed_rank_mode(r, f_min, f_max, m, sigma, beta, num_samples, num_workers
    ↪ ):
    """
    Compute stats on the size of the kernel of Gaussian matrices with fixed
    rank r and in varying number field QQ(zeta_f).

    :param r:
    :param f_min:
    :param f_max:
    :param m:
    :param sigma:
    :param beta:
    :param num_samples:
    :param num_workers:
    """
    print(
        f"Fixed rank mode, parameters: r={r}, f_min={f_min}, f_max={f_max}, "
        "m={m}, sigma={sigma}, beta={beta}, num_samples={num_samples}"
    )
    for f in range(f_min, f_max):
        K, _ = CyclotomicField(f, "zeta").objgen()
        d = K.degree()
        OK = K.ring_of_integers()
        basis_roi = OK.basis()

        current_m = int(eval(m))
        current_sigma = float(eval(sigma))

        places = K.places()
        basis_canonical_emb = matrix(
            RDF, [list(nf_element_to_rd_vector(b, places)) for b in basis_roi
                ↪ ]
        )

        out = run_batch(
            r,
            current_m,
            current_sigma,
            beta,
            OK,
            basis_canonical_emb,
            basis_roi,
            places,
            num_samples,
            num_workers,
        )
        print_result(f, d, current_m, r, beta, out)


def fixed_conductor_mode(f, r_min, r_max, m, sigma, beta, num_samples,
    ↪ num_workers):
    """
    Compute stats on the size of the kernel of Gaussian matrices with
    varying rank and fixed number field QQ(zeta_f).

    :param f:
    :param r_min:
```

```
    :param r_max:
    :param m:
    :param sigma:
    :param beta:
    :param num_samples:
    :param num_workers:
    """
    print(
        f"Fixed conductor mode, parameters: f={f}, r_min={r_min}, f_max={
            ↪ r_max}, "
        f"m={m}, sigma={sigma}, beta={beta}, num_samples={num_samples}"
    )

    K, _ = CyclotomicField(f, "zeta").objgen()
    d = K.degree()
    OK = K.ring_of_integers()
    basis_roi = OK.basis()
    places = K.places()
    basis_canonical_emb = matrix(
        RDF, [list(nf_element_to_rd_vector(b, places)) for b in basis_roi]
    )

    for r in range(r_min, r_max):

        current_m = int(eval(m))
        current_sigma = float(eval(sigma))

        out = run_batch(
            r,
            current_m,
            current_sigma,
            beta,
            OK,
            basis_canonical_emb,
            basis_roi,
            places,
            num_samples,
            num_workers,
        )
        print_result(f, d, current_m, r, beta, out)


def main():
    global number_trial
    parser = argparse.ArgumentParser(
        formatter_class=argparse.RawDescriptionHelpFormatter,
        description="""Calculate statistics on the size in infinity norm of
            ↪ the
canonical embedding of a beta-reduced basis of the kernel of some Gaussian
matrix on the ROI of a cyclotomic number field. In rank mode, the number
field is fixed and r_min <=r < r_max. In conductor mode, the rank is fixed
and f_min <=f < f_max. In fixed mode, r and f are fixed.

Format of the output: conductor degree rank width (ZZ-dimension of the kernel
    ↪ ,
BKZ block-size) min, 1st quartile, median, 3rd quartile, maximum
""",
    )
    parser.add_argument(
        "-f", type=int, help="The conductor of the number field (rank or
            ↪ fixed mode)"
    )
    parser.add_argument(
        "-r",
        type=int,
        help="The rank of the module to consider (conductor or fixed mode)",
    )
```

```
parser.add_argument(
    "-m",
    type=str,
    help="Formula for m in function of r, f and d (in python syntax).
        ↪ Default 2*r*log(d)",
    default="2*r*log(d)",
)

parser.add_argument(
    "--sigma",
    type=str,
    help="Formula for sigma in function of r, f and d (in python syntax).
        ↪  Default 2*sqrt(rd)",
    default="2*sqrt(r*d)",
)

parser.add_argument(
    "--f_min",
    type=int,
    help="The min conductor of the number field (in conductor mode)",
)
parser.add_argument(
    "--f_max",
    type=int,
    help="The max conductor of the number field (in conductor mode)",
)

parser.add_argument(
    "--r_min", type=int, help="The min rank of the module (in rank mode)"
)
parser.add_argument(
    "--r_max", type=int, help="The max rank of the module (in rank mode)"
)

parser.add_argument(
    "--beta",
    type=float,
    help="The block-size coefficient for BKZ (default 0.5). The final
        ↪ block-size will be beta*m*r",
    default=0.5,
)

parser.add_argument(
    "--samples",
    type=int,
    default=100,
    help="Number of sample to use (default: 100)",
)

parser.add_argument(
    "--cores", type=int, default=0, help="Number of cores to use (default
        ↪ : all)"
)
if len(argv) == 1:
    parser.print_help()
    exit(1)

args = parser.parse_args()

num_workers = args.cores if args.cores > 0 else None
number_trial = args.samples
beta = args.beta
num_samples = args.samples
m = args.m
sigma = args.sigma

if args.r is not None and args.f is not None:
    f = args.f
```

```
        r = args.r
        fixed_mode(f, r, m, sigma, beta, num_samples, num_workers)
    elif args.r is not None:
        r = args.r
        f_min = args.f_min
        f_max = args.f_max
        fixed_rank_mode(r, f_min, f_max, m, sigma, beta, num_samples,
            ↪ num_workers)
    elif args.f is not None:
        f = args.f
        r_min = args.r_min
        r_max = args.r_max
        fixed_conductor_mode(f, r_min, r_max, m, sigma, beta, num_samples,
            ↪ num_workers)


if __name__ == "__main__":
    main()
```

C.3 Parametrising singular value inequalities

```
from sage.symbolic.relation import solve

"""
    This script simplifies the tail bounds on Chi random variables defined in
        ↪  Section 6, Lemma 6.1 of the paper.

    We first write down the expressions obtain from the Gershgorin circle
        ↪ theorem that dominate other values in the maximum (resp. minimum).
        ↪  They are stated as a comment. Then we manually write down the
        ↪ corresponding tail bounds defined in variables f1 to f7.

    Next we simplify the tail bounds only making them larger (resp. smaller).
        ↪  Lastly, we plot the difference between our values and the desired
        ↪  upper (resp. lower) bound 1 + 1/x (resp. 1 - 1/x).

    We check that the inequalities work by plotting them. The larger we make
        ↪ m the closer the values are to the bound, hence we parametrise it
        ↪ with c and give the user the option to plot for different values
        ↪ of c. In the paper we use c = 7 and x \geq 4 since it satisfies
        ↪ all inequalities.

    This script has a range of different plots we draw via show(plot(...)).
        ↪ To change the plot displayed please comment out the current plot
        ↪ and uncomment the one required. In all plots the bound is
        ↪ satisfied whenever the line is above zero.
"""

var("m,k,r,c,x")
r = k   # We set r = k for simplicity, in the paper we set m = c^2 * x^2 *
    ↪ max(r,k).
m = c^2 * x^2 * k

# We only need the inequality plots for positive parameters.
assume(c>1)
assume(x>1)

# THE UPPER TAIL

# The random variables we analyse:

# x_(m-1)^2 + y_(r-1)^2 + x_(m) * y_(r-1) + x_(m-1) * y_(r-2)

# Their norm divided by m is smaller than:

f1 = (m - 1)/m + 2 * sqrt((m-1)*k/m^2) + 2*k/m + (r - 1)/m + 2 * sqrt((r-1)*k
    ↪ /m^2) + 2*k/m + sqrt(1 + 2* sqrt(k/m) + 2*k/m) * sqrt((r - 1)/m + 2 *
```

```
    ↪ sqrt((r-1)*k/m^2) + 2*k/m) + sqrt((m-1)/m + 2 * sqrt((m-1)*k/m^2) + 2*
    ↪ k/m) * sqrt((r - 2)/m + 2 * sqrt((r-2)*k/m^2) + 2*k/m)

# We simplify the expression making the bound looser

f1 = f1.subs({(m-1):m})
f1 = f1.subs({(k-1):k})
f1 = f1.subs({(k-2):k})
f1 = f1.canonicalize_radical()

# Next random variables we analyse:

# x_(m)^2 / m + x_(m) * y_(r-1) / m <=

f2 = 1 + 2 * sqrt(k/m) + 2*k/m + sqrt(1 + 2* sqrt(k/m) + 2*k/m) * sqrt((r -
    ↪ 1)/m + 2 * sqrt((r-1)*k/m^2) + 2*k/m)

# We simplify the expression

f2 = f2.subs({(k-1):k})
f2 = f2.canonicalize_radical()

# The plot of f1 and f2 compared with 1 + 1/x.

show(plot(((1 + 1/x) - f2).subs(c=7), 1, 10) + plot(((1 + 1/x) - f1).subs(c
    ↪ =7), 1, 10, color = "green"))

# THE LOWER TAIL

"""
    Here O \cdot Z \cdot Z^\transpose \cdot O^\transpose has different form
        ↪ depending on r so we consider cases r >= 4, r = 3 and r = 2.
        ↪ Otherwise the lower tail is analysed in the same way.
"""

# r >= 4

# Random variable we analyse:

# x_(m)^2 /m - x_(m) * y_(r-1) /m >=

f3 = 1 - 2 * sqrt(k/m) - sqrt(1 + 2 * sqrt(k/m) + 2 * k/m) * sqrt((r - 1)/m +
    ↪  2 * sqrt((r-1)*k/m^2) + 2 * k/m)

# We simplify the expression

f3 = f3.subs({(k-1):k})
f3 = f3.canonicalize_radical()

# Random variable we analyse:

# x_(m - r + 2)^2 / m + y_(2)^2 / m - x_(m - r + 2)  * y_(1) / m - x_(m - r +
    ↪  3) * y_(2) / m >=
# we expect this one to be the smallest

f4 = 1 - (r - 4)/m - 2 * sqrt((m - r + 4)*k/m^2) - sqrt((m - r + 2)/m + 2 *
    ↪ sqrt((m - r + 2)*k/m^2) + 2 * k/m) * sqrt(1/m + 2 * sqrt(k/m^2) + 2 *
    ↪ k/m) - sqrt((m - r + 3)/m + 2 * sqrt((m - r + 3)*k/m^2) + 2 * k/m) *
    ↪ sqrt(2/m + 2 * sqrt(2*k/m^2) + 2 * k/m)

# We simplify the expression

f4 = f4.subs({(k-4):k})
f4 = f4.subs({(m - r + 4):m})
f4 = f4.subs({(m - r + 2):m})
f4 = f4.subs({(m - r + 3):m})
f4 = f4.subs(k = 4)
f4 = f4.canonicalize_radical()
```

```
# Random variable we analyse:

# x_(m - r + 1)^2 / m + y_(1)^2 / m - x_(m - r + 2) * y_(1) / m >=

f5 = 1 - (r - 2)/m - 2 * sqrt((m - r + 2)*k/m^2) - sqrt((m - r+ 2)/m + 2 *
    ↪ sqrt((m - r + 2)*k/m^2) + 2 * k/m) * sqrt(1/m + 2 * sqrt(k/m^2) + 2 *
    ↪ k/m)

# We simplify the expression

f5 = f5.subs({(k-2):k})
f5 = f5.subs({(m - r + 2):m})
f5 = f5.subs(k = 4)
f5 = f5.canonicalize_radical()

# c = 6 works for x > 4

# Uncomment to show the plot one at a time.

#show(plot((f3 - (1 - 1/x)).subs(c=6), 1, 10) + plot((f4 - (1 - 1/x)).subs(c
    ↪ = 6), 1, 10, color = "green") + plot((f5 - (1 - 1/x)).subs(c=6), 1,
    ↪ 10, color = "red"))

# special case for r = 3 (since we have m - r + 4) f6 = f4 but the bounds
    ↪ after are different

f6 = (m - r + 4)/m - 2 * sqrt((m - r + 4)*k/m^2) - sqrt((m - r+ 2)/m + 2 *
    ↪ sqrt((m - r + 2)*k/m^2) + 2 * k/m) * sqrt(1/m + 2 * sqrt(k/m^2) + 2 *
    ↪ k/m) - sqrt((m - r + 3)/m + 2 * sqrt((m-r+3)*k/m^2) + 2 * k/m) * sqrt
    ↪ (2/m + 2 * sqrt(2*k/m^2) + 2 * k/m)

# We simplify the expression

f6 = f6.subs({(m - r + 4):(m+1)})
f6 = f6.subs({(m - r + 2):m})
f6 = f6.subs({(m - r + 3):m})
f6 = f6.subs(k = 3)
f6 = f6.canonicalize_radical()

# since k = 3 now the combined effect makes c = 6 require x > 8 which is a
    ↪ bit worse than for r > 4

# Uncomment to show the plot one at a time.

#show(plot((f3 - (1 - 1/x)).subs(c=7), 1, 10) + plot((f6 - (1 - 1/x)).subs(c
    ↪ =7), 1, 10, color = "green") + plot((f5 - (1 - 1/x)).subs(c=7), 1, 10,
    ↪  color = "red"))

# for r = 2 the matrix looks simpler and we only get x_(m)^2 / m - x_(m) * y_
    ↪ (r-1)) / m >=

f7 = 1 - 2 * sqrt(k/m) - sqrt( 1 + 2 * sqrt(k/m) + 2 * k/m) * sqrt((r - 1)/m
    ↪ + 2 * sqrt((r-1)*k/m^2) + 2 * k/m)

# We simplify the expression

f7 = f7.subs({(k-1):k})
f7 = f7.canonicalize_radical()

# Uncomment to show the plot one at a time.

#show(plot((f7 - (1 - 1/x)).subs(c=5), 1, 10))
```