Gaussian Kernel Lattices and Smoothing Bounds from Theta Integrals

Samed Düzlü[1], Nihar Gargava[2], and Erkan Tairi[3]

[1] Technical University of Munich
samed.duzlu@tum.de

[2] Institut de Mathématiques d'Orsay, Université Paris-Saclay
nihar.gargava@universite-paris-saclay.fr

[3] University of California, Berkeley
erkan.tairi@berkeley.edu

Abstract. A Gaussian leftover hash lemma (LHL) states that, for a matrix $\boldsymbol{X}$ with discrete Gaussian columns and a sufficiently wide Gaussian vector $\boldsymbol{v}$, the product $\boldsymbol{Xv}$ is close to a discrete Gaussian. Over the integers this is a classical tool, and Albrecht–Felderhoff–Lai–Lapiha–Woo (EUROCRYPT'26) recently extended it to modules over a number field K of degree d. In every version, the lemma rests on two facts about the kernel lattice of $\boldsymbol{X}$: that $\boldsymbol{X}$ is surjective, and that the smoothing parameter of the kernel is small.

We prove both facts with a single analytic estimate. An exact identity expresses the Gaussian mass of the dual kernel, multiplied by the index of the image of $\boldsymbol{X}$, as an integral of one-dimensional periodic Gaussian sums over a torus. Bounding this integral below two forces the index to be one, which is surjectivity, and bounds the smoothing parameter at the same time. No separate surjectivity argument and no GRH are needed.

For smoothing error $2^{-\lambda}$ the smoothing parameter is $O_K(\sqrt{\lambda})$ with explicit constants, once $\boldsymbol{X}$ has $m \gtrsim dr\ln(srm)$ columns, and a deterministic lower bound shows that this order is optimal. The width s of the matrix enters only through the number of columns. Over power-of-two cyclotomic fields the theorem even applies to matrices whose integer coefficients follow a discrete Gaussian of width one. In the SIS-to-k-SIS reduction of Albrecht et al. (EUROCRYPT'26), the new bound lowers the hint width from $d^{39/2}\lambda^{9/2}$ to $d^{3/2}\lambda^{3/2}$ and the norm loss from $d^{41/2}\lambda^{5}$ to $d^{3/2}\lambda^{2}$, up to logarithmic factors. Additionally, we also formalized the main results of the paper in Lean.

Table of Contents

1 Introduction . . . 3
- 1.1 Our contributions . . . 4
- 1.2 Related work . . . 7
- 1.3 Parameters and comparison . . . 7
- 1.4 Open problems . . . 8
- 1.5 Organization . . . 9

2 Technical Overview . . . 9
- 2.1 Lattices, theta functions and Poisson summation . . . 10
- 2.2 Smoothing parameter, discrete Gaussians and statistical distance . . . 12
- 2.3 The Gaussian leftover hash lemma . . . 13
- 2.4 An exact theta integral . . . 13
- 2.5 Exploiting the theta integral . . . 15
- 2.6 The main result at constant width . . . 15
- 2.7 The one-column estimate . . . 16
- 2.8 Proof of the main theorem . . . 17
- 2.9 The other regimes and where they are treated . . . 19

3 Preliminaries . . . 20
- 3.1 Canonical geometry and notation . . . 20
- 3.2 Coefficient coordinates and basis constants . . . 20
- 3.3 Prime-power cyclotomic fields . . . 21
- 3.4 Metric change . . . 22
- 3.5 Discrete Gaussian tools . . . 22
- 3.6 Spectral estimates and exact shaping . . . 23

4 The Dual Kernel Lattice . . . 25
- 4.1 Covolume and the description of the dual . . . 25
- 4.2 The Gaussian-heuristic scale and a lower bound . . . 26
- 4.3 Two regimes . . . 27

5 Kernel Smoothing from an Exact Theta Integral . . . 27
- 5.1 The image index and the exact theta integral . . . 28
- 5.2 The main term and the remainder . . . 29
- 5.3 The finite theorem at polynomial width . . . 32
- 5.4 Constant coefficient width in power-of-two fields . . . 34
- 5.5 A matching lower bound, and the failure of the Gaussian heuristic . . . 37

6 Gaussian Leftover Hash Lemmas . . . 38
- 6.1 Anisotropic and spherical leftover hash lemmas . . . 38
- 6.2 Parameters of the SIS-to-k-SIS reduction . . . 40

7 Comparison with Prior Work . . . 46
- 7.1 Comparison with earlier bounds . . . 46
- 7.2 Sources of the compared bounds . . . 48

1 Introduction

The *leftover hash lemma* (LHL) is a fundamental information-theoretic result in cryptography first established by Impagliazzo, Levin, and Luby [ILL89]. It formalizes the principle that a sufficiently high-entropy source can be compressed into a distribution that is statistically close to uniform using a randomly chosen universal hash function. In particular, the resulting output remains close to uniform even when the hash function is revealed, and the statement is robust to arbitrary side information about the source. This robustness is particularly useful when the source may be subject to arbitrary leakage or when its precise distribution is inconvenient to describe. In cryptographic applications specifically, the LHL can be used in a security reduction as an information-theoretic argument to switch elements that come from a certain distribution by uniformly random, independent values.

The LHL appears quite naturally in lattice-based cryptography. A basic example is the linear map

$$H_{\boldsymbol{A}} : \{0,1\}^m \longrightarrow \mathbb{Z}_q^r, \qquad H_{\boldsymbol{A}}(\boldsymbol{x}) = \boldsymbol{A}\boldsymbol{x} \bmod q,$$

where q is prime, $m > r$, and $\boldsymbol{A} \xleftarrow{\$} \mathbb{Z}_q^{r\times m}$ is uniformly random. We can view the columns of $\boldsymbol{A}$ as a family of subset-sum instances in $\mathbb{Z}_q^r$. If $\boldsymbol{x}$ comes from a distribution with sufficiently high min-entropy, the LHL says that the pairs $(\boldsymbol{A}, \boldsymbol{A}\boldsymbol{x})$ and $(\boldsymbol{A}, \boldsymbol{u})$ with uniform $\boldsymbol{u} \xleftarrow{\$} \mathbb{Z}_q^r$, are statistically close. In particular, this statement does not rely on lattice-based hardness assumptions like SIS. We note that an attacker would be able to distinguish the cases when $\boldsymbol{A}$ is not surjective and the second component does not lie in the range of $\boldsymbol{A}$. It is known that in this setting, the overwhelming majority of matrices are surjective, and the LHL carries this statement even further.

This LHL version is already a useful ingredient to cryptographic applications. For example, it is used in the analysis of lattice-based identification and Fiat-Shamir protocols, where it can show that linear images of sufficiently high-entropy masking values are statistically close to uniform. Most prominently, the LHL is utilized in this form by Lyubashevsky in [Lyu09].

There are two closely related ways in which this phenomenon enters lattice-based cryptography. The first keeps the finite quotient $\mathbb{Z}_q^r$, while the second replaces the finite quotient by an infinite lattice setting.

The finite setting: subset sums and SIS. The subset-sum interpretation is perhaps the most direct route and the perspective that was initially adopted by Ajtai [Ajt96], and appeared again in Regev's work [Reg05]. Finding a preimage of $\boldsymbol{y} \in \mathbb{Z}_q^r$ under $H_{\boldsymbol{A}}$ amounts to finding a binary vector $\boldsymbol{x}$ such that $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{y} \pmod q$. When m is larger than $r \log q$, the LHL implies that a random target $\boldsymbol{y}$ has, with overwhelming probability, many binary preimages. In particular, this statement ensures the existence of strong trapdoor functions from lattices coined *preimage sampleable functions* in Gentry, Peikert, and Vaikuntanathan's work [GPV08].

A different perspective arises by enlarging the domain from $\{0,1\}^m$ to a set of short integer vectors, $B_\beta = \{\boldsymbol{x} \in \mathbb{Z}^m : \|\boldsymbol{x}\| \le \beta\}$. Given a matrix $\boldsymbol{A}$, finding a short preimage of a target $\boldsymbol{y}$ amounts to an inhomogeneous SIS problem. In this version, we sample elements from a sufficiently wide discrete Gaussian on $\mathbb{Z}^m$ and ask whether the image becomes close to uniform. Two natural objects appear in this situation. First, the kernel lattice $\Lambda_q^\perp(\boldsymbol{A}) = \{\boldsymbol{x} \in \mathbb{Z}^m : \boldsymbol{A}\boldsymbol{x} = \boldsymbol{0} \bmod q\}$. Second, the *smoothing parameter*, which gives the exact threshold for the width of a discrete Gaussian to yield a uniform distribution on the codomain modulo the kernel. Informally, once the Gaussian parameter is larger than $\eta_\varepsilon(\Lambda_q^\perp(\boldsymbol{A}))$, its image under $\boldsymbol{x} \mapsto \boldsymbol{A}\boldsymbol{x} \bmod q$ is close to uniform. This is one of the standard bridges between Gaussian sampling, smoothing parameters, and the statistical regularity of lattice maps.

In [JLWG25], Jin, Liu, Wang, and Gu provide a generalization of this setting where a pair of a lattice Λ and a sublattice $\Lambda' \subseteq \Lambda$ are studied. Their LHL version shows uniformity in Λ/Λ' with a discrete Gaussian input distribution.

Another extension of this finite setting to structured algebraic lattices has been elaborated by Boudgoust and Lapiha in [BL26]. They provide a systematic treatment of the LHL over cyclotomic rings with lattices reduced modulo an integer prime. Boudgoust and Lapiha study two approaches, an ideal counting approach that works for general input distributions, and a smoothing argument, which is tailored to the discrete Gaussian distribution.

From uniformity to a Gaussian distribution. Moving from finite domains to integer lattices requires adaptations. Since no uniform distribution is available over the infinite set $\mathbb{Z}^k$, the target distribution cannot be uniform. Agrawal, Gentry, Halevi, and Sahai [AGHS13] introduced a discrete Gaussian leftover hash lemma over infinite domains, in which uniform distributions are replaced by discrete Gaussian

distributions. Their result was subsequently improved by Aggarwal and Regev [AR16], and again by Kirshanova, Nguyen, Stehlé and Wallet [KNSW20].

As with the target distribution, this setting needs to properly choose the matrix $\boldsymbol{X} \in \mathbb{Z}^{r \times m}$. These are sampled columnwise, independently from a discrete Gaussian distribution. The input vector $\boldsymbol{v} \in \mathbb{Z}^m$ is sampled from another discrete Gaussian. The Gaussian parameters governing $\boldsymbol{X}$ and $\boldsymbol{v}$ are a priori unrelated. Under suitable conditions on these parameters, the distribution of $\boldsymbol{X}\boldsymbol{v}$ is statistically close to an appropriate discrete Gaussian whose parameters depend on the distribution of $\boldsymbol{X}$, rather than on its particular realization.

This Gaussian leftover hash version is proved in two successive steps. First, fix $\boldsymbol{X}$ and let $\boldsymbol{v} \xleftarrow{\$} D_{\mathbb{Z}^m, \sqrt{\boldsymbol{\Sigma}}}$ be a discrete Gaussian over $\mathbb{Z}^m$ with parameter $\sqrt{\boldsymbol{\Sigma}}$, that is, with covariance proportional to $\boldsymbol{\Sigma}$. For comparison, if $\boldsymbol{v}$ were a continuous Gaussian with covariance matrix $\boldsymbol{\Sigma}$, then $\boldsymbol{X}\boldsymbol{v}$ would again be Gaussian with parameter $\sqrt{\boldsymbol{X}\boldsymbol{\Sigma}\boldsymbol{X}^\mathsf{T}}$. The analogous statement does not hold exactly for discrete Gaussians. Nevertheless, under suitable smoothing conditions, the pushforward of $\boldsymbol{v}$ under $\boldsymbol{X}$ is statistically close to a discrete Gaussian with the corresponding transformed parameter $\sqrt{\boldsymbol{X}\boldsymbol{\Sigma}\boldsymbol{X}^\mathsf{T}}$. Second, when $\boldsymbol{X}$ is itself sampled from an appropriate discrete Gaussian distribution, this Gaussian distribution, which depends on $\boldsymbol{X}$, can be shown to be close to a Gaussian whose parameters depend only on the Gaussian parameter of the distribution from which $\boldsymbol{X}$ was sampled. The resulting distribution therefore loses its dependence on the particular realization of $\boldsymbol{X}$, yielding the Gaussian analogue of the leftover hash phenomenon over integer lattices. This proof outline repeats in the later variants of the Gaussian LHL results.

Gaussian leftover hashing over number fields. The integer statement was subsequently generalized from $\mathbb{Z}$-lattices to modules over rings of integers of number fields by Albrecht, Felderhoff, Lai, Lapiha, and Woo [AFL+26]. At a high level, their argument follows the same structure already visible over the integers. First, for a fixed matrix $\boldsymbol{X}$, the pushforward of the Gaussian $\boldsymbol{v}$ is compared with a Gaussian whose parameter is determined by $\boldsymbol{X}$. Second, when $\boldsymbol{X}$ is itself sampled from an appropriate Gaussian distribution, the resulting distribution is shown to become essentially independent of the particular geometry of $\boldsymbol{X}$. Their results apply to parameter sets with $m \geq r\,(\ln(rd))^{1+o(1)}$, where d is the degree of the number field.

One important shortcoming of the result is the Gaussian width of the distribution of the matrices $\boldsymbol{X}$. The idea of their proof is to bound the smoothing parameter of the kernel $\Lambda^\perp(\boldsymbol{X}) = \{\boldsymbol{v} \mid \boldsymbol{X}\boldsymbol{v} = \boldsymbol{0}\}$. To this end, they construct a matrix $\boldsymbol{U}$ of controlled norm satisfying $\boldsymbol{X}\boldsymbol{U} = f\boldsymbol{I}$, where f is the conductor of the number field, and use $\boldsymbol{U}$ to obtain sufficiently short vectors in $\Lambda^\perp(\boldsymbol{X})$. While the resulting bound on the norm of $\boldsymbol{U}$ is sufficient for their proof, it is not expected to be tight and leaves room for improvement. This loss propagates to the parameter bounds in their Gaussian LHL, resulting in a comparatively large lower bound on the Gaussian parameter $\boldsymbol{S}_{\boldsymbol{X}}$ from which the matrix $\boldsymbol{X}$ is sampled.

From a technical viewpoint, there are two important open problems. First, the surjectivity of the Gaussian sampled matrix $\boldsymbol{X}$ depends on the Generalized Riemann Hypothesis (GRH). Second, the step showing that the discrete Gaussian target distribution, whose parameter $\sqrt{\boldsymbol{X}\boldsymbol{\Sigma}\boldsymbol{X}^\mathsf{T}}$ a priori depends on the realization of $\boldsymbol{X}$, is close to a spherical discrete Gaussian that depends on $\boldsymbol{X}$ only through the width s of its distribution and the dimensions is expressed in terms of the Rényi divergence, rather than the commonly used statistical distance.

Technical aspects of the proofs of Gaussian LHL. A recurring technical ingredient in Gaussian leftover hash lemmas over infinite domains is the smoothing parameter of the orthogonal lattice $\Lambda^\perp(\boldsymbol{X}) = \{\boldsymbol{v} \in \mathbb{Z}^m \mid \boldsymbol{X}\boldsymbol{v} = \boldsymbol{0}\}$. Roughly, control over the output distribution $\boldsymbol{X}\boldsymbol{v}$ is given, once we have a good bound on the smoothing parameter of this lattice. Establishing sharper bounds for this smoothing parameter therefore results in better estimations of the resulting Gaussian distribution on the target side. Our work targets this aspect and approaches the computation of the smoothing parameter from a novel, *analytical* perspective.

1.1 Our contributions

Our main contribution is an analytic approach to the smoothing parameter of random kernel lattices that works directly with the Gaussian mass of the dual kernel, rather than first estimating its shortest vector. The starting point is an exact *theta-integral identity*. To the best of our knowledge, an exact theta integral of this kind has not been used before to bound smoothing parameters, or as a matter of fact, anywhere else in lattice-based cryptography.

For a lattice $\Lambda \subseteq \mathbb{R}^n$, a point $\boldsymbol{x} \in \mathbb{R}^n$ and $t > 0$, write $\Theta_t(\Lambda, \boldsymbol{x}) := \sum_{\boldsymbol{v}\in\Lambda} \exp(-\pi t^2 \, ||\boldsymbol{v} - \boldsymbol{x}||^2)$ for the theta function of Λ, and $\Theta_t(\Lambda)$ when $\boldsymbol{x} = \boldsymbol{0}$. For a matrix $\boldsymbol{A} \in \mathbb{Z}^{r\times m}$ of full row rank, with $L_{\boldsymbol{A}} = \ker_{\mathbb{Z}} \boldsymbol{A}$, the identity relates the theta function of the dual $L^*_{\boldsymbol{A}}$ to an integral of the ambient theta function $\Theta_t(\mathbb{Z}^m, \cdot)$ over the real row space V of $\boldsymbol{A}$ modulo the lattice Λ_{row} generated by its rows:

$$I_{\boldsymbol{A}} \, \Theta_t(L^*_{\boldsymbol{A}}) = t^r \int_{V/\Lambda_{\mathrm{row}}} \Theta_t(\mathbb{Z}^m, \boldsymbol{x}) \, d\boldsymbol{x}, \qquad I_{\boldsymbol{A}} := [\mathbb{Z}^r : \boldsymbol{A}\mathbb{Z}^m].$$

Thus, the geometry of the dual kernel is encoded in an integral involving the much simpler ambient lattice $\mathbb{Z}^m$.

A useful feature of this identity is that surjectivity and smoothing can be established by the same estimate. The index $I_{\boldsymbol{A}}$ is a positive integer, while the zero vector contributes 1 to $\Theta_t(L^*_{\boldsymbol{A}})$. Consequently, showing that the theta integral is smaller than $1+\varepsilon$ forces the index to be one and simultaneously bounds the Gaussian mass of the nonzero vectors of $L^*_{\boldsymbol{A}}$, that is, $\eta_\varepsilon(L_{\boldsymbol{A}}) \le t$. In particular, the argument does not require a separate surjectivity theorem or a conversion from a lower bound on the first minimum of the dual kernel. To our knowledge, this gives the first proof of surjectivity without requiring GRH in the number field case.

We carry out this program for random matrices $\boldsymbol{X} \in \mathbb{Z}^{r\times m}$ whose columns are independent discrete Gaussian vectors of any fixed width $s \ge 1$. The theta integral is controlled by separating a small neighborhood of the origin from its complement. Near the origin, the zero Gaussian gives the main contribution and the remaining terms are controlled directly. Away from the origin, the key probabilistic input is a one-column contraction estimate. The expected theta contribution of a random column is bounded uniformly by a constant $q(s) < 1$. Independence of the columns then amplifies this fixed contraction to exponential decay $q(s)^m$, making the contribution away from the origin negligible.

Combining the exact theta identity with this two-region estimate yields our main result in this setting (Theorem 2.7). For every fixed $s \ge 1$, if $m \gg_s r \ln r$ and $\boldsymbol{X}$ has independent $D_{\mathbb{Z},s}$ entries, then, except with probability $O(2^{-r})$,

$$\boldsymbol{X}\mathbb{Z}^m = \mathbb{Z}^r \qquad \text{and} \qquad \eta_{2^{-r}}(\ker_{\mathbb{Z}} \boldsymbol{X}) \ll_s \sqrt{r + \ln m}.$$

In the natural regime $\ln m = O(r)$, this gives

$$\eta_{2^{-r}}(\ker_{\mathbb{Z}} \boldsymbol{X}) = O_s(\sqrt{r}).$$

The argument therefore reduces the smoothing problem for these random kernel lattices to a local Gaussian estimate near the origin and a dimension-uniform anti-concentration estimate for a single random column away from it.

The statement above is the rational, constant-width case, and it is the setting of the technical overview in Section 2, where every idea of the proof already appears. The method is not limited to it. To pass to a number field K of degree d, we fix an integral basis of $\mathcal{O}_K$ whose first element is 1 and work in its coefficient coordinates. There $\boldsymbol{X} \in \mathcal{O}_K^{r\times m}$ becomes an integer matrix of size $dr \times dm$ whose integer kernel is the coefficient image of the module kernel, so the exact theta identity applies to it without change. The d integer columns that come from one ring column are dependent, but the one-column estimate needs only a single witness coordinate per ring column, the one that belongs to the basis element 1, and independence is used only across the m ring columns. Finally, the kernel lattice in the canonical embedding is the image of the coefficient kernel under a fixed linear map, so the smoothing bound transfers at the cost of a constant of the chosen basis. The constant-width argument additionally needs the integer coefficients of a ring Gaussian to be independent, which holds for power-of-two cyclotomic fields. The two informal theorems below state our results in their general form, for modules over a number field K and for both width regimes, and their proofs are given in Sections 5 and 6.

Theorem 1.1 (Kernel smoothing over number fields, informal Theorems 5.6 and 5.9 and Proposition 5.14). *Let K be a number field of degree d, let $\lambda \ge 1$ be an accuracy parameter, and let $\boldsymbol{X} \in \mathcal{O}_K^{r\times m}$ have independent columns from the discrete Gaussian $D_{\mathcal{O}_K^r,s}$, with $m \gg_K dr\ln(srm) + \lambda$ columns. Write $L_{\boldsymbol{X}}$ for the kernel lattice of $\boldsymbol{X}$.*

(i) Polynomial width. *If $s \gg_K \sqrt{dr}$, then, except with probability $2^{-\lambda}$,*

$$\boldsymbol{X}\mathcal{O}_K^m = \mathcal{O}_K^r \qquad \textit{and} \qquad \eta_{2^{-\lambda}}(L_{\boldsymbol{X}}) \ll_K \sqrt{dr + \lambda + \ln m}.$$

(ii) Constant width. *If K is a power-of-two cyclotomic field, the same conclusion holds for every $s \geq \sqrt{d}$, that is, as soon as the integer coefficients of $\boldsymbol{X}$ have width at least one, with $\eta_{2^{-\lambda}}(L_{\boldsymbol{X}}) \ll \sqrt{d\,(dr + \lambda + \ln m)}$ and absolute implied constants.*

(iii) Optimality. *For every $\boldsymbol{X}$ with $m > r$ and every $0 < \varepsilon < 1$ one has $\eta_\varepsilon(L_{\boldsymbol{X}}) \gg_K \sqrt{\ln(1/\varepsilon)}$, and $\eta_\varepsilon(L_{\boldsymbol{X}}) \geq \sqrt{d \ln(2/\varepsilon)/\pi}$ for power-of-two cyclotomic fields.*

In particular, for $r = \lceil \lambda/d \rceil$ and $\ln m = O(\lambda)$ the smoothing parameter is pinned: $\eta_{2^{-\lambda}}(L_{\boldsymbol{X}}) = \Theta_K(\sqrt{\lambda})$. Surjectivity is part of the conclusion and needs neither GRH nor a separate generation theorem.

From smoothing to the leftover hash lemma. A Gaussian LHL follows from Theorem 1.1 through the standard pushforward lemma (Lemma 3.4). That lemma has two hypotheses: the matrix $\boldsymbol{X}$ must be surjective, and the parameter of the input Gaussian must dominate the smoothing parameter of the kernel $L_{\boldsymbol{X}}$. Both hold on the single event of Theorem 1.1, so the error of the LHL is the failure probability of that event plus the error of the pushforward, and no hypothesis such as GRH enters through surjectivity. The target Gaussian a priori depends on $\boldsymbol{X}$. If the input parameter may be chosen after $\boldsymbol{X}$, it can be shaped so that the target becomes spherical.

Theorem 1.2 (Gaussian LHL over number fields, informal Corollaries 6.1 and 6.2). *In the setting of Theorem 1.1, in either width regime, let B be the smoothing bound of part (i) or (ii), and write $\widetilde{\Phi}(\boldsymbol{X})$ for the real matrix of $\boldsymbol{X}$ in the canonical embedding.*

(i) Ellipsoidal target. *For every input parameter $\sqrt{\boldsymbol{\Sigma}}$ with smallest singular value at least B and every center $\boldsymbol{c} \in \mathcal{O}_K^m$, which may depend on $\boldsymbol{X}$, the pairs*

$$\left(\boldsymbol{X},\ \boldsymbol{X} \cdot D_{\mathcal{O}_K^m, \sqrt{\boldsymbol{\Sigma}}, \boldsymbol{c}}\right) \qquad \text{and} \qquad \left(\boldsymbol{X},\ D_{\mathcal{O}_K^r, \sqrt{\widetilde{\Phi}(\boldsymbol{X})\boldsymbol{\Sigma}\widetilde{\Phi}(\boldsymbol{X})^T}, \boldsymbol{X}\boldsymbol{c}}\right)$$

are within statistical distance $2^{-\lambda}$.

(ii) Spherical target. *For every output width $w \gg_K B\, s\sqrt{m\,(dr + \lambda + \ln m)}$ there is an input parameter $\boldsymbol{S}_{\boldsymbol{X}}$, computed from $\boldsymbol{X}$, such that $\left(\boldsymbol{X},\ \boldsymbol{X} \cdot D_{\mathcal{O}_K^m, \boldsymbol{S}_{\boldsymbol{X}}, \boldsymbol{c}}\right)$ is within statistical distance $2^{-\lambda}$ of $\left(\boldsymbol{X},\ D_{\mathcal{O}_K^r, w, \boldsymbol{X}\boldsymbol{c}}\right)$. Over power-of-two cyclotomic fields $w \gg B\, s\sqrt{m}$ suffices.*

Applying part (ii) to k input vectors at once gives k simultaneous Gaussian hints (Corollary 6.3), which is what the SIS-to-k-SIS reduction of [AFL+26, Section 7] needs. Section 1.3 compares the resulting parameters with the original analysis.

Regimes and explicit bounds. Theorem 1.1 lives in the *wide regime* $m \gtrsim dr \ln(srm)$, shown on the right of Figure 1. There the smoothing parameter is pinned between the deterministic lower bound of part (iii) and the theta-integral bounds of parts (i) and (ii), which do not grow with m. The precise versions of these two parts carry explicit constants. Write $s_- := s/\alpha_K$ for the *coefficient width* of the columns. Theorem 5.6 gives the bound $B_{\mathrm{geo}} := 64\,\alpha_K \kappa_K \mu_K \sqrt{dr + \ln(2m/\delta)}$ when $s_- \geq 8\sqrt{dr}$, with $\delta = 2^{-\lambda-4}$, and over power-of-two cyclotomic fields Theorem 5.9 gives $B_{\mathrm{cw}} := 640\sqrt{d(dr + \ln(2m/\delta))}$ when $s_- \geq 1$, with $\delta = 2^{-\lambda-6}$. Here $\alpha_K, \kappa_K, \mu_K$ are constants of the chosen integral basis (Section 3.2), equal to $\sqrt{d}, 1, 1$ for power-of-two cyclotomics. In the *narrow regime*, where $m - r$ is fixed and s grows, the lower bound of Corollary 4.4 grows like $s^{r/(m-r)}$ and no matching upper bound is known. We return to it in Section 1.4.

Formal verification. The results of Sections 3 to 5 and of Section 6.1 have been formalized in Lean 4 with Mathlib, and the source code is publicly available.[4] It contains machine-checked proofs of the exact theta integral, of both finite smoothing theorems with their corollaries, of the lower bounds and the spectral estimates, and of the leftover hash lemmas, including the numerical Gaussian-shaping algorithm with proved accuracy and polynomial cost. The SIS-to-k-SIS reduction of Section 6.2 is formalized as an explicit finite program whose success probability is bounded in both geometric parameter regimes. Only the asymptotic family of Corollary 6.9, which relies on a prime search, and the comparisons of Section 7 are outside its scope. An axiom audit of every declaration reports only the standard foundations of Lean, with no unproved step and no additional axiom.

[4] https://github.com/etairi/geometric-gaussian-lhl

1.2 Related work

Over $\mathbb{Z}$, two families of arguments give upper bounds on $\eta_\varepsilon(L_{\boldsymbol{X}})$. Agrawal, Gentry, Halevi and Sahai [AGHS13] and Aggarwal and Regev [AR16] construct short integer preimages of the standard basis vectors, which give surjectivity and a short basis of the kernel. Kirshanova, Nguyen, Stehlé and Wallet [KNSW20] instead bound the first minimum of the dual kernel from below, through an auxiliary construction-A lattice and a union bound over its residues, and then convert that bound into a smoothing bound. Over a number field, Albrecht, Felderhoff, Lai, Lapiha and Woo [AFL+26] construct a short right inverse of $\boldsymbol{X}$ by a pigeonhole argument and obtain a bound that is polynomial in all parameters but carries large powers of d. Their surjectivity statement relies on GRH. The bounds of the first family and of [AFL+26] grow linearly with the width s, and the bound of the second family loses a factor of order $\sqrt{(m-r)/r}$ at moderate errors, because a first minimum is converted into a smoothing parameter.

Other related results of LHL type concern quotients of lattices. This was treated by Jin, Liu, Wang, and Gu [JLWG25] for a lattice and a sublattice, generalizing classical lattice-LHL. A systematic treatment of the mod q case over cyclotomic rings, where the analysis depends on how q splits, was presented by Boudgoust and Lapiha [BL26].

1.3 Parameters and comparison

We compare our bounds, and what they imply for the leftover hash lemma and for the SIS-to-k-SIS reduction, with those of [AFL+26]. To make the comparison concrete we fix one family of parameters over power-of-two cyclotomic fields, indexed by the accuracy parameter λ:

$$\begin{gathered} d = \Theta(\lambda^{\alpha}), \quad r = k = \Theta(\lambda^{1-\alpha}), \quad m = \Theta(\lambda L), \\ L = \ln(2\lambda), \quad \nu = \lambda^{-w}, \quad w \to \infty,\ w \le L, \end{gathered} \tag{1}$$

Here $K = \mathbb{Q}(\zeta_{2d})$ is the power-of-two cyclotomic field of conductor $2d$, whose degree is d, r is the number of rows of $\boldsymbol{X}$, which equals the number k of hints in the reduction, and m is the number of columns. The exponent $0 < \alpha < 1$ is fixed and balances the degree against the rank, and ν is the common value to which the smoothing, matrix-failure and LHL error budgets of [AFL+26] are set. Table 1 gives the comparison for kernel smoothing and the spherical Gaussian LHL, and Table 2 gives it for the complete SIS-to-k-SIS reduction.

The rows of the two tables are the following quantities. The width s_1 is the Gaussian width of the matrix $\boldsymbol{X}$, written s above, and in the reduction it is the width of the first block of hints. The leftover hash lemma is applied to an input Gaussian with parameter $\sqrt{\boldsymbol{\Sigma}}$, a square root of the covariance shape $\boldsymbol{\Sigma}$, so that a spherical Gaussian of width s corresponds to $\sqrt{\boldsymbol{\Sigma}} = s\boldsymbol{I}$. Its smallest singular value $s_{\min}(\sqrt{\boldsymbol{\Sigma}})$ has to dominate the smoothing parameter of the kernel. The quantity s_{out} is the width of the spherical output Gaussian obtained after shaping the input. The error Δ_{LHL} is the statistical distance between the joint law of the matrix and the output and its target law. In the reduction, k is the number of hints, which equals r in the family above, s_2 is the width of the hint columns, q is the modulus of the SIS instance, and β_0 and β_1 are the norm bounds of the source SIS instance and of the k-SIS instance, so that β_0/β_1 is the norm loss of the reduction.

Each column of Table 1 uses the smallest width s_1 admitted by its own theorem. The polynomial-width theorem admits $s_1 \ge 8\sqrt{d \cdot dr} = \Theta(d^{1/2}\lambda^{1/2})$ because $dr = \Theta(\lambda)$, and the constant-width theorem admits $s_1 = \sqrt{d}$, the smallest width of all the theorems compared. In the reduction, by contrast, s_1 is fixed by the imported modular window instead of by the kernel theorem, and the two geometric bounds then give the same parameters (Remark 6.10). The derivations for our columns are in Section 6.

The entries for [AFL+26] are obtained as follows. With $N = d(m-r)$, the power-of-two statements of [AFL+26, Corollaries 4.3, 4.13, 5.2 and 6.1, Lemma 6.6] give

$$\eta_\varepsilon(L_{\boldsymbol{X}}) = O\big(\gamma_N(\varepsilon) m^{3/2} d^{18} s_1\big), \qquad s_{\min}(\sqrt{\boldsymbol{\Sigma}}) = \omega\big(\gamma_N(\varepsilon) m^{3/2} d^{18} s_1\big),$$
$$s_{\mathrm{out}} = \omega\big(\gamma_N(\varepsilon) m^{2} d^{18} s_1^2\big),$$

where $\gamma_N(\varepsilon) = \sqrt{\ln(2N(1+1/\varepsilon))/\pi}$ is the factor of Equation (8), the first statement is unconditional, and the other two use GRH through the full-image step. For a conductor $2^a = 2d$ the factor 2^{18a} in those statements becomes $\Theta(d^{18})$. In Table 1 the smoothing-only width is the smallest one allowed by

quantity	**[AFL+26]**	**this paper (B_{geo})**	**this paper (B_{cw})**
assumption	unconditional (smoothing), GRH (LHL)	unconditional	unconditional
columns m	$r\log(drs)\log\log(drs)$	$7.5\,(dr\ln W + 2\ln(1/\delta))$, $\ln W = \Theta(L)$	$49.5\,(dr\ln W + 2\ln(1/\delta))$, $\ln W = \Theta(L)$
s_1 (smoothing only)	$\Theta(d^{1/2}L^{1/2})$	$\Theta(d^{1/2}\lambda^{1/2})$	$\Theta(d^{1/2})$
$\eta_\nu(L_{\boldsymbol{X}})$ at that s_1	$O(d^{37/2}\lambda^{3/2}L^{5/2}\sqrt{w})$	$\Theta(d^{1/2}\lambda^{1/2})$	$\Theta(d^{1/2}\lambda^{1/2})$
s_1 (complete LHL)	$\Theta(\lambda d^{1/2}L^{1/2}\sqrt{w})$	$\Theta(d^{1/2}\lambda^{1/2})$	$\Theta(d^{1/2})$
$s_{\min}(\sqrt{\boldsymbol{\Sigma}})$	$\omega(d^{37/2}\lambda^{5/2}L^{5/2}w)$	$\Theta(d^{1/2}\lambda^{1/2})$	$\Theta(d^{1/2}\lambda^{1/2})$
s_{out}	$\omega(d^{19}\lambda^4L^{7/2}w^{3/2})$	$\Theta(d\lambda^{3/2}L^{1/2})$	$\Theta(d\lambda L^{1/2})$
Δ_{LHL}	$O(m\nu + m2^{-\Theta(\lambda)})$	$O(2^{-\lambda})$	$O(2^{-\lambda})$

Table 1. Kernel-smoothing and spherical Gaussian-LHL parameters in the family Equation (1), specialized to power-of-two cyclotomics. All entries are sufficient theorem scales. The column-count row gives the shape of each theorem's hypothesis on m, and in this family all are $O(\lambda L)$ or smaller. Here W is the polynomial in s, m/r and $dr + \lambda + \ln m$ defined in Equation (27) for the first geometric column and in Equation (40) for the second. The kernel-smoothing bound of [AFL+26] is unconditional, but its LHL rows use GRH through surjectivity. The two geometric columns use Theorem 5.6 at accuracy λ, $\delta = 2^{-\lambda-4}$, and Theorem 5.9 at coefficient width one, $\delta = 2^{-\lambda-6}$. Their bounds carry no factor depending on the error budget ν because they are direct theta estimates valid at every error $\varepsilon \geq 2^{-\lambda}$. The smaller admissible width $s_1 = \sqrt{d}$ of the last column lowers s_{out} by $\lambda^{1/2}$. Derivations are in Section 6.2.

quantity	[AFL+26] (GRH)	this paper with B_{geo} or B_{cw}
s_1	$\Theta(\lambda d^{1/2}L^{1/2}\sqrt{w})$	$\Theta(d\lambda^{1/2}L^{1/2})$
smoothing bound at s_1	$O(d^{37/2}\lambda^{5/2}L^{5/2}w)$	$\Theta(d^{1/2}\lambda^{1/2})$
s_2	$\omega(d^{39/2}\lambda^{9/2}L^{7/2}w^{3/2})$	$\Theta(d^{3/2}\lambda^{3/2}L)$
β_0/β_1	$\omega(d^{41/2}\lambda^5L^4w^{3/2})$	$\Theta(d^{3/2}\lambda^2L^{3/2})$
$\ln q$	$\Theta(dL)$	$\Theta(dL)$

Table 2. Parameters of the complete SIS-to-k-SIS reduction in the family Equation (1). Both columns include the imported algebraic reduction of [AFL+26, Section 7], and the last uses the hint generator of Lemma 6.4 with $B = B_{\mathrm{geo}}$ or $B = B_{\mathrm{cw}}$ (Corollary 6.9). The reduction loses an additive $2^{-\Omega(\lambda)} + \mathsf{negl}(\lambda)$ in the adversary's advantage, which accounts for all statistical errors and failure events. The two geometric bounds give the same entries because s_1 is fixed by the imported modular window and not by the kernel theorem (Remark 6.10). With $d = \lambda^\alpha$, the hint width is $\lambda^{3/2+3\alpha/2}$ and the norm loss $\lambda^{2+3\alpha/2}$ up to logarithms, against $\lambda^{9/2+39\alpha/2}$ and $\lambda^{5+41\alpha/2}$ for the original analysis.

the smoothing theorem of [AFL+26], whereas the complete LHL needs the larger width of its GRH-dependent surjectivity step. Its s_{out} entry combines the preimage requirement with exact shaping and the spectral estimate of its Section 6 rather than the coarser width of its Section 7. In Table 2 we use the width condition stated in [AFL+26, Section 7], and the norm-loss entry reports the printed theorem, whose multiplier $O(d\sqrt{m})$ is replaced by $O(\sqrt{m})$ in our column through Lemma 6.4. All entries are sufficient theorem scales, not lower bounds on the true optimum.

1.4 Open problems

Figure 1 summarizes what is known. The smoothing parameter of the kernel is determined up to a constant only in the wide regime, where the upper and lower bounds match. We record the larger questions that remain, in decreasing order of what we understand about them.

The narrow regime. Fix the number $m - r$ of extra columns and let the width s grow. The kernel then has few dimensions in which to absorb a large covolume, and for surjective matrices whose singular values are all of order $s\sqrt{m}$, Corollary 4.4 shows that the smoothing parameter must grow at least like $s^{r/(m-r)}$. We do not know whether this lower bound is tight. Our method, like the union-bound arguments of [AGHS13, KNSW20], succeeds by playing the m independent columns against a number of

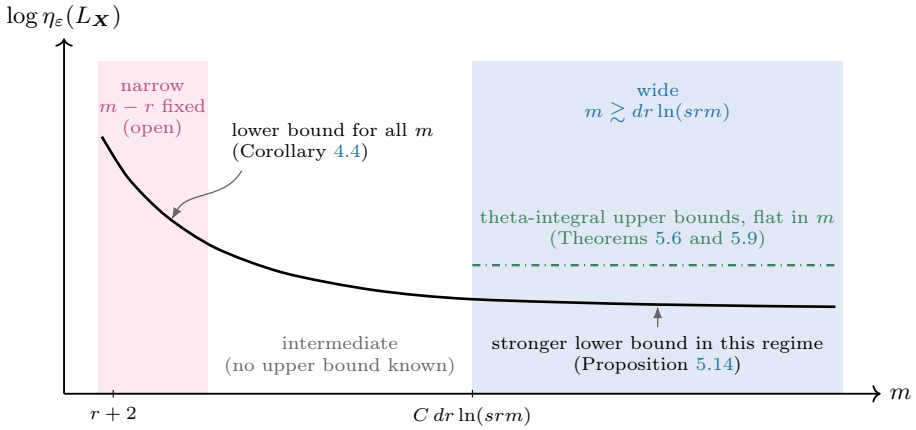


Fig. 1. Schematic picture of the smoothing parameter of the kernel as a function of the number of columns m, for a fixed field, rank and width. The solid curve is a lower bound and the dash-dotted line an upper bound. They match up to a constant only in the wide regime (Corollary 5.15).

bad directions that grows polynomially with s, which requires $m \gtrsim dr \ln s$ columns (Remark 4.5). With a fixed number of columns and a growing width this trade cannot be won, so a different idea is needed. Whether $\eta_\varepsilon(L_{\boldsymbol{X}}) = \Theta(s^{r/(m-r)})$ holds with high probability in this regime is open, and this paper does not address surjectivity there either.

The intermediate range. Between the two regimes, for $r < m \lesssim dr \ln s$, the only statements we have are the lower bounds of Propositions 4.3 and 5.14. Every known upper bound, here and in [AGHS13, AR16, KNSW20, AFL+26], needs a column count of order at least $r \ln s$ (Section 7.2). How the smoothing parameter passes from the constant scale of the wide regime to the polynomial scale of the narrow one, and where that transition happens, is open.

Constant width beyond power-of-two cyclotomics. The constant-width theorem uses two features of the power basis of $\mathbb{Q}(\zeta_{2^a})$: the coefficients of a ring Gaussian are independent, and multiplication by a basis element permutes them up to sign. In a general field the coefficient Gaussian is ellipsoidal and the d columns of a ring block are dependent images of one coefficient vector, so the only tool available is the integer displacement of Lemma 5.4, whose norm $\sqrt{dr}$ forces a width of order $\sqrt{dr}$. Whether every number field, or every integral basis, admits a smoothing bound of the optimal order at bounded coefficient width is open.

1.5 Organization

Section 2 presents the whole argument in its simplest setting, the constant-width regime over $\mathbb{Q}$, and maps each step to the general statement that carries it out. Section 3 contains the preliminaries that are needed to understand the sections that follow it. Section 4 describes the dual kernel lattice, proves the lower bound, and separates the two regimes. Section 5 gives the geometric route: the exact theta integral, the finite and asymptotic smoothing theorems at polynomial width, the constant-width theorem for power-of-two fields, and the matching lower bound. Section 6 derives the resulting leftover hash lemmas, the hint generator, and the reduction parameters. Section 7 compares the bounds with earlier work and records the sources of the compared bounds. All proofs are given in the body of the paper.

2 Technical Overview

This section presents the argument of the paper in its simplest setting. The field is $\mathbb{Q}$, the matrix $\boldsymbol{X} \in \mathbb{Z}^{r \times m}$ has independent entries from the discrete Gaussian $D_{\mathbb{Z},s}$, and the width $s \geq 1$ is a fixed constant, so that most entries are 0 or ± 1. This is the constant-width regime of Theorem 5.9, specialized to the rational case (Remark 5.12). We chose it because it is the regime in which the theta-integral method shows most clearly what it gains over earlier arguments. No width of order $\sqrt{r}$ is needed, surjectivity of

$\boldsymbol{X}$ comes out of the same estimate as the smoothing bound, and the resulting order $\sqrt{r}$ is optimal. The proofs in this section are complete, but they keep the constants implicit.

The other regimes and the general setting are treated in Sections 3 to 6. Section 5 proves the polynomial-width theorem over an arbitrary number field (Theorem 5.6) and the constant-width theorem over power-of-two cyclotomics (Theorem 5.9), both with explicit constants. Section 4 gives the lower bounds that show that the order $\sqrt{r}$ cannot be improved, and it delimits the wide and narrow regimes of Figure 1. Section 6 derives the leftover hash lemmas and the reduction. Along the way we point to the statement of the paper that carries out each step in general, and Figure 2 draws the whole route.

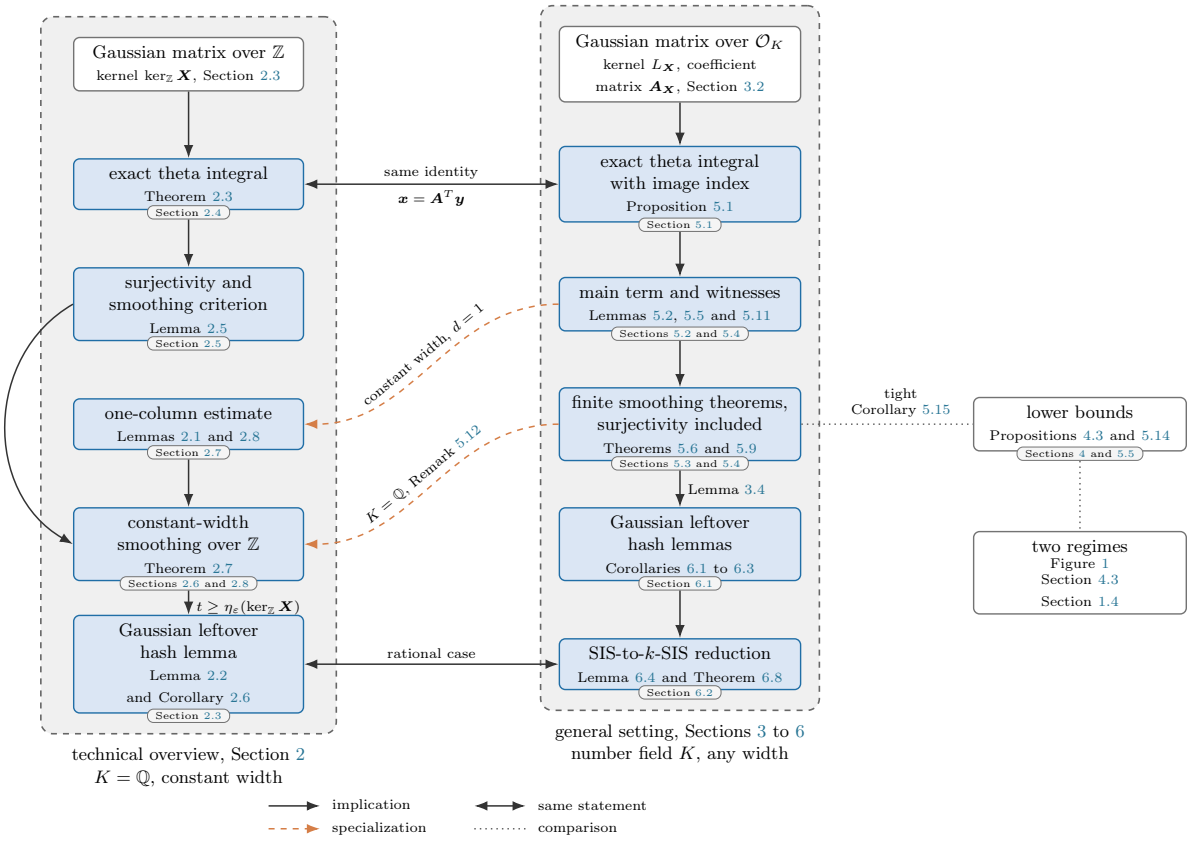


Fig. 2. Roadmap of the paper. The left column is the argument of this section in its simplest form, and the right column is the same argument over a number field at either width, followed by its consequences. Each box names the result and the subsection that proves it.

Throughout this section we use theta functions rather than the Gaussian mass ρ of Section 3. The two notations are related by $\Theta_\alpha(\Lambda, \boldsymbol{x}) = \rho_{1/\alpha}(\Lambda - \boldsymbol{x})$, and the one-dimensional function $\Theta_t(\mathbb{Z}, u)$ below is the periodic Gaussian sum $\varphi_t(u)$ of Section 5.

2.1 Lattices, theta functions and Poisson summation

For us, a lattice $\Lambda \subseteq \mathbb{R}^n$ is a discrete additive subgroup of $\mathbb{R}^n$. We do not require it to have maximal $\mathbb{Z}$-rank, and its rank k may be any integer with $0 \leq k \leq n$. The relevant Euclidean space for Λ is its real span $V_\Lambda := \operatorname{span}_{\mathbb{R}}(\Lambda)$. If $\boldsymbol{b}_1, \ldots, \boldsymbol{b}_k$ is a $\mathbb{Z}$-basis of Λ, a fundamental domain is, for example, $\{\sum_{j=1}^k t_j \boldsymbol{b}_j : 0 \leq t_j < 1\}$. We define the covolume of Λ intrinsically in its span by

$$\operatorname{covol}(\Lambda) := \operatorname{vol}_{V_\Lambda}(V_\Lambda / \Lambda) = \sqrt{\det\big(\langle \boldsymbol{b}_i, \boldsymbol{b}_j \rangle\big)_{i,j=1}^{k}}.$$

The intrinsic dual lattice is

$$\Lambda^* := \{\boldsymbol{y} \in V_\Lambda : \langle \boldsymbol{y}, \boldsymbol{z} \rangle \in \mathbb{Z} \text{ for every } \boldsymbol{z} \in \Lambda\}.$$

It is a full-rank lattice in $V_\Lambda \subseteq \mathbb{R}^n$.

For a lattice $\Lambda \subseteq \mathbb{R}^n$, a point $\boldsymbol{x} \in \mathbb{R}^n$ and $\alpha > 0$ we put

$$\Theta_\alpha(\Lambda, \boldsymbol{x}) := \sum_{\boldsymbol{v} \in \Lambda} \exp\bigl(-\pi\alpha^2 \, ||\boldsymbol{v} - \boldsymbol{x}||^2\bigr).$$

The subscript α is an inverse width. A discrete Gaussian of width s has $\alpha = 1/s$. When $\boldsymbol{x} = \boldsymbol{0}$ we simply write $\Theta_\alpha(\Lambda)$. Over the integers the theta function becomes

$$\Theta_t(\mathbb{Z}, u) = \sum_{n \in \mathbb{Z}} e^{-\pi t^2 (n-u)^2}.$$

This function of u has period one, and the theta function of $\mathbb{Z}^n$ is the product of n such factors, one for each coordinate. Figure 3 shows the case $n = 2$. For small t the function is nearly flat, and increasing t gives narrower peaks around the integer points.

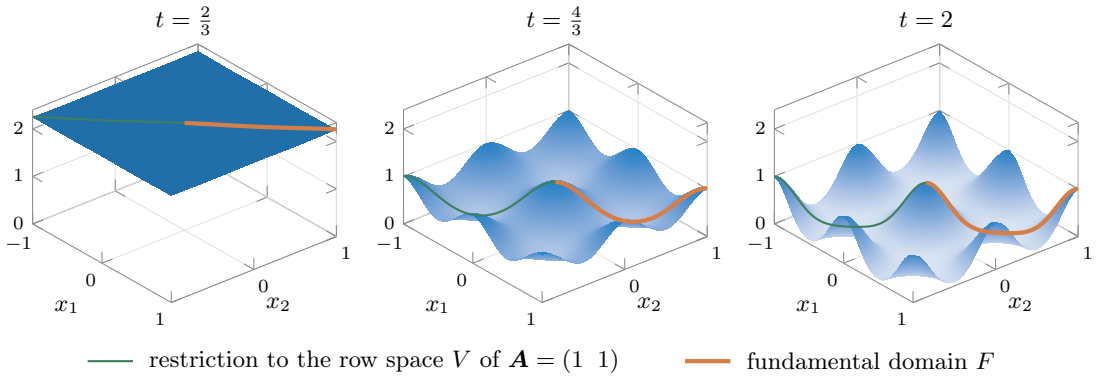


Fig. 3. The theta function $\Theta_t(\mathbb{Z}^2, \boldsymbol{x}) = \Theta_t(\mathbb{Z}, x_1)\,\Theta_t(\mathbb{Z}, x_2)$ on the same scale for three values of t. For small t it is nearly constant at height $1/t^2$, and as t grows it concentrates on the integer points. The curves show its restriction to the row space of $\boldsymbol{A} = (1 \;\; 1)$, with the notation of Figure 5. The integral in Theorem 2.3 runs over the orange part, and the right-hand side of Equation (3) is about 2.12, 1.12 and 1.004 in the three panels.

We also recall the Poisson summation formula [MR07, Lemma 2.8] in the span of a lattice. Let V_Λ be the real span of Λ and put $k = \operatorname{rank} \Lambda$. With the intrinsic Fourier transform

$$\widehat{f}(\boldsymbol{y}) = \int_{V_\Lambda} f(\boldsymbol{u}) e^{-2\pi i \langle \boldsymbol{u}, \boldsymbol{y} \rangle} \, d\boldsymbol{u},$$

for every Schwartz function $f \colon V_\Lambda \to \mathbb{C}$ and every $\boldsymbol{x} \in V_\Lambda$ one has

$$\sum_{\boldsymbol{v} \in \Lambda} f(\boldsymbol{v} - \boldsymbol{x}) = \frac{1}{\operatorname{covol}(\Lambda)} \sum_{\boldsymbol{w} \in \Lambda^*} \widehat{f}(\boldsymbol{w}) e^{-2\pi i \langle \boldsymbol{w}, \boldsymbol{x} \rangle}.$$

Applied to $f(\boldsymbol{u}) = e^{-\pi\alpha^2 ||\boldsymbol{u}||^2}$ this gives, for $\boldsymbol{x} \in V_\Lambda$,

$$\Theta_\alpha(\Lambda, \boldsymbol{x}) = \frac{1}{\alpha^k \operatorname{covol}(\Lambda)} \sum_{\boldsymbol{w} \in \Lambda^*} e^{-\pi ||\boldsymbol{w}||^2/\alpha^2} e^{-2\pi i \langle \boldsymbol{w}, \boldsymbol{x} \rangle}. \tag{2}$$

For a general $\boldsymbol{x} \in \mathbb{R}^n$, write $\boldsymbol{x} = \boldsymbol{x}_V + \boldsymbol{x}_\perp$ with $\boldsymbol{x}_V \in V_\Lambda$ and $\boldsymbol{x}_\perp \in V_\Lambda^\perp$. Then the same identity holds with an additional factor $e^{-\pi\alpha^2 ||\boldsymbol{x}_\perp||^2}$ and with $\boldsymbol{x}$ in the character replaced by $\boldsymbol{x}_V$.

2.2 Smoothing parameter, discrete Gaussians and statistical distance

For a lattice $\Lambda \subseteq \mathbb{R}^n$ and $0 < \varepsilon < 1$ the smoothing parameter is

$$\eta_\varepsilon(\Lambda) = \inf\{t > 0 : \Theta_t(\Lambda^*) \le 1 + \varepsilon\}.$$

This is the definition of Section 3, since $\Theta_t(\Lambda^*) - 1$ is the Gaussian mass $\rho_{1/t}$ of the nonzero dual vectors. After Poisson summation in V_Λ the condition can also be written in terms of the theta function of Λ itself. Let $k = \operatorname{rank}\Lambda$. Since $\operatorname{covol}(\Lambda^*) = \operatorname{covol}(\Lambda)^{-1}$, Equation (2) gives

$$\Theta_t(\Lambda^*) = t^{-k} \operatorname{covol}(\Lambda)\, \Theta_{1/t}(\Lambda),$$

and consequently

$$\eta_\varepsilon(\Lambda) = \inf\bigl\{t > 0 : t^{-k} \operatorname{covol}(\Lambda)\, \Theta_{1/t}(\Lambda) \le 1 + \varepsilon\bigr\}.$$

Let $\Lambda \subseteq \mathbb{R}^n$ be a lattice and let $\boldsymbol{S}\colon V_\Lambda \to V_\Lambda$ be symmetric positive definite. For a center $\boldsymbol{c} \in V_\Lambda$, the discrete Gaussian $D_{\Lambda,\boldsymbol{S},\boldsymbol{c}}$ assigns to $\boldsymbol{u} \in \Lambda$ probability proportional to $e^{-\pi\left|\left|\boldsymbol{S}^{-1}(\boldsymbol{u}-\boldsymbol{c})\right|\right|^2}$. More explicitly,

$$\mathbb{P}_{\boldsymbol{z}\leftarrow D_{\Lambda,\boldsymbol{S},\boldsymbol{c}}}[\boldsymbol{z} = \boldsymbol{u}] = \frac{e^{-\pi\left|\left|\boldsymbol{S}^{-1}(\boldsymbol{u}-\boldsymbol{c})\right|\right|^2}}{\Theta_1(\boldsymbol{S}^{-1}\Lambda, \boldsymbol{S}^{-1}\boldsymbol{c})}, \qquad \boldsymbol{u} \in \Lambda.$$

We write $D_{\Lambda,\boldsymbol{S}}$ when $\boldsymbol{c} = \boldsymbol{0}$, and $D_{\Lambda,s}$ when $\boldsymbol{S} = s\boldsymbol{I}$. The matrix $\boldsymbol{S}$ determines the covariance shape $\boldsymbol{S}\boldsymbol{S}^T$, which is the form used in the Gaussian leftover hash lemma below.

Let P and Q be two probability distributions on a countable set Ω. Their statistical distance is

$$\mathrm{SD}(P, Q) := \frac{1}{2}\sum_{u\in\Omega} |P(u) - Q(u)| = \sup_{T\subseteq\Omega} |P(T) - Q(T)|\,.$$

It measures how distinguishable the two distributions are. The distance is 0 exactly when they are the same, and it is at most 1 in general.

The following lemma is the only property of the width-s Gaussian that the constant-width argument needs. It says that a unit translate of $D_{\mathbb{Z},s}$ still overlaps $D_{\mathbb{Z},s}$ by a fixed amount, however small s is. Figure 4 shows this overlap for two widths.

Lemma 2.1 (Translation bound for a discrete Gaussian). *Let $s > 0$, let $Z \leftarrow D_{\mathbb{Z},s}$, and let $k \in \mathbb{Z}$. Then*

$$\mathrm{SD}(Z, Z + k) \le \min\Bigl\{1, \frac{|k|}{s}\Bigr\}.$$

Moreover $\mathrm{SD}(Z, Z + 1) = \mathbb{P}[Z = 0] = \Theta_{1/s}(\mathbb{Z})^{-1}$, which is strictly less than one.

Proof. Let $p(n) = \mathbb{P}[Z = n] = e^{-\pi n^2/s^2}/\Theta_{1/s}(\mathbb{Z})$. The sequence $p(n)$ is symmetric and decreases as $|n|$ grows. Thus

$$\mathrm{SD}(Z, Z + 1) = \frac{1}{2}\sum_{n\in\mathbb{Z}} |p(n) - p(n-1)| = p(0) = \frac{1}{\Theta_{1/s}(\mathbb{Z})},$$

and this is less than one because the sum $\Theta_{1/s}(\mathbb{Z})$ contains the term 1 for $n = 0$ and further positive terms. By the triangle inequality, for $k \ge 0$,

$$\mathrm{SD}(Z, Z + k) \le \sum_{j=0}^{k-1} \mathrm{SD}(Z + j, Z + j + 1) = \frac{k}{\Theta_{1/s}(\mathbb{Z})}.$$

The same bound with $|k|$ holds for negative k. Poisson summation gives $\Theta_{1/s}(\mathbb{Z}) = s\,\Theta_s(\mathbb{Z}) \ge s$, which proves the displayed bound. □

In the general development the translation bound appears as Equation (23), and the exact value $\mathrm{SD}(Z, Z + 1) = \Theta_{1/s}(\mathbb{Z})^{-1}$ is what the proof of Lemma 5.11 uses.

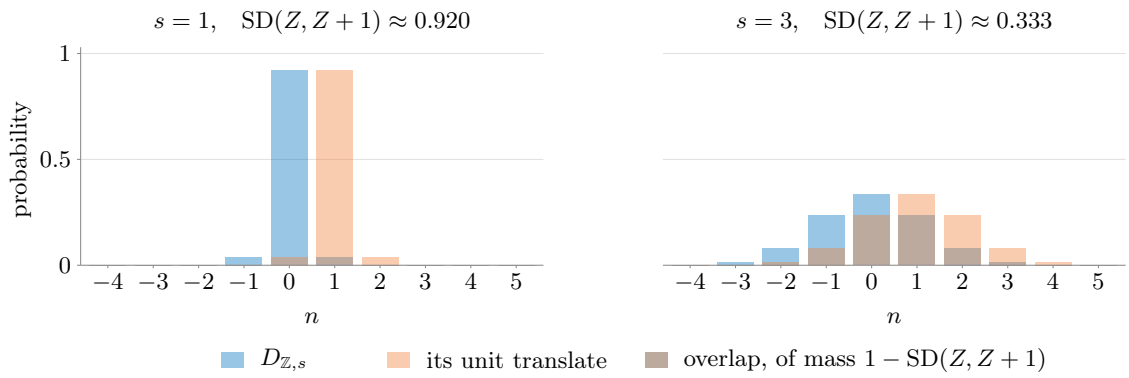


Fig. 4. The law of $Z \leftarrow D_{\mathbb{Z},s}$ and its unit translate, for two widths. The region where the two histograms overlap has mass $1 - \mathrm{SD}(Z, Z+1) = 1 - \mathbb{P}[Z = 0]$. By Lemma 2.1 it is positive for every width, and it is still about 0.08 at $s = 1$.

2.3 The Gaussian leftover hash lemma

Lemma 2.2 (Gaussian leftover hash lemma over $\mathbb{Z}$). *Let $\boldsymbol{X} \in \mathbb{Z}^{r \times m}$ have full row rank and put $L_{\boldsymbol{X}} := \ker_{\mathbb{Z}} \boldsymbol{X}$. If $\boldsymbol{X}\mathbb{Z}^m = \mathbb{Z}^r$ and $\boldsymbol{z} \leftarrow D_{\mathbb{Z}^m,t}$, then $\boldsymbol{X}\boldsymbol{z}$ is close to the discrete Gaussian predicted by the real map $\boldsymbol{X}$. That is, for every $0 < \varepsilon < 1$ and $t \geq \eta_\varepsilon(L_{\boldsymbol{X}})$,*

$$\mathrm{SD}\Big(\boldsymbol{X}\boldsymbol{z},\, D_{\mathbb{Z}^r,\, t(\boldsymbol{X}\boldsymbol{X}^T)^{1/2}}\Big) \leq \frac{\varepsilon}{1-\varepsilon}.$$

This is the rational case of Lemma 3.4 with the input parameter $t\boldsymbol{I}$. Thus the analytic problem is to bound the smoothing parameter of the integer kernel. In this section we fix r, let m grow, and sample the entries independently as $X_{ij} \leftarrow D_{\mathbb{Z},s}$, so that for every $u \in \mathbb{Z}$

$$\mathbb{P}[X_{ij} = u] = \frac{e^{-\pi u^2/s^2}}{\Theta_{1/s}(\mathbb{Z})}.$$

The question is the following. How large must m be, and how small can we prove $\eta_\varepsilon(\ker_{\mathbb{Z}} \boldsymbol{X})$ to be, with high probability over $\boldsymbol{X}$, and can $\boldsymbol{X}\mathbb{Z}^m = \mathbb{Z}^r$ be established on the same event?

2.4 An exact theta integral

Consider integers $1 < r < m$ and a matrix $\boldsymbol{A} \in \mathbb{Z}^{r \times m}$ of full rank, that is rank $\boldsymbol{A} = r$. Let $L_{\boldsymbol{A}} = \ker_{\mathbb{Z}} \boldsymbol{A}$, a lattice of rank $m - r$. Denote by

$$V = \mathbb{R}\boldsymbol{a}_1 + \mathbb{R}\boldsymbol{a}_2 + \cdots + \mathbb{R}\boldsymbol{a}_r \subseteq \mathbb{R}^m$$

the r-dimensional subspace of $\mathbb{R}^m$ spanned by the rows $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_r$ of $\boldsymbol{A}$, and put

$$\Lambda_{\mathrm{row}} = \mathbb{Z}\boldsymbol{a}_1 + \mathbb{Z}\boldsymbol{a}_2 + \cdots + \mathbb{Z}\boldsymbol{a}_r \subseteq \mathbb{Z}^m \cap V.$$

We consider the average

$$\int_{V/\Lambda_{\mathrm{row}}} \Theta_t(\mathbb{Z}^m, \boldsymbol{x})\, d\boldsymbol{x}$$

of $\Theta_t(\mathbb{Z}^m, \boldsymbol{x})$ over a fundamental domain of V/Λ_{row}. The integrand is indeed Λ_{row}-invariant, since $\Lambda_{\mathrm{row}} \subseteq \mathbb{Z}^m$, and the integration is with respect to the Lebesgue measure of $V \subseteq \mathbb{R}^m$. The remarkable fact is that this integral is again the theta function of a lattice.

Theorem 2.3 (Exact theta integral). *With the notation above, for every $t > 0$,*

$$I_{\boldsymbol{A}} \cdot \Theta_t(L_{\boldsymbol{A}}^*) = t^r \int_{V/\Lambda_{\mathrm{row}}} \Theta_t(\mathbb{Z}^m, \boldsymbol{x})\, d\boldsymbol{x}. \tag{3}$$

Here $I_{\boldsymbol{A}} = [\mathbb{Z}^r : \boldsymbol{A}\mathbb{Z}^m]$ is the index of the image, and $L_{\boldsymbol{A}}^*$ is the intrinsic dual of $L_{\boldsymbol{A}}$ defined in Section 2.1.

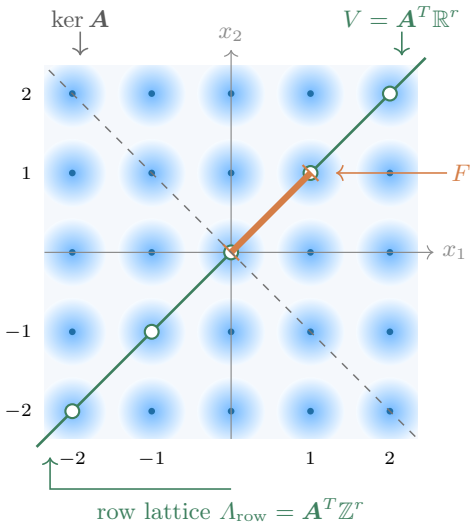


Fig. 5. For $\boldsymbol{A} = (1 \;\; 1)$, the green diagonal is the row space V, its marked points form Λ_{row}, the orange segment is a fundamental domain F, and the dashed line is $\ker \boldsymbol{A}$. The exact theta identity integrates the theta function of $\mathbb{Z}^2$ over F.

Figure 5 illustrates the objects of the identity for the matrix $\boldsymbol{A} = (1 \;\; 1)$. The theorem is Proposition 5.1, written on the row space rather than on the torus $\mathbb{R}^r/\mathbb{Z}^r$. The two forms are related by the parametrization $\boldsymbol{x} = \boldsymbol{A}^T\boldsymbol{y}$, under which $d\boldsymbol{x} = \det(\boldsymbol{A}\boldsymbol{A}^T)^{1/2}\, d\boldsymbol{y}$ and $\Theta_t(\mathbb{Z}^m, \boldsymbol{A}^T\boldsymbol{y}) = \prod_{j=1}^m \Theta_t(\mathbb{Z}, \langle \boldsymbol{a}^{(j)}, \boldsymbol{y}\rangle)$ with $\boldsymbol{a}^{(j)}$ the j-th column of $\boldsymbol{A}$. Section 5.1 proves the identity by unfolding the integral over the cosets of the row lattice. A second proof applies Poisson summation twice, once before and once after integrating, and we sketch it because it explains where the index comes from.

Proof (Proof sketch). Poisson summation on $\mathbb{Z}^m$ gives

$$\Theta_t(\mathbb{Z}^m, \boldsymbol{x}) = t^{-m} \sum_{\boldsymbol{v}\in\mathbb{Z}^m} e^{-\pi||\boldsymbol{v}||^2/t^2} e^{-2\pi i\langle \boldsymbol{v},\boldsymbol{x}\rangle}.$$

For $\boldsymbol{v} \in \mathbb{Z}^m$ the character $\boldsymbol{x} \mapsto e^{-2\pi i\langle \boldsymbol{v},\boldsymbol{x}\rangle}$ is Λ_{row}-periodic on V, and its integral over V/Λ_{row} vanishes unless the character is trivial on V. Many terms are therefore killed by the integration, and the surviving vectors are exactly those in $\mathbb{Z}^m \cap V^\perp = \{\boldsymbol{v} \in \mathbb{Z}^m : \boldsymbol{A}\boldsymbol{v} = \boldsymbol{0}\} = L_{\boldsymbol{A}}$. Since $\text{vol}(V/\Lambda_{\text{row}}) = \text{covol}(\Lambda_{\text{row}})$, we obtain

$$\int_{V/\Lambda_{\text{row}}} \Theta_t(\mathbb{Z}^m, \boldsymbol{x})\; d\boldsymbol{x} = \frac{\text{covol}(\Lambda_{\text{row}})}{t^m}\, \Theta_{1/t}(L_{\boldsymbol{A}}).$$

Now apply Poisson summation once more, to the lattice $L_{\boldsymbol{A}}$ in the Euclidean space $V^\perp$ of dimension $m - r$, in which it has full rank. This gives

$$\Theta_{1/t}(L_{\boldsymbol{A}}) = \frac{t^{m-r}}{\text{covol}(L_{\boldsymbol{A}})}\, \Theta_t(L_{\boldsymbol{A}}^*).$$

Finally $\text{covol}(\Lambda_{\text{row}}) = I_{\boldsymbol{A}}\, \text{covol}(L_{\boldsymbol{A}})$. To see this, let $R_{\text{sat}} = V \cap \mathbb{Z}^m$ be the saturation of the row lattice. Since R_{sat} is primitive, a basis of it extends to a basis of $\mathbb{Z}^m$. Therefore, $1 = \text{covol}(\mathbb{Z}^m) = \text{covol}(R_{\text{sat}})\, \text{covol}(\boldsymbol{P}\mathbb{Z}^m)$, where $\boldsymbol{P}$ is the orthogonal projection onto $V^\perp = \ker \boldsymbol{A}$. By Lemma 4.2 we have $\boldsymbol{P}\mathbb{Z}^m = L_{\boldsymbol{A}}^*$, whose covolume is $1/\,\text{covol}(L_{\boldsymbol{A}})$. Hence, $\text{covol}(R_{\text{sat}}) = \text{covol}(L_{\boldsymbol{A}})$, and the Smith normal form of $\boldsymbol{A}$ gives $[R_{\text{sat}} : \Lambda_{\text{row}}] = [\mathbb{Z}^r : \boldsymbol{A}\mathbb{Z}^m] = I_{\boldsymbol{A}}$. □

Remark 2.4. One can also describe $L_{\boldsymbol{A}}^*$ as the orthogonal projection of $\mathbb{Z}^m$ onto the real kernel $\ker \boldsymbol{A} \subseteq \mathbb{R}^m$. This is Lemma 4.2.

The main lesson is that we started with the theta function of a lattice of rank m and integrated away r dimensions to obtain the theta function of a lattice of rank $m - r$. This is in some sense dual to the mechanism of the leftover hash lemma. There we average over the $m - r$ dimensions of the kernel and obtain r output coordinates, whereas here we average over the r output coordinates and obtain a theta series in $m - r$ dimensions.

2.5 Exploiting the theta integral

What do we do with Theorem 2.3? Suppose that we managed to show that

$$t^r \int_{V/\Lambda_{\mathrm{row}}} \Theta_t(\mathbb{Z}^m, \boldsymbol{x})\, d\boldsymbol{x} \le 1+\varepsilon \qquad \text{for some } \varepsilon < 1.$$

Then Theorem 2.3 gives $I_{\boldsymbol{A}}\, \Theta_t(L_{\boldsymbol{A}}^*) \le 1+\varepsilon$. This has two consequences. First, because $\Theta_t(L_{\boldsymbol{A}}^*) \ge 1$, we immediately conclude that $I_{\boldsymbol{A}} \le 1+\varepsilon < 2$, and since $I_{\boldsymbol{A}}$ is a positive integer, $I_{\boldsymbol{A}} = 1$. Hence the map $\boldsymbol{A}\colon \mathbb{Z}^m \to \mathbb{Z}^r$ is surjective. Second, with $I_{\boldsymbol{A}} = 1$ the inequality reads $\Theta_t(L_{\boldsymbol{A}}^*) \le 1+\varepsilon$, which is a bound on the smoothing parameter. For the matrix $\boldsymbol{A} = (1 \;\; 1)$ of Figures 3 and 5, the left-hand side is about 2.12 at $t = 2/3$ and about 1.12 at $t = 4/3$. The argument therefore applies at the second value and not at the first.

Lemma 2.5 (Theta integral criterion). *Suppose that, for some $t > 0$ and $0 < \varepsilon < 1$,*

$$t^r \int_{V/\Lambda_{\mathrm{row}}} \Theta_t(\mathbb{Z}^m, \boldsymbol{x})\, d\boldsymbol{x} \le 1+\varepsilon.$$

Then $\boldsymbol{A}\mathbb{Z}^m = \mathbb{Z}^r$ and $\eta_\varepsilon(L_{\boldsymbol{A}}) \le t$.

The same deduction is made after Proposition 5.1 in Section 5.1, and it is the assembly step in the proofs of Theorems 5.6 and 5.9. This lemma plugs directly into the Gaussian leftover hash lemma of Lemma 2.2. If its hypothesis holds for t, then $\eta_\varepsilon(L_{\boldsymbol{A}}) \le t$, so the leftover hash lemma applies to an input sampled from $D_{\mathbb{Z}^m,t}$.

Corollary 2.6 (Theta integral criterion and Gaussian LHL). *Let $\boldsymbol{A} \in \mathbb{Z}^{r\times m}$ be a random matrix. For fixed $t > 0$ and $0 < \varepsilon < 1$ let $\mathcal{E}_{t,\varepsilon}$ be the event that the hypothesis of Lemma 2.5 is satisfied. Then whenever $\mathcal{E}_{t,\varepsilon}$ holds, one has*

$$\boldsymbol{A}\mathbb{Z}^m = \mathbb{Z}^r \qquad \text{and} \qquad \mathrm{SD}\Big(\boldsymbol{A}\boldsymbol{z},\, D_{\mathbb{Z}^r,\, t(\boldsymbol{A}\boldsymbol{A}^T)^{1/2}}\Big) \le \frac{\varepsilon}{1-\varepsilon},$$

where $\boldsymbol{z} \leftarrow D_{\mathbb{Z}^m,t}$.

If $\mathcal{E}_{t,\varepsilon}$ has probability at least $1-\tau$, the same comparison holds for the joint law of $(\boldsymbol{A}, \boldsymbol{A}\boldsymbol{z})$ up to the additional error τ. This is how the leftover hash lemmas of Section 6 are obtained from the smoothing theorems, starting with Corollary 6.1.

2.6 The main result at constant width

We now state the result that this section proves. The matrix width s and the input width t play different roles. The first controls the random map, while the second controls the Gaussian source that is passed through that map.

Theorem 2.7 (Constant coefficient width over the integers). *Fix a constant $s \ge 1$, let $\varepsilon = 2^{-r}$, and take $m \gg_s r \ln r$. Sample the entries of $\boldsymbol{X} \in \mathbb{Z}^{r\times m}$ independently from $D_{\mathbb{Z},s}$. Then, with probability $1 - O(2^{-r})$,*

$$\boldsymbol{X}\mathbb{Z}^m = \mathbb{Z}^r \qquad \text{and} \qquad \eta_\varepsilon(\ker_{\mathbb{Z}} \boldsymbol{X}) \ll_s \sqrt{r + \ln m}.$$

When $\ln m = O(r)$ this gives the bound $O(\sqrt{r})$, which is optimal in order by Proposition 5.14 and Corollary 5.15. Theorem 5.9 together with Remark 5.12 gives the same statement with explicit constants that are uniform in $s \ge 1$, and Theorem 5.9 itself proves it over power-of-two cyclotomic fields, where the d coefficients of a ring element are independent width-s Gaussians. We do not assume that $\boldsymbol{X}\mathbb{Z}^m = \mathbb{Z}^r$. Surjectivity is deduced from the theta integral bound.

Let us first explain the proof before introducing the constants. By Lemma 2.5, for a matrix $\boldsymbol{X}$ of full row rank we only need to make the normalized theta integral at most $1+\varepsilon$, with t of order $\sqrt{r + \ln m}$. Put $\mathcal{Q} = [-1/2, 1/2)^r$ and write $\boldsymbol{x} = \boldsymbol{X}^T\boldsymbol{y}$ with $\boldsymbol{y} \in \mathcal{Q}$. We split the integral on the right-hand side of Equation (3) into a small ball around zero in the $\boldsymbol{y}$ coordinates and its complement. The radius will be chosen in the proof.

Near the origin, the Gaussian centered at zero dominates the theta function, and the other Gaussian terms contribute only a small error. This makes the contribution near the origin at most $1 + \varepsilon/2$.

Outside the ball, things get harder. Lemma 2.8 shows that, for each fixed $\boldsymbol{y}$ outside the ball, the expected contribution of one column is at most a constant $q < 1$ depending only on s. Independence makes the expected product at most q^m. Once the column norms are controlled, sufficiently many columns make q^m small enough to absorb the remaining factors in the integral, and Markov's inequality bounds the contribution outside the ball by $\varepsilon/2$ with high probability. The proof also checks that large columns and matrices of rank less than r are rare.

2.7 The one-column estimate

Lemma 2.8 (One-column estimate). *Fix $s \geq 1$. There are constants $C(s) > 0$, $c(s) > 0$ and $q = q(s) < 1$ such that, for every $\boldsymbol{y} \in \mathcal{Q}$,*

$$t \geq C(s)\sqrt{r} \quad \text{and} \quad t\,||\boldsymbol{y}|| \geq c(s) \quad \Longrightarrow \quad \mathbb{E}_{\boldsymbol{Z} \leftarrow D_{\mathbb{Z}^r,s}} \Theta_t(\mathbb{Z}, \langle \boldsymbol{Z}, \boldsymbol{y} \rangle) \leq q(s).$$

The constants depend only on s, not on r, t or $\boldsymbol{y}$.

The proof distinguishes two kinds of directions $\boldsymbol{y}$, exactly as Lemma 5.11 does in general. If some coordinate of $\boldsymbol{y}$ is large, a unit translation of one coordinate of $\boldsymbol{Z}$ moves the phase $\langle \boldsymbol{Z}, \boldsymbol{y} \rangle$ away from the integers, and Lemma 2.1 makes this translation cheap. If all coordinates are small, the phase is a sum of many small independent terms and a moment computation shows that it stays away from the integers with positive probability. For this overview we give an elementary moment proof with constants depending on s. The uniform explicit constants in Lemma 5.11, and their Lean formalization, use the Berry–Esseen theorem [Ber41, Ess42].

Proof. Write $\boldsymbol{Z} = (Z_1, \ldots, Z_r)$ and choose $C(s) \geq 8$. It is enough to show that $\langle \boldsymbol{Z}, \boldsymbol{y} \rangle$ has a fixed positive probability of being at distance at least $1/t$ from the integers. Indeed, Poisson summation and Gaussian tail sums give, for $t \geq 8$,

$$\begin{aligned} &\Theta_t(\mathbb{Z}, u) \leq 1 + 3e^{-\pi t^2} && \text{for every } u, \\ &\Theta_t(\mathbb{Z}, u) \leq \tfrac{1}{8} && \text{if } \operatorname{dist}(u, \mathbb{Z}) \geq \tfrac{1}{t}. \end{aligned} \tag{4}$$

Case 1: $\operatorname{dist}(y_i, \mathbb{Z}) \geq 2/t$ *for some* i. Let

$$E_{\boldsymbol{y}} = \Big\{ \boldsymbol{z} \in \mathbb{Z}^r : \operatorname{dist}(\langle \boldsymbol{z}, \boldsymbol{y} \rangle, \mathbb{Z}) < \frac{1}{t} \Big\}.$$

The inner products $\langle \boldsymbol{z}, \boldsymbol{y} \rangle$ and $\langle \boldsymbol{z} + \boldsymbol{e}_i, \boldsymbol{y} \rangle$ differ by y_i, so they cannot both be within $1/t$ of integers. The events $\{\boldsymbol{Z} \in E_{\boldsymbol{y}}\}$ and $\{\boldsymbol{Z} + \boldsymbol{e}_i \in E_{\boldsymbol{y}}\}$ use the same sample $\boldsymbol{Z}$ and are disjoint, so

$$\mathbb{P}[\boldsymbol{Z} \in E_{\boldsymbol{y}}] + \mathbb{P}[\boldsymbol{Z} + \boldsymbol{e}_i \in E_{\boldsymbol{y}}] \leq 1.$$

By the definition of statistical distance,

$$\mathbb{P}[\boldsymbol{Z} \in E_{\boldsymbol{y}}] - \mathbb{P}[\boldsymbol{Z} + \boldsymbol{e}_i \in E_{\boldsymbol{y}}] \leq \mathrm{SD}(\boldsymbol{Z}, \boldsymbol{Z} + \boldsymbol{e}_i) = \mathrm{SD}\big(D_{\mathbb{Z},s}, D_{\mathbb{Z},s} + 1\big).$$

The equality holds because the coordinates are independent and only coordinate i changes. Adding this to the preceding bound gives $2\mathbb{P}[\boldsymbol{Z} \in E_{\boldsymbol{y}}] \leq 1 + \mathrm{SD}(D_{\mathbb{Z},s}, D_{\mathbb{Z},s} + 1)$, and Lemma 2.1 yields

$$\mathbb{P}\Big[\operatorname{dist}(\langle \boldsymbol{Z}, \boldsymbol{y} \rangle, \mathbb{Z}) \geq \frac{1}{t}\Big] \geq \frac{1 - \mathrm{SD}(D_{\mathbb{Z},s}, D_{\mathbb{Z},s} + 1)}{2} > 0.$$

This lower bound depends only on s.

Case 2: $\operatorname{dist}(y_i, \mathbb{Z}) < 2/t$ *for every* i. Since $\boldsymbol{y} \in \mathcal{Q}$ and $t \geq 8$, we have $|y_i| < 2/t$ and hence

$$||\boldsymbol{y}||^2 < \frac{4r}{t^2}, \qquad t\,||\boldsymbol{y}|| < 2\sqrt{r}. \tag{5}$$

We show that the inner product $\langle \boldsymbol{Z}, \boldsymbol{y} \rangle$ has a fixed positive probability of lying between $1/t$ and $1/4$ in absolute value. In this range its closest integer is zero.

The second and fourth moments of Z_1 are finite constants depending only on s, and $\mathbb{E}Z_1^2 > 0$. Independence and symmetry give

$$\mathbb{E}\,\langle \boldsymbol{Z}, \boldsymbol{y}\rangle^2 = \mathbb{E}Z_1^2\ ||\boldsymbol{y}||^2 \asymp_s ||\boldsymbol{y}||^2,$$
$$\mathbb{E}\,\langle \boldsymbol{Z}, \boldsymbol{y}\rangle^4 = \mathbb{E}Z_1^4 \sum_i y_i^4 + 6(\mathbb{E}Z_1^2)^2 \sum_{i<j} y_i^2 y_j^2 \ll_s ||\boldsymbol{y}||^4.$$

Split the second moment according to whether $|\langle \boldsymbol{Z}, \boldsymbol{y}\rangle| < 1/t$. The Cauchy–Schwarz inequality gives

$$\mathbb{E}\,\langle \boldsymbol{Z}, \boldsymbol{y}\rangle^2 \le \frac{1}{t^2} + \left(\mathbb{E}\,\langle \boldsymbol{Z}, \boldsymbol{y}\rangle^4\right)^{1/2} \mathbb{P}\Big[|\langle \boldsymbol{Z}, \boldsymbol{y}\rangle| \ge \frac{1}{t}\Big]^{1/2}. \tag{6}$$

Since $\mathbb{E}Z_1^2 > 0$ depends only on s, we choose $c(s)$ so that the hypothesis $t\,||\boldsymbol{y}|| \ge c(s)$ gives

$$t\,||\boldsymbol{y}|| \ge c(s) \ge \sqrt{\frac{2}{\mathbb{E}Z_1^2}}, \qquad \text{which gives} \qquad \frac{1}{t^2} \le \frac{1}{2}\,\mathbb{E}\,\langle \boldsymbol{Z}, \boldsymbol{y}\rangle^2. \tag{7}$$

If the upper bound in Equation (5) and the lower bound in Equation (7) cannot hold simultaneously, Case 2 is empty. Rearranging Equation (6) now gives

$$\mathbb{P}\Big[|\langle \boldsymbol{Z}, \boldsymbol{y}\rangle| \ge \frac{1}{t}\Big] \ge \frac{(\mathbb{E}\,\langle \boldsymbol{Z}, \boldsymbol{y}\rangle^2)^2}{4\,\mathbb{E}\,\langle \boldsymbol{Z}, \boldsymbol{y}\rangle^4} \gg_s 1.$$

On the other hand, Markov's inequality gives

$$\mathbb{P}\Big[|\langle \boldsymbol{Z}, \boldsymbol{y}\rangle| > \frac{1}{4}\Big] \le 16\,\mathbb{E}\,\langle \boldsymbol{Z}, \boldsymbol{y}\rangle^2 \ll_s \frac{r}{t^2} \le \frac{1}{C(s)^2}.$$

By taking $C(s)$ large enough, this Markov bound is less than half the preceding lower bound. Subtracting the two probabilities gives

$$\mathbb{P}\Big[\operatorname{dist}(\langle \boldsymbol{Z}, \boldsymbol{y}\rangle, \mathbb{Z}) \ge \frac{1}{t}\Big] \ge \mathbb{P}\Big[\frac{1}{t} \le |\langle \boldsymbol{Z}, \boldsymbol{y}\rangle| \le \frac{1}{4}\Big] \gg_s 1.$$

In both cases the inner product stays away from the integers with a probability bounded below in terms of s alone. The theta bounds Equation (4) now give

$$\mathbb{E}\,\Theta_t(\mathbb{Z}, \langle \boldsymbol{Z}, \boldsymbol{y}\rangle) \le 1 + 3e^{-\pi t^2} - \frac{7}{8}\,\mathbb{P}\Big[\operatorname{dist}(\langle \boldsymbol{Z}, \boldsymbol{y}\rangle, \mathbb{Z}) \ge \frac{1}{t}\Big] \le q(s) < 1$$

after increasing $C(s)$ if necessary. □

2.8 Proof of the main theorem

Proof (Proof of Theorem 2.7). Choose $t \asymp_s \sqrt{r + \ln m}$ with a sufficiently large implicit constant. Write $L = \ker_{\mathbb{Z}} \boldsymbol{X}$. When $\boldsymbol{X}$ has full row rank, the exact theta integral identity with $\boldsymbol{A} = \boldsymbol{X}$ gives

$$I_{\boldsymbol{X}}\,\Theta_t(L^*) = t^r \int_{V/\Lambda_{\mathrm{row}}} \Theta_t(\mathbb{Z}^m, \boldsymbol{x})\ d\boldsymbol{x},$$

where $I_{\boldsymbol{X}} = [\mathbb{Z}^r : \boldsymbol{X}\mathbb{Z}^m]$. We bound the probability that $\boldsymbol{X}$ has rank less than r at the end.

Use the fundamental domain $F = \boldsymbol{X}^T\mathcal{Q}$ for Λ_{row} and write $\Delta_{\boldsymbol{X}} = \sqrt{\det(\boldsymbol{X}\boldsymbol{X}^T)}$. Under the parametrization $\boldsymbol{x} = \boldsymbol{X}^T\boldsymbol{y}$ we have $d\boldsymbol{x} = \Delta_{\boldsymbol{X}}\ d\boldsymbol{y}$ and

$$\Theta_t(\mathbb{Z}^m, \boldsymbol{X}^T\boldsymbol{y}) = \prod_{j=1}^m \Theta_t\big(\mathbb{Z}, \langle \boldsymbol{x}_j, \boldsymbol{y}\rangle\big),$$

where $\boldsymbol{x}_j$ is the j-th column of $\boldsymbol{X}$.

The contribution from the ball B_δ. We split the integral into a small ball in the $\boldsymbol{y}$ coordinates and its complement. To choose the radius, let G be the event that every column of $\boldsymbol{X}$ has norm at most $C_s\sqrt{r+\ln m}$, where C_s is a sufficiently large constant depending only on s. The discrete Gaussian tail bound Equation (25) and a union bound over the m columns give

$$\mathbb{P}[G^c] \le m\, 2^{r/2} e^{-\pi C_s^2 (r+\ln m)/(2s^2)} = O(2^{-r}).$$

Take

$$\delta = \frac{1}{4C_s\sqrt{r+\ln m}}, \qquad B_\delta = \{\boldsymbol{y} \in \mathcal{Q} : ||\boldsymbol{y}|| \le \delta\}.$$

On G, every $\boldsymbol{y} \in B_\delta$ satisfies $|\langle \boldsymbol{x}_j, \boldsymbol{y}\rangle| \le ||\boldsymbol{x}_j||\,||\boldsymbol{y}|| \le 1/4$ for every j. Thus each coordinate of $\boldsymbol{X}^T\boldsymbol{y}$ is close to zero. For $\boldsymbol{v} \in \mathbb{Z}^m$,

$$\left|\left|\boldsymbol{v} - \boldsymbol{X}^T\boldsymbol{y}\right|\right|^2 - \left|\left|\boldsymbol{X}^T\boldsymbol{y}\right|\right|^2 \ge \sum_{j=1}^m \Big(v_j^2 - \frac{|v_j|}{2}\Big) \ge \frac{1}{2}\,||\boldsymbol{v}||^2.$$

Summing over $\boldsymbol{v}$ gives

$$\Theta_t(\mathbb{Z}^m, \boldsymbol{X}^T\boldsymbol{y}) \le e^{-\pi t^2 ||\boldsymbol{X}^T\boldsymbol{y}||^2} \big(1 + 3e^{-\pi t^2/2}\big)^m.$$

What does the Gaussian near the origin contribute? We bound its integral over B_δ by its integral over all of $\mathbb{R}^r$, which is exactly normalized:

$$t^r \Delta_{\boldsymbol{X}} \int_{B_\delta} e^{-\pi t^2 ||\boldsymbol{X}^T\boldsymbol{y}||^2}\, d\boldsymbol{y} \le t^r \Delta_{\boldsymbol{X}} \int_{\mathbb{R}^r} e^{-\pi t^2 ||\boldsymbol{X}^T\boldsymbol{y}||^2}\, d\boldsymbol{y} = 1.$$

Consequently, on G, the contribution near the origin is

$$t^r \Delta_{\boldsymbol{X}} \int_{B_\delta} \Theta_t(\mathbb{Z}^m, \boldsymbol{X}^T\boldsymbol{y})\, d\boldsymbol{y} \le \big(1 + 3e^{-\pi t^2/2}\big)^m < 1 + \frac{\varepsilon}{2}.$$

The last inequality follows by choosing the constant in t large enough, since $\varepsilon = 2^{-r}$. This is Lemma 5.2 in the rational case, and the point is that the whole-space normalization is exact however badly conditioned $\boldsymbol{X}$ may be.

The contribution outside B_δ. Choose the constant in t so that $t \ge C(s)\sqrt{r}$ and $t\delta \ge c(s)$. Away from the ball we have $t\,||\boldsymbol{y}|| > t\delta \ge c(s)$, so Lemma 2.8 applies. Denote the contribution outside the ball by

$$\mathcal{R}_{\boldsymbol{X}} = t^r \Delta_{\boldsymbol{X}} \int_{\mathcal{Q}\setminus B_\delta} \prod_{j=1}^m \Theta_t\big(\mathbb{Z}, \langle \boldsymbol{x}_j, \boldsymbol{y}\rangle\big)\, d\boldsymbol{y},$$

and set $\mathcal{R}_{\boldsymbol{X}} = 0$ when $\boldsymbol{X}$ has rank less than r. On G, the inequality between the arithmetic and geometric means of the eigenvalues of $\boldsymbol{X}\boldsymbol{X}^T$ gives

$$\Delta_{\boldsymbol{X}} \le \Big(\frac{1}{r}\sum_{j=1}^m ||\boldsymbol{x}_j||^2\Big)^{r/2} \le \Big(C_s\sqrt{\frac{m(r+\ln m)}{r}}\Big)^r.$$

The domain $\mathcal{Q}\setminus B_\delta$ is fixed independently of $\boldsymbol{X}$. Thus we can bound $\Delta_{\boldsymbol{X}}$ on G, drop the restriction to G inside the nonnegative expectation, and use the independence of the columns. Fubini's theorem and Lemma 2.8 give

$$\begin{aligned}\mathbb{E}[\mathbf{1}_G \mathcal{R}_{\boldsymbol{X}}] &\le \Big(C_s t\sqrt{\frac{m(r+\ln m)}{r}}\Big)^r \int_{\mathcal{Q}\setminus B_\delta} \prod_{j=1}^m \mathbb{E}\,\Theta_t\big(\mathbb{Z}, \langle \boldsymbol{x}_j, \boldsymbol{y}\rangle\big)\, d\boldsymbol{y} \\ &\le \Big(C_s t\sqrt{\frac{m(r+\ln m)}{r}}\Big)^r q^m.\end{aligned}$$

Here $\mathbf{1}_G$ equals one on G and zero otherwise. The logarithm of the factor before q^m is $O_s(r\ln m)$. To compare it with the decay of q^m, take $m \ge Kr\ln r$, where $K \ge 2$ will depend only on s. Since $\ln m/m$ decreases for $m > e$, for $r \ge 2$ we have

$$\frac{r\ln m}{m} \le \frac{\ln(Kr\ln r)}{K\ln r} \le \frac{2 + \ln K/\ln 2}{K}.$$

The last bound tends to zero as K increases. Choosing K sufficiently large therefore bounds the expectation by $q^{m/2} \le 2^{-2r}$. Markov's inequality gives

$$\mathbb{P}\Big[G \text{ and } \mathcal{R}_{\boldsymbol{X}} > \frac{\varepsilon}{2}\Big] \le \frac{2}{\varepsilon}\mathbb{E}[\mathbf{1}_G \mathcal{R}_{\boldsymbol{X}}] = O(2^{-r}).$$

The tail, determinant and remainder estimates are those of Equations (35) to (37) in the general proof.

Rank. It remains to bound the probability that $\boldsymbol{X}$ has rank less than r. The largest probability mass at any point for one entry is $p_* = \Theta_{1/s}(\mathbb{Z})^{-1} < 1$. For a fixed proper subspace $W \subsetneq \mathbb{R}^r$, let $\boldsymbol{Z} \leftarrow D_{\mathbb{Z}^r,s}$. Choose a nonzero vector $\boldsymbol{a}$ orthogonal to W and an index i with $a_i \neq 0$. The condition $\boldsymbol{Z} \in W$ implies

$$Z_i = -\frac{1}{a_i}\sum_{j \neq i} a_j Z_j.$$

Conditioning on all coordinates except Z_i fixes the right-hand side, and by independence each possible value of Z_i still has probability at most p_*. Thus $\mathbb{P}[\boldsymbol{Z} \in W] \le p_*$. Now fix any $r-1$ columns of $\boldsymbol{X}$ and condition on their values. Their span has dimension at most $r-1$, and the other $m-r+1$ columns remain independent, so the probability that they all lie in that span is at most p_*^{m-r+1}. If $\boldsymbol{X}$ has rank less than r, some $r-1$ columns span all the columns. A union bound over these choices gives

$$\mathbb{P}[\operatorname{rank} \boldsymbol{X} < r] \le \binom{m}{r-1} p_*^{m-r+1} = O(2^{-r})$$

when $m \gg_s r \ln r$. The general versions of this step are Equations (38) and (45).

Assembly. Except with probability $O(2^{-r})$, the matrix has full row rank and the right-hand side of Equation (3) is at most $1+\varepsilon$. The theta integral criterion of Lemma 2.5 then gives $I_{\boldsymbol{X}} = 1$ and $\Theta_t(L^*) \le 1+\varepsilon$. Hence $\boldsymbol{X}\mathbb{Z}^m = \mathbb{Z}^r$ and $\eta_\varepsilon(L) \le t \ll_s \sqrt{r + \ln m}$. □

2.9 The other regimes and where they are treated

The argument above has three moving parts, the exact identity, the main term near the origin, and the one-column estimate away from it. Only the last part depends on the width regime. At polynomial width the one-column estimate is obtained by displacing the direction $\boldsymbol{y}$ by an integer vector in all r coordinates at once (Lemmas 5.4 and 5.5), which is why that regime needs a width of order $\sqrt{r}$. At constant width the displacement is replaced by the one-coordinate translation and the moment argument of Lemma 2.8. Table 3 summarizes the two regimes over $\mathbb{Q}$.

Regime	Width and columns	Conclusion with high probability	Where
Polynomial	$s = c_X r^p$ with $p \ge \frac{1}{2}$, and $m \asymp r \ln r$.	With failure probability $O(2^{-r})$, $\boldsymbol{X}\mathbb{Z}^m = \mathbb{Z}^r$ and $\eta_{2^{-r}}(\ker_{\mathbb{Z}} \boldsymbol{X}) = O_p(\sqrt{r})$.	Theorems 5.6 and 5.8
Constant	$s \ge 1$ is a fixed constant, in particular $s = 1$, and $m \asymp r \ln r$.	With failure probability $O(2^{-r})$, $\boldsymbol{X}\mathbb{Z}^m = \mathbb{Z}^r$ and $\eta_{2^{-r}}(\ker_{\mathbb{Z}} \boldsymbol{X}) = O_s(\sqrt{r})$.	Theorem 2.7, Theorem 5.9 and Remark 5.12

Table 3. The two width regimes over $\mathbb{Q}$ in which the theta integral proves surjectivity and smoothing on the same event. In both, the order $\sqrt{r}$ is optimal by Proposition 5.14. The narrow regime, in which $m - r$ is fixed and $s \to \infty$, is not reached by the method and is discussed in Sections 1.4 and 4.3.

Over a number field of degree d the same three parts appear, with $\boldsymbol{X}$ replaced by its integer coefficient matrix of size $dr \times dm$. Each ring column then contributes d dependent integer columns, and the one-column estimate uses one witness coordinate per ring column, so that only the m ring columns need to be independent (Sections 3.2 and 5.2).

Remark 2.9 (Relation to random matrix theory). A similar pattern appears in the study of random matrices. Rudelson and Vershynin [RV08] separate directions according to their arithmetic structure and use small-ball estimates, that is, bounds on the probability that a random linear form lands in a short interval, to control the smallest singular value. Tao and Vu [TV09] develop the complementary inverse principle, that an unusually concentrated linear form must come from structured coefficients. Those papers concern the invertibility of random matrices rather than theta integrals, but the case distinction in Lemma 2.8 is of the same kind. As far as we know, this way of arguing has not appeared in the analysis of Gaussian leftover hash lemmas before.

3 Preliminaries

3.1 Canonical geometry and notation

Let K be a number field of degree d with ring of integers $\mathcal{O}_K$, discriminant Δ_K, and Minkowski completion $K_{\mathbb{R}} = K \otimes_{\mathbb{Q}} \mathbb{R}$. We use all d complex embeddings and equip $K_{\mathbb{R}}$ with the t_2 inner product

$$\langle x, y \rangle = \sum_{\tau\colon K \hookrightarrow \mathbb{C}} \tau(x)\overline{\tau(y)}, \qquad ||x||^2 = \langle x, x \rangle ,$$

extended coordinatewise to $K_{\mathbb{R}}^m$. The canonical embedding is denoted by Φ and is used coordinatewise on powers of K. Under the usual one-representative-per-conjugate-pair model with the complex coordinates scaled by $\sqrt{2}$, the space $K_{\mathbb{R}}^m$ is isometric to $\mathbb{R}^{dm}$, and all real matrices below use this identification. The lattice $\Phi(\mathcal{O}_K)$ has covolume $|\Delta_K|^{1/2}$, and $\mathrm{rd}(K) := |\Delta_K|^{1/d}$ is the root discriminant. All logarithms are natural unless a base is shown.

For a lattice $L \subset \mathbb{R}^M$ of rank N, a center $\boldsymbol{c}$ and a nonsingular matrix $\boldsymbol{S}$ we write

$$\rho_{\boldsymbol{S},\boldsymbol{c}}(\boldsymbol{x}) := \exp\bigl(-\pi(\boldsymbol{x}-\boldsymbol{c})^T(\boldsymbol{S}\boldsymbol{S}^T)^{-1}(\boldsymbol{x}-\boldsymbol{c})\bigr),$$
$$D_{L,\boldsymbol{S},\boldsymbol{c}}(\boldsymbol{x}) := \frac{\rho_{\boldsymbol{S},\boldsymbol{c}}(\boldsymbol{x})}{\rho_{\boldsymbol{S},\boldsymbol{c}}(L)} \quad (\boldsymbol{x} \in L).$$

We omit $\boldsymbol{c}$ when it is zero and write ρ_s and $D_{L,s}$ when $\boldsymbol{S} = s\boldsymbol{I}$. For an $\mathcal{O}_K$-module we use the same notation before and after the canonical embedding, so that $D_{\mathcal{O}_K^r,s}$ is the pullback of $D_{\Phi(\mathcal{O}_K^r),s}$. The dual lattice is $L^* = \{\boldsymbol{y} \in \operatorname{span} L : \langle \boldsymbol{y}, L\rangle \subseteq \mathbb{Z}\}$, and the smoothing parameter is

$$\eta_\varepsilon(L) := \inf\left\{t > 0 : \rho_{1/t}(L^* \setminus \{\boldsymbol{0}\}) \le \varepsilon\right\}.$$

Successive minima $\lambda_i(L)$ refer to the Euclidean norm, and $\lambda_1^\infty(L)$ is the first minimum for the *ambient* ℓ_∞ norm of $\mathbb{R}^M$. The covering radius is $\mu(L)$. We put

$$\gamma_h(\varepsilon) := \sqrt{\frac{\ln\bigl(2h(1+\varepsilon^{-1})\bigr)}{\pi}}, \qquad \text{so that} \qquad \eta_\varepsilon(L) \le \gamma_N(\varepsilon)\,\lambda_N(L) \tag{8}$$

for every lattice of rank N, by [MR07, Lemma 3.3]. In particular $\eta_\varepsilon(\mathbb{Z}^M) \le \gamma_M(\varepsilon)$. Statistical distance is denoted SD.

For $\boldsymbol{X} \in \mathcal{O}_K^{r\times m}$ let $\widetilde{\Phi}(\boldsymbol{X}) \in \mathbb{R}^{dr\times dm}$ be the real matrix of $\boldsymbol{y} \mapsto \boldsymbol{X}\boldsymbol{y}$ in canonical coordinates, and

$$\Lambda^\perp(\boldsymbol{X}) := \{\boldsymbol{v} \in \mathcal{O}_K^m : \boldsymbol{X}\boldsymbol{v} = \boldsymbol{0}\}, \qquad L_{\boldsymbol{X}} := \Phi(\Lambda^\perp(\boldsymbol{X})) \subset \mathbb{R}^{dm}.$$

We write $n := m - r$ and $N := dn$. When $\widetilde{\Phi}(\boldsymbol{X})$ has full row rank, $L_{\boldsymbol{X}}$ has rank N, and in general its rank is at least N.

3.2 Coefficient coordinates and basis constants

Fix an integral basis $\boldsymbol{\omega} = (\omega_1, \dots, \omega_d)$ of $\mathcal{O}_K$ with $\omega_1 = 1$. Such a basis exists because $\mathcal{O}_K/\mathbb{Z}$ is torsion free. Let $c\colon \mathcal{O}_K \to \mathbb{Z}^d$ be the coefficient isomorphism and $c_j\colon \mathcal{O}_K^j \to \mathbb{Z}^{dj}$ its coordinatewise extension. Let

$\boldsymbol{B}_K\colon \mathbb{R}^d \to K_{\mathbb{R}}$ be the linear map $\boldsymbol{B}_K \boldsymbol{z} = \Phi(\sum_k z_k \omega_k)$, so that $\Phi(a) = \boldsymbol{B}_K c(a)$, and let $\boldsymbol{G}_K = \boldsymbol{B}_K^T \boldsymbol{B}_K$ be the Gram matrix of the basis. We define the basis constants

$$\alpha_K := ||\boldsymbol{B}_K||_{\mathrm{op}} = \lambda_{\max}(\boldsymbol{G}_K)^{1/2}, \qquad \beta_K := \left|\left|\boldsymbol{B}_K^{-1}\right|\right|_{\mathrm{op}} = \lambda_{\min}(\boldsymbol{G}_K)^{-1/2},$$
$$\kappa_K := \alpha_K \beta_K.$$

Let $\boldsymbol{T}_k \in \mathbb{Z}^{d\times d}$ be the matrix of multiplication by ω_k in coefficient coordinates, so that $c(\omega_k a) = \boldsymbol{T}_k c(a)$ and $\boldsymbol{T}_1 = \boldsymbol{I}_d$, and put

$$\mu_K := \max_{1\le k\le d} ||\boldsymbol{T}_k||_{\mathrm{op}}.$$

Since $\det \boldsymbol{G}_K = |\Delta_K|$, the discriminant controls the product of the singular values of $\boldsymbol{B}_K$ but not the condition number κ_K, which depends on the chosen basis.

For $\boldsymbol{X} = [\boldsymbol{x}_1|\cdots|\boldsymbol{x}_m]$ let $\boldsymbol{A}_{\boldsymbol{X}} \in \mathbb{Z}^{dr\times dm}$ be the *coefficient matrix*, defined by $c_r(\boldsymbol{X}\boldsymbol{y}) = \boldsymbol{A}_{\boldsymbol{X}} c_m(\boldsymbol{y})$ for all $\boldsymbol{y} \in \mathcal{O}_K^m$. Index its columns by pairs $(j,k) \in [m]\times[d]$. Then

$$\boldsymbol{A}_{\boldsymbol{X}} \boldsymbol{e}_{(j,k)} = c_r(\omega_k \boldsymbol{x}_j) = (\boldsymbol{I}_r \otimes \boldsymbol{T}_k)\, c_r(\boldsymbol{x}_j), \tag{9}$$

so the $k=1$ columns form the matrix $\boldsymbol{C}_{\boldsymbol{X}} = [c_r(\boldsymbol{x}_1)|\cdots|c_r(\boldsymbol{x}_m)]$ of coefficient columns, and every column of $\boldsymbol{A}_{\boldsymbol{X}}$ has norm at most $\mu_K\, ||c_r(\boldsymbol{x}_j)||$. Moreover

$$\ker_{\mathbb{Z}} \boldsymbol{A}_{\boldsymbol{X}} = c_m(\Lambda^{\perp}(\boldsymbol{X})), \qquad L_{\boldsymbol{X}} = \boldsymbol{B}_K^{\oplus m}\big(\ker_{\mathbb{Z}} \boldsymbol{A}_{\boldsymbol{X}}\big), \tag{10}$$

where $\boldsymbol{B}_K^{\oplus m}$ has operator norm α_K and its inverse has operator norm β_K.

The coefficient law. Throughout, $\boldsymbol{X}$ has independent columns $\boldsymbol{x}_j \leftarrow D_{\mathcal{O}_K^r, s}$. Since $||\Phi(\boldsymbol{x})||^2 = c_r(\boldsymbol{x})^T(\boldsymbol{I}_r \otimes \boldsymbol{G}_K)c_r(\boldsymbol{x})$, the coefficient vectors are independent ellipsoidal integer Gaussians:

$$\begin{gathered} c_r(\boldsymbol{x}_j) \leftarrow D_{\mathbb{Z}^R, \boldsymbol{S}}, \qquad \boldsymbol{S} := s\,(\boldsymbol{I}_r \otimes \boldsymbol{G}_K)^{-1/2}, \qquad R := dr, \\ s_- := s_{\min}(\boldsymbol{S}) = s/\alpha_K, \qquad s_+ := s_{\max}(\boldsymbol{S}) = s\beta_K. \end{gathered} \tag{11}$$

Equivalently $\boldsymbol{S}^{-1} c_r(\boldsymbol{x}_j) \leftarrow D_{\boldsymbol{S}^{-1}\mathbb{Z}^R, 1}$. We call s_- the *coefficient width*. It governs every anti-concentration statement below.

3.3 Prime-power cyclotomic fields

Let p be a prime, $\ell \ge 1$, $f = p^\ell$, $K = \mathbb{Q}(\zeta_f)$, $h = p^{\ell-1}$ and $d = \varphi(f) = (p-1)h$. Put $\kappa_p := 1$ if $p = 2$ and $\kappa_p := p$ if p is odd. The following proposition makes the basis constants explicit for the power basis.

Proposition 3.1 (Power-basis geometry for prime-power cyclotomics). *For the integral power basis $\boldsymbol{\omega} = (1, \zeta_f, \ldots, \zeta_f^{d-1})$, a simultaneous permutation of rows and columns puts the Gram matrix in the form*

$$\boldsymbol{G}_K = \boldsymbol{I}_h \otimes \big(p^\ell \boldsymbol{I}_{p-1} - h\mathbf{1}_{p-1}\mathbf{1}_{p-1}^T\big). \tag{12}$$

The eigenvalue h has multiplicity h and the eigenvalue p^ℓ has multiplicity $h(p-2)$. Consequently

$$\alpha_K = \sqrt{\frac{\kappa_p d}{p-1}}, \qquad \beta_K = \sqrt{\frac{p-1}{d}}, \qquad \kappa_K = \sqrt{\kappa_p}, \qquad \mu_K \le \sqrt{\kappa_p},$$

and for $p = 2$ the last inequality is an equality, $\mu_K = 1$.

Proof. For $0 \le j,k < d$ the Gram entry is the Ramanujan sum $(\boldsymbol{G}_K)_{j,k} = \sum_{u\in(\mathbb{Z}/f\mathbb{Z})^\times} \zeta_f^{u(j-k)} = c_f(j-k)$, and for $f = p^\ell$ one has $c_f(a) = d$ if $f \mid a$, $c_f(a) = -h$ if $h \mid a$ and $f \nmid a$, and $c_f(a) = 0$ otherwise. Group the exponents $0,\ldots,d-1$ according to their residue modulo h. Each class $\{u, u+h, \ldots, u+(p-2)h\}$ has Gram block with diagonal d and off-diagonal $-h$, which is the block in Equation (12). Its eigenvalue on $\mathbf{1}_{p-1}$ is $p^\ell - h(p-1) = h$ and on the orthogonal complement it is p^ℓ. This gives the spectrum and the first three identities. For μ_K, let $\boldsymbol{U}_j$ be multiplication by ζ_f^j in canonical coordinates. It is orthogonal because every conjugate of ζ_f^j has modulus one, and $\boldsymbol{B}_K \boldsymbol{T}_j = \boldsymbol{U}_j \boldsymbol{B}_K$. Hence $||\boldsymbol{T}_j||_{\mathrm{op}} \le \left|\left|\boldsymbol{B}_K^{-1}\right|\right|_{\mathrm{op}} ||\boldsymbol{B}_K||_{\mathrm{op}} = \sqrt{\lambda_{\max}(\boldsymbol{G}_K)/\lambda_{\min}(\boldsymbol{G}_K)} = \sqrt{\kappa_p}$. For $p = 2$ this bound is one and $\boldsymbol{T}_1 = \boldsymbol{I}_d$ attains it. □

For $p = 2$ the blocks have size one, so $\boldsymbol{G}_K = d\,\boldsymbol{I}_d$. Thus the power basis is orthogonal for the t_2 form, $\Phi(\mathcal{O}_K) = \sqrt{d}\,\boldsymbol{U}\mathbb{Z}^d$ for an orthogonal $\boldsymbol{U}$, and multiplication by any power of ζ_f is a signed permutation of coefficients.

Corollary 3.2 (Product coefficient Gaussian). *Let $K = \mathbb{Q}(\zeta_{2^a})$ of degree d. If $x \leftarrow D_{\mathcal{O}_K,s}$ and $x = \sum_{j<d} Z_j\zeta^j$, then the Z_j are independent with $Z_j \leftarrow D_{\mathbb{Z},s/\sqrt{d}}$. The same holds coordinatewise for $D_{\mathcal{O}_K^r,s}$, so that $\boldsymbol{S} = (s/\sqrt{d})\boldsymbol{I}_R$ in Equation (11).*

Proof. By orthogonality the Gaussian weight factors as $\prod_j \exp(-\pi d Z_j^2/s^2)$, and so does the normalizing constant. □

3.4 Metric change

The following lemma transports a smoothing bound through a linear isomorphism. It is used to pass from the coefficient kernel $\ker_{\mathbb{Z}} \boldsymbol{A}_{\boldsymbol{X}}$ to the canonical kernel $L_{\boldsymbol{X}}$.

Lemma 3.3 (Metric change). *Let $\boldsymbol{T}$ be a linear isomorphism from $\operatorname{span} L$ onto $\operatorname{span}(\boldsymbol{T}L)$. Then $\eta_\varepsilon(\boldsymbol{T}L) \le ||\boldsymbol{T}||_{\mathrm{op}}\ \eta_\varepsilon(L)$.*

Proof. The dual of $\boldsymbol{T}L$ is $\boldsymbol{T}^{-*}L^*$, and $\left|\left|\boldsymbol{T}^{-*}\boldsymbol{y}\right|\right| \ge ||\boldsymbol{y}|| \,/\, ||\boldsymbol{T}||_{\mathrm{op}}$. Compare the two dual Gaussian sums term by term. □

3.5 Discrete Gaussian tools

The Gaussian pushforward lemma converts a smoothing bound into a leftover hash lemma.

Lemma 3.4 (Pushforward [GMPW20, Theorem 3.1], in the form of [AFL+26, Lemma 2.8]). *Let $0 < \varepsilon < 1$, let $\boldsymbol{X}\mathcal{O}_K^m = \mathcal{O}_K^r$, let $\boldsymbol{S} \in \mathbb{R}^{dm\times dm}$ be positive definite, and let $\boldsymbol{c} \in \mathcal{O}_K^m$. If $s_{\min}(\boldsymbol{S}) \ge \eta_\varepsilon(L_{\boldsymbol{X}})$ then*

$$\mathrm{SD}\Big(\boldsymbol{X}\cdot D_{\mathcal{O}_K^m,\boldsymbol{S},\boldsymbol{c}},\ D_{\mathcal{O}_K^r,\sqrt{\widetilde{\Phi}(\boldsymbol{X})\boldsymbol{S}\boldsymbol{S}^T\widetilde{\Phi}(\boldsymbol{X})^T},\boldsymbol{X}\boldsymbol{c}}\Big) \le \frac{\varepsilon}{1-\varepsilon}.$$

The parameter matrix in the target is a real matrix in canonical coordinates, namely $\sqrt{\widetilde{\Phi}(\boldsymbol{X})\boldsymbol{S}\boldsymbol{S}^T\widetilde{\Phi}(\boldsymbol{X})^T}$, where $\widetilde{\Phi}(\boldsymbol{X})$ is the real matrix of $\boldsymbol{X}$. It cannot be written as $\boldsymbol{X}\boldsymbol{S}\boldsymbol{S}^T\boldsymbol{X}^T$, because $\boldsymbol{X}$ has entries in $\mathcal{O}_K$ while $\boldsymbol{S}$ is a real $dm \times dm$ matrix, so that product is not defined.

We also need an elementary stability fact to pass from an exact Gaussian parameter to a finite-precision implementation.

Lemma 3.5 (Stability under parameter approximation). *Let $L \subset \mathbb{R}^n$ be a full-rank lattice, let $\boldsymbol{S}$ and $\widehat{\boldsymbol{S}}$ be positive definite, and put*

$$\boldsymbol{Q} = (\boldsymbol{S}\boldsymbol{S}^T)^{-1}, \qquad \widehat{\boldsymbol{Q}} = (\widehat{\boldsymbol{S}}\widehat{\boldsymbol{S}}^T)^{-1}.$$

If, for some $0 < \xi < 1/2$ and $0 < \theta \le 1/2$,

$$\left|\left|\boldsymbol{S}^T(\widehat{\boldsymbol{Q}} - \boldsymbol{Q})\boldsymbol{S}\right|\right|_{\mathrm{op}} \le \theta, \tag{13}$$

then, for an absolute constant C_{stab},

$$\mathrm{SD}\Big(D_{L,\boldsymbol{S}}, D_{L,\widehat{\boldsymbol{S}}}\Big) \le C_{\mathrm{stab}}\left(\xi + \theta\big(n + \ln(1/\xi)\big)\right). \tag{14}$$

Consequently, statistical distance at most τ is obtained by taking $\xi = \Theta(\tau)$ and $\theta = \Theta(\tau/(n + \ln(1/\tau)))$.

Proof. For $\boldsymbol{X} \leftarrow D_{L,\boldsymbol{S}}$ and $\boldsymbol{Y} = \boldsymbol{S}^{-1}\boldsymbol{X}$, completing the square and applying Poisson summation to the shifted lattice theta series give

$$\mathbb{E}\exp(\langle \boldsymbol{t}, \boldsymbol{Y}\rangle) \le \exp\left(\frac{||\boldsymbol{t}||_2^2}{4\pi}\right) \qquad (\boldsymbol{t} \in \mathbb{R}^n).$$

A $1/2$-net of the unit sphere therefore shows that, with $H = n + \ln(1/\xi)$,

$$\mathbb{P}\Big[||\boldsymbol{Y}||_2^2 > C_0 H\Big] \le \xi$$

for an absolute C_0. The same argument applies to $D_{L,\widehat{\boldsymbol{S}}}$. Moreover, Equation (13) implies

$$(1-\theta)\boldsymbol{Q} \preceq \widehat{\boldsymbol{Q}} \preceq (1+\theta)\boldsymbol{Q}.$$

Indeed, the norm bound in Equation (13) first gives $-\theta\boldsymbol{I} \preceq \boldsymbol{S}^T(\widehat{\boldsymbol{Q}} - \boldsymbol{Q})\boldsymbol{S} \preceq \theta\boldsymbol{I}$. Congruence by $\boldsymbol{S}^{-T}$ on the left and $\boldsymbol{S}^{-1}$ on the right then gives the displayed comparison. Consequently the common ellipsoid

$$E = \{\boldsymbol{x} : \boldsymbol{x}^T\boldsymbol{Q}\boldsymbol{x} \le 2C_0H\}$$

has probability at least $1-\xi$ under both distributions: for the second one, $\boldsymbol{x} \notin E$ implies $\boldsymbol{x}^T\widehat{\boldsymbol{Q}}\boldsymbol{x} > C_0H$ because $\theta \le 1/2$.

Let $w, \widehat{w}$ be the two unnormalized weights and let $Z, \widehat{Z}$ be their total masses. On E,

$$e^{-a}w(\boldsymbol{x}) \le \widehat{w}(\boldsymbol{x}) \le e^a w(\boldsymbol{x}), \qquad a := 2\pi C_0\theta H.$$

Using that each distribution gives E mass at least $1-\xi$ yields

$$(1-\xi)e^{-a} \le \frac{\widehat{Z}}{Z} \le \frac{e^a}{1-\xi}.$$

Thus, on E, the ratio of the normalized probabilities lies between $(1-\xi)e^{-2a}$ and $e^{2a}/(1-\xi)$. Adding the two discarded tails gives $\mathrm{SD} = O(\xi + a)$ when $\xi + a$ is bounded by a small absolute constant. When it is not, the trivial bound $\mathrm{SD} \le 1$ gives the same conclusion after enlarging the absolute constant. This proves Equation (14). The final assertion follows by substitution. □

3.6 Spectral estimates and exact shaping

This subsection collects the spectral facts behind the spherical leftover hash lemma. The first controls the largest singular value of $\widetilde{\Phi}(\boldsymbol{X})$ over any number field, the second does the same for a matrix of ellipsoidal Gaussian columns, the third gives a two-sided estimate for power-of-two cyclotomics, and the fourth shapes a Gaussian parameter so that its image is spherical.

Lemma 3.6 (Field-independent upper singular value). *Let K be any number field of degree d, let $m, r \ge 1$, and sample $\boldsymbol{X} \leftarrow D_{\mathcal{O}_K^r,s}^m$. For $0 < \delta < 1$ put*

$$U_{\mathrm{sp}}(s,m,r,d,\delta) := \frac{16s}{9}\sqrt{\frac{2(r+m)\ln 9 + \ln(2d/\delta)}{\pi}}.$$

Then $s_{\max}(\widetilde{\Phi}(\boldsymbol{X})) \le U_{\mathrm{sp}}(s,m,r,d,\delta)$ with probability at least $1-\delta$. In particular, if $m \ge r$ and $m \ge \ln(2d/\delta)$, then $s_{\max}(\widetilde{\Phi}(\boldsymbol{X})) \le 4s\sqrt{m}$ with probability at least $1-\delta$.

Proof. Let $L = \Phi(\mathcal{O}_K^r)$, let $\boldsymbol{x} \leftarrow D_{\mathcal{O}_K^r,s}$ and $\boldsymbol{Y} = \Phi(\boldsymbol{x})$. Completing the square, for every $\boldsymbol{t} \in \mathbb{R}^{dr}$,

$$\mathbb{E}e^{\langle \boldsymbol{t},\boldsymbol{Y}\rangle} = e^{s^2||\boldsymbol{t}||^2/(4\pi)}\frac{\rho_s(L - s^2\boldsymbol{t}/(2\pi))}{\rho_s(L)},$$

and Poisson summation [MR07, Lemma 2.8] gives $\rho_s(L-\boldsymbol{c}) = \frac{s^{dr}}{\det L}\sum_{\boldsymbol{w}\in L^*} e^{-\pi s^2||\boldsymbol{w}||^2}e^{2\pi i\langle \boldsymbol{w},\boldsymbol{c}\rangle}$, so $\rho_s(L-\boldsymbol{c}) \le \rho_s(L)$ and

$$\mathbb{E}e^{\langle \boldsymbol{t},\boldsymbol{Y}\rangle} \le e^{s^2||\boldsymbol{t}||^2/(4\pi)}. \tag{15}$$

In the scaled real Minkowski model $\widetilde{\Phi}(\boldsymbol{X})$ is orthogonally equivalent to the direct sum of the real-place blocks $\tau(\boldsymbol{X}) \in \mathbb{R}^{r\times m}$ and the realified complex-place blocks $\mathcal{R}(\tau(\boldsymbol{X})) \in \mathbb{R}^{2r\times 2m}$, one for each conjugate pair. For a real block and unit vectors $\boldsymbol{u} \in \mathbb{R}^r$, $\boldsymbol{v} \in \mathbb{R}^m$, the bilinear form $\boldsymbol{u}^T\tau(\boldsymbol{X})\boldsymbol{v}$ is a sum over the m independent columns of real linear functionals of their embeddings whose coefficient vectors have squared norms summing to $\sum_j v_j^2 \,||\boldsymbol{u}||^2 = 1$. For a complex block write $\boldsymbol{v} = (a_j, b_j)_j$ and let $\boldsymbol{J}$ be the complex structure on $\mathbb{R}^{2r}$. The contribution of the j-th column is $\langle a_j\boldsymbol{u} - b_j\boldsymbol{J}\boldsymbol{u}, \boldsymbol{q}_j\rangle$ for the unscaled real form $\boldsymbol{q}_j$ of the column, whose scaled coefficient vector is $(a_j\boldsymbol{u} - b_j\boldsymbol{J}\boldsymbol{u})/\sqrt{2}$, and since $\langle \boldsymbol{u}, \boldsymbol{J}\boldsymbol{u}\rangle = 0$ the squared norms sum to $\frac{1}{2}\sum_j(a_j^2 + b_j^2) \le 1$. In either case, independence and Equation (15) give $\mathbb{E}\exp(\theta\boldsymbol{u}^T\boldsymbol{M}_\tau\boldsymbol{v}) \le e^{s^2\theta^2/(4\pi)}$ for every $\theta \in \mathbb{R}$, hence $\mathbb{P}[\left|\boldsymbol{u}^T\boldsymbol{M}_\tau\boldsymbol{v}\right| \ge t] \le 2e^{-\pi t^2/s^2}$ by Chernoff. Take 1/4-nets of the two unit spheres, of sizes at most 9^p and 9^q for a $p \times q$ block [Ver26, Corollary 4.2.11]. Applying [Ver26, Lemma 4.4.1] once on each side gives $||\boldsymbol{M}_\tau||_{\mathrm{op}} \le \frac{16}{9}\max\left|\boldsymbol{u}^T\boldsymbol{M}_\tau\boldsymbol{v}\right|$ over the nets. Every block has $p + q \le 2(r + m)$ and there are at most d blocks. With $t = s\sqrt{(2(r + m)\ln 9 + \ln(2d/\delta))/\pi}$ a union bound proves the claim. If $m \ge r$ and $m \ge \ln(2d/\delta)$ then $2(r + m)\ln 9 + \ln(2d/\delta) \le (4\ln 9 + 1)m$ and the constant is below 4. □

The same argument applies to an ellipsoidal Gaussian parameter and will later control the operator norm of the matrix of sampled hints.

Lemma 3.7 (Operator norm of ellipsoidal ring-Gaussian columns). *Let K be any number field of degree d, let $\boldsymbol{S} \in \mathbb{R}^{dm\times dm}$ be positive definite, and sample the columns of $\boldsymbol{R} \in \mathcal{O}_K^{m\times k}$ independently from $D_{\mathcal{O}_K^m,\boldsymbol{S}}$. For every $0 < \tau < 1$, except with probability τ,*

$$s_{\max}(\widetilde{\Phi}(\boldsymbol{R})) \le \frac{16 s_{\max}(\boldsymbol{S})}{9}\sqrt{\frac{2(m + k)\ln 9 + \ln(2d/\tau)}{\pi}}. \tag{16}$$

In particular, if

$$m \ge k + \ln(2d/\tau), \tag{17}$$

then, except with probability τ,

$$s_{\max}(\widetilde{\Phi}(\boldsymbol{R})) \le 4 s_{\max}(\boldsymbol{S})\sqrt{m}. \tag{18}$$

Proof. Put $L = \Phi(\mathcal{O}_K^m)$ and $\boldsymbol{C} = \boldsymbol{S}\boldsymbol{S}^T$. For $\boldsymbol{r} \leftarrow D_{\mathcal{O}_K^m,\boldsymbol{S}}$, completing the square gives, for every $\boldsymbol{t} \in \mathbb{R}^{dm}$,

$$\mathbb{E}\exp\big(\langle \boldsymbol{t}, \Phi(\boldsymbol{r})\rangle\big) = \exp\left(\frac{\boldsymbol{t}^T\boldsymbol{C}\boldsymbol{t}}{4\pi}\right)\frac{\rho_{\boldsymbol{S}}(L - \boldsymbol{C}\boldsymbol{t}/(2\pi))}{\rho_{\boldsymbol{S}}(L)}.$$

Poisson summation for the anisotropic theta series shows that its shifted mass is at most its mass at the origin. Consequently,

$$\mathbb{E}\exp\big(\langle \boldsymbol{t}, \Phi(\boldsymbol{r})\rangle\big) \le \exp\left(\frac{\boldsymbol{t}^T\boldsymbol{S}\boldsymbol{S}^T\boldsymbol{t}}{4\pi}\right) \le \exp\left(\frac{s_{\max}(\boldsymbol{S})^2\,||\boldsymbol{t}||_2^2}{4\pi}\right).$$

For every real or realified complex block of $\widetilde{\Phi}(\boldsymbol{R})$, the bilinear-form calculation in the proof of Lemma 3.6 therefore gives

$$\mathbb{P}\big[\left|\boldsymbol{u}^T\boldsymbol{M}_\tau\boldsymbol{v}\right| \ge t\big] \le 2\exp\left(-\frac{\pi t^2}{s_{\max}(\boldsymbol{S})^2}\right)$$

for unit $\boldsymbol{u}, \boldsymbol{v}$. The same 1/4-net argument, now with block dimensions at most $2m \times 2k$, proves Equation (16). Under Equation (17), one has $m \ge k$ and $m \ge \ln(2d/\tau)$, so the numerical simplification from Lemma 3.6 gives Equation (18). □

Lemma 3.8 (Power-of-two singular values). *There are absolute constants $c_{\mathrm{sp}}, c'_{\mathrm{sp}}, c''_{\mathrm{sp}} > 0$ with the following property. Let $m, r \ge 1$. Let $K = \mathbb{Q}(\zeta_{2^a})$ have degree d, let $\boldsymbol{X} \leftarrow D_{\mathcal{O}_K^r,s}^m$ with $s/\sqrt{d} \ge 1$, and let $0 < \delta < 1$. If $m \ge c_{\mathrm{sp}}(r + \ln(2d/\delta))$, then with probability at least $1 - \delta$,*

$$c'_{\mathrm{sp}} s\sqrt{m} \le s_{\min}(\widetilde{\Phi}(\boldsymbol{X})) \le s_{\max}(\widetilde{\Phi}(\boldsymbol{X})) \le c''_{\mathrm{sp}} s\sqrt{m}. \tag{19}$$

Proof. Put $\sigma = s/\sqrt{d}$. By Corollary 3.2 every power-basis coefficient is an independent $Z \leftarrow D_{\mathbb{Z},\sigma}$. Poisson summation gives $\rho_\sigma(\mathbb{Z}-u) \le \rho_\sigma(\mathbb{Z})$, so completing the square yields $\mathbb{E}e^{tZ} \le e^{\sigma^2 t^2/(4\pi)}$. Moreover $\mathbb{E}Z^2 \asymp \sigma^2$ uniformly for $\sigma \ge 1$. The upper bound follows from the moment generating function. For the lower bound, note that $\rho_\sigma(\mathbb{Z}) \ll \sigma$, while the integers z with $\sigma/2 \le |z| \le \sigma$ carry unnormalized Gaussian mass of order σ and each of them contributes $z^2 \ge \sigma^2/4$. If $d = 1$, then $K = \mathbb{Q}$ and each column has independent $D_{\mathbb{Z},s}$ coordinates. Writing $v = \mathbb{E}Z^2$, the columns divided by $\sqrt{v}$ are centered and isotropic in $\mathbb{R}^r$: their covariance is $\boldsymbol{I}_r$. Independence and the displayed moment-generating bound show that every unit linear functional is subgaussian with an absolute parameter, since $v \asymp s^2$. Apply the two-sided sample-covariance theorem [MPTJ07, Corollary 2.6] to these real columns with deviation parameter $1/2$. Their singular values are absolute multiples of $\sqrt{m}$, with failure at most e^{-cm} for $m \ge Cr$. Rescaling by $\sqrt{v} \asymp s$ and increasing c_{sp} makes this failure at most δ. This proves the degree-one case. Henceforth assume $d \ge 2$.

Fix a complex embedding $\tau(\zeta) = \zeta^u$. For $x = \sum_\ell Z_\ell \zeta^\ell$ we have $\tau(x) = \sum_\ell Z_\ell \zeta^{u\ell}$ and $\sum_\ell \zeta^{2u\ell} = 0$, so the vector $(\Re\tau(x), \Im\tau(x))$ is centered with covariance $(d\mathbb{E}Z^2/2)\boldsymbol{I}_2$, and every real linear functional of its normalization is subgaussian with an absolute parameter. The same holds for a column of $\tau(\boldsymbol{X}) \in \mathbb{C}^{r\times m}$ viewed in $\mathbb{R}^{2r}$. The two-sided sample-covariance theorem for independent isotropic subgaussian vectors [MPTJ07, Corollary 2.6] with deviation parameter $1/2$ gives, for one conjugate pair, singular values that are absolute multiples of $\sqrt{m}$ times the coordinate scale, with failure probability e^{-cm} once $m \ge Cr$, and increasing c_{sp} makes $e^{-cm} \le 2\delta/d$. The realification satisfies $\mathcal{R}(\tau(\boldsymbol{X}))\mathcal{R}(\tau(\boldsymbol{X}))^T = \sum_j (\boldsymbol{v}_j\boldsymbol{v}_j^T + \boldsymbol{J}\boldsymbol{v}_j\boldsymbol{v}_j^T\boldsymbol{J}^T)$, so the same bounds hold for it, and $d\mathbb{E}Z^2 \asymp s^2$ turns them into absolute multiples of $s\sqrt{m}$. A union bound over the $d/2$ conjugate pairs proves the claim, since $\widetilde{\Phi}(\boldsymbol{X})$ is orthogonally block-diagonalized into these realifications. □

4 The Dual Kernel Lattice

4.1 Covolume and the description of the dual

Proposition 4.1 (Kernel covolume). *If $\boldsymbol{X}\mathcal{O}_K^m = \mathcal{O}_K^r$, then $\operatorname{rank} L_{\boldsymbol{X}} = N$ and*

$$\det(L_{\boldsymbol{X}}) = |\Delta_K|^{n/2} \det(\widetilde{\Phi}(\boldsymbol{X})\widetilde{\Phi}(\boldsymbol{X})^T)^{1/2},$$

and hence

$$\det(L_{\boldsymbol{X}})^{1/N} = \mathrm{rd}(K)^{1/2} \det(\widetilde{\Phi}(\boldsymbol{X})\widetilde{\Phi}(\boldsymbol{X})^T)^{1/(2N)} \le \mathrm{rd}(K)^{1/2} s_{\max}(\widetilde{\Phi}(\boldsymbol{X}))^{r/n}.$$

Proof. Put $\Lambda_m = \Phi(\mathcal{O}_K^m)$ and $\Lambda_r = \Phi(\mathcal{O}_K^r)$, of covolumes $|\Delta_K|^{m/2}$ and $|\Delta_K|^{r/2}$, and $W = \ker \widetilde{\Phi}(\boldsymbol{X})$. Surjectivity makes $\widetilde{\Phi}(\boldsymbol{X})$ map Λ_m onto Λ_r and makes $L_{\boldsymbol{X}} = \Lambda_m \cap W$ a primitive sublattice. Lift a basis of Λ_r to Λ_m and project the lifts orthogonally to $W^\perp$. Together with a basis of $L_{\boldsymbol{X}}$ they span a fundamental domain of Λ_m, and $\widetilde{\Phi}(\boldsymbol{X})$ maps the projected parallelotope onto a fundamental domain of Λ_r with Jacobian $\det(\widetilde{\Phi}(\boldsymbol{X})\widetilde{\Phi}(\boldsymbol{X})^T)^{1/2}$. Therefore $|\Delta_K|^{m/2} = \det(L_{\boldsymbol{X}})\,|\Delta_K|^{r/2}\det(\widetilde{\Phi}(\boldsymbol{X})\widetilde{\Phi}(\boldsymbol{X})^T)^{-1/2}$. The inequality is Hadamard's. □

Now let $E := \ker \boldsymbol{A}_{\boldsymbol{X}} \subseteq \mathbb{R}^{dm}$ be the coefficient kernel space. It is a rational subspace, so $L := \ker_{\mathbb{Z}} \boldsymbol{A}_{\boldsymbol{X}} = \mathbb{Z}^{dm} \cap E$ is a primitive sublattice of $\mathbb{Z}^{dm}$ of rank $\dim E$. Write π_E for the orthogonal projection onto E.

Lemma 4.2 (Dual of a primitive kernel). *$L^* = \pi_E(\mathbb{Z}^{dm}) = (\mathbb{Z}^{dm} + \boldsymbol{A}_{\boldsymbol{X}}^T\mathbb{R}^{dr}) \cap E$. Every $\boldsymbol{u} \in L^*$ can thus be written as $\boldsymbol{u} = \boldsymbol{k} + \boldsymbol{A}_{\boldsymbol{X}}^T\boldsymbol{y}$ with $\boldsymbol{k} \in \mathbb{Z}^{dm}$ and $\boldsymbol{y} \in \mathbb{R}^{dr}$.*

Proof. For $\boldsymbol{k} \in \mathbb{Z}^{dm}$ and $\boldsymbol{v} \in L \subseteq E$ we have $\langle \pi_E\boldsymbol{k}, \boldsymbol{v}\rangle = \langle \boldsymbol{k}, \boldsymbol{v}\rangle \in \mathbb{Z}$, so $\pi_E(\mathbb{Z}^{dm}) \subseteq L^*$. Conversely let $\boldsymbol{y} \in L^*$. Since L is primitive, a $\mathbb{Z}$-basis $\boldsymbol{b}_1, \dots, \boldsymbol{b}_{N'}$ of L extends to a basis $\boldsymbol{b}_1, \dots, \boldsymbol{b}_{dm}$ of $\mathbb{Z}^{dm}$. Let $\boldsymbol{b}_1^*, \dots, \boldsymbol{b}_{dm}^*$ be the dual basis, which is again a basis of $\mathbb{Z}^{dm}$. Put $\boldsymbol{k} := \sum_{i\le N'} \langle \boldsymbol{y}, \boldsymbol{b}_i\rangle \boldsymbol{b}_i^* \in \mathbb{Z}^{dm}$. For $\boldsymbol{v} = \sum_{i\le N'} c_i\boldsymbol{b}_i \in L$ we get $\langle \boldsymbol{k}, \boldsymbol{v}\rangle = \sum_i c_i \langle \boldsymbol{y}, \boldsymbol{b}_i\rangle = \langle \boldsymbol{y}, \boldsymbol{v}\rangle$, so $\boldsymbol{k} - \boldsymbol{y}$ is orthogonal to L, hence to E, and $\pi_E\boldsymbol{k} = \pi_E\boldsymbol{y} = \boldsymbol{y}$. The second description follows from $E^\perp = \boldsymbol{A}_{\boldsymbol{X}}^T\mathbb{R}^{dr}$. □

In canonical coordinates put $\boldsymbol{T} := \boldsymbol{B}_K^{\oplus m}$ and $F := \boldsymbol{T}E = \operatorname{span} L_{\boldsymbol{X}} = \ker \widetilde{\Phi}(\boldsymbol{X})$. Writing π_F for the orthogonal projection onto F, the intrinsic dual is

$$L_{\boldsymbol{X}}^* = \pi_F\big(\boldsymbol{T}^{-T} L^*\big).$$

Indeed, the restriction $\boldsymbol{T}|_E : E \to F$ is an isomorphism, and its inverse adjoint is $\pi_F \boldsymbol{T}^{-T}|_E$. For $\boldsymbol{u} \in E$ and $\boldsymbol{v} \in E$, $\left\langle \pi_F \boldsymbol{T}^{-T}\boldsymbol{u}, \boldsymbol{T}\boldsymbol{v} \right\rangle = \langle \boldsymbol{u}, \boldsymbol{v} \rangle$, which identifies the two intrinsic dual lattices. The projection is necessary because the ambient inverse transpose need not map E into F.

4.2 The Gaussian-heuristic scale and a lower bound

For a lattice L of rank N put

$$\operatorname{gh}(L) := \sqrt{\frac{N}{2\pi e}}\, \det(L)^{1/N},$$

the Gaussian-heuristic prediction for $\lambda_1(L)$. Note that $\operatorname{gh}(L^*)\operatorname{gh}(L) = N/(2\pi e)$.

Proposition 4.3 (Lower bound on the smoothing parameter). *For every lattice L of rank $N \geq 2$ and every $\varepsilon \in (0,1)$,*

$$\eta_\varepsilon(L) \;\geq\; \sqrt{\frac{\ln(2/\varepsilon)}{\pi}}\, \frac{1}{\lambda_1(L^*)} \;\geq\; 0.7\sqrt{\frac{\ln(1/\varepsilon)}{N}}\, \det(L)^{1/N}.$$

Proof. If $\boldsymbol{y}$ is a shortest nonzero dual vector then $\pm\boldsymbol{y}$ contribute $2e^{-\pi t^2 \lambda_1(L^*)^2}$ to $\rho_{1/t}(L^* \setminus \{\boldsymbol{0}\})$. At $t = \eta_\varepsilon(L)$ this sum is at most ε, which gives the first inequality. Minkowski's theorem gives $\lambda_1(L^*) \leq 2V_N^{-1/N} \det(L^*)^{1/N} = 2V_N^{-1/N} \det(L)^{-1/N}$, where $V_N = \pi^{N/2}/\Gamma(N/2+1)$ is the volume of the unit ball. We have $\Gamma(x+1) \leq x^x$ for every $x \in \{1, \frac{3}{2}, 2, \frac{5}{2}, \dots\}$. This holds at $x = 1$ and at $x = \frac{3}{2}$, where $\Gamma(\frac{5}{2}) = \frac{3}{4}\sqrt{\pi} < 1.33 < (\frac{3}{2})^{3/2}$, and it passes from x to $x+1$ because $\Gamma(x+2) = (x+1)\Gamma(x+1) \leq (x+1)x^x \leq (x+1)^{x+1}$. With $x = N/2$ this gives $V_N \geq (2\pi/N)^{N/2}$, and therefore, $V_N^{1/N}\sqrt{N} \geq \sqrt{2\pi} > 2.5$ for all $N \geq 2$. Hence $\eta_\varepsilon(L) \geq \sqrt{\ln(2/\varepsilon)/\pi} \cdot \frac{2.5}{2\sqrt{N}} \det(L)^{1/N} \geq 0.7\sqrt{\ln(1/\varepsilon)/N}\, \det(L)^{1/N}$. ∎

Corollary 4.4 (Natural scale of the kernel). *Assume $N = d(m-r) \geq 2$ and $0 < \varepsilon < 1$. On the surjectivity event,*

$$\eta_\varepsilon(L_{\boldsymbol{X}}) \;\geq\; 0.7\sqrt{\frac{\ln(1/\varepsilon)}{N}}\, \operatorname{rd}(K)^{1/2}\, s_{\min}(\widetilde{\Phi}(\boldsymbol{X}))^{r/n}.$$

For power-of-two cyclotomics, let $r \geq 1$, $\boldsymbol{X} \leftarrow D^m_{\mathcal{O}_K^r, s}$, $s/\sqrt{d} \geq 1$, and $0 < \delta < 1$. If $m \geq c_{\mathrm{sp}}(r + \ln(2d/\delta))$, then there is a spectral event $\mathcal{S}$ of probability at least $1-\delta$ such that every $\boldsymbol{X} \in \mathcal{S}$ with $\boldsymbol{X}\mathcal{O}_K^m = \mathcal{O}_K^r$ satisfies

$$\eta_{2^{-N}}(L_{\boldsymbol{X}}) \geq c\sqrt{d}\,(s\sqrt{m})^{r/n}$$

for an absolute constant $c > 0$. Here c_{sp} may be enlarged to be at least 2. This is an implication on $\mathcal{S}$, not a conditional probability estimate after conditioning on integral surjectivity. If integral surjectivity fails with probability at most δ_{surj}, then surjectivity and this lower bound hold jointly with probability at least $1 - \delta - \delta_{\mathrm{surj}}$.

Proof. Combine Proposition 4.3 with $\det(\widetilde{\Phi}(\boldsymbol{X})\widetilde{\Phi}(\boldsymbol{X})^T)^{1/2} \geq s_{\min}(\widetilde{\Phi}(\boldsymbol{X}))^{dr}$ in Proposition 4.1. For the spectral assertion take $\mathcal{S}$ from Lemma 3.8. On its intersection with the integral-surjectivity event, substitute $\operatorname{rd}(K) = d$, $\varepsilon = 2^{-N}$ and $s_{\min}(\widetilde{\Phi}(\boldsymbol{X})) \geq c'_{\mathrm{sp}} s\sqrt{m}$ into the first assertion. The column condition gives $m \geq 2r$, hence $0 < r/n \leq 1$, so $(c'_{\mathrm{sp}})^{r/n} \geq \min\{1, c'_{\mathrm{sp}}\}$. Thus one may take $c = 0.7\sqrt{\ln 2}\min\{1, c'_{\mathrm{sp}}\}$. The joint probability estimate is a union bound and does not require independence. ∎

Proposition 5.14 below gives the second lower bound

$$\eta_\varepsilon(L_{\boldsymbol{X}}) \geq \beta_K^{-1}\sqrt{\ln(2/\varepsilon)/\pi} \qquad (m > r),$$

which is independent of s and m. The two lower bounds govern different regimes: the projected-dual bound is the stronger one in the wide regime, where $\det(L_{\boldsymbol{X}})^{1/N}$ is bounded, and the natural-scale bound is the stronger one in the narrow regime, where it grows with s.

4.3 Two regimes

Corollary 4.4 and the projected-dual bound of Proposition 5.14 delimit two regimes, drawn in Figure 1.

Wide regime. If $m \geq C\, dr \ln(srm)$ then $(s\sqrt{m})^{r/n} = e^{O(1/C)}$, so the natural scale is $\Theta(\mathrm{rd}(K)^{1/2})$, and the Minkowski lower bound of Proposition 4.3 is not sharp here. Proposition 5.14 shows that $L^*_{\boldsymbol{X}}$ always contains a vector of norm at most β_K, so that $\eta_{2^{-\lambda}}(L_{\boldsymbol{X}}) \geq \beta_K^{-1}\sqrt{\lambda \ln 2/\pi}$ for every λ. The geometric theorems of Section 5 match this lower bound up to an absolute constant: $\eta_{2^{-\lambda}}(L_{\boldsymbol{X}}) = \Theta_K(\sqrt{\lambda})$ once $m \gtrsim dr \ln(srm)$ and $\lambda \asymp dr$ (Corollary 5.15). In coefficient coordinates the picture is as follows. The projected standard basis gives $\lambda_1(L^*) \leq 1$, and the theta estimate gives a matching lower bound of the form $c/(\kappa_K \mu_K)$ (Remark 5.16). The Gaussian-heuristic prediction $\lambda_1(L^*) \approx \mathrm{gh}(L^*) = \Theta(\sqrt{N})$ therefore fails by a factor of order $\sqrt{N}$.

Narrow regime. Fix $m > r$ with $N \geq 2$ and let $s \to \infty$. For integrally surjective matrices whose nonzero singular values are all of order $s\sqrt{m}$, Corollary 4.4 gives $\eta_\varepsilon(L_{\boldsymbol{X}}) \geq c\, s^{r/n}$ for fixed $0 < \varepsilon < 1$. Such kernels have covolume of order $(s\sqrt{m})^{dr}$ concentrated in $N = nd$ dimensions. The geometric route of Section 5 does not reach this regime, and neither does any union-bound argument in the style of [AGHS13, KNSW20]. We leave it open (Remark 4.5 and Section 1.4).

Remark 4.5 (Why union bounds cannot reach the narrow regime). The dual-minimum method of [AGHS13, KNSW20], transported to coefficient coordinates, certifies a lower bound on $\lambda_1(L^*)$ through an auxiliary construction-A lattice $\boldsymbol{A}_{\boldsymbol{X}}^T\mathbb{Z}^{dr} + q\mathbb{Z}^{dm}$. Every such argument approximates a dual vector $\boldsymbol{u} = \boldsymbol{k} + \boldsymbol{A}_{\boldsymbol{X}}^T\boldsymbol{y}$ by rounding $\boldsymbol{y}$ to $q^{-1}\mathbb{Z}^{dr}$. The rounding error is of order $s_{\min}(\boldsymbol{A}_{\boldsymbol{X}})/q$ in the worst direction. The argument then applies a union bound over the q^{dr} residues $\boldsymbol{z} \bmod q$, with a per-column anti-concentration probability $p < 1$. These two steps pull in opposite directions. The rounding error is negligible only if $q \gtrsim s\sqrt{m}/\rho$, where ρ is the target lower bound on $\lambda_1(L^*_{\boldsymbol{X}})$, so q must be large, while the union bound needs $m \geq dr \ln q/\ln(1/p)$, so $\ln q$ must be at most a constant times $m/(dr)$. In the narrow regime the natural target is $\rho \asymp \sqrt{N}(s\sqrt{m})^{-r/n}$, so $\ln q \gtrsim (1 + r/n)\ln s$, which is incompatible with a fixed m as $s \to \infty$. The obstruction is not an artifact of a specific lemma. For fixed m the number q^{dr} of residues to be excluded grows polynomially in s while only m independent witnesses are available, and the dual lattice genuinely has vectors of norm about $s^{-r/n}$. The theta integral of Section 5 is subject to the same obstruction in a different guise: its remainder is bounded by $W^R(7/8)^m$ with W polynomial in s, so it too needs $m \gtrsim dr \ln s$ columns. The narrow regime therefore requires a first-moment analysis of the dual kernel ensemble, which this paper does not provide (Section 1.4).

Remark 4.6 (A refuted conjecture). An earlier version of this paper conjectured that for power-of-two cyclotomics, in the wide regime, $\lambda_1(L^*_{\boldsymbol{X}}) \geq c_0\, \mathrm{gh}(L^*_{\boldsymbol{X}})$ with overwhelming probability, or equivalently $\eta_{2^{-N}}(L_{\boldsymbol{X}}) \leq C\sqrt{d}$. This is false. By Proposition 5.14 the dual kernel contains a vector of norm at most $\beta_K = 1/\sqrt{d}$ for every $\boldsymbol{X}$, so $\lambda_1(L^*_{\boldsymbol{X}}) = O(\mathrm{gh}(L^*_{\boldsymbol{X}})/\sqrt{N})$ and $\eta_{2^{-N}}(L_{\boldsymbol{X}}) \geq \sqrt{dN \ln 2/\pi}$ deterministically. The correct statement is that $\eta_{2^{-\lambda}}(L_{\boldsymbol{X}}) = \Theta_K(\sqrt{\lambda})$ in the wide regime (Corollary 5.15), and that the first minimum of the dual coefficient kernel lies in an interval $[c/(\kappa_K\mu_K), 1]$ with an absolute constant c on the event of Theorem 5.6 (Remark 5.16). Only the constants remain open.

5 Kernel Smoothing from an Exact Theta Integral

The dual-minimum method of Agrawal, Gentry, Halevi and Sahai [AGHS13] and of Kirshanova, Nguyen, Stehlé and Wallet [KNSW20], developed over $\mathbb{Z}$, proceeds in two steps. It first certifies a lower bound on the first minimum of the dual kernel by a union bound over the residues of an auxiliary construction-A lattice, and it then converts that bound into a smoothing bound through the transference lemmas of Banaszczyk [Ban93, Lemma 1.5] and Peikert [Pei08, Lemma 3.5]. The conversion counts dual vectors in shells, and it costs a factor of order $\sqrt{(m-r)/r}$ at error $2^{-\lambda}$, $\lambda \asymp dr$, unless the error is as small as e^{-N} (Remark 5.17). Surjectivity has to be established separately, for instance through the generation theorem of Aggarwal and Regev [AR16], and the anti-concentration step needs a coefficient width of order $\sqrt{R}$. This section takes a more direct route. An exact integral identity expresses the full dual Gaussian mass of the coefficient kernel, multiplied by the index of the image of $\boldsymbol{A}_{\boldsymbol{X}}$, as a torus integral of one-dimensional periodic Gaussian sums. A Gaussian integral near the origin gives a main term whose coefficient is exactly one, and the independence of the m ring columns controls the remainder. Because an

integer index below two equals one, the same estimate proves surjectivity and smoothing together. The resulting bound is $O_K(\sqrt{\lambda})$ at error $2^{-\lambda}$ for polynomial width and growing rank. A deterministic lower bound shows that this order cannot be improved. For power-of-two cyclotomics the width can be lowered to coefficient width one. The finite theorems and their complete proofs are in Sections 5.3 and 5.4.

Throughout this section $\boldsymbol{X} \leftarrow D^m_{\mathcal{O}_K^r,s}$, $R = dr$, $M = dm$, and $\lambda \geq 1$ is an accuracy parameter. Total variation distance and statistical distance are the same quantity, written SD. We use the coefficient matrix $\boldsymbol{A}_{\boldsymbol{X}} \in \mathbb{Z}^{R\times M}$ of Section 3.2 and the two identities

$$L_{\boldsymbol{X}} = \boldsymbol{B}_K^{\oplus m} \ker_{\mathbb{Z}} \boldsymbol{A}_{\boldsymbol{X}}, \qquad \widetilde{\Phi}(\boldsymbol{X}) = \boldsymbol{B}_K^{\oplus r}\, \boldsymbol{A}_{\boldsymbol{X}}\, (\boldsymbol{B}_K^{-1})^{\oplus m}, \tag{20}$$

the first of which is Equation (10). The second holds because both sides are the matrix of $\boldsymbol{y} \mapsto \boldsymbol{X}\boldsymbol{y}$ in canonical coordinates: passing to coefficient coordinates costs $(\boldsymbol{B}_K^{-1})^{\oplus m}$, the coefficient matrix acts, and $\boldsymbol{B}_K^{\oplus r}$ returns to canonical coordinates.

5.1 The image index and the exact theta integral

Let $\boldsymbol{A} \in \mathbb{Z}^{R\times M}$ be any integer matrix of full row rank and put

$$L_{\boldsymbol{A}} := \ker_{\mathbb{Z}} \boldsymbol{A}, \qquad D_{\boldsymbol{A}} := \det(\boldsymbol{A}\boldsymbol{A}^T)^{1/2}, \qquad I_{\boldsymbol{A}} := [\mathbb{Z}^R : \boldsymbol{A}\mathbb{Z}^M].$$

The positive integer $I_{\boldsymbol{A}}$ is the *image index*. Full real row rank does not force $I_{\boldsymbol{A}} = 1$, as the matrix $(2\ \ 0)$ shows. For $t > 0$ let

$$\varphi_t(u) := \sum_{k\in\mathbb{Z}} e^{-\pi t^2 (k-u)^2}, \qquad Q_R := [-\tfrac{1}{2}, \tfrac{1}{2}]^R,$$

so that φ_t is one-periodic and integration over Q_R represents integration over $\mathbb{R}^R/\mathbb{Z}^R$ with total mass one. Recall that $\rho_{1/t}(L^*_{\boldsymbol{A}}) = \sum_{\boldsymbol{v}\in L^*_{\boldsymbol{A}}} e^{-\pi t^2 ||\boldsymbol{v}||^2}$, so that $\eta_\varepsilon(L_{\boldsymbol{A}}) \leq t$ exactly when $\rho_{1/t}(L^*_{\boldsymbol{A}}) - 1 \leq \varepsilon$.

Proposition 5.1 (Exact theta integral). *For every full-row-rank $\boldsymbol{A} \in \mathbb{Z}^{R\times M}$ and every $t > 0$,*

$$I_{\boldsymbol{A}}\, \rho_{1/t}(L^*_{\boldsymbol{A}}) \;=\; t^R D_{\boldsymbol{A}} \int_{Q_R} \prod_{\ell=1}^{M} \varphi_t\big((\boldsymbol{A}^T\boldsymbol{y})_\ell\big)\, d\boldsymbol{y}. \tag{21}$$

Proof. Let $V = \boldsymbol{A}^T\mathbb{R}^R$ be the row space and let $\boldsymbol{P}$ be the orthogonal projection onto $V^\perp = \ker \boldsymbol{A}$. The row lattice is $R_0 = \boldsymbol{A}^T\mathbb{Z}^R$ and its saturation is $R_{\mathrm{sat}} = V \cap \mathbb{Z}^M$. Since $L_{\boldsymbol{A}}$ is a primitive sublattice of $\mathbb{Z}^M$, the proof of Lemma 4.2, which uses only primitivity, gives $L^*_{\boldsymbol{A}} = \boldsymbol{P}(\mathbb{Z}^M)$.

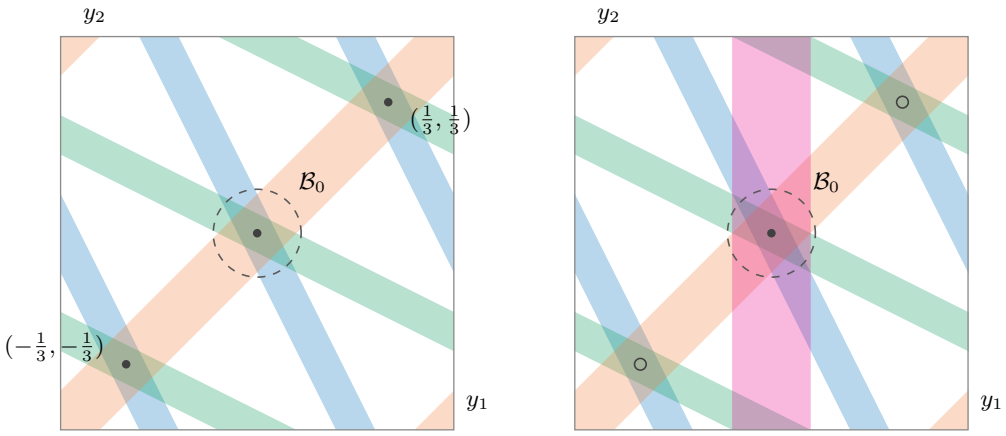
Expanding the product in Equation (21) gives $\sum_{\boldsymbol{k}\in\mathbb{Z}^M} e^{-\pi t^2 \left|\left|\boldsymbol{k}-\boldsymbol{A}^T\boldsymbol{y}\right|\right|^2}$. Because $\boldsymbol{A}^T$ is injective, the classes of $\mathbb{Z}^M$ modulo R_0 are the sets $\boldsymbol{k}_0 + \boldsymbol{A}^T\mathbb{Z}^R$, and for $\boldsymbol{k} = \boldsymbol{k}_0 + \boldsymbol{A}^T\boldsymbol{z}$ we have $\left|\left|\boldsymbol{k} - \boldsymbol{A}^T\boldsymbol{y}\right|\right| = \left|\left|\boldsymbol{k}_0 - \boldsymbol{A}^T(\boldsymbol{y}-\boldsymbol{z})\right|\right|$. As $\boldsymbol{z}$ ranges over $\mathbb{Z}^R$ and $\boldsymbol{y}$ over Q_R, the difference $\boldsymbol{y}-\boldsymbol{z}$ ranges over $\mathbb{R}^R$, so the contribution of one class unfolds to $\int_{\mathbb{R}^R} e^{-\pi t^2 \left|\left|\boldsymbol{k}_0-\boldsymbol{A}^T\boldsymbol{y}\right|\right|^2} d\boldsymbol{y}$. Orthogonal decomposition gives $\left|\left|\boldsymbol{k}_0 - \boldsymbol{A}^T\boldsymbol{y}\right|\right|^2 = ||\boldsymbol{P}\boldsymbol{k}_0||^2 + \left|\left|(\boldsymbol{I}-\boldsymbol{P})\boldsymbol{k}_0 - \boldsymbol{A}^T\boldsymbol{y}\right|\right|^2$, and $\boldsymbol{y} \mapsto \boldsymbol{A}^T\boldsymbol{y}$ is a linear bijection of $\mathbb{R}^R$ onto V with Jacobian $D_{\boldsymbol{A}}$. Hence

$$\int_{\mathbb{R}^R} e^{-\pi t^2 \left|\left|\boldsymbol{k}_0-\boldsymbol{A}^T\boldsymbol{y}\right|\right|^2} d\boldsymbol{y} = \frac{e^{-\pi t^2 ||\boldsymbol{P}\boldsymbol{k}_0||^2}}{t^R D_{\boldsymbol{A}}}.$$

The vector $\boldsymbol{P}\boldsymbol{k}_0$ depends on $\boldsymbol{k}_0$ only modulo R_{sat}, and $\boldsymbol{k}_0 \mapsto \boldsymbol{P}\boldsymbol{k}_0$ induces a bijection $\mathbb{Z}^M/R_{\mathrm{sat}} \to \boldsymbol{P}(\mathbb{Z}^M)$. Each class modulo R_{sat} consists of $[R_{\mathrm{sat}} : R_0]$ classes modulo R_0, and the Smith normal form of $\boldsymbol{A}$ shows $[R_{\mathrm{sat}} : R_0] = [\mathbb{Z}^R : \boldsymbol{A}\mathbb{Z}^M] = I_{\boldsymbol{A}}$: if $\boldsymbol{A} = \boldsymbol{U}\,(\mathrm{diag}(d_1,\dots,d_R) \mid \boldsymbol{0})\,\boldsymbol{V}$ with unimodular $\boldsymbol{U},\boldsymbol{V}$, both indices equal $d_1\cdots d_R$. Therefore every vector of $L^*_{\boldsymbol{A}}$ occurs exactly $I_{\boldsymbol{A}}$ times among the classes modulo R_0. All summands are nonnegative, so Tonelli's theorem justifies the exchange of sum and integral, and the right-hand side of Equation (21) equals $I_{\boldsymbol{A}} \sum_{\boldsymbol{v}\in L^*_{\boldsymbol{A}}} e^{-\pi t^2||\boldsymbol{v}||^2}$. □

Expanding each φ_t in its Fourier series, $\varphi_t(u) = t^{-1}\sum_{k\in\mathbb{Z}} e^{-\pi k^2/t^2} e^{2\pi i k u}$, and integrating over the torus selects exactly the $\boldsymbol{z} \in \mathbb{Z}^M$ with $\boldsymbol{A}\boldsymbol{z} = \boldsymbol{0}$. Thus Equation (21) is Poisson summation on the kernel,

written so that the index of the image stays visible. The Fourier expansion also shows in passing that $\varphi_t(u) \le \varphi_t(0)$, since all of its coefficients are positive. The index is useful rather than inconvenient. Suppose the right-hand side of Equation (21) is at most $1+\varepsilon < 2$ for some $\varepsilon \in (0,1)$. The zero vector alone contributes one to $\rho_{1/t}(L_{\boldsymbol{A}}^*)$, so $\rho_{1/t}(L_{\boldsymbol{A}}^*) \ge 1$, and the positive integer $I_{\boldsymbol{A}}$ is then below two. Hence $I_{\boldsymbol{A}} = 1$, which says $\boldsymbol{A}\mathbb{Z}^M = \mathbb{Z}^R$. With $I_{\boldsymbol{A}} = 1$ the left-hand side is $\rho_{1/t}(L_{\boldsymbol{A}}^*)$ itself, so the nonzero dual vectors carry mass $\rho_{1/t}(L_{\boldsymbol{A}}^* \setminus \{\boldsymbol{0}\}) \le \varepsilon$, which says $\eta_\varepsilon(L_{\boldsymbol{A}}) \le t$. Figure 6 draws the integrand for $R = 2$: it is a product of periodic stripe patterns, one per column, and its spikes are the points of $(\boldsymbol{A}\mathbb{Z}^M)^*/\mathbb{Z}^R$, of which there are exactly $I_{\boldsymbol{A}}$.



(a) $\boldsymbol{A} = \left(\begin{smallmatrix} 2 & 1 & 1 \\ 1 & 2 & -1 \end{smallmatrix}\right)$, image index $I_{\boldsymbol{A}} = 3$ (b) $\boldsymbol{A} = \left(\begin{smallmatrix} 2 & 1 & 1 & 1 \\ 1 & 2 & -1 & 0 \end{smallmatrix}\right)$, image index $I_{\boldsymbol{A}} = 1$

Fig. 6. The integrand of the exact theta integral Equation (21) for $R = 2$, drawn to scale in the torus Q_R with $t = 10$. Each column $\boldsymbol{a}_\ell$ contributes the factor $\varphi_t(\langle \boldsymbol{a}_\ell, \boldsymbol{y}\rangle)$, which is close to one only in the stripes around the lines $\langle \boldsymbol{a}_\ell, \boldsymbol{y}\rangle \in \mathbb{Z}$. Blue, green, orange and pink belong to the columns $(2,1)$, $(1,2)$, $(1,-1)$ and $(1,0)$, in this order. The product is concentrated at the points where all stripes meet, and the torus contains exactly $I_{\boldsymbol{A}}$ such spikes, marked by filled dots. In (a) the three columns leave three spikes. In (b) the fourth column, in pink, has stripes that miss $\pm(\frac{1}{3},\frac{1}{3})$ and brings the index down to one. The open circles mark these two points, which are no longer spikes. The dashed circle is the ball $\mathcal{B}_0$ of Lemma 5.2.

5.2 The main term and the remainder

Lemma 5.2 (The Gaussian main term is exactly normalized). *Suppose every column of $\boldsymbol{A}$ has Euclidean norm at most $B_0 \ge 1$ and let $t \ge 1$. For $|u| \le 1/4$,*

$$\varphi_t(u) \le e^{-\pi t^2 u^2}\left(1 + 3e^{-\pi t^2/2}\right), \tag{22}$$

and the contribution of the ball $\mathcal{B}_0 := \{\boldsymbol{y} : ||\boldsymbol{y}|| \le 1/(4B_0)\} \subseteq Q_R$ to the right-hand side of Equation (21) is at most

$$\left(1 + 3e^{-\pi t^2/2}\right)^M t^R D_{\boldsymbol{A}} \int_{\mathbb{R}^R} e^{-\pi t^2 ||\boldsymbol{A}^T \boldsymbol{y}||^2} d\boldsymbol{y} = \left(1 + 3e^{-\pi t^2/2}\right)^M.$$

Proof. For $|u| \le 1/4$ and $k \ne 0$ one has $(k-u)^2 - u^2 = k^2 - 2ku \ge k^2 - |k|/2$, and the exponents $k^2 - |k|/2$ for $|k| = 1, 2, 3, \ldots$ increase by at least $5/2$ at each step. Separating $k = 0$,

$$\varphi_t(u) \le e^{-\pi t^2 u^2}\left(1 + \frac{2e^{-\pi t^2/2}}{1 - e^{-3\pi t^2/2}}\right) \le e^{-\pi t^2 u^2}\left(1 + 3e^{-\pi t^2/2}\right)$$

because $2/(1 - e^{-3\pi/2}) < 3$. If $||\boldsymbol{y}|| \le 1/(4B_0)$, every coordinate of $\boldsymbol{A}^T\boldsymbol{y}$ has absolute value at most $1/4$, so Equation (22) applies to every factor. The ball integral is bounded by the whole-space integral, and $\int_{\mathbb{R}^R} e^{-\pi t^2 \boldsymbol{y}^T \boldsymbol{A}\boldsymbol{A}^T \boldsymbol{y}} d\boldsymbol{y} = \det(t^2 \boldsymbol{A}\boldsymbol{A}^T)^{-1/2} = t^{-R} D_{\boldsymbol{A}}^{-1}$. □

The ball integral was bounded by a whole-space Gaussian integral, not declared equal to it. The point is that the whole-space normalization is exact however badly conditioned $\boldsymbol{A}$ may be, as long as it has full row rank. No approximation of $D_{\boldsymbol{A}}$ enters the main term, and no fundamental-cell error of the kind that appears when a lattice sum is compared with a continuous integral is incurred. The following three lemmas control the rest of the torus at polynomial width.

Lemma 5.3 (Elementary Gaussian estimates). *Let $\boldsymbol{x} \leftarrow D_{\mathbb{Z}^R,\boldsymbol{S}}$ with $\boldsymbol{S}$ positive definite symmetric, and put $s_0 = s_{\min}(\boldsymbol{S})$, $s_1 = s_{\max}(\boldsymbol{S})$. For $\boldsymbol{v} \in \mathbb{Z}^R$, $\boldsymbol{h} \in \mathbb{R}^R$ and $h_0 > 0$,*

$$\mathrm{SD}(\boldsymbol{x}, \boldsymbol{x}+\boldsymbol{v}) \le \sqrt{\pi/2}\, \left|\left|\boldsymbol{S}^{-1}\boldsymbol{v}\right|\right|, \tag{23}$$

$$\mathbb{E}e^{\langle \boldsymbol{h}, \boldsymbol{x}\rangle} \le e^{||\boldsymbol{S}\boldsymbol{h}||^2/(4\pi)}, \tag{24}$$

$$\mathbb{P}[||\boldsymbol{x}|| > s_1 h_0] \le 2^{R/2} e^{-\pi h_0^2/2}. \tag{25}$$

For every proper real linear subspace $V \subset \mathbb{R}^R$, $\mathbb{P}[\boldsymbol{x} \in V] \le 2/s_0$, with no additional restriction on $s_0 > 0$.

Proof. For distributions P, Q with positive masses on the same countable set, the relative entropy is $D_{\mathrm{KL}}(P\|Q) = \sum_{\boldsymbol{z}} P(\boldsymbol{z}) \ln(P(\boldsymbol{z})/Q(\boldsymbol{z}))$ and Pinsker's inequality states $\mathrm{SD}(P,Q) \le \sqrt{D_{\mathrm{KL}}(P\|Q)/2}$. The law of $\boldsymbol{x}+\boldsymbol{v}$ is $D_{\mathbb{Z}^R,\boldsymbol{S},\boldsymbol{v}}$, and an integer translation preserves the Gaussian partition function, so the two normalizing constants agree. Direct expansion gives

$$\begin{aligned} D_{\mathrm{KL}}(\boldsymbol{x} \,\|\, \boldsymbol{x}+\boldsymbol{v}) &= \pi\, \mathbb{E}\left[\left|\left|\boldsymbol{S}^{-1}(\boldsymbol{x}-\boldsymbol{v})\right|\right|^2 - \left|\left|\boldsymbol{S}^{-1}\boldsymbol{x}\right|\right|^2\right] \\ &= \pi \left|\left|\boldsymbol{S}^{-1}\boldsymbol{v}\right|\right|^2 - 2\pi \left\langle \boldsymbol{S}^{-1}\mathbb{E}\boldsymbol{x}, \boldsymbol{S}^{-1}\boldsymbol{v}\right\rangle = \pi \left|\left|\boldsymbol{S}^{-1}\boldsymbol{v}\right|\right|^2, \end{aligned}$$

because the centered Gaussian has mean zero. This proves Equation (23).

For Equation (24), completing the square gives $\mathbb{E}e^{\langle \boldsymbol{h},\boldsymbol{x}\rangle} = e^{||\boldsymbol{S}\boldsymbol{h}||^2/(4\pi)} \rho_{\boldsymbol{S}}(\mathbb{Z}^R - \boldsymbol{c})/\rho_{\boldsymbol{S}}(\mathbb{Z}^R)$ with $\boldsymbol{c} = \boldsymbol{S}\boldsymbol{S}^T\boldsymbol{h}/(2\pi)$, and the ratio is at most one by Poisson summation, since the Fourier coefficients of a centered Gaussian are nonnegative. This is the computation in the proof of Lemma 3.7.

For Equation (25), let $\boldsymbol{g}$ be an independent continuous Gaussian vector on $\mathbb{R}^R$ with covariance $(\pi/s_1^2)\boldsymbol{I}_R$, so that $\mathbb{E}_{\boldsymbol{g}} e^{\langle \boldsymbol{g},\boldsymbol{x}\rangle} = e^{\pi||\boldsymbol{x}||^2/(2s_1^2)}$ for every fixed $\boldsymbol{x}$. Exchanging the expectations and applying Equation (24),

$$\mathbb{E}_{\boldsymbol{x}} e^{\pi||\boldsymbol{x}||^2/(2s_1^2)} = \mathbb{E}_{\boldsymbol{g}}\mathbb{E}_{\boldsymbol{x}} e^{\langle \boldsymbol{g},\boldsymbol{x}\rangle} \le \mathbb{E}_{\boldsymbol{g}} e^{||\boldsymbol{S}\boldsymbol{g}||^2/(4\pi)} \le \mathbb{E}_{\boldsymbol{g}} e^{s_1^2||\boldsymbol{g}||^2/(4\pi)} = 2^{R/2},$$

because $s_1^2 \left|\left|\boldsymbol{g}\right|\right|^2/(4\pi)$ is one quarter of a χ^2 variable with R degrees of freedom, whose exponential moment at $1/4$ is $(1-1/2)^{-R/2}$. Markov's inequality gives $\mathbb{P}[||\boldsymbol{x}|| > s_1 h_0] \le 2^{R/2} e^{-\pi h_0^2/2}$.

For the last assertion, choose a nonzero vector $\boldsymbol{h} \in V^\perp$ and a coordinate i with $h_i \neq 0$. The signed coordinate vector $\boldsymbol{v} = \mathrm{sgn}(h_i)\boldsymbol{e}_i \in \mathbb{Z}^R$ has $||\boldsymbol{v}|| = 1$ and $\langle \boldsymbol{h}, \boldsymbol{v}\rangle > 0$. For the half-space $\mathcal{H} := \{\boldsymbol{z} \in \mathbb{R}^R : \langle \boldsymbol{h}, \boldsymbol{z}\rangle \le 0\}$, the sets V and $\mathcal{H} - \boldsymbol{v}$ are disjoint and both contained in $\mathcal{H}$. Therefore

$$\begin{aligned} \mathbb{P}[\boldsymbol{x} \in V] &\le \mathbb{P}[\boldsymbol{x} \in \mathcal{H}] - \mathbb{P}[\boldsymbol{x}+\boldsymbol{v} \in \mathcal{H}] \\ &\le \mathrm{SD}(\boldsymbol{x}, \boldsymbol{x}+\boldsymbol{v}) \le \sqrt{\pi/2}\, \left|\left|\boldsymbol{S}^{-1}\boldsymbol{v}\right|\right| \le \frac{\sqrt{\pi/2}}{s_0} \le \frac{2}{s_0}, \end{aligned}$$

where we used Equation (23) and $\left|\left|\boldsymbol{S}^{-1}\right|\right|_{\mathrm{op}} = s_0^{-1}$. □

Lemma 5.4 (An integer displacement). *Let $0 < \tau \le 1/8$ and $\boldsymbol{0} \neq \boldsymbol{y} \in Q_R$. There is $\boldsymbol{v} \in \mathbb{Z}^R$ with*

$$2\tau \le \langle \boldsymbol{y}, \boldsymbol{v}\rangle \le 1 - 2\tau, \qquad ||\boldsymbol{v}|| \le \frac{2\tau}{||\boldsymbol{y}||} + \sqrt{R}. \tag{26}$$

In particular the sets $E_{\boldsymbol{y}} := \{\boldsymbol{x} \in \mathbb{Z}^R : \mathrm{dist}(\langle \boldsymbol{y}, \boldsymbol{x}\rangle, \mathbb{Z}) < \tau\}$ and $E_{\boldsymbol{y}} + \boldsymbol{v}$ are disjoint.

Proof. If $|y_i| \ge 2\tau$ for some i, take $\boldsymbol{v} = \mathrm{sgn}(y_i)\boldsymbol{e}_i$. Then $\langle \boldsymbol{y}, \boldsymbol{v}\rangle = |y_i| \in [2\tau, 1/2]$, which lies in $[2\tau, 1-2\tau]$ because $\tau \le 1/8$, and $||\boldsymbol{v}|| = 1 \le \sqrt{R}$. Otherwise all $|y_i| < 2\tau$. Put $c = 2\tau/\left|\left|\boldsymbol{y}\right|\right|^2$ and initially $v_i = \mathrm{sgn}(y_i)\lfloor c\,|y_i| \rfloor$. Then $\langle \boldsymbol{y}, \boldsymbol{v}\rangle \ge \sum_i |y_i|\,(c\,|y_i| - 1) = 2\tau - ||\boldsymbol{y}||_1$, so the deficit below 2τ is at most $||\boldsymbol{y}||_1$. Increase the coordinates one at a time by one unit in the sign of y_i, each coordinate at most once,

until the inner product reaches 2τ. This is possible because the total available increase is $||\boldsymbol{y}||_1$, which covers the deficit. Each step adds $|y_i| < 2\tau$, so the final value satisfies $2\tau \le \langle \boldsymbol{y}, \boldsymbol{v} \rangle < 4\tau \le 1 - 2\tau$. The final absolute coordinates are at most $c\,|y_i| + 1$, which gives $||\boldsymbol{v}|| \le c\,||\boldsymbol{y}|| + \sqrt{R} = 2\tau/\,||\boldsymbol{y}|| + \sqrt{R}$. If $\boldsymbol{x} \in E_{\boldsymbol{y}} \cap (E_{\boldsymbol{y}} + \boldsymbol{v})$, then $\langle \boldsymbol{y}, \boldsymbol{x} \rangle$ and $\langle \boldsymbol{y}, \boldsymbol{x} - \boldsymbol{v} \rangle$ are both within τ of integers, so $\langle \boldsymbol{y}, \boldsymbol{v} \rangle$ is within 2τ of an integer. This contradicts $2\tau \le \langle \boldsymbol{y}, \boldsymbol{v} \rangle \le 1 - 2\tau$. □

For the third lemma we fix the parameters of the polynomial-width theorem. Put

$$\begin{aligned} \delta &:= 2^{-\lambda-4}, \qquad H := R + \ln \frac{2m}{\delta}, \qquad U := s_+ \sqrt{H}, \\ t &:= 64\,\kappa_K \mu_K \sqrt{H}, \qquad W := t\,\mu_K\, U \sqrt{M/R}, \end{aligned} \tag{27}$$

and assume

$$s_- \ge 8\sqrt{R}, \qquad m \ln \tfrac{8}{7} \;\ge\; R \ln W + 2 \ln \tfrac{1}{\delta}. \tag{28}$$

Here $s_- = s/\alpha_K$ and $s_+ = s\beta_K$ are the extreme widths of the coefficient Gaussian Equation (11), so U is the column-norm threshold of the tail event below and $s_+/s_- = \kappa_K$.

Lemma 5.5 (Expected contribution of one ring column). *Assume Equation (28). For every $\boldsymbol{y} \in Q_R$ with $||\boldsymbol{y}|| > 1/(4\mu_K U)$,*

$$\mathbb{E}_{\boldsymbol{x}_j} \prod_{k=1}^{d} \varphi_t\big(\langle \boldsymbol{y}, (\boldsymbol{I}_r \otimes \boldsymbol{T}_k) c_r(\boldsymbol{x}_j) \rangle\big) \le \frac{7}{8}. \tag{29}$$

Proof. Since $t = 64\kappa_K \mu_K \sqrt{H} \ge 64$, we may take $\tau = 1/t \le 1/8$ in Lemma 5.4 and obtain $\boldsymbol{v} \in \mathbb{Z}^R$ with Equation (26). By Equation (27) and Equation (28),

$$||\boldsymbol{v}|| \le \frac{2\tau}{||\boldsymbol{y}||} + \sqrt{R} \le \frac{8\mu_K U}{t} + \sqrt{R} = \frac{s_+}{8\kappa_K} + \sqrt{R} \le \frac{s_-}{8} + \frac{s_-}{8} = \frac{s_-}{4},$$

using $s_+/\kappa_K = s_-$. The coefficient column $c_r(\boldsymbol{x}_j)$ has law $D_{\mathbb{Z}^R, \boldsymbol{S}}$ with $s_{\min}(\boldsymbol{S}) = s_-$, so Equation (23) gives $\mathrm{SD}(c_r(\boldsymbol{x}_j), c_r(\boldsymbol{x}_j) + \boldsymbol{v}) \le \sqrt{\pi/2}/4 < 1/2$. Because $E_{\boldsymbol{y}}$ and $E_{\boldsymbol{y}} + \boldsymbol{v}$ are disjoint, $\mathbb{P}[c_r(\boldsymbol{x}_j) \in E_{\boldsymbol{y}}] + \mathbb{P}[c_r(\boldsymbol{x}_j) \in E_{\boldsymbol{y}} + \boldsymbol{v}] \le 1$, and the second probability is at least the first minus the statistical distance. Hence

$$\mathbb{P}[c_r(\boldsymbol{x}_j) \notin E_{\boldsymbol{y}}] \;\ge\; \tfrac{1}{2}\big(1 - \mathrm{SD}\big) \;\ge\; \tfrac{1}{4}.$$

For $t \ge 8$ the elementary bounds

$$\begin{aligned} \varphi_t(0) &\le 1 + \frac{2e^{-\pi t^2}}{1 - e^{-3\pi t^2}} \le 1 + 3e^{-\pi t^2}, \\ \mathrm{dist}(u, \mathbb{Z}) \ge \frac{1}{t} \;&\Longrightarrow\; \varphi_t(u) \le e^{-\pi} + \frac{2e^{-\pi t^2/4}}{1 - e^{-2\pi t^2}} < \frac{1}{8} \end{aligned}$$

follow by separating the nearest integer and bounding the distances to the other integers from below by $k - 1/2$ for $k \ge 1$. Every factor in Equation (29) is at most $\varphi_t(0)$, since $\varphi_t(u) \le \varphi_t(0)$ by the Fourier expansion of φ_t. On the event $c_r(\boldsymbol{x}_j) \notin E_{\boldsymbol{y}}$ the factor $k = 1$ is at most $1/8$, because $\boldsymbol{T}_1 = \boldsymbol{I}_d$ makes it $\varphi_t(\langle \boldsymbol{y}, c_r(\boldsymbol{x}_j) \rangle)$. Therefore the expectation in Equation (29) is at most $\varphi_t(0)^d\big(1 - \frac{7}{8}\mathbb{P}[c_r(\boldsymbol{x}_j) \notin E_{\boldsymbol{y}}]\big) \le \frac{25}{32}\varphi_t(0)^d$. Finally $d \le R \le t^2$ and $t \ge 8$ give $\varphi_t(0)^d \le \exp(3t^2 e^{-\pi t^2}) < 28/25$, and $\frac{25}{32} \cdot \frac{28}{25} = \frac{7}{8}$. No independence within a ring column was used. □

The d coordinates that one ring column contributes to $\boldsymbol{A}_{\boldsymbol{X}}^T \boldsymbol{y}$ are the inner products in Equation (29), by Equation (9). They may be dependent. The proof used the witness coordinate $k = 1$, for which $\boldsymbol{T}_1 = \boldsymbol{I}_d$, to obtain a saving, and bounded the other $d - 1$ factors by $\varphi_t(0)$. Independence is used only between the m ring columns. The displacement of Lemma 5.4 moves all R coordinates at once and therefore has norm up to $\sqrt{R}$. The translation bound Equation (23) is only useful when this norm is small compared with the width, so it needs $s_- \gtrsim \sqrt{R}$. This is where the width condition in Equation (28) comes from. Geometrically, $E_{\boldsymbol{y}}$ consists of the integer points in the strips of thickness $2\tau/\,||\boldsymbol{y}||$ around the hyperplanes $\langle \boldsymbol{y}, \boldsymbol{x} \rangle \in \mathbb{Z}$, which are spaced $1/\,||\boldsymbol{y}||$ apart, and the displacement moves these strips into the gaps between them.

5.3 The finite theorem at polynomial width

Theorem 5.6 (Finite geometric smoothing bound). *Let K, the integral basis, and the constants $\alpha_K, \beta_K, \kappa_K, \mu_K$ be as in Section 3.2, let $r \geq 1$, $m > r$, $\lambda \geq 1$, define δ, H, U, t, W by Equation (27), and assume Equation (28). Sample $\boldsymbol{X} \leftarrow D^m_{\mathcal{O}_K^r, s}$. There is an event of probability at least $1 - 3\delta$ on which, simultaneously,*

$$\boldsymbol{X}\mathcal{O}_K^m = \mathcal{O}_K^r, \tag{30}$$

$$\rho_{1/t}\big((\ker_{\mathbb{Z}} \boldsymbol{A}_{\boldsymbol{X}})^* \setminus \{\boldsymbol{0}\}\big) \leq 2\delta, \tag{31}$$

$$\eta_{2\delta}(L_{\boldsymbol{X}}) \leq B_{\mathrm{geo}} := \alpha_K t = 64\,\alpha_K \kappa_K \mu_K \sqrt{R + \ln(2m/\delta)}, \tag{32}$$

$$||c_r(\boldsymbol{x}_j)|| \leq U \quad (1 \leq j \leq m). \tag{33}$$

On the same event, for every $\varepsilon \in (0,1)$,

$$\eta_\varepsilon(L_{\boldsymbol{X}}) \leq B_{\mathrm{geo}} \sqrt{\max\Big\{1, \frac{\ln(1/\varepsilon)}{\ln(1/(2\delta))}\Big\}}. \tag{34}$$

The proof assembles Proposition 5.1, Lemma 5.2 and Lemma 5.5 with three exceptional events: a column-norm tail event, a Markov bound for the remainder, and an elementary full-rank failure. The constant 64 is what remained after the inequalities were balanced, not a natural constant. The two conditions Equation (28) are explicit, so the theorem can be evaluated numerically. The width condition is $s_- \geq 8\sqrt{R}$, and a sufficient column condition is $m \geq 7.5\,(R \ln W + 2\ln(1/\delta))$, since $1/\ln(8/7) < 7.5$. Here W is polynomial in s, and no surjectivity input is needed. The hypotheses of the dual-minimum theorem of [KNSW20] over $\mathbb{Q}$, which also needs a width of order $\sqrt{r}$ and a column count of order $r \ln s$, are recorded in Section 7.2.

Proof. Let G be the event that $||c_r(\boldsymbol{x}_j)|| \leq U$ for every j. Since $s_{\max}(\boldsymbol{S}) = s_+$ and $U = s_+\sqrt{H}$, the tail bound Equation (25) with $h_0 = \sqrt{H}$ and a union bound over the m columns give

$$\mathbb{P}[G^c] \leq m\, 2^{R/2} e^{-\pi H/2} \leq \delta, \tag{35}$$

because $\pi H/2 \geq \frac{\pi}{2}\big(R + \ln m + \ln(1/\delta)\big) \geq \frac{R\ln 2}{2} + \ln m + \ln(1/\delta)$. On G every one of the M columns of $\boldsymbol{A}_{\boldsymbol{X}}$ has norm at most $\mu_K U$ by Equation (9). If $\boldsymbol{A}_{\boldsymbol{X}}$ has full row rank, the arithmetic-geometric mean inequality applied to the eigenvalues of $\boldsymbol{A}_{\boldsymbol{X}}\boldsymbol{A}_{\boldsymbol{X}}^T$ yields

$$D^2_{\boldsymbol{A}_{\boldsymbol{X}}} = \det(\boldsymbol{A}_{\boldsymbol{X}}\boldsymbol{A}_{\boldsymbol{X}}^T) \leq \Big(\frac{||\boldsymbol{A}_{\boldsymbol{X}}||_F^2}{R}\Big)^R \leq \Big(\frac{M\mu_K^2 U^2}{R}\Big)^R, \qquad \text{so that} \quad t^R D_{\boldsymbol{A}_{\boldsymbol{X}}} \leq W^R. \tag{36}$$

Let $\mathcal{B}_0 = \{\boldsymbol{y} : ||\boldsymbol{y}|| \leq 1/(4\mu_K U)\}$, which lies inside Q_R because $U \geq s_- \geq 8$. For full-row-rank $\boldsymbol{A}_{\boldsymbol{X}}$ on G, Lemma 5.2 with $B_0 = \mu_K U$ bounds the contribution of $\mathcal{B}_0$ to the right-hand side of Equation (21) by $(1 + 3e^{-\pi t^2/2})^M \leq \exp(3Me^{-\pi t^2/2})$. Now $M = dm \leq Rm \leq e^R m$ and $H \geq \ln(1/\delta) \geq 5\ln 2 > \ln 3$, so $\ln(6M/\delta) \leq \ln 6 + R + \ln m + \ln(1/\delta) = H + \ln 3 \leq 2H$. Since $t^2 \geq 4096H$, we get $3Me^{-\pi t^2/2} \leq \delta/2$ and $e^{\delta/2} \leq 1 + \delta$. Thus the near-origin contribution is at most $1 + \delta$.

Define the normalized remainder

$$J := t^R D_{\boldsymbol{A}_{\boldsymbol{X}}} \int_{Q_R \setminus \mathcal{B}_0} \prod_{\ell=1}^{M} \varphi_t\big((\boldsymbol{A}_{\boldsymbol{X}}^T \boldsymbol{y})_\ell\big)\, d\boldsymbol{y}$$

when $\boldsymbol{A}_{\boldsymbol{X}}$ has full row rank and $J := 0$ otherwise. On G use Equation (36) and then drop the indicator of G inside the nonnegative expectation. The m ring columns remain independent, the integrand factors over them by Equation (9), and every $\boldsymbol{y} \notin \mathcal{B}_0$ satisfies $||\boldsymbol{y}|| > 1/(4\mu_K U)$, so Lemma 5.5 gives

$$\begin{aligned} \mathbb{E}[\mathbf{1}_G J] &\leq W^R \int_{Q_R \setminus \mathcal{B}_0} \prod_{j=1}^{m} \mathbb{E}_{\boldsymbol{x}_j} \prod_{k=1}^{d} \varphi_t\big(\langle \boldsymbol{y}, (\boldsymbol{I}_r \otimes \boldsymbol{T}_k) c_r(\boldsymbol{x}_j)\rangle\big)\, d\boldsymbol{y} \\ &\leq W^R \Big(\frac{7}{8}\Big)^m \leq \delta^2, \end{aligned} \tag{37}$$

where the last step is the column condition in Equation (28). Markov's inequality gives $\mathbb{P}[G \cap \{J > \delta\}] \le \delta$.

Rank deficiency is excluded as follows. The coefficient column matrix $\boldsymbol{C}_{\boldsymbol{X}} = [c_r(\boldsymbol{x}_1) | \cdots | c_r(\boldsymbol{x}_m)]$ is a submatrix of $\boldsymbol{A}_{\boldsymbol{X}}$, so it suffices that $\boldsymbol{C}_{\boldsymbol{X}}$ has rank R. By the last part of Lemma 5.3 with $s_0 = s_- \ge 8\sqrt{R} \ge 8$, an independent coefficient column lies in any fixed proper subspace with probability at most $1/4$. If $\operatorname{rank} \boldsymbol{C}_{\boldsymbol{X}} < R$, some set of $R-1$ columns spans all of them, after padding a spanning subset if necessary. Conditioning on those columns and using the independence of the remaining $m - R + 1$ columns,

$$\mathbb{P}[\operatorname{rank} \boldsymbol{C}_{\boldsymbol{X}} < R] \le \binom{m}{R-1} 4^{-(m-R+1)} \le 2^{-m+2R-2} \le \delta. \tag{38}$$

Here $W \ge t \ge 64$ and the column condition imply $m \ge (R \ln 2 + 2\ln(1/\delta))/\ln(8/7) \ge 2R + \log_2(1/\delta)$, which gives the last inequality.

Outside the three events in Equations (35), (37) and (38), the matrix $\boldsymbol{A}_{\boldsymbol{X}}$ has full row rank and Proposition 5.1 gives

$$I_{\boldsymbol{A}_{\boldsymbol{X}}}\, \rho_{1/t}\big((\ker_{\mathbb{Z}} \boldsymbol{A}_{\boldsymbol{X}})^*\big) \le (1+\delta) + \delta = 1 + 2\delta < 2.$$

The integer $I_{\boldsymbol{A}_{\boldsymbol{X}}}$ must equal one, so $\boldsymbol{A}_{\boldsymbol{X}} \mathbb{Z}^M = \mathbb{Z}^R$, which is Equation (30), and $\rho_{1/t}((\ker_{\mathbb{Z}} \boldsymbol{A}_{\boldsymbol{X}})^* \setminus \{\boldsymbol{0}\}) \le 2\delta$, which is Equation (31). The column bound Equation (33) is the event G. For Equation (32), $L_{\boldsymbol{X}} = \boldsymbol{B}_K^{\oplus m} \ker_{\mathbb{Z}} \boldsymbol{A}_{\boldsymbol{X}}$ by Equation (20), and Lemma 3.3 gives $\eta_{2\delta}(L_{\boldsymbol{X}}) \le \alpha_K\, \eta_{2\delta}(\ker_{\mathbb{Z}} \boldsymbol{A}_{\boldsymbol{X}}) \le \alpha_K t$. Finally, for $q \ge 1$ and any lattice L, positivity and the bound $e^{-\pi t^2 ||\boldsymbol{v}||^2} \le 1$ give

$$\rho_{1/(t\sqrt{q})}(L^* \setminus \{\boldsymbol{0}\}) = \sum_{\boldsymbol{v} \ne \boldsymbol{0}} \big(e^{-\pi t^2 ||\boldsymbol{v}||^2}\big)^q \le \Big(\sum_{\boldsymbol{v} \ne \boldsymbol{0}} e^{-\pi t^2 ||\boldsymbol{v}||^2}\Big)^q. \tag{39}$$

With $q = \max\{1, \ln(1/\varepsilon)/\ln(1/(2\delta))\}$ the right-hand side is at most $(2\delta)^q \le \varepsilon$ for $L = \ker_{\mathbb{Z}} \boldsymbol{A}_{\boldsymbol{X}}$, so $\eta_\varepsilon(\ker_{\mathbb{Z}} \boldsymbol{A}_{\boldsymbol{X}}) \le t\sqrt{q}$, and the metric change gives Equation (34). The exceptional event does not depend on ε. □

Corollary 5.7 (Power-of-two cyclotomics). *Let $K = \mathbb{Q}(\zeta_{2^a})$ of degree d with the power basis, so that $\alpha_K = \sqrt{d}$, $\beta_K = 1/\sqrt{d}$ and $\kappa_K = \mu_K = 1$. If*

$$s \ge 8\sqrt{dR} \qquad \text{and} \qquad m \ln \tfrac{8}{7} \ge R \ln W + 2 \ln \tfrac{1}{\delta},$$
$$\text{where} \qquad W = \frac{64\, s\, (R + \ln(2m/\delta))}{\sqrt{d}} \sqrt{\frac{m}{r}},$$

then, except with probability 3δ, $\boldsymbol{X}$ is surjective and

$$\eta_{2\delta}(L_{\boldsymbol{X}}) \le 64 \sqrt{d\,\big(dr + \ln(2m/\delta)\big)}.$$

For $K = \mathbb{Q}$ take $d = 1$.

Proof. Insert the constants of Proposition 3.1 into Equations (27) and (28): $s_- = s_+ = s/\sqrt{d}$, $U = s\sqrt{H}/\sqrt{d}$ and $W = tU\sqrt{M/R} = 64\sqrt{H} \cdot s\sqrt{H}/\sqrt{d} \cdot \sqrt{m/r}$. □

Theorem 5.8 (Polynomial width and growing rank over a fixed field). *Fix K and a real number $p \ge 1/2$. There are constants c_X, c_m, C, depending only on K and p, such that the following holds for all sufficiently large λ. Put*

$$r = \lceil \lambda/d \rceil, \quad R = dr, \quad s = c_X R^p, \quad m = \lceil c_m R \ln(2R) \rceil, \quad \delta = 2^{-\lambda-4},$$

and sample $\boldsymbol{X} \leftarrow D^m_{\mathcal{O}_K^r, s}$. Outside an event of probability at most 3δ,

$$\boldsymbol{X}\mathcal{O}_K^m = \mathcal{O}_K^r \qquad \text{and} \qquad \eta_{2\delta}(L_{\boldsymbol{X}}) \le C\sqrt{\lambda}.$$

In particular the same bound holds at the larger error $2^{-\lambda}$.

Proof. Since $r = \lceil \lambda/d \rceil$ we have $\lambda \le R < \lambda + d$. Take $c_X \ge 8\alpha_K$, so that $s_- = c_X R^p/\alpha_K \ge 8\sqrt{R}$. For fixed K, p, c_X, c_m the quantities in Equation (27) satisfy $H = \Theta_K(R)$, since $\ln(1/\delta) = (\lambda+4)\ln 2 \le 2R \ln 2$ for $\lambda \ge 4$ and $\ln m = O_K(\ln R)$, and $W = tU\mu_K\sqrt{m/r} = O_{K,p,c_X,c_m}\big(R^{p+1}\sqrt{\ln R}\big)$. Hence the right-hand side of the column condition is $(p+1)R\ln R + O(R \ln\ln R + R)$, while the left-hand side is at least $c_m \ln(8/7)\, R \ln(2R)$. Choosing $c_m > (p+1)/\ln(8/7)$ with a fixed positive margin makes both conditions in Equation (28) hold for all large λ, and Theorem 5.6 gives $\eta_{2\delta}(L_{\boldsymbol{X}}) \le 64\alpha_K\kappa_K\mu_K\sqrt{H} = O_{K,p}(\sqrt{\lambda})$. □

The restriction $r = \lceil \lambda/d \rceil$ is a convenient normalization, not a limitation of the method: Theorem 5.6 allows arbitrary r, m, λ subject to the two explicit inequalities. The width s has not disappeared from the hypotheses. It enters the required number of columns logarithmically through W, but it no longer appears as a multiplicative factor in the smoothing bound.

5.4 Constant coefficient width in power-of-two fields

For $K = \mathbb{Q}(\zeta_{2^a})$ with the power basis, Corollary 3.2 says that the drm integer coefficients of $\boldsymbol{X}$ are independent copies of $D_{\mathbb{Z},s_-}$ with $s_- = s/\sqrt{d}$. The following theorem allows $s_- = 1$, a matrix in which about 92% of all coefficients are zero and almost all others are ± 1. The width condition of Theorem 5.6 disappears, at the price of larger constants.

Theorem 5.9 (Constant coefficient width). *Let $K = \mathbb{Q}(\zeta_{2^a})$ have degree $d \geq 2$, let $r \geq 1$, $m > r$, $\lambda \geq 1$, and let $s \geq \sqrt{d}$, so that the coefficient width $s_- = s/\sqrt{d}$ is at least one. Put*

$$\delta := 2^{-\lambda-6}, \qquad H := R + \ln \frac{2m}{\delta}, \qquad t := 640\sqrt{H}, \qquad W := 640\, s_- H \sqrt{m/r}, \tag{40}$$

and assume

$$m \ln \tfrac{50}{49} \;\geq\; R \ln W + 2 \ln \tfrac{1}{\delta}. \tag{41}$$

Sample $\boldsymbol{X} \leftarrow D^m_{\mathcal{O}_K^r,s}$. Then, with probability at least $1 - 3\delta$,

$$\boldsymbol{X}\mathcal{O}_K^m = \mathcal{O}_K^r \qquad \text{and} \qquad \eta_{2\delta}(L_{\boldsymbol{X}}) \leq B_{\mathrm{cw}} := \sqrt{d}\, t = 640\sqrt{d\,\big(dr + \ln(2m/\delta)\big)}. \tag{42}$$

On the same event $\eta_\varepsilon(L_{\boldsymbol{X}}) \leq B_{\mathrm{cw}} \sqrt{\max\{1, \ln(1/\varepsilon)/\ln(1/(2\delta))\}}$ for every $\varepsilon \in (0,1)$.

The proof, given after the two lemmas below, uses the identity of Proposition 5.1 and the main term of Lemma 5.2 unchanged. What changes is the estimate for one ring column away from the origin ball. That estimate can no longer rely on Lemma 5.4, because rounding a direction in all R coordinates costs a displacement of norm $\sqrt{R}$, and at width one the Gaussian does not survive such a shift, since the translation bound Equation (23) becomes trivial for $\|\boldsymbol{v}\| \geq 1$. Two lemmas replace it.

Lemma 5.10 (Uniform moments). *Let $Z \leftarrow D_{\mathbb{Z},\sigma}$ with $\sigma \geq 1$, and put $v_\sigma := \mathbb{E}Z^2$ and $\beta_\sigma := \mathbb{E}\,|Z|^3$. Then*

$$v_\sigma \geq \frac{\sigma^2}{16}, \qquad \beta_\sigma \leq \frac{\sigma^3}{4}, \qquad \mathbb{E}e^{uZ} \leq e^{\sigma^2 u^2/(4\pi)} \quad (u \in \mathbb{R}).$$

In the proofs below $K = \mathbb{Q}(\zeta_{2^a})$ with $d \geq 2$, the coefficient width is $\sigma := s_- = s/\sqrt{d} \geq 1$, and δ, H, t, W are those of Equation (40). By Corollary 3.2 the coefficient columns $c_r(\boldsymbol{x}_j)$ are independent with law $D_{\mathbb{Z}^R,\sigma}$, and by Proposition 3.1 multiplication by a power of ζ_{2^a} permutes coefficients up to sign, so every column of $\boldsymbol{A}_{\boldsymbol{X}}$ has the same norm as the coefficient column of its ring block.

Proof. The exponential moment is Equation (24) with $R = 1$. For the variance, differentiate the Poisson identity $\rho_\sigma(\mathbb{Z}) = \sigma\rho_{1/\sigma}(\mathbb{Z})$ with respect to σ: the left-hand side gives $(2\pi/\sigma^3)\sum_z z^2 e^{-\pi z^2/\sigma^2}$, the right-hand side gives $\rho_{1/\sigma}(\mathbb{Z}) - 4\pi\sigma^2 \sum_{j\geq 1} j^2 e^{-\pi\sigma^2 j^2}$, and dividing by $\rho_\sigma(\mathbb{Z}) = \sigma\rho_{1/\sigma}(\mathbb{Z})$ yields

$$\frac{v_\sigma}{\sigma^2} = \frac{1}{2\pi} - \frac{2\sigma^2 \sum_{j\geq 1} j^2 e^{-\pi\sigma^2 j^2}}{1 + 2\sum_{j\geq 1} e^{-\pi\sigma^2 j^2}} \;\geq\; \frac{1}{2\pi} - \frac{2e^{-\pi}}{1 - 4e^{-3\pi}} \;>\; \frac{1}{16}.$$

Here $u \mapsto ue^{-\pi j^2 u}$ is decreasing for $u \geq 1$, so $\sigma^2 e^{-\pi\sigma^2 j^2} \leq e^{-\pi j^2}$, and consecutive terms $j^2 e^{-\pi j^2}$ have ratio at most $4e^{-3\pi}$. For the third moment, the sum of a nonnegative unimodal function over the positive integers is at most its integral over $[0,\infty)$ plus its maximum. Applied to $x \mapsto x^3 e^{-\pi x^2/\sigma^2}$, whose integral is $\sigma^4/(2\pi^2)$ and whose maximum is $\sigma^3(3/(2\pi e))^{3/2}$, and using $\rho_\sigma(\mathbb{Z}) \geq \sigma$, this gives

$$\beta_\sigma \leq \frac{2}{\sigma}\Big(\frac{\sigma^4}{2\pi^2} + \sigma^3\Big(\frac{3}{2\pi e}\Big)^{3/2}\Big) \leq \sigma^3\Big(\frac{1}{\pi^2} + 2\Big(\frac{3}{2\pi e}\Big)^{3/2}\Big) < \frac{\sigma^3}{4},$$

since $\sigma^2 \leq \sigma^3$ for $\sigma \geq 1$ and the bracket is less than 0.2486. ∎

Lemma 5.11 (Uniform contraction at constant width). *Let $H \geq R \geq 1$, $\sigma \geq 1$, $t = 640\sqrt{H}$, and $\boldsymbol{Z} \leftarrow D_{\mathbb{Z}^R,\sigma}$. If $\boldsymbol{y} \in Q_R$ and $||\boldsymbol{y}|| \geq 1/(4\sigma\sqrt{H})$, then*

$$\mathbb{E}\,\varphi_t(\langle \boldsymbol{y}, \boldsymbol{Z}\rangle) \leq \frac{39}{40}. \tag{43}$$

If $\boldsymbol{T}_1 = \boldsymbol{I}_d, \boldsymbol{T}_2, \ldots, \boldsymbol{T}_d$ are the coefficient multiplication matrices of the power basis of $\mathbb{Q}(\zeta_{2^a})$ and $d \leq R$, then

$$\mathbb{E}\prod_{k=1}^{d} \varphi_t\big(\langle \boldsymbol{y}, (\boldsymbol{I}_r \otimes \boldsymbol{T}_k)\boldsymbol{Z}\rangle\big) \leq \frac{49}{50}. \tag{44}$$

The proof distinguishes two kinds of directions. If some coordinate satisfies $|y_i| \geq 8/(\sigma t)$, one translates only the i-th coordinate of $\boldsymbol{Z}$ by an integer. The distance between $D_{\mathbb{Z},\sigma}$ and its unit translate is exactly the mass at zero, $\rho_\sigma(\mathbb{Z})^{-1} \leq (1 + 2e^{-\pi})^{-1} < 0.921$, so the two laws overlap with a fixed positive probability even at $\sigma = 1$, and the phase $\langle \boldsymbol{y}, \boldsymbol{Z}\rangle$ cannot be close to an integer on both laws. If all coordinates are smaller than $8/(\sigma t)$, the sum $\langle \boldsymbol{y}, \boldsymbol{Z}\rangle$ of many small independent terms is close to a normal variable of variance $v_\sigma\,||\boldsymbol{y}||^2$, and the Berry–Esseen theorem, applied to the single bounded test function $u \mapsto e^{-\pi t^2 u^2}$, shows that its expectation is at most a constant below one. No comparison of a discrete with a continuous law in total variation is made. The uniform bound $39/40$ for one scalar factor and $49/50$ for one ring block replace the factor $7/8$ of Lemma 5.5, which is why the column constant becomes $1/\ln(50/49) \approx 49.5$ instead of $1/\ln(8/7) \approx 7.5$ and the smoothing constant 640 instead of 64.

Proof. We use the elementary bounds, valid for $t \geq 8$ and proved as in the proof of Lemma 5.5,

$$\varphi_t(u) \leq \varphi_t(0) \leq 1 + 3e^{-\pi t^2}, \qquad \operatorname{dist}(u, \mathbb{Z}) \geq 1/t \implies \varphi_t(u) \leq 1/8,$$
$$|u| \leq 1/4 \implies 0 \leq \varphi_t(u) - e^{-\pi t^2 u^2} \leq 3e^{-\pi t^2/2}.$$

Write P for the law $D_{\mathbb{Z}^R,\sigma}$ of $\boldsymbol{Z}$, a product of copies of $D_{\mathbb{Z},\sigma}$.

Case 1: some coordinate satisfies $a := |y_i| \geq 8/(\sigma t)$. If $a \geq 2/t$, take $\boldsymbol{v} = \operatorname{sgn}(y_i)\boldsymbol{e}_i$. The laws of $\boldsymbol{Z}$ and $\boldsymbol{Z} + \boldsymbol{v}$ differ only in the i-th coordinate, and the distance between the symmetric unimodal law $D_{\mathbb{Z},\sigma}$ and its unit translate is its mass at zero, by telescoping. Hence

$$\mathrm{SD}(P, P + \boldsymbol{v}) = \rho_\sigma(\mathbb{Z})^{-1} \leq a_0 := \frac{1}{1 + 2e^{-\pi}} < 0.921.$$

If $a < 2/t$, then $\sigma > 4$, because $a \geq 8/(\sigma t)$. In this case take $v_i = \operatorname{sgn}(y_i)\lceil 2/(ta)\rceil$ and all other coordinates zero. Then $|v_i| \leq 2/(ta) + 1 \leq \sigma/4 + 1 \leq \sigma/2$, and Equation (23) gives $\mathrm{SD}(P, P + \boldsymbol{v}) \leq \sqrt{\pi/2} \cdot |v_i|/\sigma \leq \sqrt{\pi/8} < a_0$. In both cases $2/t \leq \langle \boldsymbol{y}, \boldsymbol{v}\rangle \leq 1 - 2/t$: in the first $\langle \boldsymbol{y}, \boldsymbol{v}\rangle = a \in [2/t, 1/2]$, in the second $\langle \boldsymbol{y}, \boldsymbol{v}\rangle = a\lceil 2/(ta)\rceil \in [2/t, 4/t)$. Consequently the sets $E := \{\boldsymbol{z} : \operatorname{dist}(\langle \boldsymbol{y}, \boldsymbol{z}\rangle, \mathbb{Z}) < 1/t\}$ and $E + \boldsymbol{v}$ are disjoint, exactly as in Lemma 5.4, and as in the proof of Lemma 5.5,

$$P(E^c) \geq \tfrac{1}{2}(1 - a_0) =: \gamma_0 > 0.039,$$
$$\mathbb{E}\,\varphi_t(\langle \boldsymbol{y}, \boldsymbol{Z}\rangle) \leq \varphi_t(0)\big(1 - \tfrac{7}{8}\gamma_0\big) < 0.966 < \tfrac{39}{40}.$$

Case 2: $||\boldsymbol{y}||_\infty < 8/(\sigma t)$. Put $S := \langle \boldsymbol{y}, \boldsymbol{Z}\rangle = \sum_i y_i Z_i$, a sum of independent centered summands with $\operatorname{Var} S = v_\sigma\,||\boldsymbol{y}||^2$ and $\sum_i \mathbb{E}\,|y_i Z_i|^3 = \beta_\sigma \sum_i |y_i|^3 \leq \beta_\sigma\,||\boldsymbol{y}||_\infty\,||\boldsymbol{y}||^2$. Let G be a centered normal variable with the same variance. The Berry–Esseen inequality [Ber41, Ess42] for independent, not necessarily identically distributed summands, with the constant 0.5583 of Shevtsova [She13], and Lemma 5.10 give

$$\sup_u |\mathbb{P}[S \leq u] - \mathbb{P}[G \leq u]| \leq 0.56\,\frac{\beta_\sigma \sum_i |y_i|^3}{v_\sigma^{3/2}\,||\boldsymbol{y}||^3} \leq 0.56 \cdot 16 \cdot \frac{||\boldsymbol{y}||_\infty}{||\boldsymbol{y}||}$$
$$\leq \frac{224}{25} \cdot \frac{8}{\sigma t} \cdot 4\sigma\sqrt{H} = \frac{224}{25} \cdot \frac{32\sqrt{H}}{t} = \frac{56}{125},$$

using $||\boldsymbol{y}|| \geq 1/(4\sigma\sqrt{H})$ and $t = 640\sqrt{H}$. For $f(u) = e^{-\pi t^2 u^2}$ we have $\int_{\mathbb{R}} |f'(u)|\,du = 2$, so integration by parts bounds $|\mathbb{E}f(S) - \mathbb{E}f(G)|$ by $2 \cdot 56/125 = 112/125$. Moreover

$$\mathbb{E}f(G) = \big(1 + 2\pi t^2 v_\sigma\,||\boldsymbol{y}||^2\big)^{-1/2} \leq \Big(1 + \frac{2\pi t^2 \sigma^2}{16 \cdot 16\sigma^2 H}\Big)^{-1/2}$$

$$= \Big(1 + \frac{\pi\, 640^2}{128}\Big)^{-1/2} < \frac{1}{100},$$

so $\mathbb{E} f(S) \le 0.906$. The other periodic images of the Gaussian cannot be ignored. Since $\sigma^2 ||\boldsymbol{y}||^2 \le \sigma^2 R\, ||\boldsymbol{y}||_\infty^2 < 64R/t^2$, the exponential moment of Lemma 5.10 and a Chernoff bound give $\mathbb{P}[|S| > 1/4] \le 2\exp(-\pi t^2/(1024R)) \le 2e^{-400\pi}$, because $t^2 = 640^2 H \ge 640^2 R$. On $|S| \le 1/4$ use the third elementary bound, on its complement the first. Hence

$$\mathbb{E}\,\varphi_t(S) \le 0.906 + 3e^{-\pi t^2/2} + 2\big(1 + 3e^{-\pi t^2}\big)e^{-400\pi} < 0.907 < \tfrac{39}{40}.$$

This proves Equation (43). The normal approximation was used for one bounded test function, not as a total variation comparison with a continuous distribution.

For Equation (44), the factor $k = 1$ is $\varphi_t(\langle \boldsymbol{y}, \boldsymbol{Z}\rangle)$ because $\boldsymbol{T}_1 = \boldsymbol{I}_d$, and each of the other $d-1$ factors is at most $\varphi_t(0)$. Since $d \le R \le H$, $\varphi_t(0)^{d-1} \le \exp(3He^{-\pi 640^2 H}) < 1001/1000$, and $(39/40)(1001/1000) < 49/50$. No independence within a ring column is used. □

Proof. Put $U := \sigma\sqrt{H}$ and note that $W = tU\sqrt{M/R} = 640\sqrt{H}\cdot\sigma\sqrt{H}\cdot\sqrt{m/r}$ agrees with Equation (40). Write $L := \ln(1/\delta)$.

Rank. For a ring entry $x \leftarrow D_{\mathcal{O}_K,s}$ the largest atom is p_σ^d with $p_\sigma := \rho_\sigma(\mathbb{Z})^{-1} \le \rho_1(\mathbb{Z})^{-1} < 12/13$, since the d coefficients are independent. A proper K-linear subspace of K^r is contained in the kernel of a nonzero K-linear functional, which determines one ring entry of a vector in it from the others, so an independent column lies in the subspace with probability at most p_σ^d. If the columns have K-rank below r, some $r-1$ of them span all of them, after padding. Conditioning on such a set and taking a union bound,

$$\mathbb{P}[\operatorname{rank}_K \boldsymbol{X} < r] \le \binom{m}{r-1} p_\sigma^{d(m-r+1)}. \tag{45}$$

Since $W \ge 640H\sqrt{m/r} \ge 640\sqrt{mr} \ge \sqrt{m}$ and $\ln(50/49) < 1/26$, the hypothesis Equation (41) gives $m/26 > R\ln W + 2L \ge \frac{1}{2}R\ln m + 2L$, hence $m > 13R\ln m + 52L$ and in particular $m > 2r$. With $\ln(13/12) > 1/13$ and $m - r + 1 > m/2$, the logarithm of the right-hand side of Equation (45) is at most

$$r\ln m - \frac{dm}{26} < r\ln m - \frac{d^2 r}{2}\ln m - 2dL \le -2dL \le -L,$$

where $d \ge 2$ was used. So rank failure has probability at most δ. Full K-rank of $\boldsymbol{X}$ gives full real row rank of $\boldsymbol{A}_{\boldsymbol{X}}$, whose real rank is d times the K-rank.

Column norms and the main term. Let G be the event that all m coefficient columns have norm at most U. The tail bound Equation (25) with $s_1 = \sigma$ and $h_0 = \sqrt{H}$ gives $\mathbb{P}[G^c] \le m2^{R/2}e^{-\pi H/2} \le \delta$, as in Equation (35). On G all M columns of $\boldsymbol{A}_{\boldsymbol{X}}$ have norm at most U, and for full-row-rank $\boldsymbol{A}_{\boldsymbol{X}}$ the arithmetic-geometric mean inequality gives $D_{\boldsymbol{A}_{\boldsymbol{X}}} \le (U\sqrt{M/R})^R$, so $t^R D_{\boldsymbol{A}_{\boldsymbol{X}}} \le W^R$. Let $\mathcal{B}_0 = \{\boldsymbol{y} : ||\boldsymbol{y}|| \le 1/(4U)\} \subseteq Q_R$. By Lemma 5.2 with $B_0 = U$, the contribution of $\mathcal{B}_0$ to Equation (21) is at most $(1 + 3e^{-\pi t^2/2})^M \le 1 + \delta$, exactly as in the proof of Theorem 5.6, since $M \le Rm$, $\ln(6M/\delta) \le 2H$ and now $t^2 = 640^2 H$.

Remainder. Let J be the normalized integral over $Q_R \setminus \mathcal{B}_0$, set to zero for rank-deficient $\boldsymbol{A}_{\boldsymbol{X}}$. On G apply the determinant bound and drop the indicator inside the nonnegative expectation. The integrand factors over the m independent ring columns by Equation (9), and every $\boldsymbol{y} \notin \mathcal{B}_0$ satisfies $||\boldsymbol{y}|| > 1/(4U) = 1/(4\sigma\sqrt{H})$, so Equation (44) applies to each factor and

$$\mathbb{E}[\mathbf{1}_G J] \le W^R\Big(\frac{49}{50}\Big)^m \le \delta^2$$

by Equation (41). Markov's inequality excludes $J > \delta$ on G with probability at most δ.

Assembly. Outside the three exceptional events, Proposition 5.1 gives

$$I_{\boldsymbol{A}_{\boldsymbol{X}}}\, \rho_{1/t}\big((\ker_{\mathbb{Z}} \boldsymbol{A}_{\boldsymbol{X}})^*\big) \le 1 + 2\delta < 2.$$

Both factors are at least one and $I_{\boldsymbol{A}_{\boldsymbol{X}}}$ is an integer, so $I_{\boldsymbol{A}_{\boldsymbol{X}}} = 1$, which is surjectivity, and the nonzero dual mass of the coefficient kernel at parameter $1/t$ is at most 2δ. Multiplication of lengths by $\alpha_K = \sqrt{d}$ through Lemma 3.3 gives Equation (42), and Equation (39) gives the assertion for every ε. □

Remark 5.12 (The rational case). The hypothesis $d \geq 2$ is used only in the rank step, in the inequality $r \ln m - dm/26 < (r - d^2 r/2) \ln m - 2dL$. For $K = \mathbb{Q}$, that is $d = 1$ and $\boldsymbol{X} \in \mathbb{Z}^{r \times m}$ with independent $D_{\mathbb{Z},\sigma}$ entries, $\sigma \geq 1$, the step closes directly: since $W \geq 640\sqrt{m}$, the hypothesis Equation (41) gives $m \geq 49\,(r \ln 640 + \frac{r}{2} \ln m + 2L) \geq 24\, r \ln m + 316\, r + 98\, L$, hence $(m - r + 1) \ln(13/12) \geq 0.08\,(m - r) > r \ln m + L$, which is what Equation (45) needs. All other steps of the proof are unchanged, so Theorem 5.9 also holds over $\mathbb{Q}$ with the same constants.

Remark 5.13 (What the constant-width theorem does and does not change). In the range $\lambda \leq R \leq 2\lambda$ and $s_- = 1$, a sufficiently large absolute multiple of $R \ln(2R)$ satisfies Equation (41) for all large λ, uniformly in d and r, because $H = O(R)$ and $W = O(R^{3/2}\sqrt{\ln R})$ by $d \leq R$. Hence $\eta_{2^{-\lambda}}(L_{\boldsymbol{X}}) = O(\sqrt{d\lambda})$ at coefficient width one, with the same order of columns as in Corollary 5.7 and the same order of smoothing, which Proposition 5.14 below shows to be optimal including its dependence on d. The constants are conservative: at $d = 16$, $r = 8$, $\lambda = 128$ and $s_- = 1$ the theorem needs $m = 115\,004$ columns and gives $B_{\mathrm{cw}} \approx 3.9 \cdot 10^4$ (evaluated from Equations (40) and (41)), against 18 605 columns and $B_{\mathrm{geo}} \approx 3.9 \cdot 10^3$ for Corollary 5.7 at its smallest width $s_- = 8\sqrt{R}$. The theorem is stated for $d \geq 2$ because one arithmetic step of the rank estimate uses it. Remark 5.12 explains that the same proof also covers $K = \mathbb{Q}$. The matrix at width one is far from uniform modulo any integer $q \geq 2$: more generally, for $\sigma > 0$, $Z \leftarrow D_{\mathbb{Z},\sigma}$ and U_q uniform on $\mathbb{Z}/q\mathbb{Z}$, Poisson summation gives

$$\mathrm{SD}(Z \bmod q,\, U_q) \;\geq\; \frac{e^{-\pi\sigma^2/q^2}}{2\rho_{1/\sigma}(\mathbb{Z})}, \tag{46}$$

which at $\sigma = 1$ exceeds $0.459\, e^{-\pi/q^2}$, so no regularity statement for uniform matrices modulo q [GPV08, Corollary 2.8] applies in this regime. To prove Equation (46), test the first Fourier character: its coefficient under Z mod q is $\rho_\sigma(\mathbb{Z})^{-1} \sum_z e^{-\pi z^2/\sigma^2} e^{2\pi i z/q} = \rho_{1/\sigma}(\mathbb{Z})^{-1} \sum_j e^{-\pi\sigma^2(j+1/q)^2} \geq e^{-\pi\sigma^2/q^2}/\rho_{1/\sigma}(\mathbb{Z})$, while its coefficient under U_q is zero, and the statistical distance is at least half the difference.

5.5 A matching lower bound, and the failure of the Gaussian heuristic

Proposition 5.14 (Projected dual vector). *Let L be a nonzero primitive sublattice of $\Lambda_m = \Phi(\mathcal{O}_K^m)$. Then*

$$\lambda_1(L^*) \leq \beta_K, \qquad \textit{and hence} \qquad \eta_\varepsilon(L) \geq \frac{1}{\beta_K}\sqrt{\frac{\ln(2/\varepsilon)}{\pi}} \quad (0 < \varepsilon < 1).$$

In coefficient coordinates, every nonzero primitive sublattice $L' \subseteq \mathbb{Z}^M$ satisfies $\lambda_1(L'^*) \leq 1$. Both statements apply to $L_{\boldsymbol{X}}$ and to $\ker_{\mathbb{Z}} \boldsymbol{A}_{\boldsymbol{X}}$ whenever $m > r$.

Proof. The argument of Lemma 4.2 applies to a primitive sublattice of any lattice Λ and gives $L^* = \pi_{\mathrm{span}\, L}(\Lambda^*)$. The dual lattice of $\Lambda_m = \boldsymbol{B}_K^{\oplus m}\mathbb{Z}^M$ has the basis formed by the columns of $(\boldsymbol{B}_K^{-T})^{\oplus m}$, each of norm at most $\left\|\boldsymbol{B}_K^{-1}\right\|_{\mathrm{op}} = \beta_K$. At least one basis vector has a nonzero projection onto the nonzero subspace $\mathrm{span}\, L$, and this projection is a nonzero vector of L^* of norm at most β_K. The two vectors $\pm\boldsymbol{v}$ contribute $2e^{-\pi t^2\beta_K^2}$ to $\rho_{1/t}(L^* \setminus \{\mathbf{0}\})$, and at $t = \eta_\varepsilon(L)$ this is at most ε. For $\Lambda = \mathbb{Z}^M$ the dual basis is the standard basis, of norm one. A kernel is primitive, and $L_{\boldsymbol{X}} \neq 0$ because $\mathrm{rank}\, L_{\boldsymbol{X}} \geq d(m - r)$. □

Corollary 5.15 (The smoothing parameter is pinned). *In the setting of Theorem 5.8, on the good event,*

$$\frac{1}{\beta_K}\sqrt{\frac{\ln(2^{\lambda+1})}{\pi}} \;\leq\; \eta_{2^{-\lambda}}(L_{\boldsymbol{X}}) \;\leq\; C\sqrt{\lambda},$$

so $\eta_{2^{-\lambda}}(L_{\boldsymbol{X}}) = \Theta_K(\sqrt{\lambda})$. For power-of-two cyclotomics, on the event of Corollary 5.7 or of Theorem 5.9 with $R = \lambda$ and $\ln(2m) \leq \lambda/10$, the ratio of the upper to the lower bound is at most $64\sqrt{2.7\pi/\ln 2} < 225$, respectively $640\sqrt{2.7\pi/\ln 2} < 2250$.

Proof. Combine Proposition 5.14 at $\varepsilon = 2^{-\lambda}$ with Theorem 5.8. For the numerical ratios, the existing hypotheses already imply $\lambda \geq 28$. Indeed, $W \geq 1$ and either column condition give $m/7 \geq 2\ln(1/\delta) \geq 10 \ln 2 > 5$, because both $\ln(8/7)$ and $\ln(50/49)$ are less than $1/7$. In particular $m \geq 16$, so

$\lambda \geq 10\ln(2m) \geq 10\ln 32 > 28$. Insert $\beta_K = 1/\sqrt{d}$ and the two upper bounds, and use $\delta = 2^{-\lambda-4}$ or $2^{-\lambda-6}$, respectively, to obtain

$$\ln(2m/\delta) \leq \ln(2m) + (\lambda+6)\ln 2 \leq 0.1R + 0.7R + 4.2 \leq 1.7R.$$

Thus $R + \ln(2m/\delta) \leq 2.7R$, which gives both asserted ratios. □

Remark 5.16 (The Gaussian heuristic fails for kernel duals). Proposition 5.14 shows that the dual of the coefficient kernel contains a nonzero vector of norm at most one for every $\boldsymbol{X}$ with a nonzero kernel, in particular whenever $m > r$. Its Gaussian-heuristic prediction [GN08] $\mathrm{gh}(L^*) = \sqrt{N/(2\pi e)}\det(L)^{-1/N}$ is of order $\sqrt{N}$ in the wide regime, where $\det(L)^{1/N} = \Theta(1)$. The dual kernel is therefore not a Gaussian-heuristic lattice: its first minimum is smaller than the heuristic value by a factor of order $\sqrt{N}$, because it contains the projections of the ambient standard basis. The same phenomenon in canonical coordinates gives $\lambda_1(L^*_{\boldsymbol{X}}) \leq \beta_K$, and in particular the conjecture recorded in Remark 4.6 is false: for power-of-two cyclotomics $\eta_{2^{-N}}(L_{\boldsymbol{X}}) \geq \sqrt{d\,N\ln 2/\pi}$ deterministically, whereas the conjecture asked for $\eta_{2^{-N}}(L_{\boldsymbol{X}}) \leq C\sqrt{d}$. Conversely, on the event of Theorem 5.6, Equation (31) and the first inequality of Proposition 4.3 at $\varepsilon = 2\delta$ give

$$\lambda_1\big((\ker_{\mathbb{Z}} \boldsymbol{A}_{\boldsymbol{X}})^*\big) \;\geq\; \frac{1}{t}\sqrt{\frac{\ln(1/\delta)}{\pi}} = \frac{1}{64\kappa_K\mu_K}\sqrt{\frac{\ln(1/\delta)}{\pi H}},$$

and for $R = \lambda \geq 28$ and $\ln(2m) \leq \lambda/10$ one has $\ln(1/\delta) \geq R\ln 2$ and $H \leq 2.7R$ as in the proof of Corollary 5.15, so the right-hand side is at least $\sqrt{\ln 2/(2.7\pi)}/(64\kappa_K\mu_K) > 1/(224\kappa_K\mu_K)$. The first minimum of the dual coefficient kernel is therefore pinned in the interval $[1/(224\kappa_K\mu_K),\, 1]$ in the wide regime, not near the value of order $\sqrt{N}$ suggested by the Minkowski bound of Proposition 4.3, which is far from sharp for these lattices.

Remark 5.17 (Comparison with the dual-minimum method). For power-of-two cyclotomics the upper bounds and the lower bound of Proposition 5.14 compare as follows. At the error $\varepsilon = 2^{-\lambda}$ with $\lambda \asymp R = dr$, Theorems 5.6 and 5.9 give $\Theta(\sqrt{dR})$, which is tight by Corollary 5.15. At the much smaller error $\varepsilon = e^{-N}$, their ε-extensions with $\ln(1/\varepsilon)/\ln(1/(2\delta)) \approx N/\lambda$ give $\Theta(\sqrt{dN})$, which matches the lower bound $\sqrt{dN\ln 2/\pi}$ up to an absolute constant. The dual-minimum method of [AGHS13, KNSW20] first certifies a lower bound $\lambda_1(L^*) \geq c$ on the dual of the coefficient kernel and then converts it into a smoothing bound. The conversion counts dual vectors in shells, as in [Ban93, Lemma 1.5], and yields a bound of the form $\eta_\varepsilon(L) \leq c^{-1}\sqrt{(C_0N + \ln(2/\varepsilon))/\pi}$ with an absolute constant C_0. The shell-counting term C_0N dominates unless $\ln(1/\varepsilon) \gtrsim N$. At error $2^{-\lambda}$ with $\lambda \asymp R$ the bound is therefore $\Theta(\sqrt{dN})$, larger than the geometric bound by the factor $\sqrt{N/R} = \sqrt{(m-r)/r}$. The conversion through the ambient ℓ_∞ minimum [Pei08, Lemma 3.5] is worse still at this error: a construction-A argument with one scalar witness per column certifies an ℓ_∞ minimum only of order $1/\sqrt{R + \ln m}$, and the lemma multiplies this by $\sqrt{\ln(1/\varepsilon)}$ instead of adding the two quantities under one square root, which is the shape of the first bound of [KNSW20] recorded in Section 7.2. Over $\mathbb{Q}$ the effect is the factor $\sqrt{\ln r}$ in Table 4. The direct theta estimate never estimates a first minimum, and this is where the saving comes from. In the width, the theorems form a ladder: the dual-minimum theorem of [KNSW20] needs $s \geq 20\sqrt{r}$ over $\mathbb{Q}$ (Section 7.2), Theorem 5.6 needs $s_- \geq 8\sqrt{R}$, and Theorem 5.9 needs $s_- \geq 1$ in power-of-two fields.

6 Gaussian Leftover Hash Lemmas

Throughout this section, the *geometric event* means the event of Theorem 5.6, with its error $\delta = 2^{-\lambda-4}$ and bound $B = B_{\mathrm{geo}}$, or the event of Theorem 5.9, with $\delta = 2^{-\lambda-6}$ and $B = B_{\mathrm{cw}}$. On either event $\boldsymbol{X}$ is surjective, $\eta_{2\delta}(L_{\boldsymbol{X}}) \leq B$, and every coefficient column of $\boldsymbol{A}_{\boldsymbol{X}}$ has norm at most B_0, where $B_0 = \mu_K U$ in the first case and $B_0 = U = s_-\sqrt{H}$ in the second.

6.1 Anisotropic and spherical leftover hash lemmas

Corollary 6.1 (Geometric Gaussian LHL, ellipsoidal target). *Assume the hypotheses of Theorem 5.6 or of Theorem 5.9, let $\boldsymbol{S} \in \mathbb{R}^{M\times M}$ be positive definite with $s_{\min}(\boldsymbol{S}) \geq B$, and let $\boldsymbol{c} \in \mathcal{O}_K^m$ be*

fixed or any prescribed function of $\boldsymbol{X}$*. For every value of* $\boldsymbol{X}$ *define the total conditional law*

$$\mathcal{Q}_{\boldsymbol{X}} := \begin{cases} D_{\mathcal{O}_K^r, \sqrt{\widetilde{\Phi}(\boldsymbol{X})\boldsymbol{S}\boldsymbol{S}^T\widetilde{\Phi}(\boldsymbol{X})^T}, \boldsymbol{X}\boldsymbol{c}} & \text{if } \boldsymbol{X}\mathcal{O}_K^m = \mathcal{O}_K^r, \\ D_{\mathcal{O}_K^r, 1} & \text{otherwise.} \end{cases}$$

Then the joint distributions $(\boldsymbol{X}, \boldsymbol{X} \cdot D_{\mathcal{O}_K^m, \boldsymbol{S}, \boldsymbol{c}})$ *and* $(\boldsymbol{X}, \mathcal{Q}_{\boldsymbol{X}})$ *have statistical distance at most*

$$3\delta + \frac{2\delta}{1-2\delta} \;<\; 2^{-\lambda}.$$

In the family of Theorem 5.8 it suffices that $s_{\min}(\boldsymbol{S}) \geq C\sqrt{\lambda}$*, and for power-of-two cyclotomics at coefficient width one it suffices that* $s_{\min}(\boldsymbol{S}) \geq 640\sqrt{d(dr + \ln(2m/\delta))}$*.*

Proof. On the geometric event $\boldsymbol{X}$ is surjective and $\eta_{2\delta}(L_{\boldsymbol{X}}) \leq B \leq s_{\min}(\boldsymbol{S})$, so Lemma 3.4 bounds the conditional distance by $2\delta/(1-2\delta)$, uniformly in the center. Use the trivial bound one on the complementary event of probability at most 3δ. Since $\delta \leq 2^{-5}$, $3\delta + 2\delta/(1-2\delta) < 5.2\delta < 2^{-\lambda}$. □

The target depends on $\boldsymbol{X}$. This is the Gaussian analogue of leftover hashing. The output law is compared with the explicit conditional law suggested by the real linear map, and no claim is made that the output is independent of the matrix. If the same good matrix is used for J independent Gaussian inputs, the conditional errors add and the bad-matrix probability is paid once, giving $3\delta + 2J\delta/(1-2\delta)$.

Remark (Gaussian shaping). For a full-row-rank $\boldsymbol{A} \in \mathbb{R}^{p \times q}$ with $1 \leq p \leq q$ and a prescribed output width $w > 0$, put $\boldsymbol{A}^\dagger = \boldsymbol{A}^T(\boldsymbol{A}\boldsymbol{A}^T)^{-1}$ and $\boldsymbol{P} = \boldsymbol{I}_q - \boldsymbol{A}^\dagger\boldsymbol{A}$. The shaped parameter is

$$\boldsymbol{S}_{\boldsymbol{A}} = \sqrt{\boldsymbol{\Sigma}_{\boldsymbol{A}}}, \qquad \boldsymbol{\Sigma}_{\boldsymbol{A}} = w^2\left(\boldsymbol{A}^\dagger(\boldsymbol{A}^\dagger)^T + \frac{\operatorname{tr}((\boldsymbol{A}\boldsymbol{A}^T)^{-1})}{p}\boldsymbol{P}\right).$$

It is positive definite and satisfies $\boldsymbol{A}\boldsymbol{S}_{\boldsymbol{A}}\boldsymbol{S}_{\boldsymbol{A}}^T\boldsymbol{A}^T = w^2\boldsymbol{I}_p$, $s_{\min}(\boldsymbol{S}_{\boldsymbol{A}}) = w/s_{\max}(\boldsymbol{A})$ and $s_{\max}(\boldsymbol{S}_{\boldsymbol{A}}) = w/s_{\min}(\boldsymbol{A})$. The numerical shaping approximation and its polynomial-cost guarantees for finitely encoded inputs are treated in the accompanying Lean formalization.

Corollary 6.2 (Spherical output by changing the input parameter). *Assume the hypotheses of Theorem 5.6 or of Theorem 5.9, put* $V_0 := \kappa_K B_0\sqrt{M}$ *and* $w := BV_0$*. On the geometric event take the shaped parameter* $\boldsymbol{S}_{\boldsymbol{X}} = \sqrt{\boldsymbol{\Sigma}_{\widetilde{\Phi}(\boldsymbol{X})}}$ *with output width* w*. On the complementary event put* $\boldsymbol{S}_{\boldsymbol{X}} = \boldsymbol{I}_M$*. Then for every* $\boldsymbol{c} \in \mathcal{O}_K^m$ *the joint law of* $(\boldsymbol{X}, \boldsymbol{X} \cdot D_{\mathcal{O}_K^m, \boldsymbol{S}_{\boldsymbol{X}}, \boldsymbol{c}})$ *is within* $3\delta + 2\delta/(1-2\delta) < 2^{-\lambda}$ *of the law that samples* $\boldsymbol{X}$ *and then* $D_{\mathcal{O}_K^r, w, \boldsymbol{X}\boldsymbol{c}}$*. More generally, for any prescribed width* $w > 0$*, the same construction and joint-distance bound apply if* $w \geq B\, s_{\max}(\widetilde{\Phi}(\boldsymbol{X}))$ *for every* $\boldsymbol{X}$ *in the geometric event.*

For power-of-two cyclotomics, fix $0 < \delta_{\mathrm{sp}} < 1$ *and assume in addition that*

$$m \geq c_{\mathrm{sp}}\big(r + \ln(2d/\delta_{\mathrm{sp}})\big).$$

Let E_{geo} *be the geometric event and* E_{sp} *the spectral event of Lemma 3.8 with error* δ_{sp}*. Fix any width* $w \geq c''_{\mathrm{sp}} s\sqrt{m}\, B$*, independently of the sampled matrix. Use the same shaped parameter on* $E_{\mathrm{geo}} \cap E_{\mathrm{sp}}$ *and put* $\boldsymbol{S}_{\boldsymbol{X}} = \boldsymbol{I}_M$ *on its complement. With this prescription, the same two joint laws have statistical distance at most*

$$3\delta + \delta_{\mathrm{sp}} + \frac{2\delta}{1-2\delta}.$$

Taking $\delta_{\mathrm{sp}} = \delta$ *gives* $4\delta + 2\delta/(1-2\delta) < 2^{-\lambda}$ *for both choices of the geometric event.*

Proof. By Equation (20), $s_{\max}(\widetilde{\Phi}(\boldsymbol{X})) \leq \alpha_K\beta_K\,||\boldsymbol{A}_{\boldsymbol{X}}||_{\mathrm{op}} \leq \kappa_K\,||\boldsymbol{A}_{\boldsymbol{X}}||_F$, and on the geometric event every one of the M columns of $\boldsymbol{A}_{\boldsymbol{X}}$ has norm at most B_0, so $s_{\max}(\widetilde{\Phi}(\boldsymbol{X})) \leq \kappa_K B_0\sqrt{M} = V_0$. By construction, $s_{\min}(\boldsymbol{S}_{\boldsymbol{X}}) = w/s_{\max}(\widetilde{\Phi}(\boldsymbol{X})) \geq B \geq \eta_{2\delta}(L_{\boldsymbol{X}})$ and $\widetilde{\Phi}(\boldsymbol{X})\boldsymbol{S}_{\boldsymbol{X}}\boldsymbol{S}_{\boldsymbol{X}}^T\widetilde{\Phi}(\boldsymbol{X})^T = w^2\boldsymbol{I}_R$, so Lemma 3.4 gives the conditional bound $2\delta/(1-2\delta)$, and the exceptional event is paid once. For the spectral variant, both geometric theorems give $s/\sqrt{d} \geq 1$ in the power-of-two case. Together with the additional column condition, Lemma 3.8 therefore gives $s_{\max}(\widetilde{\Phi}(\boldsymbol{X})) \leq c''_{\mathrm{sp}} s\sqrt{m}$ on E_{sp}. The same conditional argument applies on $E_{\mathrm{geo}} \cap E_{\mathrm{sp}}$. A union bound gives $\Pr[(E_{\mathrm{geo}} \cap E_{\mathrm{sp}})^c] \leq 3\delta + \delta_{\mathrm{sp}}$, without any independence assumption, and the distance is at most one on this complement. Finally, $\delta \leq 2^{-5}$ implies $4\delta + 2\delta/(1-2\delta) < 6.2\delta < 2^{-\lambda}$. □

The construction shapes the input parameter so that the output is spherical. It changes the input law. Whether that change is admissible depends on the protocol, and a spherical input with a spherical target is a different question, which [OFKT23] and [AFL+26, Section 6] address with Rényi-divergence estimates.

Corollary 6.3 (k simultaneous hints). *In the deterministic-V_0 version of Corollary 6.2, set $r = k$ and let w and $\boldsymbol{S}_{\boldsymbol{X}}$ be as there. Conditional on $\boldsymbol{X}$, sample the columns of $\boldsymbol{R} \in \mathcal{O}_K^{m\times k}$ independently from $D_{\mathcal{O}_K^m,\boldsymbol{S}_{\boldsymbol{X}}}$ and put $\boldsymbol{X}_2 = \boldsymbol{X}\boldsymbol{R}+\boldsymbol{I}_k$. Then $(\boldsymbol{X}, \boldsymbol{X}_2)$ is within $3\delta+2k\delta/(1-2\delta)$ of the distribution consisting of m independent columns from $D_{\mathcal{O}_K^k,s}$ followed by independent columns $D_{\mathcal{O}_K^k,w,\boldsymbol{e}_i}$ for $i \in [k]$. For the spectral version of Corollary 6.2, with its additional column condition at $r = k$, prescribed width w, and identity fallback outside $E_{\mathrm{geo}} \cap E_{\mathrm{sp}}$, the same comparison has distance at most*

$$3\delta + \delta_{\mathrm{sp}} + \frac{2k\delta}{1-2\delta}.$$

In particular, $\delta_{\mathrm{sp}} = \delta$ gives $4\delta + 2k\delta/(1-2\delta)$.

Proof. Condition on $\boldsymbol{X}$ in the geometric event. The i-th column of $\boldsymbol{X}_2$ is $\boldsymbol{X}\boldsymbol{r}_i + \boldsymbol{e}_i$ with $\boldsymbol{r}_i \leftarrow D_{\mathcal{O}_K^m,\boldsymbol{S}_{\boldsymbol{X}}}$, and the proof of Corollary 6.2 shows that $\boldsymbol{X}\boldsymbol{r}_i$ is within $2\delta/(1-2\delta)$ of $D_{\mathcal{O}_K^k,w}$, so that $\boldsymbol{X}\boldsymbol{r}_i + \boldsymbol{e}_i$ is within the same distance of $D_{\mathcal{O}_K^k,w,\boldsymbol{e}_i}$. Statistical distance tensorizes by the triangle inequality, so the k independent columns contribute $2k\delta/(1-2\delta)$. Average over $\boldsymbol{X}$, counting the complementary event of probability at most 3δ once. For the spectral version, condition on $E_{\mathrm{geo}} \cap E_{\mathrm{sp}}$ instead. The k conditional errors still add, and the probability of its complement is at most $3\delta + \delta_{\mathrm{sp}}$, paid once for the entire matrix of hints. □

6.2 Parameters of the SIS-to-k-SIS reduction

Albrecht, Felderhoff, Lai, Lapiha and Woo [AFL+26, Section 7] reduce SIS to k-SIS by generating k Gaussian kernel hints. Their algebraic reduction needs only an appropriate hint generator, and the analytic input is used to show that its output has the prescribed Gaussian law. We now build such a generator with controlled operator norm. Throughout, K is a power-of-two cyclotomic field of degree d, $\boldsymbol{X}_1 \leftarrow D^m_{\mathcal{O}_K^k,s_1}$, and $B(\varepsilon)$ denotes any bound such that $\eta_\varepsilon(L_{\boldsymbol{X}_1}) \le B(\varepsilon)$ and $\boldsymbol{X}_1$ is surjective on an event E_B. The geometric events supply two choices, $B = B_{\mathrm{geo}}$ on the event of Theorem 5.6 and $B = B_{\mathrm{cw}}$ on the event of Theorem 5.9, both with failure probability $p_B = 3\delta$ and LHL error $\varepsilon_{\mathrm{LHL}} = 2\delta$. The proofs use only these two properties of B and the spectral estimates of Section 3.6.

What changes in the original reduction. The reduction of [AFL+26, Theorem 7.3] first simulates a public matrix with shifted Gaussian kernel hints, compares the shifted and centered games, and extracts a short SIS solution. We keep this construction and its modular arithmetic conditions. Three points in its proof require care. First, the norm used in [AFL+26, Lemma 7.1] is a maximum column norm, whereas extraction needs an operator norm. The spectral hint lemma below supplies that bound. Second, the nonzero-extraction argument in [AFL+26, Lemma 7.6] passes from a modular span to a field span. Lemma 6.5 replaces this step by an exact kernel identity and also proves exact public-matrix simulation. Third, the smoothing-error quantifier in the proof of [AFL+26, Lemma 7.5] is too strong. Lemma 6.6 gives the required comparison without smoothing. We also give the fieldwise residue-mass argument needed by the imported independence estimate. The other modular and hybrid steps are recalled from [AFL+26, Lemmas 2.14–2.16 and Theorem 7.3] with their stated hypotheses. The displayed block construction also fixes the sign of the lower block of $\boldsymbol{V}$ and uses unimodularity. The stronger claim $\det \boldsymbol{U} = 1$ is not needed, and it fails for some dimensions.

For a matrix over $\mathcal{O}_K$, $\widetilde{\Phi}(\cdot)$ denotes its real canonical matrix, and when the matrix acts on a solution vector we use the operator norm of this real matrix. This distinction is important: the matrix norm in [AFL+26] is the maximum column norm, which alone incurs a $\sqrt{m}$ loss when converted to an operator bound.

Lemma 6.4 (Spectral hint generator). *Assume the hypotheses of Lemma 3.8 with error δ_{sp} and let E_B have failure probability p_B. Let $0 < \varepsilon_{\mathrm{samp}}, \delta_{\mathrm{op}} < 1$, put $r_{\mathrm{samp}} := C_{\mathrm{samp}}\sqrt{\ln(dm/\varepsilon_{\mathrm{samp}})}$ for an absolute constant C_{samp}, and suppose $m \ge k + \ln(2d/\delta_{\mathrm{op}})$ and*

$$s_2 \ge \max\left\{s_1,\ c''_{\mathrm{sp}} s_1 \sqrt{m}\, B(\varepsilon_{\mathrm{LHL}})\right\}, \qquad B(\varepsilon_{\mathrm{LHL}}) \ge 4r_{\mathrm{samp}}\sqrt{d}. \tag{47}$$

There is a probabilistic polynomial-time procedure that outputs $(\boldsymbol{X}_1, \boldsymbol{X}_2) \in \mathcal{O}_K^{k\times m} \times \mathcal{O}_K^{k\times k}$ *and a unimodular* $\boldsymbol{U} \in \mathcal{O}_K^{(m+k)\times(m+k)}$ *with the following three properties.*

(i) The pair $(\boldsymbol{X}_1, \boldsymbol{X}_2)$ *is within statistical distance* $p_B + \delta_{\mathrm{sp}} + k\varepsilon_{\mathrm{LHL}}/(1-\varepsilon_{\mathrm{LHL}}) + k\varepsilon_{\mathrm{samp}}$ *of* m *independent columns from* $D_{\mathcal{O}_K^k, s_1}$ *followed by independent columns* $D_{\mathcal{O}_K^k, s_2, \boldsymbol{e}_i}$, $i \in [k]$.

(ii) The last m *columns* $\boldsymbol{V}$ *of* $\boldsymbol{U}$ *satisfy* $(\boldsymbol{X}_1, \boldsymbol{X}_2)\boldsymbol{V} = \boldsymbol{0}$.

(iii) Except with probability $\delta_{\mathrm{op}} + k\varepsilon_{\mathrm{samp}}$ *in addition to the preceding bad events,* $s_{\max}(\widetilde{\Phi}(\boldsymbol{V}^T)) \le C_U s_2 \sqrt{m}$ *for an absolute constant* C_U.

Proof. We first describe the procedure. Define the exact target $\boldsymbol{S}^\star_{\boldsymbol{X}_1}$ by the shaping formula with $\boldsymbol{A} = \widetilde{\Phi}(\boldsymbol{X}_1)$ and output width s_2 when $\widetilde{\Phi}(\boldsymbol{X}_1)$ has full row rank, and by $4r_{\mathrm{samp}}\sqrt{d}\,\boldsymbol{I}_{dm}$ otherwise. Use a positive-definite approximation $\widehat{\boldsymbol{S}}_{\boldsymbol{X}_1}$, accepted only when Lemma 3.5 certifies

$$\mathrm{SD}\big(D_{\mathcal{O}_K^m, \boldsymbol{S}^\star_{\boldsymbol{X}_1}}, D_{\mathcal{O}_K^m, \widehat{\boldsymbol{S}}_{\boldsymbol{X}_1}}\big) \le \varepsilon_{\mathrm{samp}}/2$$

and its smallest singular value is certified to be at least $2r_{\mathrm{samp}}\sqrt{d}$. If either certificate fails, use the fallback $\widehat{\boldsymbol{S}}_{\boldsymbol{X}_1} = 2r_{\mathrm{samp}}\sqrt{d}\,\boldsymbol{I}_{dm}$. Sample the columns of $\boldsymbol{R} \in \mathcal{O}_K^{m\times k}$ independently, within distance $\varepsilon_{\mathrm{samp}}/2$ per column, from $D_{\mathcal{O}_K^m, \widehat{\boldsymbol{S}}_{\boldsymbol{X}_1}}$, and set

$$\boldsymbol{X}_2 = \boldsymbol{X}_1\boldsymbol{R} + \boldsymbol{I}_k, \qquad \boldsymbol{U} = \begin{pmatrix} -\boldsymbol{R} & -\boldsymbol{I}_m - \boldsymbol{R}\boldsymbol{X}_1 \\ \boldsymbol{I}_k & \boldsymbol{X}_1 \end{pmatrix}.$$

Let $\boldsymbol{V} \in \mathcal{O}_K^{(m+k)\times m}$ be the last m columns of $\boldsymbol{U}$.

On the intersection of E_B and the spectral event of Lemma 3.8, the two certificates pass. Thus Lemma 3.5 and the sampler guarantee that an implemented column can be coupled to a column from $D_{\mathcal{O}_K^m, \boldsymbol{S}^\star_{\boldsymbol{X}_1}}$ with failure probability at most $\varepsilon_{\mathrm{samp}}$. Item (i) now follows from the argument of Corollaries 6.2 and 6.3, in which the geometric event is replaced by E_B and the bound there by B, together with a k-column hybrid for the sampling error. For item (ii), factor

$$\boldsymbol{U} = \begin{pmatrix} \boldsymbol{I}_m & -\boldsymbol{R} \\ \boldsymbol{0} & \boldsymbol{I}_k \end{pmatrix} \begin{pmatrix} \boldsymbol{0} & -\boldsymbol{I}_m \\ \boldsymbol{I}_k & \boldsymbol{X}_1 \end{pmatrix}.$$

Both factors are unimodular, and $\det \boldsymbol{U} = (-1)^{m(k+1)} \in \{\pm 1\}$. Direct multiplication gives $(\boldsymbol{X}_1, \boldsymbol{X}_2)\boldsymbol{U} = (\boldsymbol{I}_k, \boldsymbol{0}_{k\times m})$, which proves the kernel identity.

For item (iii), the shaping identity and Equation (19) give

$$s_{\max}(\boldsymbol{S}^\star_{\boldsymbol{X}_1}) = \frac{s_2}{s_{\min}(\widetilde{\Phi}(\boldsymbol{X}_1))} \le \frac{s_2}{c'_{\mathrm{sp}} s_1 \sqrt{m}}.$$

Apply Lemma 3.7 with $\tau = \delta_{\mathrm{op}}$, conditionally on $\boldsymbol{X}_1$, to ideal columns from the exact target distribution, and transfer the resulting event to the implemented columns by the preceding coupling, at cost at most $k\varepsilon_{\mathrm{samp}}$. Since $s_{\max}(\widetilde{\Phi}(\boldsymbol{X}_1)) = O(s_1\sqrt{m})$ and $s_2 \ge s_1$, submultiplicativity gives

$$\begin{aligned} s_{\max}(\widetilde{\Phi}(\boldsymbol{V}^T)) = s_{\max}(\widetilde{\Phi}(\boldsymbol{V})) &\le 1 + s_{\max}(\widetilde{\Phi}(\boldsymbol{R}))s_{\max}(\widetilde{\Phi}(\boldsymbol{X}_1)) + s_{\max}(\widetilde{\Phi}(\boldsymbol{X}_1)) \\ &\le C_U s_2 \sqrt{m}. \end{aligned}$$

For each complex block $\boldsymbol{A}$, $\left\|\boldsymbol{A}^T\right\|_2 = \left\|\overline{\boldsymbol{A}^*}\right\|_2 = \|\boldsymbol{A}\|_2$, so transposition preserves the singular values relevant to the canonical real matrices. This direct operator estimate also repairs two gaps in [AFL+26, Lemma 7.1]: its displayed Ω must point in the upper-bound direction, and its max-column bound does not by itself control multiplication by a solution vector.

For the power basis, the canonical basis matrix $\boldsymbol{B}_1$ of $\Phi(\mathcal{O}_K^m)$ satisfies $\boldsymbol{B}_1\boldsymbol{B}_1^T = d\boldsymbol{I}_{dm}$, so $\widehat{\boldsymbol{S}}_{\boldsymbol{X}_1}\widehat{\boldsymbol{S}}_{\boldsymbol{X}_1}^T \succ r_{\mathrm{samp}}^2 \boldsymbol{B}_1\boldsymbol{B}_1^T$ on both branches. The generic ellipsoidal sampler [Pei10, Algorithm 1 and Theorem 4.2], with the tail parameter in its proof made explicit, consequently samples each column within $\varepsilon_{\mathrm{samp}}/2$. The factor-four margin in Equation (47) leaves the certified factor-two sampling margin after approximation. On the good event, a sufficiently accurate approximation satisfies both certificates. This is also the Gaussian-shape extension contemplated in [AFL+26, Remark 7.4]. □

Exact algebra in the reduction. Write $\boldsymbol{H} = (\boldsymbol{X}_1, \boldsymbol{X}_1\boldsymbol{R}+\boldsymbol{I}_k)^T$ for the column hints and $\boldsymbol{E} = \boldsymbol{V}^T$ for the extraction map. The next lemma repairs the nonzero-extraction argument in [AFL+26, Lemma 7.6] and makes the last hybrid in [AFL+26, Lemma 7.2] exact.

Lemma 6.5 (Exact extraction and uniform simulation). *Let $\mathcal{R}$ be a commutative ring, $\boldsymbol{X}_1 \in \mathcal{R}^{k\times m}$ and $\boldsymbol{R} \in \mathcal{R}^{m\times k}$, and put*

$$\boldsymbol{H} = \begin{pmatrix} \boldsymbol{X}_1^T \\ \boldsymbol{I}_k + \boldsymbol{R}^T\boldsymbol{X}_1^T \end{pmatrix}, \qquad \boldsymbol{E} = \begin{pmatrix} -\boldsymbol{I}_m - \boldsymbol{X}_1^T\boldsymbol{R}^T & \boldsymbol{X}_1^T \end{pmatrix}.$$

Then $\ker \boldsymbol{E} = \operatorname{Im} \boldsymbol{H}$. Moreover, for every n, the map

$$\mathcal{R}^{n\times m} \longrightarrow \left\{ \boldsymbol{B} \in \mathcal{R}^{n\times(m+k)} : \boldsymbol{B}\boldsymbol{H} = \boldsymbol{0} \right\}, \qquad \boldsymbol{A} \longmapsto \boldsymbol{A}\boldsymbol{E}$$

is a bijection. In particular, when $\mathcal{R}$ is finite, it sends the uniform distribution to the uniform distribution on the displayed set.

Proof. Multiplication gives $\boldsymbol{E}\boldsymbol{H} = \boldsymbol{0}$. If $\boldsymbol{E}\binom{\boldsymbol{a}}{\boldsymbol{b}} = \boldsymbol{0}$, set $\boldsymbol{c} = \boldsymbol{b} - \boldsymbol{R}^T\boldsymbol{a}$. The equation says $\boldsymbol{a} = \boldsymbol{X}_1^T\boldsymbol{c}$, and hence $\boldsymbol{b} = (\boldsymbol{I}_k + \boldsymbol{R}^T\boldsymbol{X}_1^T)\boldsymbol{c}$. Thus $\binom{\boldsymbol{a}}{\boldsymbol{b}} = \boldsymbol{H}\boldsymbol{c}$, proving the kernel identity over $\mathcal{R}$. For the bijection, define

$$\boldsymbol{J} = \begin{pmatrix} -\boldsymbol{I}_m \\ -\boldsymbol{R}^T \end{pmatrix}, \qquad \boldsymbol{L} = \begin{pmatrix} -\boldsymbol{R}^T & \boldsymbol{I}_k \end{pmatrix}.$$

Direct multiplication gives $\boldsymbol{E}\boldsymbol{J} = \boldsymbol{I}_m$ and $\boldsymbol{J}\boldsymbol{E} + \boldsymbol{H}\boldsymbol{L} = \boldsymbol{I}_{m+k}$. The inverse map is $\boldsymbol{B} \mapsto \boldsymbol{B}\boldsymbol{J}$: if $\boldsymbol{B}\boldsymbol{H} = \boldsymbol{0}$, then $\boldsymbol{B} = \boldsymbol{B}\boldsymbol{J}\boldsymbol{E}$, while $(\boldsymbol{A}\boldsymbol{E})\boldsymbol{J} = \boldsymbol{A}$. □

Apply the lemma over K to see that a vector outside the K-span of the hint columns has nonzero image under $\boldsymbol{E}$. This is an exact statement, with no exceptional modular-dependence event. In particular, it avoids the implication from membership in a span modulo q to membership in a K-linear span used in [AFL+26, Lemma 7.6]. Apply the same lemma over $\mathcal{O}_K/q\mathcal{O}_K$ to obtain the exact final public-matrix simulation.

The centered-to-shifted comparison. The second repair concerns [AFL+26, Lemma 7.5]. Its proof invokes a fixed-width smoothing guarantee for every sufficiently small error, whereas [AFL+26, Lemma 2.16] supplies a particular error. The following direct calculation removes the smoothing condition from this comparison altogether.

Lemma 6.6 (Gaussian shift comparison). *Let $L \subset \mathbb{R}^N$ be a full-rank lattice, let $\boldsymbol{S} \in \mathbb{R}^{N\times N}$ be nonsingular, and let $\boldsymbol{c} \in \mathbb{R}^N$. Put $P_0 = D_{L,\boldsymbol{S}}$, $P_c = D_{L,\boldsymbol{S},\boldsymbol{c}}$ and $Z(\boldsymbol{c}) = \rho_{\boldsymbol{S},\boldsymbol{c}}(L)$. Then*

$$\sum_{\boldsymbol{x}\in L} \frac{P_0(\boldsymbol{x})^2}{P_c(\boldsymbol{x})} = e^{2\pi||\boldsymbol{S}^{-1}\boldsymbol{c}||^2} \frac{Z(\boldsymbol{c})Z(-\boldsymbol{c})}{Z(\boldsymbol{0})^2} \le e^{2\pi||\boldsymbol{S}^{-1}\boldsymbol{c}||^2}.$$

Consequently, for any function $f : L \to [0,1]$, representing the success probability of a possibly randomized test,

$$\mathbb{E}_{P_c} f \;\ge\; e^{-2\pi||\boldsymbol{S}^{-1}\boldsymbol{c}||^2} (\mathbb{E}_{P_0} f)^2.$$

Proof. Completing the square gives

$$\frac{\rho_{\boldsymbol{S}}(\boldsymbol{x})^2}{\rho_{\boldsymbol{S},\boldsymbol{c}}(\boldsymbol{x})} = e^{2\pi||\boldsymbol{S}^{-1}\boldsymbol{c}||^2} \rho_{\boldsymbol{S},-\boldsymbol{c}}(\boldsymbol{x}).$$

Summing and inserting the normalizing factors proves the identity. Poisson summation expresses $Z(\boldsymbol{c})$ as a sum of nonnegative dual Gaussian weights multiplied by complex numbers of modulus one, so $Z(\boldsymbol{c}) \le Z(\boldsymbol{0})$. The same holds for $-\boldsymbol{c}$. Finally, Cauchy–Schwarz and $0 \le f \le 1$ give

$$(\mathbb{E}_{P_0} f)^2 \le \mathbb{E}_{P_c} f \cdot \sum_{\boldsymbol{x}\in L} \frac{P_0(\boldsymbol{x})^2 f(\boldsymbol{x})}{P_c(\boldsymbol{x})} \le \mathbb{E}_{P_c} f \cdot e^{2\pi||\boldsymbol{S}^{-1}\boldsymbol{c}||^2}.$$

□

For a public matrix $\boldsymbol{B}$, define its embedded modular kernel by

$$L_{\boldsymbol{B},q} := \Phi\left\{\boldsymbol{z} \in \mathcal{O}_K^{m+k} : \boldsymbol{B}\boldsymbol{z} = \boldsymbol{0} \pmod{q}\right\}.$$

Use $\boldsymbol{S} = \operatorname{diag}(s_1\boldsymbol{I}_{dm}, s_2\boldsymbol{I}_{dk})$ and centers $\Phi(\boldsymbol{e}_{m+i})$, $i \in [k]$, where $\boldsymbol{e}_{m+i}$ is the corresponding unit vector of $\mathcal{O}_K^{m+k}$. The moment in the lemma multiplies over independent columns. The total shift energy is $\sum_{i\in[k]} \left|\left|\boldsymbol{S}^{-1}\Phi(\boldsymbol{e}_{m+i})\right|\right|^2 = dk/s_2^2$, since $||\Phi(1)||^2 = d$ and the hint block of $\boldsymbol{S}$ is $s_2\boldsymbol{I}_{dk}$. The arithmetic window below implies $s_2^2 \geq dk$, so the exponent $2\pi\, dk/s_2^2$ is at most 2π. The bound is uniform in $\boldsymbol{B}$. Averaging over its uniform law and using Jensen's inequality therefore gives

$$\mathrm{Adv}_{\mathcal{B}}^{\text{Shift-}k\text{-SIS}} \geq e^{-2\pi}\left(\mathrm{Adv}_{\mathcal{B}}^{k\text{-SIS}}\right)^2. \tag{48}$$

Here Shift-k-SIS is the intermediate game of [AFL+26, Theorem 7.3]: its public matrix and winning predicate are unchanged, and only the i-th hint center changes from zero to $\boldsymbol{e}_{m+i}$. This comparison includes the adversary's randomness and incurs no smoothing-failure loss.

The modular estimates used by reverse sampling. We retain the reverse-sampling argument of [AFL+26, Lemma 2.14], using its independence and regularity inputs from [AFL+26, Lemmas 2.15 and 2.16]. There is a point to repair in the supporting residue-mass proof [ALLW25, Lemma 25]: full rank separately over the residue fields does not imply a common invertible minor over their product. For example, the column with entries $(1, 0)$ and $(0, 1)$ over $F \times F$ has rank one in each factor, but neither entry is a unit. The independence estimate needs only the prime-ideal case. The following version supplies it and keeps track of the ambient dimension.

Lemma 6.7 (Prime-ideal residue mass). *Let $m \geq 64k \geq 64$, put $M = m+k$, and let $\mathfrak{p}$ be a nonzero prime ideal of $\mathcal{O}_K$, with $N\mathfrak{p} := |\mathcal{O}_K/\mathfrak{p}|$. For $\boldsymbol{W} \in (\mathcal{O}_K/\mathfrak{p})^{M\times k}$, set*

$$L_{\boldsymbol{W},\mathfrak{p}} := \Phi\left\{\boldsymbol{z} \in \mathcal{O}_K^M : \boldsymbol{z} \bmod \mathfrak{p} \in \operatorname{Im}\boldsymbol{W}\right\}.$$

If $\boldsymbol{S} \in \mathbb{R}^{Md\times Md}$ is nonsingular, $\sigma > 0$, $s_{\max}(\boldsymbol{S}) \leq \sigma$ and $\sigma\sqrt{m} \leq (N\mathfrak{p})^{1/d}$, then

$$\rho_{\boldsymbol{S}}(L_{\boldsymbol{W},\mathfrak{p}}) \leq \frac{\rho_\sigma(\mathcal{O}_K^k)}{1 - e^{-md/32}}.$$

Proof. Let $r = \operatorname{rank}\boldsymbol{W} \leq k$. Since $\mathcal{O}_K/\mathfrak{p}$ is a field, some projection onto r coordinates is injective on $\operatorname{Im}\boldsymbol{W}$. Every nonzero $a \in \mathfrak{p}$ satisfies $||\Phi(a)|| \geq \sqrt{d}\,(N\mathfrak{p})^{1/d}$: the ideal norm divides $|N_{K/\mathbb{Q}}(a)|$, and the bound follows by arithmetic–geometric mean over the embeddings. Put $R_0 = \sigma\sqrt{md}/2$. The selected coordinate projection is injective on $L_{\boldsymbol{W},\mathfrak{p}} \cap \{\boldsymbol{x} : ||\boldsymbol{x}|| < R_0\}$. Indeed, two points with the same selected coordinates have difference in $\Phi(\mathfrak{p}^M)$. A nonzero difference would have length at least $\sqrt{d}\,(N\mathfrak{p})^{1/d} \geq 2R_0$, whereas their distance is strictly less than $2R_0$.

For every point, $\rho_{\boldsymbol{S}}(\boldsymbol{x}) \leq e^{-\pi||\boldsymbol{x}||^2/\sigma^2}$, which is at most the spherical weight of its projection. Injectivity therefore bounds the mass in this open ball by $\rho_\sigma(\mathcal{O}_K^r) \leq \rho_\sigma(\mathcal{O}_K^k)$. The lattice $L_{\boldsymbol{W},\mathfrak{p}}$ contains $\Phi(\mathfrak{p}^M)$ and is therefore full rank. For $\boldsymbol{Y} \leftarrow D_{L_{\boldsymbol{W},\mathfrak{p}},\boldsymbol{S}}$, exponential Markov and Poisson summation give

$$\begin{aligned}\Pr[||\boldsymbol{Y}|| \geq R_0] &\leq e^{-\pi R_0^2/(2\sigma^2)}\frac{\rho_{\sqrt{2}\boldsymbol{S}}(L_{\boldsymbol{W},\mathfrak{p}})}{\rho_{\boldsymbol{S}}(L_{\boldsymbol{W},\mathfrak{p}})} \\ &\leq 2^{Md/2}e^{-\pi md/8} \leq e^{-md/32}.\end{aligned}$$

The middle inequality follows by comparing the nonnegative dual Gaussian sums, and the last one uses $k \leq m/64$. Thus at least a $1 - e^{-md/32}$ fraction of the total mass lies in the open ball, proving the claim. The boundary is included in the tail. □

Apply this estimate separately to each prime factor of $q\mathcal{O}_K$. A dependence test conditions on previously sampled columns and tests a fresh column against their span. The tested column must not itself be included in that span. When a left inverse over $\mathcal{O}_K/q\mathcal{O}_K$ is needed, construct it in each residue field and combine the entries by the Chinese remainder theorem. These arguments avoid a common-minor assumption. The width ceiling remains $s_2\sqrt{m} \leq q^{1/g}$ even though the hints have $m + k$ coordinates, as the lemma explicitly shows.

For regularity, the distinction between m and $M = m+k$ also matters. Here is the quantitative form of the uniform-matrix estimate used below. Under the arithmetic and dimension conditions of Theorem 6.8, write $n = n_{\mathrm{SIS}}$ and put

$$t = 8d\sqrt{m}\, q^{n/m+2/(dm)}, \qquad \varepsilon_0 = 2^{-\lfloor dm/4 \rfloor}, \qquad C = 1 + 2^{1-3dm}.$$

For uniform $\boldsymbol{B} \in (\mathcal{O}_K/q\mathcal{O}_K)^{n\times M}$, the coset-averaging argument underlying [AFL+26, Lemma 2.16], keeping the M columns explicit, gives

$$\begin{aligned}\mathbb{E}_{\boldsymbol{B}}\rho_{1/t}(L^*_{\boldsymbol{B},q} \setminus \{\mathbf{0}\}) &\le (C^M - 1)2^g + q^{dn} C^M \left(\frac{4d\sqrt{m}}{t}\right)^{dM} \\ &\le 2 \cdot 2^{-dm}.\end{aligned}$$

In this calculation the zero vector is removed from every dual coset before averaging, so repeated representatives for a rank-deficient public matrix cause no problem. Grouping the residue frequencies by their image ideals contributes 2^g, the number of ideals in the product of the g residue fields. The remaining frequencies number q^{dn}. The Gaussian mass correction for each coordinate is C. For the last inequality, use $g \le d$, $m \le M \le 2dm$, $dm \ge 64$, and

$$C^M \le 2, \qquad (C^M - 1)2^g \le 2^{-dm}, \qquad q^{dn}\left(\frac{4d\sqrt{m}}{t}\right)^{dM} \le \frac{2^{-dm}}{q^2}.$$

These are the same width and mass estimates as in the regularity argument, with the column count retained explicitly. Markov's inequality and $2\cdot 2^{-dm} \le \varepsilon_0^2$ show that $\eta_{\varepsilon_0}(L_{\boldsymbol{B},q}) \le t$ except with probability at most ε_0. Moreover, $\varepsilon_0 \le 2^{-2k}$, and $dk \ge \lambda$ gives $\varepsilon_0 \le 2^{-16\lambda}$. Since $s_1, s_2 \ge t$, this also supplies the ellipsoidal smoothing condition for the diagonal hint parameter.

Thus regularity is invoked at one certified error, not at every smaller error, and for a fully uniform public matrix. A theorem for a matrix with a fixed identity block cannot be substituted directly. It is needed for reverse sampling, despite being unnecessary for Equation (48). These modular estimates, rather than the Gaussian LHL, are the reason for retaining the arithmetic window in the next theorem.

We state the imported arithmetic conditions of the reduction explicitly, because omitting them would turn the result into only a Gaussian-component estimate.

Theorem 6.8 (SIS-to-k-SIS with spectral hints). *Let K be a power-of-two cyclotomic field of degree d, let q be an unramified rational prime that splits into g prime ideals of $\mathcal{O}_K$, and let n_{SIS} be the number of rows of the source SIS instance. Suppose*

$$\begin{gathered} dk \ge \lambda, \qquad m \ge 64\max\{\lambda/\ln(dk), k\}, \qquad m \le \mathsf{poly}(d), \\ 1 \le n_{\mathrm{SIS}} \le m, \qquad d(m-k) \ge c\lambda, \qquad \frac{n_{\mathrm{SIS}}\, g}{q^{d(m-n_{\mathrm{SIS}}+1)/g}} \le \mathsf{negl}(\lambda), \end{gathered}$$

assume the hypotheses of Lemma 6.4, and suppose that for some $u \ge 1$ and $e_q := n_{\mathrm{SIS}}/m + 2/(dm)$ both s_1 and s_2 lie in the window

$$\max\left\{8d\sqrt{m}\, q^{e_q},\ 2\sqrt{d}\,(u^k q^{n_{\mathrm{SIS}}})^{1/(m-k)}\right\} \ \le\ s_i \ \le\ \min\left\{q^{1/g}/\sqrt{m},\ us_1\right\}$$

for $i = 1, 2$. If the errors of Lemma 6.4 are negligible and $\beta_0 \ge C_\beta s_2\sqrt{m}\,\beta_1$, then the reduction of [AFL+26, Theorem 7.3] with the hint generator of Lemma 6.4 maps SIS with norm bound β_0 to k-SIS with Gaussian parameter matrix $\mathrm{diag}(s_1\boldsymbol{I}_{dm}, s_2\boldsymbol{I}_{dk})$ and norm bound β_1, with $\mathrm{Adv}^{\mathrm{SIS}}_{\mathcal{A}}(\lambda) \ge (\mathrm{Adv}^{k\text{-SIS}}_{\mathcal{B}}(\lambda))^2/O(1) - \mathsf{negl}(\lambda)$.

Proof. Write δ_{hint} for the distribution error in Lemma 6.4(i), and δ_{norm} for its total operator-norm failure probability. Let δ_{rev} denote the error of the reverse-sampling comparison in [AFL+26, Lemma 2.14]. The displayed arithmetic and dimension conditions give its regularity and independence hypotheses through the estimates recalled above from [AFL+26, Lemmas 2.15 and 2.16], with the prime-ideal argument supplied in Lemma 6.7. Thus δ_{rev} is negligible. Here regularity is invoked with one admissible error, and the independence test is made over each residue field before applying a union bound over the g prime factors.

Given a uniform SIS matrix $\boldsymbol{A} \in (\mathcal{O}_K/q\mathcal{O}_K)^{n_{\mathrm{SIS}} \times m}$, run the hint generator, set $\boldsymbol{H} = (\boldsymbol{X}_1, \boldsymbol{X}_2)^T$, $\boldsymbol{E} = \boldsymbol{V}^T$, and give $(\boldsymbol{B}, \boldsymbol{H})$ to the adversary, where $\boldsymbol{B} = \boldsymbol{A}\boldsymbol{E} \bmod q$. By Lemma 6.5, conditioned on the generated hints the public matrix is exactly uniform subject to $\boldsymbol{B}\boldsymbol{H} = \boldsymbol{0}$. Consequently the hybrid argument of [AFL+26, Lemma 7.2], with Lemma 6.4 replacing its Gaussian LHL input, compares these adversary inputs with the shifted game at distance at most $\delta_{\mathrm{hint}} + \delta_{\mathrm{rev}}$.

On a valid adversarial answer $\boldsymbol{z}$, return $\boldsymbol{z}^* = \boldsymbol{E}\boldsymbol{z}$. The equation $\boldsymbol{A}\boldsymbol{z}^* = \boldsymbol{B}\boldsymbol{z} = \boldsymbol{0} \pmod q$ holds, and $\boldsymbol{z}^* \neq \boldsymbol{0}$ follows from Lemma 6.5, since a valid k-SIS answer lies outside the K-span of the hints. On the operator-norm event,

$$||\Phi(\boldsymbol{z}^*)|| \le s_{\max}(\widetilde{\Phi}(\boldsymbol{E}))\, ||\Phi(\boldsymbol{z})|| \le C_U s_2 \sqrt{m}\, ||\Phi(\boldsymbol{z})||\,.$$

Choose $C_\beta \ge C_U$. The extraction therefore preserves the required norm bound, with either strict bounds throughout or nonstrict bounds throughout. Its only additional failure probability is δ_{norm}, paid once even for an adaptive answer. Combining this with Equation (48) gives

$$\mathrm{Adv}^{\mathrm{SIS}}_{\mathcal{A}} \ge e^{-2\pi}\big(\mathrm{Adv}^{k\text{-SIS}}_{\mathcal{B}}\big)^2 - \delta_{\mathrm{hint}} - \delta_{\mathrm{rev}} - \delta_{\mathrm{norm}}.$$

Any negligible error in sampling the initial Gaussian matrix is added once and absorbed in the negligible term. This proves the stated advantage and the improved norm loss. □

The hint generator and the wrapper accept any bound B with $\eta_\varepsilon(L_{\boldsymbol{X}_1}) \le B$ and $\boldsymbol{X}_1$ surjective on an event E_B. The family below lives in power-of-two fields, so both geometric theorems apply.

Corollary 6.9 (Asymptotic parameters with B_{geo} or B_{cw}). *Let $0 < \alpha < 1$ be fixed, let $L = \ln(2\lambda)$, and let $w = w(\lambda)$ be polynomial-time computable with $w \to \infty$ and $w \le L$. Put*

$$d = 2^{\lceil \alpha \log_2 \lambda \rceil}, \quad K_\lambda = \mathbb{Q}(\zeta_{2d}), \quad k = n_{\mathrm{SIS}} = \lceil \lambda/d \rceil, \quad m = \lceil C_m \lambda L \rceil, \quad \nu = \lambda^{-w},$$

and

$$s_1 = C_1 d\sqrt{m}, \qquad s_2 = C_2\, d^{3/2}\lambda^{3/2}L, \qquad P = C_q s_2 \sqrt{m} = \Theta(d^{3/2}\lambda^2 L^{3/2}),$$

with sufficiently large absolute constants, choose a rational prime q with $\lceil P^d \rceil \le q \le \lfloor 2P^d \rfloor$ and $q \equiv 1 \pmod{2d}$, and set $u = \lceil s_2/s_1 \rceil$. Take the accuracy parameter of the chosen kernel theorem equal to λ and the error parameters $\delta_{\mathrm{sp}}, \delta_{\mathrm{op}}, \varepsilon_{\mathrm{samp}}$ equal to ν. For all sufficiently large λ these parameters satisfy every hypothesis of Theorem 6.8 with $B = B_{\mathrm{geo}}$ and with $B = B_{\mathrm{cw}}$, with $g = d$, and one may take $\beta_0 = C_\beta s_2 \sqrt{m}\, \beta_1$. In both cases

$$B = \Theta\big(d^{1/2}\lambda^{1/2}\big), \qquad s_2 = \Theta\big(d^{3/2}\lambda^{3/2}L\big),$$
$$\frac{\beta_0}{\beta_1} = \Theta\big(d^{3/2}\lambda^2 L^{3/2}\big), \qquad \ln q = \Theta(dL),$$

and the reduction loses an additive $2^{-\Omega(\lambda)} + \mathsf{negl}(\lambda)$ in the adversary's advantage, which accounts for all statistical errors and failure events.

Proof. We have $dk = \Theta(\lambda)$ and $m = \Theta(\lambda L)$, and $m = O(d^{1/\alpha+1})$ because $\alpha > 0$ is fixed, so the condition $m \le \mathsf{poly}(d)$ holds. With $R = dk = \Theta(\lambda)$ and $s_1 = C_1 d\sqrt{m}$ the coefficient width is $s_- = s_1/\sqrt{d} = C_1\sqrt{dm}$, which exceeds $8\sqrt{R}$ for large λ and is at least one. In Equations (27) and (40) one has $H = \Theta(\lambda)$ and $W = \mathsf{poly}(\lambda)$, so $\ln W = \Theta(L)$ and both column conditions, $m\ln(8/7) \ge R\ln W + 2\ln(1/\delta)$ and $m\ln(50/49) \ge R\ln W + 2\ln(1/\delta)$, are of the form $m \ge \Theta(\lambda L)$ and hold for C_m large. Hence Theorem 5.6 and Theorem 5.9 apply, with failure probabilities $3\cdot 2^{-\lambda-4}$ and $3\cdot 2^{-\lambda-6}$, and $B_{\mathrm{geo}} = 64\sqrt{dH}$ and $B_{\mathrm{cw}} = 640\sqrt{dH}$ are both $\Theta(d^{1/2}\lambda^{1/2})$. The spectral and operator column conditions hold because $k + \ln(2d/\nu) = O(\lambda^{1-\alpha} + wL) = o(m)$, and $m \ge 64\max\{\lambda/\ln(dk), k\}$ for C_m large. With $r_{\mathrm{samp}} = C_{\mathrm{samp}}\sqrt{\ln(dm/\nu)} = \Theta(\sqrt{wL})$ we have $B \gg 4r_{\mathrm{samp}}\sqrt{d}$, and since $s_1\sqrt{m} = \Theta(d\lambda L)$ the shaping threshold

$$c''_{\mathrm{sp}} s_1 \sqrt{m}\, B = \Theta(dm \cdot d^{1/2}\lambda^{1/2}) = \Theta(d^{3/2}\lambda^{3/2}L)$$

is met by s_2 for C_2 large, which is Equation (47).

The congruence makes q unramified and completely split, so $g = d$, by the splitting law for cyclotomic fields [Was97, Chapter 2]. Such a prime exists for large λ by the Siegel–Walfisz theorem

[Dav00, Chapter 20], which applies uniformly because the modulus $2d$ is at most $\ln x$ for $x = P^d$: the interval $[x, 2x]$ contains about $x/(\varphi(2d)\ln x) = x/(d\ln x)$ primes congruent to 1 modulo $2d$, so a candidate drawn uniformly from the integers congruent to 1 modulo $2d$ in this interval is prime with probability $\Theta(1/\ln x)$, and random search with a polynomial-time primality test finds one in $O(\ln x) = O(dL)$ expected trials. Capping the search after $C\lambda \ln x$ trials keeps the reduction strict polynomial time and adds a setup failure $e^{-\Omega(\lambda)}$. Let $e_q = k/m + 2/(dm)$. From $q^{1/d} = \Theta(P)$ we get $\ln(q^{e_q}) = (dk/m + 2/m)\ln(q^{1/d}) = O(1)$, so the first lower bound of the window is absorbed by s_1 for C_1 large. Also $(k\ln u + k\ln q)/(m-k) = O(1)$, so $2\sqrt{d}(u^k q^k)^{1/(m-k)} = O(\sqrt{d}) \le s_1$. By definition $s_2 \le u s_1$, and $q^{1/d}/\sqrt{m} \ge s_2$ follows from $q^{1/d} \ge P = C_q s_2 \sqrt{m}$. Finally $n_{\mathrm{SIS}} g / q^{d(m-n_{\mathrm{SIS}}+1)/g} = kd/q^{m-k+1}$ is negligible and $d(m-k) = \Omega(\lambda)$.

The base law of $\boldsymbol{X}_1$ is efficiently sampled through Corollary 3.2 within statistical distance ν. The error budget collects the kernel failure 3δ, the spectral and operator failures 2ν, the k LHL errors $2k\delta/(1-2\delta)$, the sampling errors $2k\nu$, the modular-smoothing error $2^{-\Omega(dm)}$ of the imported reduction, and the capped prime search $e^{-\Omega(\lambda)}$. Since $w \to \infty$ their sum is $2^{-\Omega(\lambda)} + \mathsf{negl}(\lambda)$. Lastly $\beta_0/\beta_1 = C_\beta s_2 \sqrt{m} = \Theta(d^{3/2}\lambda^{3/2}L \cdot \lambda^{1/2}L^{1/2})$. □

The kernel bounds B_{geo} and B_{cw}, both $\Theta(\sqrt{d\lambda})$, have the optimal order at exponential error. For the power-of-two field, Proposition 5.14 gives

$$\eta_\varepsilon(L_{\boldsymbol{X}_1}) \ge \sqrt{(d/\pi)\ln(2/\varepsilon)} = \Omega(\sqrt{d\lambda}) \qquad (\varepsilon = 2^{-\Theta(\lambda)})$$

for every valid smoothing bound. This optimality concerns the kernel parameter, not the norm loss of every possible SIS reduction, and the algebraic reduction architecture remains that of [AFL+26, Section 7].

Remark 6.10 (The width of the hints is not chosen by the kernel theorem). In the family of Corollary 6.9 the width $s_1 = \Theta(d\sqrt{m})$ is forced from below by the modular-smoothing window $s_1 \ge 8d\sqrt{m}\, q^{e_q}$ of Theorem 6.8, which is what makes the k-SIS matrix uniform modulo q in the exchange between sampling the public matrix and sampling its hints, and the rows of $\boldsymbol{X}_1$ are the first block of the hints. This floor lies above the width requirements of both kernel theorems, $8\sqrt{d\lambda}$ and $\sqrt{d}$, so neither of them binds, and the smaller admissible width of Theorem 5.9 brings nothing here. Only the kernel bound enters, through $s_2 \asymp s_1\sqrt{m}\, B$ and $\beta_0/\beta_1 \asymp s_1 m B$, which is why Corollary 6.9 gives the same orders for B_{geo} and B_{cw}. For concrete parameters Theorem 5.6 is the better choice in the reduction, with the constant 64 in place of 640 and about seven times fewer columns. An improvement from the smaller width would require a reduction whose hints do not pass through modular regularity, which we do not have.

The standalone parameters of the two geometric columns of Table 1 follow from Corollary 6.2. For Theorem 5.6 at its smallest admissible width $s_1 = 8\sqrt{dR} = \Theta(d^{1/2}\lambda^{1/2})$ one has $s_1\sqrt{m} = \Theta(d^{1/2}\lambda L^{1/2})$ and $s_{\mathrm{out}} = \Theta(s_1\sqrt{m}\, B_{\mathrm{geo}}) = \Theta(d\lambda^{3/2}L^{1/2})$. For Theorem 5.9 at $s_1 = \sqrt{d}$ one has $s_1\sqrt{m} = \Theta(d^{1/2}\lambda^{1/2}L^{1/2})$ and $s_{\mathrm{out}} = \Theta(d\lambda L^{1/2})$, smaller by $\lambda^{1/2}$. Both columns have joint error $2^{-\Omega(\lambda)}$. The geometric bounds at any error $\varepsilon \ge 2^{-\lambda}$ are the same B_{geo} or B_{cw}, so no factor depending on the error budget appears. Each column may choose the modulus required by its own conditions, and converting the asymptotic gain into concrete key or signature sizes requires analyzing a particular construction.

7 Comparison with Prior Work

7.1 Comparison with earlier bounds

The rational case. Let $K = \mathbb{Q}$, so that $L_{\boldsymbol{X}} = \ker_{\mathbb{Z}} \boldsymbol{X}$ and all basis constants equal one, and fix $p \ge 1/2$ and sufficiently large constants so that

$$s = c_X r^p, \qquad m = \lceil c_m r \ln(2r) \rceil, \qquad \lambda = r, \qquad \varepsilon = 2^{-r}. \tag{49}$$

Table 4 records the sufficient smoothing bounds of the rational theorems in this range. The exact source hypotheses are collected in Section 7.2. The first row uses the sharper estimate inside the proof of [AGHS13, Corollary 3], before its final simplification. The second combines [AR16, Lemma 4.2 and Theorem 3.2] rather than the looser summary in that paper's final theorem. The third row requires care.

source	bound on $\eta_\varepsilon(L_{\boldsymbol{X}})$	failure probability
Agrawal, Gentry, Halevi, Sahai [AGHS13]	$O_p(r^{5/2}\ln r)$	$e^{-\Omega(r)}$
Aggarwal, Regev [AR16]	$O_p(r^{p+3/2}\ln r)$	2^{-r}
Kirshanova, Nguyen, Stehlé, Wallet [KNSW20]	$O_p(\sqrt{r\ln r})$	$2^{-\Omega(r)}$
Theorem 5.8	$O_p(\sqrt{r})$	$3\cdot 2^{-r-4}$

Table 4. Sufficient smoothing bounds over $\mathbb{Q}$ in the common range Equation (49). No row claims that its bound is necessary.

The theorem of [KNSW20] gives the small bound $96\sqrt{\ln(1/\varepsilon')}$ only when $\varepsilon' \le e^{-(m-r)}$. Applying it at $\varepsilon' = \min\{\varepsilon, e^{-(m-r)}\}$ and using monotonicity of η in ε gives

$$\eta_\varepsilon(L_{\boldsymbol{X}}) \le 96\sqrt{\max\{\ln(1/\varepsilon),\, m-r\}}, \tag{50}$$

which is $O_p(\sqrt{r\ln r})$ in Equation (49). Our bound removes the factor $\sqrt{\ln r}$ from this comparison. Comparing with their first displayed bound alone would exaggerate the improvement. The constants hidden in earlier exceptional probabilities need not give error below $2^{-\lambda}$ when one sets $r=\lambda$ numerically. The table therefore compares orders of growth, not concrete values.

Modules over number fields. For a fixed field, the polynomial smoothing bound of [AFL+26, Theorem 4.2] has sufficient scale

$$O_K\Big(s\,m^{3/2}\sqrt{\ln\big(2d(m-r)(1+\varepsilon^{-1})\big)}\Big),$$

and its column count contains a term $T\ln T$ with $T=\Theta_K(r\ln(rs))$ as well as a term proportional to $\ln(1/\varepsilon)$ (Section 7.2). For fixed K, polynomial s and $\varepsilon = 2^{-\Theta(r)}$ the sufficient column count is therefore of order $r(\ln r)^2$, so that theorem cannot simply be evaluated at $m = c_m r\ln r$ and placed in Table 4. At the larger common choice $m = C_K r(\ln r)^2$ its displayed scale is $O_K(r^{p+2}(\ln r)^3)$ for $s = c_X r^p$, while Theorem 5.6 still gives $O_K(\sqrt{r})$ there. This comparison fixes the field and does not summarize the separate dependence of [AFL+26] on degree, discriminant or roots of unity. Their surjectivity and smoothing results also have separate hypotheses, and GRH enters only some of them.

Constant width. In the power-of-two power basis, Theorem 5.9 lowers the coefficient-width requirement of the theta route from $s_- \ge 8\sqrt{R}$ to $s_- \ge 1$ while keeping $m = O(R\ln R)$ columns and the optimal smoothing order $\sqrt{d\lambda}$. The polynomial smoothing theorem of [AFL+26] also admits a constant coefficient width, namely $\sqrt{2\pi}$ in the canonical convention $s \ge \sqrt{2\pi d}$, but with the bound and the column count of the previous paragraph. Constant-width Gaussian sources are familiar in finite and modular leftover hashing, where Jin, Liu, Wang and Gu [JLWG25] improve regularity lemmas for discrete Gaussians modulo sublattices under explicit entropy and modulus conditions. Their random object is a uniform public matrix modulo q and a Gaussian source. Ours is the narrow Gaussian matrix $\boldsymbol{X}$ itself, and the conclusion controls its ordinary integer kernel and an ellipsoidal Gaussian output before any quotient is taken. Describing Theorem 5.9 as the first constant-width Gaussian leftover hash lemma would be incorrect. What is new is the combination of constant matrix width, $O(R\ln R)$ columns, optimal smoothing order and unconditional surjectivity, in a regime where the matrix is provably far from uniform modulo every integer $q \ge 2$ by Equation (46).

What is different in the proof. The classical Gaussian LHL reduces the distribution question to kernel smoothing [AGHS13], and the earlier proofs then control short kernel vectors, short preimages, or minima of auxiliary lattices. Construction-A lattices are central to [AGHS13, KNSW20], Aggarwal and Regev use short integral preimages of the standard basis, and [AFL+26] extends that strategy to modules. The geometric route estimates the full dual theta series through Equation (21), whose index factor proves surjectivity at the same time. Its ingredients are Poisson summation, Gaussian integration, elementary lattice geometry, and translation estimates in the spirit of the modular anti-concentration arguments of [AGHS13, KNSW20] and of the disjoint-translate argument of [AR16]. The constant-width theorem replaces the multidimensional displacement by an exact one-coordinate overlap and a Berry–Esseen estimate. The broader Fourier strategy, a Gaussian comparison near the origin, an expected Fourier bound elsewhere, and Markov's inequality, also appears in the local limit theorems for random matrices of Franks and Saks [FS20], whose results concern signed sums and different quantitative guarantees. No priority is claimed for the ingredients, only for their present combination.

7.2 Sources of the compared bounds

For the rational comparisons in Table 4, $\boldsymbol{X} \in \mathbb{Z}^{r\times m}$ has independent $D_{\mathbb{Z},s}$ entries, $n = m - r$, and $\gamma_n(\varepsilon)$ is the function of Equation (8). Changing a logarithm base in an asymptotic bound changes only its constant. In the finite hypotheses below we nevertheless retain $\log_2$ where the source uses it. The statements in this subsection are readings of the cited theorems, recorded so that the table can be checked against them.

Agrawal, Gentry, Halevi and Sahai. In the spherical specialization of [AGHS13, Theorem 2], sufficient conditions are $s \geq \eta_\varepsilon(\mathbb{Z}^r)$ and $m \geq 10r\log_2\big(8(mr)^{3/2}s\big)$, with the source's negligible-error convention. The proof of its Corollary 3 gives $\eta_\varepsilon(L_{\boldsymbol{X}}) \leq 4mr\,\gamma_n(\varepsilon)$ before replacing the square-root logarithm by a larger expression. This is the bound used in Table 4. The theorem's exceptional event is exponentially small.

Aggarwal and Regev. Fix an auxiliary $\varepsilon_0 \in (0, 1/1000)$. Combining [AR16, Lemma 4.2 and Theorem 3.2], for $r \geq 100$ the conditions $s \geq 9\gamma_r(\varepsilon_0)$ and $m > 30r\log_2(sr)$ give, with failure probability at most 2^{-r}, the sufficient smoothing bound $\eta_\varepsilon(L_{\boldsymbol{X}}) \leq \big(1 + 2sr\sqrt{30\log_2 m\,\log_2(sr)}\big)\gamma_n(\varepsilon)$. This is the bound used in Table 4. The auxiliary error controls the matrix-generation argument. The smoothing error ε can be chosen separately.

Kirshanova, Nguyen, Stehlé and Wallet. The explicit theorem of [KNSW20] assumes $r \geq 60$, $s \geq 20\sqrt{r}$ and $m \geq 1355\,r\ln s$. It gives $\eta_\varepsilon(L_{\boldsymbol{X}}) \leq 77\sqrt{(r+\ln m)\ln(2m/\varepsilon)}$ with failure $2^{-\Omega(r)}$, and $\eta_\varepsilon(L_{\boldsymbol{X}}) \leq 96\sqrt{\ln(1/\varepsilon)}$ when $\varepsilon \leq e^{-n}$. Monotonicity yields Equation (50). The second bound is essential to the fair comparison in Section 7.1.

The module theorem. Let $f \geq 2$ satisfy $\mathbb{Q}(\zeta_f) \subseteq K$, where $f = 2$ is always available, and specialize [AFL+26, Theorem 4.2] to the spherical column parameter $s\boldsymbol{I}_R$. Put $T = r\log_2\big(dr\,|\Delta_K|^{1/(2d)}\,s\big)$. Using base-two logarithms and enlarging the absolute constant if needed, its sufficient column condition is

$$m \;\geq\; C\,T\log_2 T + \frac{2\log_2(1/\varepsilon)}{\log_2(f/1.8)},$$

and the width must also meet $s \geq \max\{\eta_{1/2}(\Lambda_r), \sqrt{2\pi d}\}$ in the canonical convention. For fixed K its smoothing conclusion is the scale quoted in Section 7.1, with field-dependent factors absorbed into O_K. With $s = c_X r^p$ and $\ln(1/\varepsilon) = \Theta(r)$ one has $T = \Theta_K(r\ln r)$, which gives the sufficient $r(\ln r)^2$ column scale. These calculations concern the polynomial smoothing theorem of [AFL+26], not its surjectivity or small-rank results.

Disclosure on AI Usage

The main results in this paper were discovered by frontier AI models (GPT-6 and Fable 5.1), while the authors formulated the research questions, proposed several directions in this discovery process, and several optimizations. In particular, the exact theta integral was discovered by AI. The authors have verified the results and take full responsibility for the correctness of all content. The introduction and technical overview sections were written by the authors, with AI assistance (GPT-6) limited to polishing the text. The remaining sections, and in particular the appendices, were developed through iterative interactions with GPT-6 and Fable 5.1. The authors repeatedly reviewed the generated drafts and directed further revisions through follow-up discussions. All figures and tables were generated using Fable 5.1 from prompts provided by the authors. Similarly, the Lean formalization was done using GPT-6 following the prompts provided by the authors.

Acknowledgments

The research of Nihar Gargava is funded from the SNSF grant 225437 titled "Random Geometry with Arithmetic Constraints". The research of Erkan Tairi is supported by the Austrian Science Fund (FWF) Project J4879-N, by Bakar Funds, Peder Sather Funds and Input Output Global.

References

AFL+26. Martin R. Albrecht, Joël Felderhoff, Russell W. F. Lai, Oleksandra Lapiha, and Ivy K. Y. Woo. A gaussian leftover hash lemma for modules over number fields. In Joan Daemen and Emmanuel Thomé, editors, *EUROCRYPT 2026, Part IV*, volume 16545 of *LNCS*, pages 124–153. Springer, Cham, May 2026. 4, 6, 7, 8, 9, 22, 40, 41, 42, 43, 44, 45, 46, 47, 48

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. 3, 7, 8, 9, 27, 38, 46, 47, 48

Ajt96. Miklós Ajtai. Generating hard instances of lattice problems (extended abstract). In *28th ACM STOC*, pages 99–108. ACM Press, May 1996. 3

ALLW25. Martin R. Albrecht, Russell W. F. Lai, Oleksandra Lapiha, and Ivy K. Y. Woo. Partial lattice trapdoors: How to split lattice trapdoors, literally. Cryptology ePrint Archive, Report 2025/367, 2025. 43

AR16. Divesh Aggarwal and Oded Regev. A note on discrete gaussian combinations of lattice vectors. *Chicago Journal of Theoretical Computer Science*, 2016. Preprint: https://arxiv.org/abs/1308.2405. 4, 7, 9, 27, 46, 47, 48

Ban93. Wojciech Banaszczyk. New bounds in some transference theorems in the geometry of numbers. *Mathematische Annalen*, 296(1):625–635, 1993. 27, 38

Ber41. Andrew C. Berry. The accuracy of the Gaussian approximation to the sum of independent variates. *Transactions of the American Mathematical Society*, 49(1):122–136, 1941. 16, 35

BL26. Katharina Boudgoust and Oleksandra Lapiha. Leftover hash lemma(s) over cyclotomic rings. In Joan Daemen and Emmanuel Thomé, editors, *EUROCRYPT 2026, Part IV*, volume 16545 of *LNCS*, pages 62–91. Springer, Cham, May 2026. 3, 7

Dav00. Harold Davenport. *Multiplicative Number Theory*, volume 74 of *Graduate Texts in Mathematics*. Springer, 3 edition, 2000. Revised by Hugh L. Montgomery. 46

Ess42. Carl-Gustav Esseen. On the Liapunoff limit of error in the theory of probability. *Arkiv för Matematik, Astronomi och Fysik*, A28:1–19, 1942. 16, 35

FS20. Cole Franks and Michael Saks. On the discrepancy of random matrices with many columns. *Random Structures & Algorithms*, 57(1):64–96, 2020. arXiv:1807.04318. 47

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. 22

GN08. Nicolas Gama and Phong Q. Nguyen. Predicting lattice reduction. In Nigel P. Smart, editor, *EUROCRYPT 2008*, volume 4965 of *LNCS*, pages 31–51. Springer, Berlin, Heidelberg, April 2008. 38

GPV08. Craig Gentry, Chris Peikert, and Vinod Vaikuntanathan. Trapdoors for hard lattices and new cryptographic constructions. In Richard E. Ladner and Cynthia Dwork, editors, *40th ACM STOC*, pages 197–206. ACM Press, May 2008. 3, 37

ILL89. Russell Impagliazzo, Leonid A. Levin, and Michael Luby. Pseudo-random generation from one-way functions (extended abstracts). In *21st ACM STOC*, pages 12–24. ACM Press, May 1989. 3

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. 3, 7, 47

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. 4, 7, 8, 9, 27, 32, 38, 47, 48

Lyu09. Vadim Lyubashevsky. Fiat-Shamir with aborts: Applications to lattice and factoring-based signatures. In Mitsuru Matsui, editor, *ASIACRYPT 2009*, volume 5912 of *LNCS*, pages 598–616. Springer, Berlin, Heidelberg, December 2009. 3

MPTJ07. Shahar Mendelson, Alain Pajor, and Nicole Tomczak-Jaegermann. Reconstruction and subgaussian operators in asymptotic geometric analysis. *Geometric and Functional Analysis*, 17(4):1248–1282, 2007. 25

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. 11, 20, 23

OFKT23. Hiroki Okada, Kazuhide Fukushima, Shinsaku Kiyomoto, and Tsuyoshi Takagi. Spherical gaussian leftover hash lemma via the Rényi divergence. In Mehdi Tibouchi and Xiaofeng Wang, editors, *ACNS 2023, Part I*, volume 13905 of *LNCS*, pages 695–724. Springer, Cham, June 2023. 40

Pei08. Chris Peikert. Limits on the hardness of lattice problems in ℓ_p norms. *Computational Complexity*, 17(2):300–351, 2008. 27, 38

Pei10. Chris Peikert. An efficient and parallel Gaussian sampler for lattices. In Tal Rabin, editor, *CRYPTO 2010*, volume 6223 of *LNCS*, pages 80–97. Springer, Berlin, Heidelberg, August 2010. 41

Reg05. Oded Regev. On lattices, learning with errors, random linear codes, and cryptography. In Harold N. Gabow and Ronald Fagin, editors, *37th ACM STOC*, pages 84–93. ACM Press, May 2005. 3

RV08. Mark Rudelson and Roman Vershynin. The Littlewood–Offord problem and invertibility of random matrices. *Advances in Mathematics*, 218(2):600–633, 2008. 20

She13. Irina G. Shevtsova. On the absolute constants in the Berry–Esseen inequality and its structural and nonuniform improvements. *Informatika i Ee Primeneniya*, 7(1):124–125, 2013. 35

TV09. Terence Tao and Van H. Vu. Inverse Littlewood–Offord theorems and the condition number of random discrete matrices. *Annals of Mathematics*, 169(2):595–632, 2009. 20

Ver26. Roman Vershynin. *High-Dimensional Probability: An Introduction with Applications in Data Science*. Cambridge University Press, 2 edition, 2026. 24

Was97. Lawrence C. Washington. *Introduction to Cyclotomic Fields*, volume 83 of *Graduate Texts in Mathematics*. Springer, 2 edition, 1997. 45