

Provable Subexponential Algorithms for Module-SIS and Module-LWE

William Youmans

Clemson University, Clemson, USA

Abstract. Ducas, Engelberts and Loyer (CRYPTO 2025) turned Wagner’s generalized birthday algorithm into a provable subexponential-time Gaussian sampler for unstructured short integer solutions. Recent preprints of Han, Gao and Hu, and of Chen, Jiang, Wang and Zhang, obtain low-sample decision learning-with-errors algorithms from this framework. We extend it to module matrices over power-of-two cyclotomic rings, for every residue degree of the prime modulus.

For fixed module ranks $k, \ell \geq 1$, let $n = kd$, $m = (k + \ell)d$, and suppose $\ln(q/\sqrt{n}) = (\ln n)^{b+o(1)}$. For every fixed $b, c, \xi > 0$, we solve Module-SIS with coefficient infinity-norm bound $q/(\ln n)^c$ in worst-case time and space $2^{(1/(2b)+\xi)n/\ln \ln n}$. For matrix-independent errors satisfying $\|e\| \leq \sigma\sqrt{m}$ with probability $1 - o(1)$ and $\ln(2 + \sigma) = (\ln n)^{a+o(1)}$, $0 \leq a < b$, decision and search Module-LWE take $2^{(1/(2(b-a))+\xi)n/\ln \ln n}$. Success is $1 - o(1)$ using only the given $k + \ell$ ring samples; error coordinates need not be independent.

The structural ingredient is a Chinese-remainder-support uncertainty bound: the loss of randomness of sparse dual labels is compensated by an arithmetic lower bound on their length. Scalar row perturbations give search over the base field, and small integer shears remove a singular-block loss. Projection, the Han–Gao–Hu martingale analysis, and a finite enumeration sampler give a common algorithmic framework.

We also prove sample-preserving scalar search-LWE and narrow Module-LWE algorithms. In particular, for polynomial prime $q = n^{\kappa+o(1)}$, $\kappa > 1/2$, one ring sample and growing rank suffice for decision and search in time and space $2^{(1/2+\xi)n/\ln \ln n}$ when the sample supplies $(1 + o(1))n/\ln \ln n$ scalar responses. Here “narrow” restricts both secret and error. Separate heuristic estimates in the model of Ducas, Engelberts and Loyer suggest that Wagner-style attacks do not reach the NIST security targets of ML-KEM and ML-DSA.

Keywords: Wagner’s algorithm · Module-SIS · Module-LWE · Discrete Gaussian sampling · Cryptanalysis

AI disclosure: predominantly AI-generated work. Under the IACR Generative AI Policy (<https://iacr.org/policy/AIPolicy.html>), this work is designated as *predominantly AI-generated work*. It is the result of chats with the large language models OpenAI GPT-5.5, GPT-5.6 Sol and GPT-6 Astra and Anthropic Claude Opus 5.5, beginning on 24 July 2026. The models proposed approaches, developed the mathematical arguments and proofs, wrote the checking and estimation code, and drafted the manuscript. The author initiated and directed these conversations and remains responsible for all mathematical claims, proofs, references and conclusions.

Note on concurrent work. Gao, Han and Hu [GHH26] recently posted provable subexponential algorithms for growing parameter families associated with the NIST third-round lattice candidates, built on the same Wagner-style Gaussian sampling framework [DEL25, HGH26]. Both works recover short secrets in families shaped like ML-KEM

E-mail: youmans.wj@gmail.com (William Youmans)

public keys in time $2^{(1/2+o(1))n/\ln \ln n}$, with different coordinate tests and geometric analyses. Their module results assume a constant residue degree, and hence $q = \Omega(n)$. In exchange, they allow errors that depend on the public matrix, such as rounding errors and the compressed public keys of ML-DSA, and they also cover power-of-two moduli and NTRU-type schemes. We allow every residue degree and all polynomial moduli $q = n^\kappa$ with $\kappa > 1/2$, give worst-case rather than expected bounds, and also treat Module-SIS and Module-LWE with arbitrary secrets. The two works reach compatible conclusions about the standards. [GHH26] states that its results do not reduce the concrete security of the specified parameter sets. Our heuristic estimates (Appendix A) place Wagner-style attacks well above the NIST security targets of ML-KEM and ML-DSA.

1 Introduction

The Short Integer Solution problem (SIS) [Ajt96] and the Learning with Errors problem (LWE) [Reg05] underlie most of lattice-based cryptography. In SIS, one seeks a nonzero short integer vector satisfying homogeneous linear equations modulo an integer. In LWE, one is given linear equations with small errors and seeks either to recover the hidden vector (search) or to distinguish these noisy equations from uniformly random data (decision). Their module variants, Module-SIS (MSIS) and Module-LWE (MLWE) [LS15], replace the uniformly random public matrix over $\mathbb{Z}_q$ by a uniformly random matrix over the ring $R_q = \mathbb{Z}_q[x]/(x^d + 1)$. They are the problems underlying the NIST standards ML-DSA [Nat24a] and ML-KEM [Nat24b]. Expanding every ring element into its coefficient vector turns an MSIS or MLWE instance of module rank k into an SIS or LWE instance of dimension $n = kd$. The resulting matrix consists of negacyclic $d \times d$ blocks, so it is far from uniform.

Wagner, BKW, and the sample-amplification gap. The Blum–Kalai–Wasserman algorithm (BKW) [BKW03] solves LWE by repeatedly combining samples that collide on a block of coordinates. Wagner’s generalized birthday algorithm [Wag02] does the same on the dual (SIS) side. With lazy modulus switching [AFFP14, KF15, GJMS17], Kirchen and Fouque [KF15] claimed a slightly subexponential algorithm for LWE with narrow errors and polynomial modulus from $\Theta(n)$ samples. Herold, Kirshanova and May [HKM18] found a gap in the sample-amplification step and could only repair it with $\Theta(n \ln n)$ samples.

Ducas, Engelberts and Loyer [DEL25] (DEL) isolated the Wagner step and recast it as an approximate discrete Gaussian sampler on a chain of q -ary lattices. Each step lifts Gaussian vectors to an auxiliary superlattice and subtracts pairs that lie in the same coset (Figure 1). This avoids sample amplification and yields a provable algorithm for SIS^∞ with norm bound $q/\text{polylog}(n)$ in time $\exp(O(n/\ln \ln n))$. DEL also explained why the sampler does not directly give an LWE algorithm [DEL25, Sect. 4.3]. The small Fourier bias requires much higher sampling accuracy than the SIS application, and at that accuracy the coordinate superlattices $(q/p)\mathbb{Z}^b$ have an unfavorable smoothing bound [DEL25, Sect. 4.3].

Two recent preprints resolve this for unstructured LWE. Han, Gao and Hu [HGH26] (HGH) use random prime-index superlattices $\frac{q}{p}\Lambda_P(a^\top)$, whose smoothing is much better, and they analyse the Fourier test on the dependent output list by a martingale argument. They obtain decision LWE in time $2^{(1/2+o(1))n/\ln \ln n}$ from $n + \Theta(n/\ln \ln n)$ samples, and more generally a modulus–error trade-off. Chen, Jiang, Wang and Zhang [CJWZ26] (CJWZ) work with the relation lattice of $[-A \mid I_n]$ and solve decision binary LWE from only $n/\ln \ln n$ samples, in time $2^{(8+o(1))n/\ln \ln n}$.

Structured matrices. All three analyses use the independence of the matrix entries in an essential way. They bound the smoothing parameters of the lattices in the chain by averaging over the matrix, and this averaging uses that $A^\top a$ is uniform in $\mathbb{Z}_q^{m-n}$

for every nonzero label a . HGH state that their analysis does not cover Module-LWE, and they identify the algebraic dependencies as the main obstacle to any implication for ML-KEM [HGH26]. For a module matrix the averaging genuinely fails. If q splits completely in R and $\alpha \in R_q^k$ is nonzero modulo only one prime above q , then $\mathbf{A}^\dagger \alpha$ takes only q^ℓ values instead of $q^{\ell d}$.

1.1 Contributions

In the parameter regimes below, we recover the same leading exponent as for unstructured matrices, despite a weaker smoothing scale. Throughout, $R = \mathbb{Z}[x]/(x^d + 1)$ with d a power of two, q is an odd prime, $\mathbf{A} \in R_q^{k \times (k+\ell)}$ is uniform, and $n = kd$ and $m = (k + \ell)d$ are the scalar dimensions. All norms are coefficient norms.

Theorem 1 (Main results, informal). *Fix integers $k, \ell \geq 1$, reals $0 \leq a < b$, and $\xi > 0$, and let q be prime with $\ln(q/\sqrt{n}) = (\ln n)^{b+o(1)}$. There are randomized algorithms that use only the $k + \ell$ given ring samples, succeed with probability $1 - o(1)$, and solve the following problems.*

1. Module-SIS $^\infty$ (Theorem 5): *given $\mathbf{A}$, find $\mathbf{z} \neq \mathbf{0}$ with $\mathbf{Az} = \mathbf{0} \bmod q$ and $\|\mathbf{z}\|_\infty \leq q/(\ln n)^c$, for any fixed $c > 0$. The worst-case time is $2^{(1/(2b)+\xi)n/\ln \ln n}$.*
2. Decision and search Module-LWE (Theorems 6 and 7): *distinguish $(\mathbf{A}, \mathbf{A}^\dagger \mathbf{s} + \mathbf{e})$ from uniform, and recover $\mathbf{s}$. This holds for every secret, and for every error independent of $\mathbf{A}$ with $\|\mathbf{e}\| \leq \sigma\sqrt{m}$ with probability $1 - o(1)$, where $\ln(2 + \sigma) = (\ln n)^{a+o(1)}$. The worst-case time is $2^{(1/(2(b-a))+\xi)n/\ln \ln n}$, and there is no condition on the residue degree of q .*

For prime $q = n^{\kappa+o(1)}$ with $\kappa > 1/2$ and bounded errors, all three running times are $2^{(1/2+\xi)n/\ln \ln n}$.

The Module-SIS output is also nonzero modulo q and has Euclidean norm at most $\sqrt{m}q/(\ln n)^c$; see Corollary 5. No inhomogeneous-MSIS result is claimed.

Theorem 2 (Further results, informal). 1. *Narrow Module-LWE with few samples (Section 6). Let $q = n^{\kappa+o(1)}$ with $\kappa > 1/2$, let $\mathbf{A} \in R_q^{k \times \ell}$ with $1 \leq \ell = O(k)$, including $\ell = o(k)$, and let the combined secret–error squared norm be $O((k + \ell)d)$. Then decision and search MLWE can be solved in time $2^{(1/2+\xi)n/\ln \Upsilon}$, where $\Upsilon = \frac{\ell}{k+\ell} \ln(q/\sqrt{\ell d}) \rightarrow \infty$. For example, $\ell = 1$ and $k \approx \ln \ln n$ uses $n/\ln \ln n$ scalar samples and time $2^{(1/2+\xi)n/\ln \ln n}$. For $d = 1$ this is unstructured narrow LWE; on the common sublinear-sample regimes it improves the leading time-exponent coefficient 8 of [CJWZ26] to $1/2$.*

2. Search-LWE without an assumed reduction (Section 7). *In every regime of [HGH26, Thm. 4.2], the secret can be recovered from $m = n + \lceil n/\ln \ln n \rceil$ samples within the same leading time bound. HGH obtain search only by assuming a sample-preserving search-to-decision reduction [HGH26, App. A].*

Our running times are worst-case bit complexities; the prior-work bounds in Table 1 are expected bounds. They are asymptotic statements about families with $n \rightarrow \infty$ (and fixed ranks in the main theorems). They do not establish an attack on the fixed parameters of ML-KEM or ML-DSA. Appendix A gives heuristic estimates for the standardized parameter sets; in the model of [DEL25], Wagner-style attacks remain well above the NIST security targets.

Table 1: Provable Gaussian–Wagner algorithms with polynomial modulus $q = n^{\kappa+o(1)}$ and narrow errors. “Samples” counts scalar samples (columns of the coefficient matrix). The time column gives $\log_2$ of the running time; prior-work bounds are expected and the new bounds are worst-case. The slack $\xi > 0$ is arbitrary. In the last row, $\Upsilon = \frac{\ell}{\kappa+\ell} \ln(q/\sqrt{\ell d})$.

Problem	Matrix	Samples	$\log_2$ time	Source
SIS $^\infty$, $\beta = q/\text{polylog}$	uniform	$n + \omega(\frac{n}{\ln \ln n})$	$(\frac{1}{2} + o(1)) \frac{n}{\ln \ln n}$	[DEL25]
decision LWE	uniform	$n + \Theta(\frac{n}{\ln \ln n})$	$(\frac{1}{2} + o(1)) \frac{n}{\ln \ln n}$	[HGH26]
decision binary LWE	uniform	$n/\ln \ln n$	$(8 + o(1)) \frac{n}{\ln \ln n}$	[CJWZ26]
search LWE	uniform	$n + \frac{n}{\ln \ln n}$	$(\frac{1}{2} + \xi) \frac{n}{\ln \ln n}$	Thm. 9
MSIS $^\infty$, $\beta = q/\text{polylog}$	module	$(1 + \frac{\ell}{k})n$	$(\frac{1}{2} + \xi) \frac{n}{\ln \ln n}$	Thm. 5
dec./search MLWE	module	$(1 + \frac{\ell}{k})n$	$(\frac{1}{2} + \xi) \frac{n}{\ln \ln n}$	Thms. 6, 7
narrow (M)LWE	module	ℓd , $1 \leq \ell = O(k)$	$(\frac{1}{2} + \xi) \frac{n}{\ln \Upsilon}$	Thm. 8

1.2 Technical Overview

The chain of lattices. Write the coefficient expansion of a systematic module matrix as $A = [A' \mid I_n] \in \mathbb{Z}_q^{n \times m}$, and let Λ_i be the q -ary lattice defined by the first n_i equations. Then $\Lambda_0 = \mathbb{Z}^{m-n}$ and $\Lambda_r = \Lambda_q^\perp(A)$. The sampler of [DEL25, HGH26] walks from Λ_0 to Λ_r . At level i , it lifts Gaussian vectors from Λ_{i-1} to $\Lambda'_i = \Lambda_i + (0 \times \Gamma_i)$, where $\Gamma_i = q\mathbb{Z}^{b_i} + \frac{q}{P_i}a_i\mathbb{Z}$ for a random $a_i \in \mathbb{F}_{P_i}^{b_i}$. It then subtracts pairs of lifted vectors in the same coset of $\Lambda'_i/\Lambda_i \cong \mathbb{Z}/P_i\mathbb{Z}$. Each level needs two smoothing conditions: one on Γ_i , which is independent of the matrix, and $s_{i-1} \geq \sqrt{2}\eta_\varepsilon(\Lambda_i)$ on the prefix lattice.

Only the last lattice matters. Every prefix lattice Λ_i is an orthogonal projection of Λ_r , and smoothing parameters do not increase under projection [CJWZ26, Lemma 2.10]. All prefix conditions therefore follow from a single bound on $\eta_\varepsilon(\Lambda_r)$. This matters for modules: the blocks of lazy modulus switching cut across ring elements, so the prefix lattices are not module lattices, but Λ_r is. We state the sampler for an *arbitrary* systematic lattice whose smoothing parameter is known (Theorem 4).

Smoothing random module lattices. The dual of $\Lambda_q^\perp(A)$ is $\mathbb{Z}^m + q^{-1}\{(A'^\top a; a) : a \in \mathbb{Z}_q^n\}$. Read the label a as a module vector $\alpha \in R_q^k$, and let $\mathcal{S}$ be its CRT support, the set of primes $\mathfrak{p}_\nu \mid q$ at which α is nonzero. Then $A'^\top \alpha$ is uniform, but only on the subgroup $V_{\mathcal{S}}^\ell$, of size $q^{\delta|\mathcal{S}|}$, where $\delta = \text{ord}_{2d}(q)$ is the residue degree (Lemma 9). An *arithmetic uncertainty principle* (Lemma 10) compensates: a nonzero $f \in R$ that vanishes modulo every $\mathfrak{p}_\nu$ with $\nu \notin \mathcal{S}$ has norm at least $\lambda_{\mathcal{S}} := q^{1-\delta|\mathcal{S}|/d}$, because $q^{d-\delta|\mathcal{S}|}$ divides its algebraic norm. The two effects cancel exactly:

$$q^{-\delta\ell|\mathcal{S}|} = \left(\frac{\lambda_{\mathcal{S}}}{q}\right)^{\ell d} \leq \left(\frac{\|a\|}{q}\right)^{m-n}.$$

After this cancellation, the first-moment computation of the scalar case goes through. The only loss is a factor $\sqrt{m-n}$ from a packing argument, which turns the modulus into the effective modulus $Q_{\text{eff}} = q/\sqrt{m-n}$ (Theorem 3).

Search over the base field. The Fourier score is invariant under $b \mapsto b + A'^\top u$, so with A fixed one cannot test a guessed secret by subtracting it. Instead we perturb one coefficient row of the systematic matrix, $A'_v = A' + e_i v^\top$ with $v \leftarrow \mathbb{F}_q^{m-n}$, and score $b + c(v; 0)$ for all $c \in \mathbb{F}_q$. The correct candidate yields an honest LWE instance for A'_v . A wrong one shifts the dual lattice of A'_v by $\gamma(v; 0)/q$ with $\gamma \neq 0$, and the coset labelled z

then has free part $A'^T z + (z_i - \gamma)v$. If $z_i \neq \gamma$, this part is uniform. If $z_i = \gamma$, then $z \neq 0$ and the module bound applies (Lemma 14). Hence q candidates per coefficient suffice for every residue degree. The perturbed matrix is no longer a module matrix, but the same dichotomy proves its smoothing.

Normalization. The last k ring columns of $\mathbf{A}$ can be invertible with probability bounded away from 1 when q splits completely and $q = O(d)$. We therefore add t times a free column to a designated one, for each $t < 3 \log n$. At each CRT factor the determinant is then a linear polynomial in t , which rules out at most one t . Hence some shear works with probability $1 - o(1)$ (Lemma 21), while norms grow by only $O(\ln n)$.

Worst-case sampling. For the lifting step we use per-level primes P_i matched to the current width, and an explicit shifted Gaussian sampler. This sampler enumerates the lattice points in a ball deterministically [DPV11, MV13], with rational weights of bounded precision (Lemma 17). The cost of a call is controlled by smoothing alone, which yields worst-case bit-complexity bounds. HGH’s martingale analysis of the Fourier test [HGH26, Lemma 4.2] then applies directly to the dependent output list.

1.3 Related Work

Algebraic attacks [AG11, ACFP14, Ste24] solve LWE with small errors in subexponential time given $n^{1+\Omega(1)}$ samples, or under regularity assumptions. Provable dual attacks [PS24] analyse Fourier tests with independent dual vectors. Discrete Gaussian sampling also underlies provable SVP and CVP algorithms [ADRS15, ADS15, ALS15], and Gao, Feng and Hu [GFH26] use random prime-index superlattices for exact SVP. Worst-case reductions for module lattice problems apply in their specified parameter regimes [LS15]. To our knowledge, no previous provable subexponential algorithm covered these MSIS and MLWE parameter families with a number of ring samples independent of n .

This work builds directly on the preprints [HGH26, CJWZ26]. We use the auxiliary lattices, lifting and combining analysis, and martingale test of [HGH26], and the relation lattice of $[-A \mid I_n]$ and the projection argument of [CJWZ26]. The required imported statements and their hypotheses are given explicitly below. We also note (Remark 2) that the combining lemma of [CJWZ26] inherits a premise of [DEL25, Lemma 3.3] that is weaker than its proof requires, although their parameters satisfy the correct condition.

Concurrent work by Gao, Han and Hu [GHH26] is compared with ours in the note preceding the introduction.

1.4 Organization

Section 2 recalls the background. Section 3 proves the smoothing bounds for module relation lattices and their scalar perturbations. Section 4 gives the worst-case sampler for arbitrary smooth systematic lattices. Section 5 derives the fixed-rank module results. Sections 6 and 7 prove the narrow-MLWE and scalar search-LWE results. Section 8 discusses limitations. The appendices contain illustrative heuristic models (Appendix A) and reproducible checks (Appendix B). The full block-schedule proof and finite implementation are in Section 4. The code for Appendices A and B is available at <https://github.com/wjyoumans/module-wagner>.

2 Preliminaries

Notation. We write $\ln$ for the natural logarithm and $\log$ for the logarithm to base two, and $[N] = \{1, \dots, N\}$. Vectors are columns; $(a; b)$ denotes vertical concatenation.

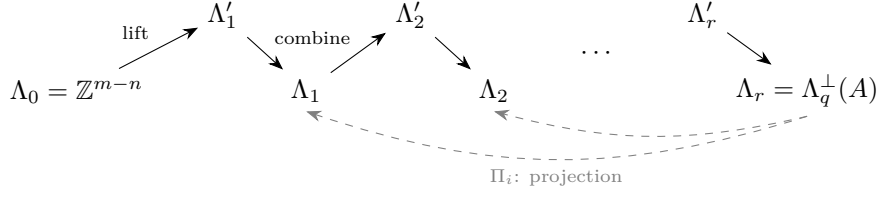


Figure 1: The Gaussian–Wagner walk of [DEL25, HGH26]. Solid arrows are the lifting and combining steps. Dashed arrows indicate that every prefix lattice Λ_i is an orthogonal projection of Λ_r , so $\eta_\varepsilon(\Lambda_i) \leq \eta_\varepsilon(\Lambda_r)$ [CJWZ26].

$\|\cdot\|$ is the Euclidean norm and $\|\cdot\|_\infty$ the infinity norm. The notation $u = e^{\pm\theta}v$ means $e^{-\theta}v \leq u \leq e^\theta v$. Elements of $\mathbb{Z}_q$ are represented by integers in $\{0, \dots, q-1\}$ whenever an integer vector is needed. All asymptotic statements refer to a family of parameters indexed by $n \rightarrow \infty$. A constant is *absolute* if it does not depend on the family.

2.1 Similarity of Distributions

We use the pointwise similarity notions of [DEL25], which go back to [ALSD21]. To avoid a clash with the residue degree δ , we call the similarity parameter θ ; it is denoted δ in [DEL25, HGH26].

Definition 1 ([DEL25, Defs. 2.1, 2.2]). Let D be a distribution on a countable set $\mathcal{X}$ and $\theta \geq 0$. A random variable X is θ -similar to D if $\Pr[X = x] = e^{\pm\theta}D(x)$ for all $x \in \mathcal{X}$. Random variables $X_1, \dots, X_N$ are *conditionally θ -similar to independent samples from D* if $\Pr[X_i = x_i \mid X_{-i} = x_{-i}] = e^{\pm\theta}D(x_i)$ for every $i \in [N]$ and every $x \in \mathcal{X}^N$ for which the conditional probability is defined.

Conditional θ -similarity is preserved under passing to sublists and implies $\Pr[X_i = x \mid X_{<i} = x_{<i}] = e^{\pm\theta}D(x)$ for every prefix in the support [DEL25, Lemma 2.1], [HGH26, Sect. 2.1]. In particular, every X_i is θ -similar to D , and hence $\Pr[X_i \in E] \leq e^\theta D(E)$ for every event E . The following martingale bound is the tool that makes the Fourier test work on dependent lists.

Lemma 1 ([HGH26, Lemma 4.2]). *Let $X_1, \dots, X_N$ be conditionally θ -similar to independent samples from D , with $0 \leq \theta \leq 1$. Then for every $h : \mathcal{X} \rightarrow [-1, 1]$ and $t > 0$,*

$$\Pr\left[\left|\frac{1}{N} \sum_{j=1}^N h(X_j) - \mathbb{E}_D h\right| > 2\theta + t\right] \leq 2 \exp(-Nt^2/8).$$

Proof. Let $a_j = \mathbb{E}[h(X_j) \mid X_{<j}]$. Averaging conditional similarity over the unobserved variables gives $|a_j - \mathbb{E}_D h| \leq e^\theta - 1 \leq 2\theta$. The sums $M_t = \sum_{j \leq t} (h(X_j) - a_j)$ form a martingale with increments of absolute value at most 2. Azuma–Hoeffding gives $\Pr[|M_N| > Nt] \leq 2e^{-Nt^2/8}$. The deterministic discrepancy between $N^{-1} \sum_j a_j$ and $\mathbb{E}_D h$ is at most 2θ . $\square$

2.2 Lattices, Gaussians and Smoothing

For a full-rank lattice $\Lambda \subset \mathbb{R}^m$, $\Lambda^* = \{y : \langle y, x \rangle \in \mathbb{Z} \ \forall x \in \Lambda\}$ is its dual and $\lambda_1(\Lambda)$ its minimum. For $A \in \mathbb{Z}_q^{n \times m}$ let

$$\Lambda_q^\perp(A) = \{x \in \mathbb{Z}^m : Ax = 0 \bmod q\}, \quad \Lambda_q(A) = A^\top \mathbb{Z}^n + q\mathbb{Z}^m,$$

so that $\Lambda_q^\perp(A)^* = q^{-1}\Lambda_q(A)$ and $\det \Lambda_q^\perp(A) \leq q^n$, with equality if A has rank n over $\mathbb{F}_q$.

Let $\rho_s(x) = \exp(-\pi\|x\|^2/s^2)$ and $\rho_s(X) = \sum_{x \in X} \rho_s(x)$. The discrete Gaussian $D_{\Lambda, s, c}$ has probability $\rho_s(x - c)/\rho_s(\Lambda - c)$ at $x \in \Lambda$, and $D_{\Lambda, s} = D_{\Lambda, s, 0}$. For $\varepsilon > 0$, the smoothing parameter $\eta_\varepsilon(\Lambda)$ is the least $s > 0$ with $\rho_{1/s}(\Lambda^* \setminus \{0\}) \leq \varepsilon$ [MR07]. Since $\Lambda \subseteq \Lambda'$ implies $\Lambda'^* \subseteq \Lambda^*$, the smoothing parameter of a superlattice is at most that of the sublattice.

Lemma 2 (Poisson summation [MR07]). *For a full-rank lattice $\Lambda \subset \mathbb{R}^m$, $s > 0$ and $z \in \mathbb{R}^m$,*

$$\sum_{x \in \Lambda} \rho_s(x) e^{2\pi i \langle x, z \rangle} = \frac{s^m}{\det \Lambda} \rho_{1/s}(\Lambda^* - z).$$

In particular $\rho_s(\Lambda) \geq s^m / \det \Lambda$, and $\rho_s(\Lambda + c) \leq \rho_s(\Lambda)$ for all c .

Lemma 3 (Periodic Gaussian [AR05], [HGH26, Lemma 2.3]). *For a full-rank lattice Λ , $s > 0$ and $c \in \mathbb{R}^m$, $\rho_s(\Lambda - c)/\rho_s(\Lambda) \geq \rho_s(c)$.*

Proof. Average the sum defining $\rho_s(\Lambda - c)$ with the same sum after replacing x by $-x$. The result is $\rho_s(c) \sum_{x \in \Lambda} \rho_s(x) \cosh(2\pi \langle x, c \rangle / s^2)$. Since $\cosh \geq 1$, division by $\rho_s(\Lambda)$ proves the claim. $\square$

The following observation of [CJWZ26] is the reason why only one lattice of the chain has to be analysed.

Lemma 4 ([CJWZ26, Lemma 2.10]). *Let $\Lambda \subset \mathbb{R}^m$ be a full-rank lattice and $V \subseteq \mathbb{R}^m$ a subspace such that the orthogonal projection $\pi_V(\Lambda)$ is a lattice spanning V . Then $\eta_\varepsilon(\pi_V(\Lambda)) \leq \eta_\varepsilon(\Lambda)$ for every $\varepsilon > 0$.*

Proof. For $y \in V$ and $x \in \Lambda$ we have $\langle y, \pi_V(x) \rangle = \langle y, x \rangle$. Hence $\pi_V(\Lambda)^* = \Lambda^* \cap V$, and $\rho_{1/s}(\pi_V(\Lambda)^* \setminus \{0\}) \leq \rho_{1/s}(\Lambda^* \setminus \{0\})$ for all s . $\square$

Lemma 5 (Tail bounds [Ban93, Lemma 1.5], [Ban95, Lemma 2.10]). *For a full-rank lattice $\Lambda \subset \mathbb{R}^m$, $s > 0$ and $R > 0$, $\Pr_{X \sim D_{\Lambda, s}}[\|X\|_\infty > Rs] < 2m e^{-\pi R^2}$. Moreover, for $e \sim D_{\mathbb{Z}, \sigma}^m$, $\Pr[\|e\|^2 > \sigma^2 m / \pi] \leq (2/e)^{m/2}$ [HGH26, Lemma 2.4].*

Lemma 6 ([MR07, Lemma 3.2]). *For every full-rank lattice $\Lambda \subset \mathbb{R}^b$, $\eta_{2^{-b}}(\Lambda) \leq \sqrt{b}/\lambda_1(\Lambda^*)$. Consequently, for every $L \geq 1$, $\eta_{e^{-L}}(\Lambda) \leq \sqrt{b + 2L}/\lambda_1(\Lambda^*)$.*

Proof. The first claim is [MR07, Lemma 3.2]. For the second, put $s = \sqrt{b}/\lambda_1(\Lambda^*)$ and $s' = \sqrt{b + 2L}/\lambda_1(\Lambda^*)$. For $0 \neq y \in \Lambda^*$ we have $\|y\| \geq \lambda_1(\Lambda^*)$, hence $\rho_{1/s'}(y) = \rho_{1/s}(y) e^{-2\pi L \|y\|^2 / \lambda_1(\Lambda^*)^2} \leq e^{-2\pi L} \rho_{1/s}(y)$. Summing over $y \neq 0$ gives $\rho_{1/s'}(\Lambda^* \setminus \{0\}) \leq 2^{-b} e^{-2\pi L} \leq e^{-L}$. $\square$

Lemma 7 (Adaptive composition [HGH26, Lemma 3.3]). *Consider a randomized computation that makes Q calls to sampling subroutines, where the input to the j -th call may depend on all previous outputs. Suppose that, for every state before the j -th call, the implemented subroutine's output law is within statistical distance ξ_j of the ideal one. Then the laws of the complete implemented and ideal computations are within statistical distance $\sum_j \xi_j$.*

We will also use the elementary bound $\rho_s(\mathbb{Z}) \leq 1 + s$, which holds for all $s > 0$ because $2 \sum_{k \geq 1} e^{-\pi k^2 / s^2} \leq 2 \int_0^\infty e^{-\pi x^2 / s^2} dx = s$.

2.3 Power-of-Two Cyclotomic Rings

Let d be a power of two, $R = \mathbb{Z}[x]/(x^d + 1)$ and $R_q = R/qR$ for an odd prime q . The map $\text{coeff} : R \rightarrow \mathbb{Z}^d$ sends an element to its coefficient vector, and it extends entrywise to vectors $\mathbf{z} \in R^t$, giving $\text{coeff}(\mathbf{z}) \in \mathbb{Z}^{td}$. Norms of ring elements and ring vectors are norms of their coefficient vectors. For $a \in R$ let $M_a \in \mathbb{Z}^{d \times d}$ be the matrix of multiplication by a , that is, $\text{coeff}(ab) = M_a \text{coeff}(b)$. For a matrix $\mathbf{M} = (m_{ij}) \in R^{k \times t}$, its *coefficient*

expansion is the block matrix $M = (M_{m_{ij}}) \in \mathbb{Z}^{kd \times td}$, so that $\text{coeff}(\mathbf{M}\mathbf{z}) = M \text{coeff}(\mathbf{z})$. We use bold letters for ring objects and the same letter in plain type for coefficient expansions. Throughout, $n = kd$ and $m = (k + \ell)d$.

Let $a^\star = a(x^{-1})$ be the conjugation automorphism of R and R_q , and let $\mathbf{M}^\dagger = (\mathbf{M}^\star)^\top$, so that $(\mathbf{M}\mathbf{M}')^\dagger = \mathbf{M}'^\dagger \mathbf{M}^\dagger$.

Lemma 8. *For every $\mathbf{M} \in R^{k \times t}$, the coefficient expansion of $\mathbf{M}^\dagger$ is $M^\top$. In particular $\text{coeff}(\mathbf{M}^\dagger \mathbf{s}) = M^\top \text{coeff}(\mathbf{s})$.*

Proof. M_x is a signed permutation matrix, so $M_x^\top = M_x^{-1} = M_{x^{-1}}$. Hence $M_a^\top = \sum_i a_i M_x^{-i} = M_{a^\star}$ for $a = \sum_i a_i x^i$, and the claim follows blockwise. $\square$

CRT decomposition. Let $\delta = \text{ord}_{2d}(q)$ be the residue degree of q , $g = d/\delta$, and $Q = q^\delta$. Then $x^d + 1 = \prod_{\nu=1}^g F_\nu$ over $\mathbb{F}_q$, with distinct monic irreducible F_ν of degree δ [LN96, Thm. 2.47]. Moreover $R_q \cong \prod_{\nu=1}^g K_\nu$ with $K_\nu = \mathbb{F}_q[x]/(F_\nu) \cong \mathbb{F}_Q$. We write $a \mapsto a_\nu$ for the projections and $\iota_\nu : K_\nu \rightarrow R_q$ for the embedding with $(\iota_\nu(c))_\nu = c$ and $(\iota_\nu(c))_\mu = 0$ for $\mu \neq \nu$. The kernel $\mathfrak{p}_\nu$ of $R \rightarrow K_\nu$ is a prime ideal of norm Q . These are the primes of R above q , and they are pairwise distinct because $q \nmid 2d$ is unramified in $\mathbb{Q}(\zeta_{2d})$ [Was97].

For $\alpha \in R_q^t$ and $\nu \in [g]$, write $\alpha_\nu \in K_\nu^t$ for the componentwise projection. The CRT support of α is $\mathcal{S}(\alpha) = \{\nu : \alpha_\nu \neq 0\}$, and for $\mathcal{S} \subseteq [g]$ we set

$$V_{\mathcal{S}}^t = \{\mathbf{u} \in R_q^t : \mathbf{u}_\nu = 0 \text{ for all } \nu \notin \mathcal{S}\}, \quad |V_{\mathcal{S}}^t| = Q^{t|\mathcal{S}|} = q^{\delta t|\mathcal{S}|}.$$

The CRT support of a vector $a \in \mathbb{Z}^{td}$ or $a \in \mathbb{Z}_q^{td}$ is that of the ring vector with coefficient vector $a \bmod q$.

2.4 Module Problems

Definition 2. Let $k, \ell \geq 1$, $\beta > 0$, and let $\mathbf{A} \in R_q^{k \times (k+\ell)}$ be uniform.

- **MSIS** $_{k,\ell,q,\beta}^\infty$: given $\mathbf{A}$, find $\mathbf{z} \in R^{k+\ell} \setminus \{\mathbf{0}\}$ with $\mathbf{A}\mathbf{z} = \mathbf{0} \bmod q$ and $\|\mathbf{z}\|_\infty \leq \beta$.
- **MLWE**: for a secret $\mathbf{s} \in R_q^k$ and an error $\mathbf{e} \in R^{k+\ell}$ independent of $\mathbf{A}$, let $\mathbf{b} = \mathbf{A}^\dagger \mathbf{s} + \mathbf{e} \in R_q^{k+\ell}$. *Search* MLWE asks for $\mathbf{s}$ given $(\mathbf{A}, \mathbf{b})$. *Decision* MLWE asks to distinguish $(\mathbf{A}, \mathbf{b})$ from $(\mathbf{A}, \mathbf{u})$ with $\mathbf{u} \in R_q^{k+\ell}$ uniform and independent of $\mathbf{A}$.

By Lemma 8, $\text{coeff}(\mathbf{b}) = A^\top \text{coeff}(\mathbf{s}) + \text{coeff}(\mathbf{e})$. This is exactly the LWE convention of [GH26], with the $n \times m$ matrix A , and MSIS asks for a short nonzero vector of $\Lambda_q^\perp(A)$. The usual convention $b_j = \langle \mathbf{a}_j, \mathbf{s} \rangle + e_j$ corresponds to replacing $\mathbf{A}$ by $\mathbf{A}^\star$, which preserves uniformity. Hence the two conventions are equivalent. A matrix $[\mathbf{A}' \mid I_k]$ with $\mathbf{A}' \in R_q^{k \times \ell}$, and its coefficient expansion $[A' \mid I_n]$, are called *systematic*. Section 5.2 reduces uniform instances to systematic ones.

3 Smoothing Random Module Lattices

Throughout this section d is a power of two, $k, \ell \geq 1$, q is an odd prime, $n = kd$ and $m = (k + \ell)d$, so that $m - n = \ell d$. Let $\mathbf{A}' \in R_q^{k \times \ell}$ be uniform, let $A = [A' \mid I_n] \in \mathbb{Z}_q^{n \times m}$ be the coefficient expansion of $[\mathbf{A}' \mid I_k]$, and let $\Lambda = \Lambda_q^\perp(A)$. We use the three quantities

$$Q_{\text{eff}} = \frac{q}{\sqrt{m-n}}, \quad \Xi = \sqrt{m-n} Q_{\text{eff}}^{n/m}, \quad \Upsilon = \frac{m-n}{m} \ln Q_{\text{eff}},$$

which satisfy $\Xi = (m-n)^{(m-n)/(2m)} q^{n/m}$, $\Xi^m = (m-n)^{(m-n)/2} q^n$ and $q/\Xi = e^\Upsilon$. The value Ξ is the scale at which Λ is smooth (Corollary 2). The value Υ measures how far below q that scale is, and it controls the number of Wagner levels.

The dual lattice. Since $\Lambda^* = q^{-1}\Lambda_q(A) = \mathbb{Z}^m + q^{-1}\{(A'^\top a; a) : a \in \mathbb{Z}_q^n\}$, the dual lattice is the disjoint union of the q^n cosets $\mathbb{Z}^m + q^{-1}(A'^\top a; a)$, one for each label $a \in \mathbb{Z}_q^n$. For $t = (t_1; 0) \in \mathbb{Z}^m$ and $s > 0$, the coordinates split, so

$$\rho_{1/s}(\Lambda^* + t/q) = \sum_{a \in \mathbb{Z}_q^n} \rho_{1/s}\left(\mathbb{Z}^{m-n} + \frac{A'^\top a + t_1}{q}\right) \rho_{1/s}\left(\mathbb{Z}^n + \frac{a}{q}\right). \quad (1)$$

We will repeatedly use that $\rho_{1/s}(\mathbb{Z}^N + v/q) = \rho_p(q\mathbb{Z}^N + v)$ for $v \in \mathbb{Z}^N$ and $p = q/s$.

3.1 Randomness and Length of Labels

Identify a label $a \in \mathbb{Z}_q^n$ with the module vector $\alpha \in R_q^k$ such that $\text{coeff}(\alpha) = a$. By Lemma 8, $A'^\top a = \text{coeff}(\mathbf{A}'^\dagger \alpha)$.

Lemma 9 (Conditional image). *For every fixed $\alpha \in R_q^k$ with CRT support $\mathcal{S}$, the vector $\mathbf{A}'^\dagger \alpha$ is uniform on $V_{\mathcal{S}}^\ell$.*

Proof. The matrix $\mathbf{B} = \mathbf{A}'^*$ is uniform, so the CRT components $b_{ij,\nu} \in K_\nu$ of its entries are independent and uniform. Now $(\mathbf{A}'^\dagger \alpha)_j = \sum_i b_{ij} \alpha_i$, and its ν -component is $\sum_i b_{ij,\nu} \alpha_{i,\nu}$. For $\nu \notin \mathcal{S}$ this vanishes. For $\nu \in \mathcal{S}$ the vector $\alpha_\nu \in K_\nu^k$ is nonzero, so the linear form $(b_i)_i \mapsto \sum_i b_i \alpha_{i,\nu}$ is onto K_ν , and its value on uniform inputs is uniform. Distinct pairs (j, ν) involve disjoint sets of the variables $b_{ij,\nu}$, so these values are independent. $\square$

In the scalar case $d = 1$ every nonzero label has full support, and $A'^\top a$ is uniform on all of $\mathbb{Z}_q^{m-n}$. For module matrices a label supported on a single CRT factor only yields $q^{\delta\ell}$ possible values. The following lemma shows that such labels are necessarily long.

Lemma 10 (Arithmetic uncertainty principle). *Let $\mathcal{S} \subseteq [g]$ and let $0 \neq f \in R$ satisfy $(f \bmod q)_\nu = 0$ for all $\nu \notin \mathcal{S}$. Then $\|f\| \geq q^{1-\delta|\mathcal{S}|/d}$.*

Proof. The hypothesis says $f \in \mathfrak{p}_\nu$ for all $\nu \notin \mathcal{S}$. Since the $\mathfrak{p}_\nu$ are distinct primes, f lies in the ideal $I = \prod_{\nu \notin \mathcal{S}} \mathfrak{p}_\nu$ of norm $Q^{g-|\mathcal{S}|} = q^{d-\delta|\mathcal{S}|}$. As $f \neq 0$, I divides (f) , and therefore $|N(f)| = N((f)) \geq q^{d-\delta|\mathcal{S}|}$. Let $\sigma_1, \dots, \sigma_d$ be the complex embeddings of $R = \mathbb{Z}[\zeta_{2d}]$. The matrix $(\zeta^i)_{\zeta,i}$, where ζ runs over the primitive $2d$ -th roots of unity and $0 \leq i < d$, satisfies $V^*V = dI_d$. Hence $\sum_j |\sigma_j(f)|^2 = d\|f\|^2$. By the AM–GM inequality,

$$|N(f)|^{2/d} = \left(\prod_j |\sigma_j(f)|^2\right)^{1/d} \leq \frac{1}{d} \sum_j |\sigma_j(f)|^2 = \|f\|^2.$$

Combining the two bounds gives $\|f\| \geq q^{(d-\delta|\mathcal{S}|)/d}$. $\square$

The scaled-isometry identity is the special feature of power-of-two cyclotomics used in this arithmetic bound. For $\mathcal{S} \subseteq [g]$ and $t \geq 1$ let

$$I_{\mathcal{S}}^t = \{z \in \mathbb{Z}^{td} : (z \bmod q) \in \text{coeff}(V_{\mathcal{S}}^t)\}, \quad \lambda_{\mathcal{S}} = q^{1-\delta|\mathcal{S}|/d}.$$

This is the coefficient embedding of the t -fold Cartesian product of the ideal $\prod_{\nu \notin \mathcal{S}} \mathfrak{p}_\nu$.

Corollary 1 (Support lattices). *The lattice $I_{\mathcal{S}}^t$ has determinant $q^{t(d-\delta|\mathcal{S}|)}$, and every nonzero $z \in I_{\mathcal{S}}^t$ satisfies $\|z\| \geq \lambda_{\mathcal{S}}$. In particular every $w \in \mathbb{Z}^{td} \setminus q\mathbb{Z}^{td}$ satisfies $\|w\| \geq \lambda_{\mathcal{S}(w)}$.*

Proof. The index of $I_{\mathcal{S}}^t$ in $\mathbb{Z}^{td}$ is $|R_q^t|/|V_{\mathcal{S}}^t| = q^{td-\delta t|\mathcal{S}|}$. A nonzero $z \in I_{\mathcal{S}}^t$ has a nonzero ring component, which satisfies the hypothesis of Lemma 10. The last claim follows because $w \in I_{\mathcal{S}(w)}^t \setminus \{0\}$. $\square$

3.2 Two Counting Lemmas

Lemma 11 (Packing to theta). *Let $\Gamma \subset \mathbb{R}^N$ be a full-rank lattice with $\lambda_1(\Gamma) \geq \lambda > 0$. Then $\rho_p(\Gamma \setminus \{0\}) \leq (2p\sqrt{N}/\lambda)^N$ for every $p > 0$.*

Proof. Balls of radius $\lambda/2$ around lattice points are disjoint. Hence $\#\{v \in \Gamma : 0 < \|v\| \leq r\}$ is 0 for $r < \lambda$ and at most $(1 + 2r/\lambda)^N \leq (3r/\lambda)^N$ for $r \geq \lambda$. Writing $e^{-\pi\|v\|^2/p^2} = \int_{\|v\|}^{\infty} \frac{2\pi r}{p^2} e^{-\pi r^2/p^2} dr$ and exchanging sum and integral gives

$$\rho_p(\Gamma \setminus \{0\}) \leq \left(\frac{3}{\lambda}\right)^N \int_0^{\infty} r^N \frac{2\pi r}{p^2} e^{-\pi r^2/p^2} dr = \left(\frac{3p}{\lambda\sqrt{\pi}}\right)^N \Gamma\left(\frac{N}{2} + 1\right).$$

Finally, $\Gamma(N/2 + 1) \leq N^{N/2}$ for $N \geq 1$ and $3/\sqrt{\pi} < 2$. $\square$

Lemma 12 (Gaussian moment). *For integers $J \geq 1$, $T \geq 0$ and every $p \geq 1$,*

$$\sum_{w \in \mathbb{Z}^J} \|w\|^T \rho_p(w) \leq (3p)^{J+T} T^{T/2},$$

where $0^0 = 1$.

Proof. For $T = 0$ the sum is $\rho_p(\mathbb{Z})^J \leq (1+p)^J$. For $T > 0$, the function $r \mapsto r^T e^{-\pi r^2/(2p^2)}$ is maximized at $r^2 = Tp^2/\pi$, so $\|w\|^T \rho_p(w) \leq (Tp^2/(\pi e))^{T/2} \rho_{\sqrt{2}p}(w)$. Summing over w gives at most

$$T^{T/2} p^T (\pi e)^{-T/2} (1 + \sqrt{2}p)^J \leq T^{T/2} (3p)^{J+T}.$$

$\square$

3.3 The Expected Dual Mass

Lemma 13. *For $0 < s \leq q$ and $p = q/s$,*

$$\Sigma(s) := \sum_{a \in \mathbb{Z}_q^n \setminus \{0\}} q^{-\delta \ell |S(a)|} \rho_p(I_{S(a)}^\ell) \rho_{1/s}\left(\mathbb{Z}^n + \frac{a}{q}\right) \leq \left(\frac{3\Xi}{s}\right)^m + \left(\frac{2\Xi}{s}\right)^m.$$

Proof. We have $\rho_{1/s}(\mathbb{Z}^n + a/q) = \sum_{w \in a + q\mathbb{Z}^n} \rho_p(w)$, and $\mathcal{S}(w) = \mathcal{S}(a)$ for these w . Hence $\Sigma(s) = \sum_{w \in \mathbb{Z}^n \setminus q\mathbb{Z}^n} \rho_p(w) F(\mathcal{S}(w))$, where $F(\mathcal{S}) = q^{-\delta \ell |\mathcal{S}|} \rho_p(I_{\mathcal{S}}^\ell)$. By Corollary 1 and Lemma 11 with $N = \ell d = m - n$, we get $\rho_p(I_{\mathcal{S}}^\ell) \leq 1 + (2p\sqrt{m-n}/\lambda_{\mathcal{S}})^{m-n}$. Moreover $q^{-\delta \ell |\mathcal{S}|} = (\lambda_{\mathcal{S}}/q)^{m-n}$; this is the cancellation. Therefore

$$F(\mathcal{S}) \leq \left(\frac{\lambda_{\mathcal{S}}}{q}\right)^{m-n} + \left(\frac{2\sqrt{m-n}}{s}\right)^{m-n}.$$

For $w \notin q\mathbb{Z}^n$ we have $\lambda_{\mathcal{S}(w)} \leq \|w\|$ by Corollary 1. With Lemma 12 ($J = n$, $T = m - n$, $p \geq 1$) and $\rho_p(\mathbb{Z}^n) \leq (2p)^n$ this gives

$$\Sigma(s) \leq \frac{(3p)^m (m-n)^{\frac{m-n}{2}}}{q^{m-n}} + \left(\frac{2\sqrt{m-n}}{s}\right)^{m-n} (2p)^n = \frac{3^m \Xi^m}{s^m} + \frac{2^m \Xi^m}{s^m},$$

using $p = q/s$ and $\Xi^m = (m-n)^{(m-n)/2} q^n$. $\square$

Theorem 3 (Expected dual mass). *For every $0 < s \leq q$,*

$$\mathbb{E}_{\mathbf{A}'}[\rho_{1/s}(\Lambda^* \setminus \mathbb{Z}^m)] \leq 2 \left(\frac{3\Xi}{s}\right)^m = 2 \left(\frac{3\sqrt{m-n} Q_{\text{eff}}^{n/m}}{s}\right)^m.$$

Proof. By (1) with $t = 0$, the mass of $\Lambda^* \setminus \mathbb{Z}^m$ is the sum over $a \neq 0$ of $\rho_{1/s}(\mathbb{Z}^{m-n} + A'^\top a/q) \rho_{1/s}(\mathbb{Z}^n + a/q)$. Fix $a \neq 0$ with support $\mathcal{S}$. By Lemma 9, $A'^\top a$ is uniform on $\text{coeff}(V_S^\ell)$, so

$$\mathbb{E} \left[\rho_{1/s} \left(\mathbb{Z}^{m-n} + \frac{A'^\top a}{q} \right) \right] = q^{-\delta\ell|\mathcal{S}|} \sum_{u \in \text{coeff}(V_S^\ell)} \rho_p(q\mathbb{Z}^{m-n} + u) = q^{-\delta\ell|\mathcal{S}|} \rho_p(I_S^\ell).$$

Summing over a gives exactly $\Sigma(s)$, and Lemma 13 applies. $\square$

Corollary 2 (Smoothing). *For $L > 0$ let*

$$u(L) = \max \left\{ \sqrt{\frac{L + \ln(7m)}{\pi}}, 6e^{2L/m} \Xi \right\}.$$

If $u(L) \leq q$ and $m \geq 3$, then $\Pr[\eta_{e^{-L}}(\Lambda) > u(L)] \leq e^{-L}$. The same holds for the perturbed lattice Λ_v of Section 3.4.

Proof. Let $s = u(L)$. Then $\rho_{1/s}(\Lambda^* \setminus \{0\}) = \rho_{1/s}(\mathbb{Z}^m \setminus \{0\}) + \rho_{1/s}(\Lambda^* \setminus \mathbb{Z}^m)$, and we bound the two terms separately.

For the first term, $s^2 \geq \ln 7/\pi$ implies $\rho_{1/s}(\mathbb{Z}) - 1 \leq 2.01 e^{-\pi s^2}$, so $x := m(\rho_{1/s}(\mathbb{Z}) - 1) \leq 2.01 e^{-L}/7 < 0.29$. Hence $\rho_{1/s}(\mathbb{Z}^m \setminus \{0\}) \leq e^x - 1 \leq x e^x < e^{-L}/2$.

For the second term, $3\Xi/s \leq e^{-2L/m}/2$, so Theorem 3 (respectively Lemma 14) gives $\mathbb{E}[\rho_{1/s}(\Lambda^* \setminus \mathbb{Z}^m)] \leq 3 \cdot 2^{-m} e^{-2L}$. By Markov's inequality, this mass exceeds $e^{-L}/2$ with probability at most $6 \cdot 2^{-m} e^{-L} \leq e^{-L}$, since $m \geq 3$. $\square$

Remark 1 (The price of structure). For $d = 1$ every nonzero label has full support, and [HGH26, Lemma 2.6] gives the sharper scale $q^{n/m}$ instead of Ξ . The extra factor $(m-n)^{(m-n)/(2m)}$ comes from Lemma 11 applied to CRT-sparse labels. In the sampler it reduces the initial log-deficit $\ln(q/s_0)$ from $\frac{m-n}{m} \ln q$ to $\Upsilon = \frac{m-n}{m} \ln Q_{\text{eff}}$. For fixed ranks and $q = n^{\kappa+o(1)}$ this is $\Theta(\ln n)$ exactly when $\kappa > 1/2$, which is also the condition under which the accuracy needed for LWE is compatible with smoothing $\mathbb{Z}^{m-n}$ [HGH26].

3.4 Stability under a Scalar Row Perturbation

The search algorithm perturbs one coefficient row of the systematic matrix. The resulting matrix is no longer the expansion of a module matrix, but its dual still splits into the same labels. Fix $i \in [n]$, let $v \in \mathbb{F}_q^{m-n}$ be uniform and independent of $\mathbf{A}'$, and put

$$A'_v = A' + e_i v^\top, \quad \Lambda_v = \Lambda_q^\perp([A'_v \mid I_n]), \quad \tilde{v} = (v; 0) \in \mathbb{Z}_q^m.$$

For a label $z \in \mathbb{Z}_q^n$, the free part of the corresponding dual coset of Λ_v is $A_v'^\top z = A'^\top z + z_i v$.

Lemma 14 (Perturbed and affine dual mass). *For every $0 < s \leq q$ and every $\gamma \in \mathbb{F}_q^\times$,*

$$\mathbb{E}_{\mathbf{A}', v} [\rho_{1/s}(\Lambda_v^* \setminus \mathbb{Z}^m)], \quad \mathbb{E}_{\mathbf{A}', v} [\rho_{1/s}(\Lambda_v^* - \gamma \tilde{v}/q)] \leq 3 \left(\frac{3\Xi}{s} \right)^m.$$

Proof. Write $p = q/s$. The dual $\Lambda_v^* - \gamma \tilde{v}/q$ is the disjoint union over $z \in \mathbb{Z}_q^n$ of the cosets $\mathbb{Z}^m + q^{-1}(A'^\top z + (z_i - \gamma)v; z)$, and the first claim is the case $\gamma = 0$ with $z = 0$ removed. We split the labels into two groups.

If $z_i \neq \gamma$, then $(z_i - \gamma)v$ is uniform on $\mathbb{Z}_q^{m-n}$ and independent of $\mathbf{A}'$, so the free part is uniform. The expected mass of the free factor is then $q^{-(m-n)} \rho_p(\mathbb{Z}^{m-n})$. Summing over all labels with $\sum_z \rho_{1/s}(\mathbb{Z}^n + z/q) = \rho_p(\mathbb{Z}^n)$, these labels contribute at most $q^{-(m-n)} (2p)^m = (2q^{n/m}/s)^m \leq (2\Xi/s)^m$.

If $z_i = \gamma$, then $z \neq 0$ (for $\gamma \neq 0$, and by assumption for $\gamma = 0$), and the free part is $A'^\top z$. Its contribution is exactly the corresponding term of $\Sigma(s)$ in Lemma 13.

In total, the expectation is at most $\Sigma(s) + (2\Xi/s)^m \leq (3\Xi/s)^m + 2(2\Xi/s)^m \leq 3(3\Xi/s)^m$. $\square$

Corollary 3 (Simultaneous suppression). *Let $L, L' > 0$ and $0 < s \leq q$ with $s \geq 3(3qe^{L+L'})^{1/m} \Xi$. Then*

$$\Pr_{\mathbf{A}', v} [\exists \gamma \in \mathbb{F}_q^\times : \rho_{1/s}(\Lambda_v^* - \gamma \tilde{v}/q) > e^{-L}] \leq e^{-L'}.$$

Proof. For each fixed γ , Lemma 14 and Markov's inequality bound the event by

$$3e^L (3\Xi/s)^m \leq e^{-L'}/q.$$

A union bound over the $q - 1$ values of γ concludes. $\square$

Since $q^{1/m} = e^{(\ln q)/m} \rightarrow 1$ whenever $\ln q = o(m)$, the factor q in the union bound is harmless. Only base-field candidates are enumerated, so the residue degree δ plays no role.

4 The Gaussian–Wagner Sampler for Smooth Systematic Lattices

We use the lattice walk of [DEL25] with the random prime-index auxiliary lattices of [HGH26]. We make two changes. First, the prefix conditions are derived from a single assumption on the last lattice, by projection. Second, each level gets its own prime, matched to its width, so that an explicit shifted sampler based on deterministic enumeration has subexponential cost. This gives worst-case bit-complexity bounds.

4.1 Lattice Chain, Auxiliary Lattices, and the Ideal Walk

Let $A = [A' \mid I_n] \in \mathbb{Z}_q^{n \times m}$ be systematic, with q prime. For block sizes $b_1, \dots, b_r \geq 1$ with $\sum_i b_i = n$, let $n_i = \sum_{j \leq i} b_j$ and let $A'_{[n_i]}$ consist of the first n_i rows of A' . The prefix lattices are

$$\Lambda_0 = \mathbb{Z}^{m-n}, \quad \Lambda_i = \Lambda_q^\perp([A'_{[n_i]} \mid I_{n_i}]) \subseteq \mathbb{Z}^{m-n+n_i}, \quad \Lambda_r = \Lambda_q^\perp(A).$$

For primes P_i and $a_i \in \mathbb{F}_{P_i}^{b_i} \setminus \{0\}$, the auxiliary lattices are $\Gamma_i = \frac{q}{P_i} \Lambda_{P_i}(a_i^\top) = q\mathbb{Z}^{b_i} + \frac{q}{P_i} a_i \mathbb{Z}$, and the superlattices are $\Lambda'_i = \Lambda_i + (\{0\}^{m-n+n_{i-1}} \times \Gamma_i)$. Deleting the last b_i coordinates maps Λ'_i onto Λ_{i-1} with kernel Γ_i , which is therefore primitive, and $\Lambda'_i/\Lambda_i \cong \mathbb{Z}/P_i\mathbb{Z}$ [HGH26, Lemma 3.1]. The class of a vector of Λ'_i is computable by rational linear algebra [HGH26, Sect. 3.2].

Lemma 15 (Prefixes are projections). *For every i , Λ_i is the orthogonal projection of $\Lambda_q^\perp(A)$ onto the first $m - n + n_i$ coordinates. Hence $\eta_\varepsilon(\Lambda_i) \leq \eta_\varepsilon(\Lambda_q^\perp(A))$ for all $\varepsilon > 0$.*

Proof. Projecting $(u, v) \in \Lambda_q^\perp(A)$ keeps the first n_i equations, so it lands in Λ_i . Conversely, $(u, w) \in \Lambda_i$ is the projection of $(u, w, -A'_{>n_i} u \bmod q) \in \Lambda_q^\perp(A)$, where $A'_{>n_i}$ consists of the remaining rows. Lemma 4 concludes. $\square$

Lemma 16 (Certified auxiliary lattices; cf. [HGH26, Lemma 3.2]). *Let P be prime and $b \geq 1$ with $P^{1/b} \geq 16$ and $\sqrt{b}P^{1/b} \leq P$. For uniform $a \in \mathbb{F}_P^b \setminus \{0\}$, the event*

$$\lambda_1(\Lambda_P^\perp(a^\top)) \geq \sqrt{b}P^{1/b}/16$$

fails with probability at most 2^{1-b} , and it can be certified deterministically in time $2^{O(b)} \text{poly}(b, \ln P)$ by an exact shortest-vector computation [MV13]. On this event, $\Gamma = \frac{q}{P} \Lambda_P(a^\top)$ satisfies $\eta_{e-L}(\Gamma) \leq \frac{16q}{P^{1/b}} \sqrt{1 + 2L/b}$ for every $L \geq 1$.

Proof. Put $R_0 = \sqrt{b}P^{1/b}/16$. The unit-cube count gives

$$\#(\mathbb{Z}^b \cap B(R_0)) \leq P \left[\sqrt{2\pi e} \left(\frac{1}{16} + \frac{1}{2P^{1/b}} \right) \right]^b < P 2^{-b}.$$

No nonzero point of this ball lies in $P\mathbb{Z}^b$. For a uniform $a \in \mathbb{F}_P^b$, each such point satisfies $\langle a, x \rangle = 0$ with probability $1/P$. A union bound, and conditioning on $a \neq 0$, give 2^{1-b} . For $b = 1$ the minimum is P directly. Since $\Gamma^* = q^{-1} \Lambda_P^\perp(a^\top)$, Lemma 6 gives smoothing. The certificate uses only integers: if the exact squared minimum is λ_1^2 , test $(256\lambda_1^2)^b \geq b^b P^2$. $\square$

At level i , a vector $x = (u, w) \in \Lambda_{i-1}$ is lifted as follows. Let Δ_i be the i -th block of b_i rows of A' and $c = \Delta_i u \in \mathbb{Z}^{b_i}$. Sample z from $D_{\Gamma_i, s_{i-1}, c}$ and output $(u, w, z - c) \in \Lambda'_i$; this is the Gaussian lift of [DEL25, Lemma 3.1]. Combining buckets the lifted vectors by their class in Λ'_i/Λ_i and subtracts disjoint pairs within each bucket, keeping one third of the list [DEL25, Alg. 3].

Proposition 1 (Ideal walk [DEL25, Lemmas 3.1, 3.3], [HGH26, Lemmas 3.4, 3.5, Thm. 3.1]). *Let $0 < \varepsilon \leq 1/2$, $s_i = 2^{i/2} s_0$, and fix $a_1, \dots, a_r$. Suppose that $P_i \leq N$ and, for every $i \in [r]$,*

$$s_{i-1} \geq \eta_\varepsilon(\Gamma_i) \quad \text{and} \quad s_{i-1} \geq \sqrt{2} \eta_\varepsilon(\Lambda_i). \quad (2)$$

If the initial list consists of $3^r N$ independent samples from $D_{\mathbb{Z}^{m-n}, s_0}$ and every lift samples exactly, then the output consists of N vectors of $\Lambda_q^\perp(A)$ that are conditionally $5(4^r - 1)\varepsilon$ -similar to independent samples from $D_{\Lambda_q^\perp(A), 2^{r/2} s_0}$.

Proof. By [DEL25, Lemma 3.1], lifting turns conditional θ -similarity into conditional $(\theta + 3\varepsilon)$ -similarity. Combining turns it into $(4\theta + 3\varepsilon)$ -similarity at width $\sqrt{2}$ times larger. The combining step needs $s_{i-1} \geq \sqrt{2} \eta_\varepsilon(\Lambda_i)$, a condition on the sublattice: the convolution in the proof of [DEL25, Lemma 3.3] takes place within cosets of Λ_i , as noted in [HGH26, footnote 6]. The list sizes are $3^{r-i} N \geq N \geq P_i$ after level i , so every bucketing step has at least $3P_i$ inputs. Solving $\theta_i = 4\theta_{i-1} + 15\varepsilon$ with $\theta_0 = 0$ gives $\theta_r = 5(4^r - 1)\varepsilon$. The argument of [HGH26, Thm. 3.1] uses the size of each quotient separately, so distinct primes at different levels are allowed. $\square$

4.2 A Finite-Precision Shifted Sampler

The following lemma replaces the exact or statistical samplers of [DEL25, HGH26]. Its cost is controlled by smoothing alone.

Lemma 17 (Finite shifted sampler). *Let $\Gamma \subset \mathbb{R}^b$ be a full-rank lattice with a rational basis, let $c \in \mathbb{Q}^b$ and $s^2 \in \mathbb{Q}_{>0}$ with $s \geq \eta_{1/2}(\Gamma)$, and let $K \geq 1$ be an integer. There is a deterministic algorithm that, given $\lceil \log M \rceil + 2K + 10$ uniform random bits, where M is bounded in (4), outputs a vector of Γ whose law is within statistical distance e^{-K} of $D_{\Gamma, s, c}$. Its time and space are at most*

$$2^{O(b)} \left(1 + \left[\frac{s}{\det(\Gamma)^{1/b}} \sqrt{1 + \frac{K+2}{b}} \right]^b \right) \cdot \text{poly}(\mathcal{B}, K), \quad (3)$$

where $\mathcal{B}$ is the bit length of the input.

Algorithm 1 Worst-case Gaussian–Wagner sampler

Require: systematic $A = [A' \mid I_n]$; blocks b_i , primes P_i , width s_0 , list size N , precision K

- 1: **for** $i = 1, \dots, r$ **do**: sample $a_i \leftarrow \mathbb{F}_{P_i}^{b_i} \setminus \{0\}$; certify Lemma 16 by exact SVP; **if** it fails **return** $\perp$
- 2: $L_0 \leftarrow 3^r N$ samples from $D_{\mathbb{Z}^{m-n}, s_0}$, coordinatewise by Lemma 17 ($b = 1$)
- 3: **for** $i = 1, \dots, r$ **do**
- 4: lift every $x \in L_{i-1}$ to Λ'_i at width s_{i-1} , using Lemma 17 on Γ_i
- 5: $L_i \leftarrow$ bucket by class in $\Lambda'_i/\Lambda_i \cong \mathbb{Z}/P_i\mathbb{Z}$; subtract disjoint pairs; keep $3^{r-i}N$ differences
- 6: **return** L_r

Proof. Let $Z_t(c) = \rho_t(\Gamma - c)$. Since $t \geq s \geq \eta_{1/2}(\Gamma)$, Lemma 2 gives $\frac{t^b}{2^{\det \Gamma}} \leq Z_t(c) \leq \frac{3t^b}{2^{\det \Gamma}}$ for every c . Let $R^2 = s^2(b + K + 2)$, which is rational, and let $S = \{x \in \Gamma : \|x - c\| \leq R\}$. The set S can be computed deterministically in time $2^{O(b)}(|S| + 1)\text{poly}(\mathcal{B}, K)$ by [DPV11, Prop. 4.1], which extends the Voronoi-cell algorithm of [MV13].

Truncation. Since $\rho_s(y) \leq e^{-\pi R^2/(2s^2)} \rho_{\sqrt{2}s}(y)$ for $\|y\| > R$, we get

$$\Pr_{X \sim D_{\Gamma, s, c}}[X \notin S] \leq e^{-\pi R^2/(2s^2)} \frac{Z_{\sqrt{2}s}(c)}{Z_s(c)} \leq 3 \cdot 2^{b/2} e^{-\pi(b+K+2)/2} < e^{-K}/8.$$

In particular, S is nonempty. Its membership comparisons use the rational R^2 exactly.

Counting. Let $t = R\sqrt{2\pi/b} \geq s$. Every point of S has $\rho_t(x - c) \geq e^{-\pi R^2/t^2} = e^{-b/2}$, so

$$M := |S| \leq e^{b/2} Z_t(c) \leq \frac{3}{2} \left[\sqrt{2\pi e} \frac{s}{\det(\Gamma)^{1/b}} \sqrt{1 + \frac{K+2}{b}} \right]^b. \quad (4)$$

This bound uses smoothing only; it needs no well-conditioned basis.

Sampling. Order S deterministically, put $H_0 = b + K + 2$, and let $w_j = \rho_s(x_j - c)$. Every weight is at least $e^{-\pi H_0}$, so their sum is at least e^{-4H_0} . Compute positive dyadic approximations with absolute error 2^{-p} for

$$p = \lceil \log M \rceil + 8H_0 + 2K + 10.$$

The sum of absolute errors divided by the true total mass is less than $e^{-K}/1024$, and normalization changes the distribution by at most this amount. Clear the common denominator, obtaining positive integers k_j , and put

$$r_0 = \lceil \log M \rceil + 2K + 10, \quad a_j = \left\lfloor 2^{r_0} \frac{\sum_{h \leq j} k_h}{\sum_{h \leq M} k_h} \right\rfloor.$$

Set $a_0 = 0$ and draw exactly r_0 bits as $U \in [0, 2^{r_0})$. Output the unique category with $a_{j-1} \leq U < a_j$; empty categories are skipped. Cumulative rounding changes the law by at most $M2^{-r_0} < e^{-K}/1024$. Adding the tail proves the claim. On the bounded exponent interval $[0, \pi H_0]$, certified approximations to π and $\exp$ by range reduction and convergent series cost $\text{poly}(\mathcal{B}, K, \log M)$ per weight. No unbounded rejection or exact-real comparison is required. The irrational t used for counting occurs only in the analysis, not the algorithm. $\square$

For a finite set indexed by $0, \dots, M_0 - 1$, draw $r_0 = \lceil \log M_0 \rceil + 2K + 10$ bits as $U < 2^{r_0}$ and use the index $\lfloor M_0 U / 2^{r_0} \rfloor$. This gives distance at most $M_0 2^{-r_0} < e^{-K}$ from uniform without enumerating the set. Integer decoding implements auxiliary vectors and perturbations. These finite-bit errors are charged by Lemma 7.

4.3 The Sampler for a Smooth Systematic Lattice

Let $h = (\ln 2)/2$ and $\gamma(x) = x + 8 \ln(1 + x) + 8$. Given $n, q, f \geq 2, L \geq 1$ and $s_* \geq 1$, define

$$r = \left\lfloor 2 \log \frac{q}{fs_*} \right\rfloor, \quad s_i = 2^{(i-r)/2} \frac{q}{f}, \quad x_i = \ln \frac{q}{s_{i-1}} \quad (i \in [r]), \quad \Psi = \frac{n}{\sum_{i=1}^r \gamma(x_i)^{-1}}.$$

Thus $s_r = q/f$, $s_* \leq s_0 < \sqrt{2}s_*$, the x_i decrease with step h from $x_1 = \ln(q/s_0)$ to $x_r = \ln(\sqrt{2}f)$, and $W' := \ln(x_1/x_r)$ is the logarithmic width of the walk. The quantity Ψ is the natural logarithm of the list size; the schedule allocates $b_i \approx \Psi/\gamma(x_i)$ coordinates to level i .

Lemma 18 (Block schedule). *Fix $C_L \geq 1$, and consider a family with $n \rightarrow \infty$ such that*

(S1) $f \rightarrow \infty$;

(S2) $L \leq C_L \Psi$;

(S3) $q \geq 2fs_*$ and $(\ln q)(\ln n) = o(\Psi)$.

Let $w_i > 0$ be rational with $w_i = \gamma(x_i)^{-1}(1 \pm n^{-6})$, put $\hat{b}_i = nw_i / \sum_j w_j$, and round these numbers down, adding one to the first few blocks so that $\sum_i b_i = n$. Let $p_i = \frac{64q}{s_{i-1}} \sqrt{1 + L/b_i}$, and let P_i be the least prime that is at least $\lceil p_i^{b_i} \rceil$. Then, for all sufficiently large n , the following hold.

(a) $\Psi/(3 \ln q) \leq b_i \leq \Psi/\gamma(x_r) + 2$. In particular $b_{\min} = \omega(\ln n)$ and $b_{\max} = o(\Psi)$.

(b) $\ln P_i \leq \Psi + 6 \ln q$. Moreover $P_i^{1/b_i} \geq 64$ and $\sqrt{b_i} P_i^{1/b_i} \leq P_i$.

(c) On the event of Lemma 16, $\eta_{e-L}(\Gamma_i) < s_{i-1}/2$.

(d) If $K = O(\Psi)$, the cost (3) of one lift at level i is $e^{o(\Psi)} \text{poly}(\mathcal{B})$, uniformly in i .

(e) $\frac{hn}{W' + h/x_r} \leq \Psi \leq \frac{hn}{W' - c(x_r)}$, where $c(x) = (8 \ln x + 20)/x$, and the upper bound requires $W' > c(x_r)$.

The schedule uses the inverse-log allocation of [HGH26, Lemma B.1] and [CJWZ26] with primes matched to each level. Unlike [HGH26], it needs no prefix inequalities, because prefixes are handled by Lemma 15.

of Lemma 18. Since $s_0 \geq s_* \geq 1$, we have $x_1 \leq \ln q$, and (S1) gives $x_r = \ln(\sqrt{2}f) \rightarrow \infty$. In particular $q \rightarrow \infty$, and $\gamma(x) \leq 2x$ for all large x .

Blocks. The rational numbers $\hat{b}_i$ sum to n , and normalization of the w_i gives $|\hat{b}_i - \Psi/\gamma(x_i)| \leq 3n^{-5} < 1$ for large n . The rounding deficit is a nonnegative integer smaller than r . Thus $|b_i - \Psi/\gamma(x_i)| \leq 2$. As $\gamma(x_i) \leq \gamma(x_1) \leq 2 \ln q$ for large n , we get $b_i \geq \Psi/(2 \ln q) - 2 \geq \Psi/(3 \ln q)$, which is $\omega(\ln n)$ by (S3). Moreover $b_i \leq \Psi/\gamma(x_r) + 2 = o(\Psi)$, since $\gamma(x_r) \geq x_r \rightarrow \infty$. This proves (a). In particular, for large n every $b_i \geq 2$ and $\Psi/\gamma(x_i) \geq 4$.

Primes. By (S2) and (a), $L/b_i \leq C_L \Psi/(\Psi/\gamma(x_i) - 2) \leq 2C_L \gamma(x_i) \leq 34C_L(1 + x_i)$. Hence

$$\ln p_i = x_i + \ln 64 + \frac{1}{2} \ln(1 + L/b_i) \leq x_i + \ln 64 + \frac{1}{2} \ln(1 + 34C_L) + \frac{1}{2} \ln(1 + x_i) \leq \gamma(x_i),$$

where the last inequality holds for $x_i \geq x_r$ large enough, since $\gamma(x) - x = 8 \ln(1 + x) + 8$. By Bertrand's postulate $P_i \leq 2 \lceil p_i^{b_i} \rceil \leq 4p_i^{b_i}$, so

$$\ln P_i \leq b_i \gamma(x_i) + \ln 4 \leq \Psi + 2\gamma(x_i) + \ln 4 \leq \Psi + 6 \ln q.$$

Moreover $P_i^{1/b_i} \geq p_i \geq 64$. For $b_i \geq 2$ we have $P_i^{1/b_i} \leq \sqrt{P_i}$ and $b_i < 64^{b_i} \leq P_i$, hence $\sqrt{b_i} P_i^{1/b_i} \leq P_i$. This proves (b).

Auxiliary smoothing. On the event of Lemma 16, using $P_i^{1/b_i} \geq p_i$,

$$\eta_{e-L}(\Gamma_i) \leq \frac{16q}{p_i} \sqrt{1 + \frac{2L}{b_i}} = \frac{s_{i-1}}{4} \sqrt{\frac{1 + 2L/b_i}{1 + L/b_i}} \leq \frac{\sqrt{2}}{4} s_{i-1} < \frac{s_{i-1}}{2},$$

which proves (c).

Lifting cost. Since $\det \Gamma_i = q^{b_i}/P_i$ and $P_i^{1/b_i} \leq 4^{1/b_i} p_i \leq 2p_i$, the ratio in (3) satisfies $s_{i-1} P_i^{1/b_i}/q \leq 128\sqrt{1 + L/b_i}$. The logarithm of (3) is therefore $O(b_i(1 + \ln(1 + (L + K)/b_i))) + O(\ln \text{poly}(\mathcal{B}))$. With $L + K \leq C\Psi$, the function $b \mapsto b \ln(1 + C\Psi/b)$ is increasing, so it is at most $b_{\max} \ln(1 + C\Psi/b_{\max}) = o(\Psi)$, because $b_{\max} = o(\Psi)$ by (a). This proves (d).

The size of Ψ . Let $S = \sum_i \gamma(x_i)^{-1}$, so that $\Psi = n/S$. The points $x_i = x_r + (r - i)h$ form an arithmetic progression. Since $\gamma(x) \geq x$ and $1/x$ is decreasing,

$$S \leq \sum_{j=0}^{r-1} \frac{1}{x_r + jh} \leq \frac{1}{x_r} + \frac{1}{h} \int_{x_r}^{x_1} \frac{dx}{x} = \frac{1}{x_r} + \frac{W'}{h},$$

which gives the lower bound on Ψ . Since $1/\gamma$ is decreasing, $S \geq \frac{1}{h} \int_{x_r}^{x_1} \frac{dx}{\gamma(x)}$. Using $\frac{1}{x} - \frac{1}{\gamma(x)} \leq \frac{\gamma(x)-x}{x^2}$ and $\ln(1+t) \leq \ln t + 1/t$, we obtain

$$\int_{x_r}^{x_1} \frac{dx}{\gamma(x)} \geq W' - \int_{x_r}^{\infty} \frac{8 \ln(1+t) + 8}{t^2} dt \geq W' - \frac{8 \ln x_r + 20}{x_r} \quad (x_r \geq 1),$$

which gives the upper bound and proves (e). $\square$

Theorem 4 (Sampler for a smooth systematic lattice). *Fix $\xi \in (0, 1]$ and $C_L \geq 1$. Consider parameters $(n, m, q, f, L, s_\star)$ satisfying (S1)–(S3), with $n < m \leq n^{C_m}$ for a fixed C_m , rational $f^2, s_\star^2 > 0$, integer L , and polynomial-length parameter encodings in n . Choose a power of two N with $\ln N = (1 + \xi)\Psi + O(1) \geq (1 + \xi)\Psi$ and an integer $K = 3 \ln N + (r + 1) \ln 3 + \ln m + O(1)$ satisfying $K \geq 3 \ln N + (r + 1) \ln 3 + \ln m + 2$, and take the blocks and primes of Lemma 18. Then, for all sufficiently large n , the following holds for every systematic $A = [A' \mid I_n] \in \mathbb{Z}_q^{n \times m}$ with*

$$s_\star \geq 2 \eta_{e-L}(\Lambda_q^\perp(A)). \quad (5)$$

Algorithm 1 outputs $\perp$ with probability at most $2r2^{-b_{\min}} + N^{-2} = n^{-\omega(1)}$. Conditioned on the auxiliary vectors passing certification, its output is within statistical distance N^{-2} of a list of N vectors of $\Lambda_q^\perp(A)$ that is conditionally $5(4^r - 1)e^{-L}$ -similar to independent samples from $D_{\Lambda_q^\perp(A), q/f}$; here $4^r < q^4$. Its running time and space are at most $e^{(1+\xi+o(1))\Psi}$ in the worst case. The output depends only on A and fresh random bits.

Proof. We check the hypotheses of Proposition 1 with $\varepsilon = e^{-L} \leq 1/2$, for every certified choice of $a_1, \dots, a_r$.

- *Auxiliary lattices.* Lemma 18(b) gives the hypotheses of Lemma 16, so ideal uniform auxiliary choices fail certification with probability at most $\sum_i 2^{1-b_i} \leq 2r2^{-b_{\min}}$. After certification, Lemma 18(c) gives $\eta_\varepsilon(\Gamma_i) < s_{i-1}/2$.
- *Prefix lattices.* By Lemma 15 and (5), $\sqrt{2} \eta_\varepsilon(\Lambda_i) \leq \sqrt{2} \eta_\varepsilon(\Lambda_q^\perp(A)) \leq s_\star/\sqrt{2} < s_0 \leq s_{i-1}$.

- *List size.* By Lemma 18(b) and (S3), $\ln P_i \leq \Psi + 6 \ln q < (1 + \xi)\Psi \leq \ln N$.

Hence the ideal walk is conditionally $5(4^r - 1)e^{-L}$ -similar to $D_{\Lambda_q^\perp(A), s_r}$ with $s_r = q/f$. Since $f s_\star \geq 2$, we have $r \leq 2 \log q$, so $4^r \leq q^4$.

Implementation distance. Fix any tuple of auxiliary vectors that passes certification. The reference walk uses this same tuple and exact Gaussian calls. It has the conditional-similarity guarantee above, independently of how the tuple was selected. There are at most $3^r(m - n)N$ one-dimensional calls and $\frac{3^{r+1}-3}{2}N$ lifts. Their hypotheses hold since $s_0 \geq 1 > \eta_{1/2}(\mathbb{Z})$ and $\eta_{1/2}(\Gamma_i) \leq \eta_\varepsilon(\Gamma_i) < s_{i-1}$. Lemma 7 bounds their joint distance by $3^{r+1}mNe^{-K} \leq N^{-2}$. Averaging over the actual certified tuples preserves conditional similarity and this distance bound. Separately, the r approximate uniform auxiliary choices change the probability of failing certification by at most $re^{-K} \leq N^{-2}$. Thus no conditioning of a global statistical-distance bound on a rare event is used.

Cost. By $m \leq n^{C_m}$, (S3), and polynomial-length encodings, $K = O(\Psi)$ and all polynomial bit factors are $e^{o(\Psi)}$. Lemma 18(d) implies that each lift costs $e^{o(\Psi)}$. Certification costs $2^{O(b_{\max})} = e^{o(\Psi)}$. Scanning for the primes costs at most $re^{\Psi+6 \ln q} \text{poly}(n)$. The $3^r(m - n)N$ initial coordinates cost $\text{poly}(q, K)$ each, since $s_0 \leq q$. Bucketing costs $N \text{poly}(n)$ per level. Since $3^r \leq q^{2 \log 3} = e^{o(\Psi)}$ by (S3), the total time and space are $Ne^{o(\Psi)} = e^{(1+\xi+o(1))\Psi}$. The truncation in Lemma 17 bounds all vectors deterministically, so all bit lengths are polynomial. $\square$

By Lemma 18(e), the running time is $2^{(1+\xi+o(1))n/(2(W'-c(x_r)))}$. In our applications $W' \rightarrow \infty$ and $c(x_r) \rightarrow 0$, so the time is $2^{(1/2+\xi')n/W'}$ for every $\xi' > \xi/2$.

Computing the parameters. The displayed real formulas specify scales, not exact-real oracles. The integer r is obtained by comparing powers of two with the rational $q^2/(f^2 s_\star^2)$; then every $s_i^2 = 2^{i-r}q^2/f^2$ is rational. Certified logarithm intervals produce the w_i of Lemma 18 in polynomial time. Its bound allows the resulting $O(1)$ block errors. Since p_i^2 is rational, $\lceil p_i^{b_i} \rceil$ is found by integer square comparisons; a deterministic polynomial-time primality test while scanning a Bertrand interval finds P_i . To choose N , approximate $(1 + \xi)\Psi/\ln 2$ within $1/4$, round an upper endpoint up, and exponentiate 2. The same upper-endpoint rule gives K and L ; logarithmic ceilings in the applications allow an extra $O(1)$ upward rounding. Fixing a rational slack smaller than a requested real slack suffices.

For the applications, replace analytic squared widths and error bounds by rational upper bounds within relative error n^{-6} , and permissible SIS norms by lower bounds of the same accuracy. In particular the implemented $s_\star^2$ is an upper approximation to $4u(L)^2$. These changes contribute only $o(1)$ to logarithmic deficits and preserve the strict smoothing, tail, and Fourier margins. All reference distributions use the *implemented* widths; no closeness claim between exponentially long lists at distinct widths is needed. Compute Fourier means to absolute error at most $\tau/64$ and their thresholds to the same accuracy. A conservative lower approximation to the bias bound may be used; it retains the separation of Lemma 20.

4.4 Instantiation for Module Lattices

Corollary 4 (Module relation sampler). *Let $d, k, \ell, q, n, m, \Xi, \Upsilon$ be as in Section 3 with $m \geq 3$. Let $L \geq 1$ with $u(L) \leq q$ for u as in Corollary 2, and put $s_\star = 2u(L)$. If $\mathbf{A}' \in R_q^{k \times \ell}$ is uniform, then the coefficient expansion of $[\mathbf{A}' \mid I_k]$ satisfies (5) with probability at least $1 - e^{-L}$, and so does its perturbation $[A'_v \mid I_n]$ of Section 3.4, for uniform v . Moreover*

$$\min \left\{ \ln \frac{q}{2\sqrt{(L + \ln(7m))/\pi}}, \Upsilon - \frac{2L}{m} - \ln 12 \right\} - \frac{1}{2} \ln 2 < x_1 < \Upsilon.$$

Proof. The first claim is Corollary 2. For the second, use $s_\star \leq s_0 < \sqrt{2}s_\star$ and

$$\ln(q/(12e^{2L/m}\Xi)) = \Upsilon - 2L/m - \ln 12,$$

which holds because $q/\Xi = e^\Upsilon$. $\square$

5 Module-SIS and Module-LWE

5.1 The Fourier Test

The following facts about the Aharonov–Regev score are standard [AR05]. Items (i), (ii) and (iv) are [HGH26, Lemma 4.1 and Eq. (6)]; item (iii) is the elementary inequality that we use for affine shifts.

Lemma 19. *Let $A \in \mathbb{Z}_q^{n \times m}$ have rank n , let $\Lambda = \Lambda_q^\perp(A)$, $f \geq 1$ and $w = q/f$, and for $b \in \mathbb{R}^m$ let $\mu_A(b) = \mathbb{E}_{X \sim D_{\Lambda, w}}[\cos(2\pi\langle X, b \rangle/q)]$.*

(i) $\mu_A(b) = \rho_{1/w}(\Lambda^\star - b/q)/\rho_{1/w}(\Lambda^\star) \geq 0$, and it depends only on b modulo $\Lambda_q(A) = q\Lambda^\star$.

(ii) If $b - x \in \Lambda_q(A)$, then $\mu_A(b) \geq e^{-\pi\|x\|^2/f^2}$.

(iii) For all $b, x \in \mathbb{R}^m$, $\mu_A(b+x) \leq e^{\pi\|x\|^2/f^2} \rho_{\sqrt{2}/w}(\Lambda^\star - b/q)$.

(iv) $\Pr_{X \sim D_{\Lambda, w}}[X \in q\mathbb{Z}^m] \leq (2f)^m q^{-(m-n)}$. Consequently, for uniform $b \in \mathbb{Z}_q^m$ we have $\mathbb{E}_b[\mu_A(b)] \leq (2f)^m q^{-(m-n)}$.

Proof. (i) This follows from Lemma 2 with $z = b/q$, divided by the same identity at $z = 0$; the ratio is real. (ii) Apply Lemma 3 to $\Lambda^\star$ at width $1/w$ and $c = x/q$, using $\Lambda^\star - b/q = \Lambda^\star - x/q$. (iii) For $y \in \Lambda^\star - b/q$ we have $\|y - x/q\|^2 \geq \frac{1}{2}\|y\|^2 - \|x/q\|^2$, so $\rho_{1/w}(y - x/q) \leq e^{\pi w^2\|x\|^2/q^2} \rho_{\sqrt{2}/w}(y)$. Sum over y and use $\rho_{1/w}(\Lambda^\star) \geq 1$. (iv) We have $\rho_w(q\mathbb{Z}^m) = \rho_{1/f}(\mathbb{Z})^m \leq 2^m$, and $\rho_w(\Lambda) \geq (q/f)^m/q^n$ by Lemma 2. For uniform b , $\mathbb{E}_b \cos(2\pi\langle X, b \rangle/q) = 1[X \in q\mathbb{Z}^m]$. $\square$

For a list $\mathcal{L} = (X_1, \dots, X_N)$ let $\hat{\mu}_{\mathcal{L}}(b) = \frac{1}{N} \sum_j \cos(2\pi\langle X_j, b \rangle/q)$.

Lemma 20 ([HGH26, Thm. 4.1]). *Let A be as in Lemma 19, $0 < \tau \leq 1$, and $H \subset \mathbb{Z}_q^m$ finite. Let $\mathcal{L}$ be a random list in Λ , independent of H given A , whose law is within statistical distance ξ of a law that is conditionally θ -similar to $D_{\Lambda, q/f}$, with $\theta \leq \tau/64$. Then*

$$\Pr[\exists b \in H : |\hat{\mu}_{\mathcal{L}}(b) - \mu_A(b)| > 3\tau/32] \leq \xi + 2|H|e^{-N\tau^2/2048}.$$

Proof. Apply Lemma 1 with $t = \tau/16$ to each $h_b(x) = \cos(2\pi\langle x, b \rangle/q)$ under the reference law, take a union bound over H , and pay ξ for the change of law. $\square$

The cosines are evaluated with absolute error at most $\tau/64$ using $O(\ln(1/\tau) + \ln q)$ bits of precision, and the threshold $\tau/2$ is replaced by a rational within $\tau/64$. Outside the event of Lemma 20, an ideal score of at least τ then yields a computed score above the threshold, and an ideal score of at most $\tau/8$ yields one below it.

5.2 Normalization by Small Shears

Write $\mathbf{A} = [\mathbf{C} \mid \mathbf{G}]$ with $\mathbf{C} \in R_q^{k \times \ell}$, $\mathbf{G} \in R_q^{k \times k}$, and let $\mathbf{c}$ be the first column of $\mathbf{C}$. For integers $0 \leq t < H$ let

$$\mathbf{U}_t = I_{k+\ell} + t e_1 e_{\ell+1}^\top, \quad \mathbf{A}_t = \mathbf{A} \mathbf{U}_t = [\mathbf{C} \mid \mathbf{G}_t], \quad \mathbf{G}_t = \mathbf{G} + t \mathbf{c} \mathbf{c}_1^\top.$$

Lemma 21 (Small shears). *Let $\mathbf{A}$ be uniform and $H < q$. Then*

$$\Pr[\forall 0 \leq t < H : \mathbf{G}_t \notin \text{GL}_k(R_q)] \leq \frac{2g}{Q(Q-1)} + \left(\frac{g}{Q+1}\right)^H \leq \frac{2g}{Q(Q-1)} + 2^{-H}.$$

For each fixed t , $\mathbf{A}_t$ is uniform, and conditioned on $\mathbf{G}_t \in \text{GL}_k(R_q)$, the matrix $\mathbf{A}'_t = \mathbf{G}_t^{-1}\mathbf{C}$ is uniform and independent of $\mathbf{G}_t$. The coefficient operator norms of $\mathbf{U}_t$ and $\mathbf{U}_t^\dagger$ in both ℓ_2 and ℓ_∞ are at most $1 + t \leq H$.

Proof. Work at one CRT factor K_ν , and let D consist of the last $k-1$ columns of $\mathbf{G}_\nu$. With probability at least $1 - \sum_{j \geq 2} Q^{-j} = 1 - \frac{1}{Q(Q-1)}$, the matrix D has rank $k-1$. Then $\det(\mathbf{G}_{t,\nu}) = \alpha + t\beta$, where $\alpha = \phi(g_1)$ and $\beta = \phi(\mathbf{c}_\nu)$ for the nonzero linear form $\phi = \det(\cdot, D)$. Since the first column g_1 of $\mathbf{G}_\nu$ and $\mathbf{c}_\nu$ are independent and uniform, so are α and β . Hence the pencil vanishes identically with probability at most $\frac{1}{Q(Q-1)} + \frac{1}{Q^2} \leq \frac{2}{Q(Q-1)}$. If it does not vanish identically, it has at most one root $t \in K_\nu$, and each fixed value is that root with probability $(Q-1)/(Q^2-1) = 1/(Q+1)$. The factors are independent. If every $t < H$ is a root at some factor, then distinct factors must block distinct values of t ; a union bound over these injective assignments gives $(g/(Q+1))^H$. Finally $Q \equiv 1 \pmod{2d}$, so $Q \geq 2d+1$ and $g/(Q+1) < 1/2$. The shear is invertible over R_q , so $\mathbf{A}_t = [\mathbf{C} \mid \mathbf{G}_t]$ is uniform. In particular, $\mathbf{C}$ and $\mathbf{G}_t$ are independent, although $\mathbf{G}_t$ is defined using the first column of $\mathbf{C}$. Given an invertible $\mathbf{G}_t$, the matrix $\mathbf{G}_t^{-1}\mathbf{C}$ is therefore uniform and independent of $\mathbf{G}_t$. The norm bounds are immediate. $\square$

All our algorithms take $H = \lceil 3 \log n \rceil$, try $t = 0, 1, \dots$ in order, and continue with the first t for which $\mathbf{G}_t$ is invertible. Invertibility over R_q and the inverse are computed with determinants, a polynomial gcd with $x^d + 1$ and adjugates, without factoring $x^d + 1$. For MLWE we set $\mathbf{b}_t = \mathbf{U}_t^\dagger \mathbf{b}$, so that

$$\mathbf{b}_t = [\mathbf{A}'_t \mid I_k]^\dagger \bar{\mathbf{s}} + \mathbf{e}_t, \quad \bar{\mathbf{s}} = \mathbf{G}_t^\dagger \mathbf{s}, \quad \mathbf{e}_t = \mathbf{U}_t^\dagger \mathbf{e}, \quad \|\mathbf{e}_t\| \leq H \|\mathbf{e}\|. \quad (6)$$

For MSIS, a solution $\mathbf{z}'$ for $\mathbf{A}_t$ yields the solution $\mathbf{U}_t \mathbf{z}'$ for $\mathbf{A}$, which is nonzero and longer in ℓ_∞ by at most a factor H . We never claim that the selected matrix has the unconditional uniform law. Instead, for each fixed t we analyse the algorithm conditioned on $\mathbf{G}_t$ being invertible, and we take a union bound over the H shears. For the fixed-rank applications, $H < q$ for all sufficiently large n . Since $g \leq d$ and $Q \geq 2d+1$, the failure bound of Lemma 21 is at most $1/(2d+1) + n^{-3} = o(1)$.

5.3 Module-SIS

Theorem 5 (Module-SIS $^\infty$). *Fix integers $k, \ell \geq 1$ and reals $b, c, \xi > 0$. For d a power of two let $n = kd$, $m = (k + \ell)d$, let q be prime with $\ln(q/\sqrt{n}) = (\ln n)^{b+o(1)}$, and let $\beta = q/(\ln n)^c$. There is a randomized algorithm that, given a uniform $\mathbf{A} \in R_q^{k \times (k+\ell)}$, solves MSIS $_{k,\ell,q,\beta}^\infty$ with probability $1 - o(1)$. Its worst-case time and space are $2^{(1/(2b+\xi)n)/\ln \ln n}$.*

Proof. Algorithm. Normalize by a shear (Section 5.2). Run Algorithm 1 on the coefficient expansion of $[\mathbf{A}'_t \mid I_k]$ with the parameters of Theorem 4 for

$$f = H(\ln n)^c \sqrt{2 \ln(2m)/\pi}, \quad L = \lceil (\ln n)^2 + 8 \ln q \rceil, \quad s_\star = 2u(L),$$

and slack $\xi' = \min\{1, b\xi\}/2$ in place of ξ . If the first output vector X_1 is nonzero with $\|X_1\|_\infty \leq \beta/H$, return $\mathbf{U}_t \text{coeff}^{-1}(X_1)$.

Parameters. We have $\ln q = \frac{1}{2} \ln n + (\ln n)^{b+o(1)}$ and $\Upsilon = \frac{\ell}{k+\ell}(\ln(q/\sqrt{n}) + \frac{1}{2} \ln \frac{k}{\ell}) = (\ln n)^{b+o(1)}$. Since L is polylogarithmic and $\ln f = O(\ln \ln n) = o(\Upsilon)$, both branches of $u(L)$ are $o(q/f)$. Thus $u(L) \leq q$ and $q \geq 2fs_\star$ hold before invoking the sampler. Corollary 4 gives

$\ln x_1 = (b + o(1)) \ln \ln n$, while $x_r = \ln(\sqrt{2}f) = O(\ln \ln n)$. Hence $W' = (b + o(1)) \ln \ln n$ and $c(x_r) \rightarrow 0$. Lemma 18(e) gives $\Psi \geq (\frac{h}{b} - o(1)) \frac{n}{\ln \ln n}$, so (S2) and (S3) hold; (S1) is clear. Lemma 18(e) also gives $\Psi \leq (\frac{h}{b} + o(1)) \frac{n}{\ln \ln n}$, so the running time $e^{(1+\xi'+o(1))\Psi}$ of Theorem 4 is at most $2^{(1/(2b)+\xi)n/\ln \ln n}$.

Correctness. Fix t and condition on $\mathbf{G}_t$ being invertible, so that $\mathbf{A}'_t$ is uniform. By Corollary 4 and Theorem 4, except with probability $n^{-\omega(1)}$, the vector X_1 is within statistical distance N^{-2} of a vector that is θ -similar to $D = D_{\Lambda, q/f}$, where $\Lambda = \Lambda_q^\perp(A_t)$ and $\theta < 5q^4 e^{-L} \leq 5q^{-4} e^{-(\ln n)^2} = o(1)$. By Lemma 2, $D(0) = \rho_{q/f}(\Lambda)^{-1} \leq f^m q^{-\ell d} = e^{-\Omega(n \ln q)}$. By Lemma 5 with $R = \beta f / (Hq) = \sqrt{2 \ln(2m)/\pi}$, we get $D(\|y\|_\infty > \beta/H) < 1/(2m)$. A union bound over the H shears and Lemma 21 give success probability $1 - o(1)$. $\square$

Corollary 5 (ℓ_2 norm and MSIS $^\times$). *In the setting of Theorem 5, the same algorithm outputs, with probability $1 - o(1)$, a vector $\mathbf{z} \in R^{k+\ell}$ with $\mathbf{A}\mathbf{z} = \mathbf{0} \bmod q$, $\mathbf{z} \not\equiv \mathbf{0} \bmod q$ and $\|\mathbf{z}\| \leq \sqrt{m}q/(\ln n)^c$. In particular, it solves the module analogue of the ℓ_2 problem SIS $^\times$ defined in [DEL25, Sect. 4.2] with this bound in the same time.*

Proof. The output of Theorem 5 is nonzero and satisfies $\|\mathbf{z}\|_\infty \leq q/(\ln n)^c < q$. A nonzero integer vector whose coefficients are all smaller than q in absolute value is nonzero modulo q . Finally, $\|\mathbf{z}\| \leq \sqrt{m}\|\mathbf{z}\|_\infty$. $\square$

The nonzero-modulo- q condition is essential when the Euclidean bound exceeds q , since qe_1 is then a trivial homogeneous solution. This corollary does not solve inhomogeneous MSIS, where a prescribed target $\mathbf{t}$ requires $\mathbf{A}\mathbf{z} = \mathbf{t} \bmod q$. Our sampler produces centered homogeneous relations only. We do not supply either a coset sampler or a reduction from inhomogeneous MSIS that preserves the module distribution and the claimed parameters.

5.4 Decision Module-LWE

The two LWE theorems share their hypotheses and parameters. Fix integers $k, \ell \geq 1$ and reals $0 \leq a < b$. For d a power of two let $n = kd$, $m = (k + \ell)d$, and let q be a prime and $\sigma \geq 1$ with

$$\ln(2 + \sigma) = (\ln n)^{a+o(1)}, \quad \ln(q/\sqrt{n}) = (\ln n)^{b+o(1)}. \quad (7)$$

The secret $\mathbf{s} \in R_q^k$ is arbitrary, and $\mathbf{e} \in R^{k+\ell}$ is any random vector independent of $\mathbf{A}$ with $\Pr[\|\mathbf{e}\| > \sigma\sqrt{m}] = o(1)$. For example, $\mathbf{e} \sim D_{\mathbb{Z}, \sigma\sqrt{\pi}}^m$ qualifies by Lemma 5. With $H = \lceil 3 \log n \rceil$ we set

$$\begin{aligned} f^2 &= F(H\sigma)^2 \ln \ln n, & F &= \frac{16\pi(k+\ell)(b-a)}{k \ln 2}, \\ \tau &= e^{-\pi H^2 \sigma^2 m / f^2}, & L &= \left\lceil \ln \frac{320q^4}{\tau} \right\rceil, \end{aligned} \quad (8)$$

$s_\star = 2u(L)$, and the parameters of Theorem 4.

Lemma 22 (Parameters). *In this setting, (S1)–(S3) hold with $C_L = 1$, $W' = (b - a + o(1)) \ln \ln n$, $c(x_r) \rightarrow 0$, and $u(L) \leq q$. Moreover $5(4^r - 1)e^{-L} \leq \tau/64$, $N\tau^2 \geq e^{\Psi/2}$, and*

$$\Psi = \left(\frac{h}{b-a} + o(1) \right) \frac{n}{\ln \ln n}, \quad \ell d \ln q - m \ln(2f) - \ln(1/\tau) \geq \left(\frac{\ell}{2} - o(1) \right) d \ln q.$$

Proof. We have $\ln(1/\tau) = \pi m / (F \ln \ln n) = \frac{\ln 2}{16(b-a)} \frac{n}{\ln \ln n}$. Since $\ln q = \frac{1}{2} \ln n + (\ln n)^{b+o(1)}$ is polylogarithmic, $L = \Theta(n/\ln \ln n)$ and $2L/m = o(1)$. Both $\ln(q/\sqrt{L}) = \ln(q/\sqrt{n}) + O(\ln \ln \ln n)$ and Υ are $(\ln n)^{b+o(1)}$, so Corollary 4 gives $\ln x_1 = (b + o(1)) \ln \ln n$. Next,

Algorithm 2 Search Module-LWE**Require:** $\mathbf{A} \in R_q^{k \times (k+\ell)}$, $\mathbf{b} \in R_q^{k+\ell}$; parameters (8)

- 1: normalize by a shear: $\bar{A} = [A' \mid I_n]$ (coefficient expansion of $[\mathbf{A}'_t \mid I_k]$), $\bar{b} = \text{coeff}(\mathbf{b}_t)$
- 2: **for** $i = 1, \dots, n$ **do**
- 3: $v \leftarrow \mathbb{F}_q^{\ell d}$; $A'_v \leftarrow A' + e_i v^\top$; $\tilde{v} \leftarrow (v; 0)$
- 4: run Algorithm 1 on $[A'_v \mid I_n]$; obtain $\mathcal{L}$
- 5: $\bar{s}_i \leftarrow$ the unique $c \in \mathbb{F}_q$ with $\hat{\mu}_{\mathcal{L}}(\bar{b} + c\tilde{v}) \geq \tau/2$ (abort if none or several)
- 6: **return** $\mathbf{s} = (\mathbf{G}_t^\dagger)^{-1} \text{coeff}^{-1}(\bar{\mathbf{s}})$

$x_r = \ln(\sqrt{2}f) = \ln \sigma + O(\ln \ln n)$. If $a > 0$ then $\ln \sigma = (\ln n)^{a+o(1)}$; if $a = 0$ then $x_r \leq (\ln n)^{o(1)}$. In both cases $\ln x_r = (a + o(1)) \ln \ln n$ and $x_r \rightarrow \infty$, which gives W' and $c(x_r) \rightarrow 0$. Lemma 18(e) then gives the stated Ψ . Hence $2 \ln(1/\tau) = (\frac{1}{4} + o(1))\Psi \leq \Psi/3$ for large n , so $L \leq \Psi$ and $N\tau^2 \geq e^{(1+\xi)\Psi - \Psi/3} \geq e^{\Psi/2}$. The bound $4^r < q^4$ gives $5(4^r - 1)e^{-L} \leq \tau/64$. The remaining conditions hold because $\ln f = O((\ln n)^{a+o(1)} + \ln \ln n) = o(\ln q)$, and $\Upsilon - \ln f \rightarrow \infty$ gives $q \geq 2fs_*$. Finally, $m \ln(2f) = o((m - n) \ln q)$ and $\ln(1/\tau) = o((m - n) \ln q)$ for fixed ranks, proving the asserted uniform-case gap. $\square$

Theorem 6 (Decision Module-LWE). *In the setting of (7), for every $\xi > 0$ there is a randomized algorithm that distinguishes $(\mathbf{A}, \mathbf{A}^\dagger \mathbf{s} + \mathbf{e})$ from $(\mathbf{A}, \mathbf{u})$, with $\mathbf{u}$ uniform, with error probability $o(1)$ on both. Its worst-case time and space are $2^{(1/(2(b-a))+\xi)n/\ln \ln n}$.*

Proof. Normalize by a shear, run Algorithm 1 on $\bar{A} = [A'_t \mid I_n]$ (the coefficient expansion), and output “LWE” if and only if $\hat{\mu}_{\mathcal{L}}(\text{coeff}(\mathbf{b}_t)) \geq \tau/2$. Fix t and condition on $\mathbf{G}_t$ invertible. Then $\mathbf{A}'_t$ is uniform and independent of $(\mathbf{G}_t, \mathbf{e})$. The list depends only on $\bar{A}$ and fresh coins, so Lemma 20 applies with $H = \{\text{coeff}(\mathbf{b}_t)\}$, $\theta \leq \tau/64$ and $\xi = N^{-2}$, except with probability $e^{-L} + 2r2^{-b_{\min}} = n^{-\omega(1)}$.

On an LWE input, $\text{coeff}(\mathbf{b}_t) - \text{coeff}(\mathbf{e}_t) \in \Lambda_q(\bar{A})$ by (6), and $\|\mathbf{e}_t\| \leq H\sigma\sqrt{m}$ on the event $\|\mathbf{e}\| \leq \sigma\sqrt{m}$. Lemma 19(ii) then gives an ideal score of at least τ . On a uniform input, $\mathbf{b}_t = \mathbf{U}_t^\dagger \mathbf{b}$ is uniform and independent of $\bar{A}$ and the list, so Lemma 19(iv) and Markov’s inequality bound $\Pr[\mu_{\bar{A}} > \tau/8]$ by $8(2f)^m q^{-\ell d}/\tau = e^{-\Omega(d \ln q)}$, using Lemma 22. Hence, for each fixed t , the error is $n^{-\omega(1)}$ plus $\Pr[\|\mathbf{e}\| > \sigma\sqrt{m}]$, and the latter event is charged once. A union bound over the H shears and Lemma 21 complete the proof. The time bound follows from Theorem 4 and Lemma 22 with $\xi' = \min\{1, (b-a)\xi\}/2$. $\square$

5.5 Search Module-LWE

Writing $\bar{s} = \text{coeff}(\bar{\mathbf{s}})$, we have $[A'_v \mid I_n]^\top \bar{s} = [A' \mid I_n]^\top \bar{s} + \bar{s}_i \tilde{v}$, and therefore

$$\bar{b} + c\tilde{v} = [A'_v \mid I_n]^\top \bar{s} + \text{coeff}(\mathbf{e}_t) + (c - \bar{s}_i)\tilde{v}. \quad (9)$$

The correct candidate gives an honest LWE sample for the perturbed matrix. A wrong one shifts it by $\gamma\tilde{v}$ with $\gamma = c - \bar{s}_i \neq 0$, which is exactly the situation of Corollary 3.

Theorem 7 (Search Module-LWE). *In the setting of (7), for every $\xi > 0$, Algorithm 2 recovers $\mathbf{s}$ with probability $1 - o(1)$. Its worst-case time and space are $2^{(1/(2(b-a))+\xi)n/\ln \ln n}$. There is no condition on the residue degree of q .*

Proof. Fix t , condition on $\mathbf{G}_t$ being invertible and on $\|\mathbf{e}\| \leq \sigma\sqrt{m}$, and fix i . Let $\Lambda_v = \Lambda_q^\perp([A'_v \mid I_n])$ and $w = q/f$. Set $s' = w/\sqrt{2}$, $L_c = \lceil 2 \ln(1/\tau) + \ln 8 \rceil$ and $L'_c = 2 \ln n$.

Wrong candidates. Since $q/\Xi = e^\Upsilon$, we have

$$\frac{s'}{3qe^{L_c+L'_c} \Xi^{1/m}} = \frac{e^\Upsilon}{3\sqrt{2}f} \exp\left(-\frac{\ln 3 + \ln q + L_c + L'_c}{m}\right) \rightarrow \infty,$$

by Lemma 22. So Corollary 3 applies at width s' : except with probability n^{-2} , every $\gamma \in \mathbb{F}_q^\times$ satisfies $\rho_{1/s'}(\Lambda_v^* - \gamma\tilde{v}/q) \leq e^{-L_c} \leq \tau^2/8$. Then, for $c \neq \tilde{s}_i$, Lemma 19(i) and (iii), applied with shift $\gamma\tilde{v}$ and $x = \text{coeff}(\mathbf{e}_t)$, give

$$\mu(\bar{b} + c\tilde{v}) = \mu(\gamma\tilde{v} + \text{coeff}(\mathbf{e}_t)) \leq e^{\pi\|\mathbf{e}_t\|^2/f^2} \rho_{1/s'}(\Lambda_v^* - \gamma\tilde{v}/q) \leq \tau^{-1} \cdot \tau^2/8 = \tau/8.$$

The correct candidate. By (9) and Lemma 19(ii), $\mu(\bar{b} + \tilde{s}_i\tilde{v}) \geq e^{-\pi\|\mathbf{e}_t\|^2/f^2} \geq \tau$.

The list. The list is a function of A'_v and fresh coins, so given $(\mathbf{A}', v, \mathbf{G}_t, \mathbf{e})$ it has the law of the sampler on A'_v . By Corollary 4, the premise (5) for A'_v fails with probability at most e^{-L} . Lemma 20 with $H = \{\bar{b} + c\tilde{v} : c \in \mathbb{F}_q\}$ then shows that all q empirical scores are within $3\tau/32$ of their ideal values, except with probability $2r2^{-b_{\min}} + N^{-2} + 2qe^{-N\tau^2/2048} = n^{-\omega(1)}$. On this event exactly one candidate passes the threshold, namely $\tilde{s}_i$.

Union bound. Each pair (t, i) fails with probability at most $n^{-2} + n^{-\omega(1)}$, including the extra e^{-K} from finite-bit sampling of v . Summing over $i \in [n]$ and $t < H$ gives $O(\log n/n)$. Adding the bound of Lemma 21 and $\Pr[\|\mathbf{e}\| > \sigma\sqrt{m}]$ proves the success probability. The algorithm makes n sampler calls, each with slack $\xi' = \min\{1, (b-a)\xi\}/2$ in Theorem 4, and nq score evaluations of cost $N\text{poly}(n)$. Since $q = 2^{o(n/\ln \ln n)}$, the time bound follows from Lemma 22. Neither the algorithm nor its analysis enumerates an extension field K_ν . $\square$

Corollary 6. *For fixed $k, \ell \geq 1$, prime $q = n^{\kappa+o(1)}$ with $\kappa > 1/2$, errors with $\Pr[\|\mathbf{e}\| > C\sqrt{m}] = o(1)$ for a constant C , and every $\xi > 0$, decision and search MLWE can be solved with probability $1 - o(1)$ in worst-case time $2^{(1/2+\xi)n/\ln \ln n}$.*

Proof. This is the case $a = 0, b = 1$ of Theorems 6 and 7. $\square$

6 Narrow Module-LWE from Few Samples

Following [CJWZ26], when both secret and error are short we work with the relation lattice of $[-A \mid I_n]$. In this lattice the secret is part of the short vector, and no normalization is needed. Let $\mathbf{A} \in R_q^{k \times \ell}$ be uniform, $\mathbf{b} = \mathbf{A}^\dagger \mathbf{s} + \mathbf{e} \in R_q^\ell$ with $\mathbf{s} \in R^k$ and $\mathbf{e} \in R^\ell$, and $n = kd$, $m = (k+\ell)d$. Now $m - n = \ell d$ is the number of scalar samples, and it may be much smaller than n . Since $-\mathbf{A}$ is uniform, Section 3 and Corollary 4 apply to $\Lambda = \Lambda_q^\perp([-A \mid I_n])$ with $\mathbf{A}' = -\mathbf{A}$, and to its row perturbations $\Lambda_v = \Lambda_q^\perp([-A + e_i v^\top \mid I_n])$.

Definition 3 (Narrow MLWE). For $B \geq 1$, *narrow MLWE* asks to distinguish $(\mathbf{A}, \mathbf{A}^\dagger \mathbf{s} + \mathbf{e})$ from $(\mathbf{A}, \mathbf{u})$ with $\mathbf{u} \in R_q^\ell$ uniform, or to recover $\mathbf{s}$. The algorithm must work for every fixed pair $(\mathbf{s}, \mathbf{e})$ with $\|\mathbf{s}\|^2 + \|\mathbf{e}\|^2 \leq B^2 m$.

Averaging over $(\mathbf{s}, \mathbf{e})$ independently of $\mathbf{A}$ covers binary or ternary secrets and errors, and any distribution supported on such pairs with probability $1 - o(1)$.

Lemma 23. *Let $\hat{b} = (\text{coeff}(\mathbf{b}); 0) \in \mathbb{Z}_q^m$, and use the notation of Lemma 19 for $[-A \mid I_n]$ or for any perturbation $[-A + e_i v^\top \mid I_n]$.*

(i) *If $\mathbf{b} = \mathbf{A}^\dagger \mathbf{s} + \mathbf{e}$, then $\mu(\hat{b}) \geq \exp(-\pi(\|\mathbf{s}\|^2 + \|\mathbf{e}\|^2)/f^2)$ for the unperturbed matrix. For the perturbed matrix, $\mu(\hat{b} - c\tilde{v})$ satisfies the same bound when $c = s_i \bmod q$, where $s = \text{coeff}(\mathbf{s})$ and $\tilde{v} = (v; 0)$.*

(ii) *If $\mathbf{b}$ is uniform, then $\mathbb{E}_{\mathbf{b}}[\mu(\hat{b})] \leq (2f)^m q^{-\ell d}$.*

Proof. (i) By Lemma 8, $\hat{b} - (\text{coeff}(\mathbf{e}); s) = (\mathbf{A}^\top s; -s) = [-A \mid I_n]^\top (-s) \in \Lambda_q([-A \mid I_n])$. Moreover $[-A + e_i v^\top \mid I_n]^\top (-s) = [-A \mid I_n]^\top (-s) - s_i \tilde{v}$. Hence $\hat{b} - s_i \tilde{v} - (\text{coeff}(\mathbf{e}); s)$ lies in the q -ary lattice of the perturbed matrix. Apply Lemma 19(ii). (ii) For $X = (u; w) \in \Lambda$ we have $\mathbb{E}_{\mathbf{b}} \cos(2\pi \langle X, \hat{b} \rangle / q) = 1[u \in q\mathbb{Z}^{\ell d}]$, and $u \in q\mathbb{Z}^{\ell d}$ implies $w = Au = 0 \bmod q$, that is, $X \in q\mathbb{Z}^m$. Apply Lemma 19(iv). $\square$

For a wrong candidate $c \neq s_i$, the vector $\hat{b} - c\tilde{v}$ differs from the correct one by $(s_i - c)\tilde{v}$. This is the affine shift of Corollary 3.

Theorem 8 (Narrow MLWE with few samples). *Fix $B \geq 1$, $\kappa > 1/2$, $C_+ \geq 1$ and $\xi > 0$. Consider parameters (d, k, ℓ, q) with d a power of two, integers $k, \ell \geq 1$ with $\ell \leq C_+k$, $n = kd \rightarrow \infty$, and $q = n^{\kappa+o(1)}$ prime, such that*

$$\Upsilon = \frac{\ell}{k + \ell} \ln \frac{q}{\sqrt{\ell d}} \rightarrow \infty.$$

Then there are randomized algorithms that solve decision narrow MLWE with error probability $o(1)$ on both input distributions, and that recover $\mathbf{s} \bmod q$ with probability $1 - o(1)$. Their worst-case time and space are $2^{(1/2+\xi)n/\ln \Upsilon}$.

Three special cases are worth spelling out.

1. *Fixed ranks.* If k, ℓ are fixed, then $\Upsilon = \Theta(\ln n)$, and with $(\ell/k)n$ scalar samples the time is $2^{(1/2+\xi)n/\ln \ln n}$. The case $\ell = k$ has the shape of an ML-KEM public key.
2. *Sublinear samples.* For $d \rightarrow \infty$ through powers of two, take $\ell = 1$ and $k = \lceil \ln \ln d \rceil$. Then $d = (1 + o(1))n/\ln \ln n$ scalar samples suffice, $\Upsilon = (\kappa - \frac{1}{2} + o(1)) \ln n/(k + 1)$, and the time is $2^{(1/2+\xi)n/\ln \ln n}$.
3. *Unstructured LWE.* For $d = 1$ the ring is $\mathbb{Z}$ and $\delta = g = 1$. Theorem 8 then concerns narrow LWE with ℓ samples and secret dimension $n = k$. Here $\ln \Upsilon = \ln \tau_C + O(1)$, where $\tau_C = \frac{\ell}{n+\ell} \ln q$ is the parameter of [CJWZ26, Thm. 1.2]. On their domain, which additionally requires $q^{n/(n+\ell)} > \sqrt{n + \ell}$, their expected running time is $e^{(8 \ln 2 + o(1))n/\ln \tau_C} = 2^{(8+o(1))n/\ln \tau_C}$ for decision only. In the common sublinear-sample regime, the limiting coefficient of $n/\ln \Upsilon$ in $\log_2$ time decreases from 8 to $1/2$. The new bound is worst-case and covers search. This compares asymptotic exponents, not concrete running times.

of Theorem 8. *Parameters.* Set

$$f^2 = \frac{16\pi}{\ln 2} B^2 \frac{m}{n} \ln \Upsilon, \quad \tau = e^{-\pi B^2 m/f^2}, \quad L = \left\lceil \ln \frac{320q^4}{\tau} \right\rceil, \quad s_* = 2u(L),$$

and use the parameters of Theorem 4. Then $n \leq m \leq (1 + C_+)n$ and $\ln(1/\tau) = \frac{\ln 2}{16} \frac{n}{\ln \Upsilon}$. Since $\Upsilon \leq \ln q = O(\ln n)$, we have $n/\ln \Upsilon \gg \ln q$, so $L = (\frac{\ln 2}{16} + o(1)) \frac{n}{\ln \Upsilon}$ and $2L/m = o(1)$. By Corollary 4, $x_1 < \Upsilon$ and

$$x_1 \geq \min \left\{ \ln \frac{q}{\sqrt{n}} - O(1), \Upsilon - O(1) \right\} \geq c_\kappa \Upsilon, \quad c_\kappa = \frac{\kappa - 1/2}{2\kappa},$$

because $\ln(q/\sqrt{n}) \geq (\kappa - \frac{1}{2} - o(1)) \ln n$ and $\Upsilon \leq (\kappa + o(1)) \ln n$. Since $x_r = \ln(\sqrt{2}f) = \frac{1}{2} \ln \ln \Upsilon + O(1) \rightarrow \infty$, we get $W' = (1 - o(1)) \ln \Upsilon$, $W' < \ln \Upsilon$ and $c(x_r) \rightarrow 0$. By Lemma 18(e),

$$\frac{hn}{\ln \Upsilon + 1} \leq \Psi \leq (h + o(1)) \frac{n}{\ln \Upsilon}.$$

Hence $2 \ln(1/\tau) = \frac{h}{4} \frac{n}{\ln \Upsilon} \leq \Psi/2$, so $L \leq \Psi$ and $N\tau^2 \geq e^{\Psi/2}$. Moreover $(\ln q)(\ln n) = o(\Psi)$, $q \geq 2fs_*$, $u(L) \leq q$, and $5(4^r - 1)e^{-L} \leq \tau/64$. So Theorem 4 applies, with running time $e^{(1+\xi'+o(1))\Psi} \leq 2^{(1/2+\xi)n/\ln \Upsilon}$ for $\xi' = \min\{1, \xi\}$.

Decision. Run the sampler on $[-A \mid I_n]$ and output “LWE” if and only if $\hat{\mu}_{\mathcal{L}}(\hat{b}) \geq \tau/2$. The list depends only on $\mathbf{A}$ and fresh coins. On an LWE input, Lemma 23(i) gives an ideal score of at least τ . On a uniform input, Lemma 23(ii) and Markov’s inequality give

$\Pr[\mu(\hat{b}) > \tau/8] \leq 8e^{-\Delta}/\tau$ with $\Delta = \ell d \ln q - m \ln(2f) \geq m(\Upsilon - \ln(2f))$; here we used $\frac{\ell d}{m} \ln q \geq \Upsilon$. Thus $\Delta - \ln(1/\tau) \geq m(\Upsilon - O(\ln \ln \Upsilon)) - O(n/\ln \Upsilon) \rightarrow \infty$. Lemma 20 and the sampler failure bounds complete the proof.

Search. For each $i \in [n]$, draw $v \leftarrow \mathbb{F}_q^{\ell d}$, run the sampler on $[-A + e_i v^\top \mid I_n]$, and return the unique $c \in \mathbb{F}_q$ with $\hat{\mu}_{\mathcal{L}}(\hat{b} - c\tilde{v}) \geq \tau/2$. The correct candidate has ideal score at least τ by Lemma 23(i). For the wrong candidates, argue as in the proof of Theorem 7. Corollary 3 applies at width $q/(\sqrt{2}f)$ with $L_c = \lceil 2 \ln(1/\tau) + \ln 8 \rceil$ and $L'_c = 2 \ln n$, because $e^\Upsilon/f \rightarrow \infty$ and $(\ln q + L_c + L'_c)/m \rightarrow 0$. Then Lemma 19(iii), with $x = (\text{coeff}(\mathbf{e}); s)$ of squared norm at most $B^2 m$, bounds every wrong ideal score by $\tau/8$. Each coordinate fails with probability $n^{-2} + n^{-\omega(1)}$, including e^{-K} for sampling v ; a union bound over i gives success probability $1 - o(1)$. There are n sampler calls and $nq = \text{poly}(n)$ scores. Since $B\sqrt{m} < q/2$ for all sufficiently large n , centered lifting of the recovered residues also recovers the original integer secret. $\square$

Remark 2 (On the analysis of [CJWZ26]). We checked the parts of [CJWZ26] that we rely on: Lemma 2.10 and the use of $[-A \mid I_n]$. Their combining lemma [CJWZ26, Lemma 3.9] reproduces the premise $s \geq \sqrt{2}\eta_\varepsilon(L')$ of [DEL25, Lemma 3.3] on the *superlattice*. The convolution step in its proof needs the corresponding bound for the sublattice L , as observed for [DEL25] in [HGH26, footnote 6]. This particular premise issue is repaired within their parameter choice: on their good event, [CJWZ26, Cor. 3.13] bounds $\eta_{\varepsilon/3}(L_i) \leq s_*/4$ for all i , which implies $s_{i-1} \geq \sqrt{2}\eta_\varepsilon(L_i)$. The larger leading time coefficient in their published analysis has two sources. One is the time P^2 of an exact coset sampler with cost P per call. The other is the allocation $\ln P/b_i \geq 4 \ln(q/s_{i-1}) + O(1)$, whereas $\ln(q/s_{i-1}) + O(\ln \ln n)$ suffices. Theorem 8 removes both through Theorem 4.

7 Search-LWE Without an Assumed Reduction

HGH obtain search-LWE only by assuming a sample-preserving search-to-decision reduction [HGH26, App. A]. Regev's reduction for prime moduli does not apply directly. Their decision algorithm answers at random when the square block of A is singular, which happens with probability $\Theta(1/q)$, while the reduction makes about nq calls. We avoid this by normalizing once and then using the scalar row perturbation of Section 5.5. For uniform matrices this perturbation becomes an exact identity.

In this section $A \in \mathbb{Z}_q^{n \times m}$ is uniform and $b = A^\top s + e$. Write $A = [C \mid G]$ with $G \in \mathbb{Z}_q^{n \times n}$. If G is singular, which happens with probability at most $1/(q-1)$, we abort. Otherwise let $[A' \mid I_n] = G^{-1}A$ and $\bar{s} = G^\top s$, so that $b = [A' \mid I_n]^\top \bar{s} + e$, and A' is uniform and independent of G and e .

Lemma 24 (Scalar perturbation). *Let A' be uniform, $v \leftarrow \mathbb{Z}_q^{m-n}$ independent of it, $A'_v = A' + e_i v^\top$, $\Lambda_v = \Lambda_q^\perp([A'_v \mid I_n])$ and $\tilde{v} = (v; 0)$. Then A'_v and v are independent and uniform. Moreover, for every fixed A'_v , every $\gamma \neq 0$ and every $u > 0$,*

$$\mathbb{E}_v [\rho_{1/u}(\Lambda_v^* - \gamma \tilde{v}/q)] = q^{-(m-n)} \rho_{q/u}(\mathbb{Z})^m.$$

Proof. Independence holds because A' is uniform. With A'_v fixed, the coset of $\Lambda_v^* - \gamma \tilde{v}/q$ labelled z has free part $A'_v{}^\top z - \gamma v$, which is uniform over $\mathbb{Z}_q^{m-n}$ for every z . Summing over z as in (1) gives $q^{-(m-n)} \rho_{q/u}(\mathbb{Z}^{m-n}) \rho_{q/u}(\mathbb{Z}^n)$. $\square$

For the sampler premise we use the smoothing bound of [HGH26], applied to the full lattice.

Lemma 25 ([HGH26, Lemma 2.6]). *Let q be prime, $1 \leq n < m$, $\varepsilon' \in (0, 1/2]$, and $t = 8 \max\{\sqrt{\ln(2m/\varepsilon')}, q^{n/m} \varepsilon'^{-1/m}\} \leq q$. If A' is uniform, then $\Pr[\eta_{\varepsilon'}(\Lambda_q^\perp([A' \mid I_n])) > t] \leq 2^{1-m}$.*

Proof. Here is a direct verification for the systematic distribution. Put $p = q/t \geq 1$. Every nonzero dual label has a uniform free part, so the expected noninteger dual mass is at most $q^{-(m-n)} \rho_p(\mathbb{Z})^m \leq 2^m q^n / t^m \leq 4^{-m} \varepsilon'$. The integer dual mass is less than $\varepsilon'/2$ by $t \geq 8\sqrt{\ln(2m/\varepsilon')}$. Markov's inequality gives failure at most $2 \cdot 4^{-m} \leq 2^{1-m}$. This is the first-moment mechanism of [HGH26, Lemma 2.6], with no assumption that a module coefficient expansion is uniform. $\square$

Theorem 9 (Search-LWE in the regimes of [HGH26]). *Fix $0 \leq a < b$ and $\xi > 0$. Let q be prime and $s_e \geq 1$ with $\ln(2 + s_e) = (\ln n)^{a+o(1)}$ and $\ln(q/\sqrt{n}) = (\ln n)^{b+o(1)}$, and let $n + n/\ln \ln n \leq m \leq 2n$. Let $A \in \mathbb{Z}_q^{n \times m}$ be uniform, $s \in \mathbb{Z}_q^n$ arbitrary, and e independent of A with $\Pr[\|e\|^2 > s_e^2 m / \pi] = o(1)$ (for example $e \sim D_{\mathbb{Z}, s_e}^m$). Then s can be recovered from $(A, A^\top s + e)$ with probability $1 - o(1)$, in worst-case time and space $2^{(1/(2(b-a))+\xi)n/\ln \ln n}$.*

Additional samples may be discarded provided the retained error vector satisfies the stated norm promise; a bound scaled to a larger original sample count alone does not imply this.

Proof. Algorithm. Normalize once as above. For each $i \in [n]$, draw v , run Algorithm 1 on $[A'_v \mid I_n]$, and return the unique $c \in \mathbb{Z}_q$ with $\hat{\mu}_{\mathcal{L}}(b + c\tilde{v}) \geq \tau/2$. Finally output $s = G^{-\top} \tilde{s}$. As in (9), $b + c\tilde{v} = [A'_v \mid I_n]^\top \tilde{s} + e + (c - \tilde{s}_i)\tilde{v}$.

Parameters. Let $f^2 = Fs_e^2 \ln \ln n$ with $F = 32(b-a)/\ln 2$, $\tau = e^{-s_e^2 m / f^2}$, $L = \lceil \ln(320q^4/\tau) \rceil$, and $s_\star = 2t$, where t is as in Lemma 25 with $\varepsilon' = e^{-L}$. Then $\ln(1/\tau) = m/(F \ln \ln n) \leq \frac{\ln 2}{16(b-a)} \frac{n}{\ln \ln n}$ and $L = \Theta(n/\ln \ln n)$, since $\ln q$ is polylogarithmic. By definition of $s_\star$,

$$x_1 \geq \ln \frac{q}{\sqrt{2}s_\star} = \min \left\{ \ln \frac{q}{16\sqrt{2} \ln(2m/\varepsilon')}, \frac{m-n}{m} \ln q - \frac{L}{m} - \ln(16\sqrt{2}) \right\}.$$

The first term is $\ln(q/\sqrt{n}) + O(\ln \ln \ln n) = (\ln n)^{b+o(1)}$. In the second, $\frac{m-n}{m} \geq \frac{1}{2 \ln \ln n}$, $\ln q \geq (\ln n)^{b+o(1)}$, and $L/m = O(1/\ln \ln n)$, so it is bounded below by $(\ln n)^{b+o(1)}$ (and can be larger when $b < 1$). Together with $x_1 \leq \ln(q/s_\star) \leq \ln(q/\sqrt{L})$, this gives $\ln x_1 = (b + o(1)) \ln \ln n$. As in Lemma 22, $\ln x_r = (a + o(1)) \ln \ln n$ and $x_r \rightarrow \infty$. So $W' = (b - a + o(1)) \ln \ln n$, $\Psi = (\frac{h}{b-a} + o(1)) \frac{n}{\ln \ln n}$, $2 \ln(1/\tau) \leq \Psi/3$, and (S1)–(S3) hold.

One coordinate. The premise (5) for A'_v , which is uniform, holds except with probability 2^{1-m} by Lemma 25. The list depends only on A'_v and fresh coins. For $c = \tilde{s}_i$, Lemma 19(ii) gives an ideal score of at least $e^{-\pi\|e\|^2/f^2} \geq \tau$ on the event $\|e\|^2 \leq s_e^2 m / \pi$. For $c \neq \tilde{s}_i$, Lemma 19(iii) and Lemma 24 at width $u = q/(\sqrt{2}f)$, together with Markov's inequality and a union bound over γ , show the following. Except with probability $8q(2\sqrt{2}f)^m q^{-(m-n)}/\tau^2$, every wrong ideal score is at most $\tau^{-1} \cdot \tau^2/8 = \tau/8$. This probability is $n^{-\omega(1)}$, because $(m-n) \ln q \geq \frac{n}{\ln \ln n} (\ln n)^{\max\{1, b\} - o(1)}$, whereas $m \ln(2\sqrt{2}f) \leq 2n(\ln n)^{a+o(1)}$ and $\max\{1, b\} > a$. Lemma 20 and Theorem 4 then bound the failure probability for coordinate i by $n^{-\omega(1)}$, including e^{-K} for finite-bit sampling of v .

All coordinates. Use internal slack $\min\{1, (b-a)\xi\}/2$ in Theorem 4. A union bound over $i \in [n]$, plus $1/(q-1)$ for normalization and the error-norm event once, gives success probability $1 - o(1)$. The algorithm makes n sampler calls and nq score evaluations, and $q = 2^{o(n/\ln \ln n)}$. $\square$

Remark 3. The proof uses [HGH26, Lemmas 2.6, 3.1, 3.2, 4.1, 4.2] with explicit failure probabilities, together with Theorem 4. It does not depend on the internal quantitative bounds in the proof of [HGH26, Thm. 3.2]. Because the prefix conditions are derived by projection (Lemma 15), the surplus $m - n = n/\ln \ln n$ suffices in all regimes, without the constant $C_{\log}/(b-a)$ of [HGH26, Cor. 3.2]; the same applies to decision LWE. An alternative to the perturbation is to replace one row of A by a fresh row that keeps the last block invertible. This also gives search, at the price of an additional failure probability of $n/(q(q-1))$.

8 Discussion and Limitations

Concrete security. Theorem 8 with $\ell = k$ applies to asymptotic families with the matrix shape of an ML-KEM public key, but not automatically to a fixed standardized parameter set. Appendix A describes a reproducible heuristic grid search at the standardized dimensions. It drops the certification and distributional overhead, uses fractional blocks and a Gaussian bias model, and does not instantiate the proved algorithm. In that model every estimate exceeds the corresponding NIST security target, by at least 2^{41} for ML-KEM and 2^{104} for ML-DSA. We conclude that Wagner-style attacks do not reach these targets, consistent with the qualitative conclusion of [GHH26]. The estimates neither certify the security of ML-KEM or ML-DSA nor exclude better attacks. A separate homogeneous-MSIS calculation distinguishes the strong-unforgeability instance from a SelfTarget-shaped relaxation; neither calculation is a demonstrated forgery attack.

Modulus, ranks and samples. The condition $\ln(q/\sqrt{n}) \rightarrow \infty$ (for example $\kappa > 1/2$) is needed for two reasons. It allows smoothing $\mathbb{Z}^{m-n}$ at the accuracy required by the Fourier test [HGH26], and it absorbs the factor $(m-n)^{(m-n)/(2m)}$ in Ξ (Remark 1). Removing the square-root loss of the packing bound is an open problem. Theorem 4 and Corollary 4 do not need fixed ranks, and Section 6 uses them with $\ell = o(k)$. With fixed ranks the number of scalar samples is $(1 + \ell/k)n$, a constant factor above n . By contrast, [HGH26] needs only $n + \Theta(n/\ln \ln n)$ samples for uniform matrices.

Rings and moduli. Lemma 10 uses that the coefficient embedding of $\mathbb{Z}[x]/(x^d + 1)$ is a scaled isometry of the canonical embedding. Other cyclotomic rings would lose a ring-dependent distortion factor. Prime moduli give the equal-degree CRT decomposition used throughout. Composite or ramified moduli would require a different support analysis. No condition on the residue degree is needed.

Sampling model. Our bounds are worst-case bit complexities, obtained with deterministic enumeration (Lemma 17) and per-level primes. Two earlier implementations are useful comparisons. HGH employ the approximate sampler of [ADSD15] and track adaptive statistical errors. CJWZ use an exact coset sampler costing $P\text{poly}(n)$ per lift [CJWZ26, Thm. 3.6]. With $N = \Theta(P)$, that cost produces a quadratic rather than nearly linear dependence on the list size. Our implementation instead has $e^{o(\Psi)}$ cost per lift in the stated parameter families.

Open problems. Open problems include the case $\kappa \leq 1/2$; module samplers that use fewer than $(1 + \Omega(1))n$ scalar samples for wide secrets; non-power-of-two rings; inhomogeneous MSIS with the same module distribution and parameters; and a rigorously justified concrete analysis of module-specific Wagner variants. The illustrative models in Appendix A do not settle the latter problem.

References

- [ABD⁺21] Roberto Avanzi, Joppe Bos, Léo Ducas, Eike Kiltz, Tancrede Lepoint, Vadim Lyubashevsky, John M. Schanck, Peter Schwabe, Gregor Seiler, and Damien Stehlé. CRYSTALS-Kyber: Algorithm specifications and supporting documentation (version 3.02). Round-3 submission to the NIST Post-Quantum Cryptography Standardization Project, 2021.

- [ACFP14] Martin R. Albrecht, Carlos Cid, Jean-Charles Faugère, and Ludovic Perret. Algebraic algorithms for LWE. Cryptology ePrint Archive, Paper 2014/1018, 2014.
- [ADRSD15] Divesh Aggarwal, Daniel Dadush, Oded Regev, and Noah Stephens-Davidowitz. Solving the shortest vector problem in 2^n time using discrete Gaussian sampling. In *47th Annual ACM Symposium on Theory of Computing (STOC)*, pages 733–742. ACM, 2015. doi:[10.1145/2746539.2746606](https://doi.org/10.1145/2746539.2746606).
- [ADSD15] Divesh Aggarwal, Daniel Dadush, and Noah Stephens-Davidowitz. Solving the closest vector problem in 2^n time – the discrete Gaussian strikes again! In *56th IEEE Symposium on Foundations of Computer Science (FOCS)*, pages 563–582. IEEE, 2015. doi:[10.1109/FOCS.2015.41](https://doi.org/10.1109/FOCS.2015.41).
- [AFFP14] Martin R. Albrecht, Jean-Charles Faugère, Robert Fitzpatrick, and Ludovic Perret. Lazy modulus switching for the BKW algorithm on LWE. In *Public-Key Cryptography – PKC 2014*, volume 8383 of *LNCS*, pages 429–445. Springer, 2014. doi:[10.1007/978-3-642-54631-0_25](https://doi.org/10.1007/978-3-642-54631-0_25).
- [AG11] Sanjeev Arora and Rong Ge. New algorithms for learning in presence of errors. In *Automata, Languages and Programming – ICALP 2011, Part I*, volume 6755 of *LNCS*, pages 403–415. Springer, 2011. doi:[10.1007/978-3-642-22006-7_34](https://doi.org/10.1007/978-3-642-22006-7_34).
- [Ajt96] Miklós Ajtai. Generating hard instances of lattice problems (extended abstract). In *28th Annual ACM Symposium on Theory of Computing (STOC)*, pages 99–108. ACM, 1996. doi:[10.1145/237814.237838](https://doi.org/10.1145/237814.237838).
- [ALSD21] Divesh Aggarwal, Zeyong Li, and Noah Stephens-Davidowitz. A $2^{n/2}$ -time algorithm for $\sqrt{n}$ -SVP and $\sqrt{n}$ -Hermite SVP, and an improved time-approximation tradeoff for (H)SVP. In *Advances in Cryptology – EURO-CRYPT 2021, Part I*, volume 12696 of *LNCS*, pages 467–497. Springer, 2021. doi:[10.1007/978-3-030-77870-5_17](https://doi.org/10.1007/978-3-030-77870-5_17).
- [AR05] Dorit Aharonov and Oded Regev. Lattice problems in $\text{NP} \cap \text{coNP}$. *Journal of the ACM*, 52(5):749–765, 2005. doi:[10.1145/1089023.1089025](https://doi.org/10.1145/1089023.1089025).
- [Ban93] Wojciech Banaszczyk. New bounds in some transference theorems in the geometry of numbers. *Mathematische Annalen*, 296:625–635, 1993. doi:[10.1007/BF01445125](https://doi.org/10.1007/BF01445125).
- [Ban95] Wojciech Banaszczyk. Inequalities for convex bodies and polar reciprocal lattices in $\mathbb{R}^n$. *Discrete & Computational Geometry*, 13:217–231, 1995. doi:[10.1007/BF02574039](https://doi.org/10.1007/BF02574039).
- [BDK⁺21] Shi Bai, Léo Ducas, Eike Kiltz, Tancrede Lepoint, Vadim Lyubashevsky, Peter Schwabe, Gregor Seiler, and Damien Stehlé. CRYSTALS-Dilithium: Algorithm specifications and supporting documentation. Version 3.1, February 2021. URL: <https://pq-crystals.org/dilithium/data/dilithium-specification-round3-20210208.pdf>.
- [BKW03] Avrim Blum, Adam Kalai, and Hal Wasserman. Noise-tolerant learning, the parity problem, and the statistical query model. *Journal of the ACM*, 50(4):506–519, 2003. doi:[10.1145/792538.792543](https://doi.org/10.1145/792538.792543).
- [CJWZ26] Zhili Chen, Kaijie Jiang, Anyu Wang, and Chuanqi Zhang. A subexponential algorithm for binary LWE with sublinear samples. Cryptology ePrint Archive, Paper 2026/2289, 2026. URL: <https://eprint.iacr.org/2026/2289>.

- [DEL25] Léo Ducas, Lynn Engelberts, and Johanna Loyer. Wagner’s algorithm provably runs in subexponential time for SIS^∞ . In *Advances in Cryptology – CRYPTO 2025, Part I*, volume 16000 of *LNCS*, pages 353–384. Springer, 2025. doi:[10.1007/978-3-032-01855-7_12](https://doi.org/10.1007/978-3-032-01855-7_12).
- [DPV11] Daniel Dadush, Chris Peikert, and Santosh Vempala. Enumerative lattice algorithms in any norm via M-ellipsoid coverings. arXiv:1011.5666v4, 2011. Full version of FOCS 2011. URL: <https://arxiv.org/abs/1011.5666v4>, doi:[10.48550/arXiv.1011.5666](https://doi.org/10.48550/arXiv.1011.5666).
- [GFH26] Yiming Gao, Yansong Feng, and Honggang Hu. Solving the shortest vector problem in $2^{0.7314n+o(n)}$ time via discrete Gaussian sampling on superlattices. Cryptology ePrint Archive, Paper 2026/1587, 2026.
- [GHH26] Yiming Gao, Xuyuan Han, and Honggang Hu. Provable subexponential algorithms for NIST third-round lattice families. Cryptology ePrint Archive, Paper 2026/2386, 2026. URL: <https://eprint.iacr.org/2026/2386>.
- [GJMS17] Qian Guo, Thomas Johansson, Erik Mårtensson, and Paul Stankovski. Coded-BKW with sieving. In *Advances in Cryptology – ASIACRYPT 2017, Part I*, volume 10624 of *LNCS*, pages 323–346. Springer, 2017. doi:[10.1007/978-3-319-70694-8_12](https://doi.org/10.1007/978-3-319-70694-8_12).
- [HGH26] Xuyuan Han, Yiming Gao, and Honggang Hu. A provable subexponential-time algorithm for LWE from $n + o(n)$ samples via Wagner-style Gaussian sampling. Cryptology ePrint Archive, Paper 2026/2261, 2026. URL: <https://eprint.iacr.org/2026/2261>.
- [HKM18] Gottfried Herold, Elena Kirshanova, and Alexander May. On the asymptotic complexity of solving LWE. *Designs, Codes and Cryptography*, 86(1):55–83, 2018. doi:[10.1007/s10623-016-0326-0](https://doi.org/10.1007/s10623-016-0326-0).
- [KF15] Paul Kirchner and Pierre-Alain Fouque. An improved BKW algorithm for LWE with applications to cryptography and lattices. In *Advances in Cryptology – CRYPTO 2015, Part I*, volume 9215 of *LNCS*, pages 43–62. Springer, 2015. doi:[10.1007/978-3-662-47989-6_3](https://doi.org/10.1007/978-3-662-47989-6_3).
- [LN96] Rudolf Lidl and Harald Niederreiter. *Finite Fields*, volume 20 of *Encyclopedia of Mathematics and its Applications*. Cambridge University Press, 2 edition, 1996. doi:[10.1017/CB09780511525926](https://doi.org/10.1017/CB09780511525926).
- [LS15] Adeline Langlois and Damien Stehlé. Worst-case to average-case reductions for module lattices. *Designs, Codes and Cryptography*, 75(3):565–599, 2015. doi:[10.1007/s10623-014-9938-4](https://doi.org/10.1007/s10623-014-9938-4).
- [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. doi:[10.1137/S0097539705447360](https://doi.org/10.1137/S0097539705447360).
- [MV13] Daniele Micciancio and Panagiotis Voulgaris. A deterministic single exponential time algorithm for most lattice problems based on Voronoi cell computations. *SIAM Journal on Computing*, 42(3):1364–1391, 2013. doi:[10.1137/100811970](https://doi.org/10.1137/100811970).
- [Nat16] National Institute of Standards and Technology. Submission requirements and evaluation criteria for the post-quantum cryptography standardization process. Technical report, NIST, 2016. Call for Proposals, Section 4.A.5.

- [Nat24a] National Institute of Standards and Technology. Module-lattice-based digital signature standard. Federal Information Processing Standards Publication FIPS 204, NIST, 2024. [doi:10.6028/NIST.FIPS.204](https://doi.org/10.6028/NIST.FIPS.204).
- [Nat24b] National Institute of Standards and Technology. Module-lattice-based key-encapsulation mechanism standard. Federal Information Processing Standards Publication FIPS 203, NIST, 2024. [doi:10.6028/NIST.FIPS.203](https://doi.org/10.6028/NIST.FIPS.203).
- [PS24] Amaury Pouly and Yixin Shen. Provable dual attacks on learning with errors. In *Advances in Cryptology – EUROCRYPT 2024, Part VII*, volume 14657 of *LNCS*, pages 256–285. Springer, 2024. [doi:10.1007/978-3-031-58754-2_10](https://doi.org/10.1007/978-3-031-58754-2_10).
- [Reg05] Oded Regev. On lattices, learning with errors, random linear codes, and cryptography. In *37th Annual ACM Symposium on Theory of Computing (STOC)*, pages 84–93. ACM, 2005. [doi:10.1145/1060590.1060603](https://doi.org/10.1145/1060590.1060603).
- [Ste24] Matthias Johann Steiner. The complexity of algebraic algorithms for LWE. In *Advances in Cryptology – EUROCRYPT 2024, Part III*, volume 14653 of *LNCS*, pages 375–403. Springer, 2024. [doi:10.1007/978-3-031-58734-4_13](https://doi.org/10.1007/978-3-031-58734-4_13).
- [Wag02] David Wagner. A generalized birthday problem. In *Advances in Cryptology – CRYPTO 2002*, volume 2442 of *LNCS*, pages 288–304. Springer, 2002. [doi:10.1007/3-540-45708-9_19](https://doi.org/10.1007/3-540-45708-9_19).
- [Was97] Lawrence C. Washington. *Introduction to Cyclotomic Fields*, volume 83 of *Graduate Texts in Mathematics*. Springer, 2 edition, 1997. [doi:10.1007/978-1-4612-1934-7](https://doi.org/10.1007/978-1-4612-1934-7).

A Heuristic Estimates for ML-KEM and ML-DSA

Following [DEL25, App. A], this appendix estimates the cost of heuristic Wagner-style attacks on the lattice problems underlying ML-KEM and ML-DSA. The model drops the overhead needed for the proofs and is generous to the attacker in several respects. It is not a run of the proved algorithm and not a lower bound for every Wagner variant. The accompanying scripts (see Appendix B) implement DEL-inspired models for a dual bias criterion and for homogeneous short relations. Section A.3 compares the results with the NIST security targets.

A.1 Narrow-MLWE bias model

Model. Write $T = m - n$ for the number of scalar responses and n for the secret dimension, and use the augmented relation lattice of $[-A \mid I_n]$. For a candidate list exponent $\lambda = \log_2 N$, choose the least positive w with $2^w \binom{T}{w} > N$, initialize the modeled standard deviation to $s = \sqrt{w/T}$, and put $n_{\text{left}} = n$. A round sets

$$p = \frac{q}{s\sqrt{2\pi e}}, \quad n_{\text{left}} \leftarrow \max\{0, n_{\text{left}} - \lambda / \log_2 p\}, \quad s \leftarrow \sqrt{2}s,$$

until $p \leq 1$ or no coordinates remain. It allows sample reuse and fractional blocks and does not account for conditioning on collisions. If the secret and error coefficients have modeled variance v , the bias estimate and guessing penalty are

$$\log_2 \beta = -\frac{2\pi^2 s^2 v (T + n - n_{\text{left}})}{q^2 \ln 2}, \quad G = n_{\text{left}} \log_2 (2\eta + 1).$$

At a round passing $\lambda + 2\log_2 \beta \geq \log_2 \max\{2, 1 + G\}$, the score is $\lambda + G$. The script retains the lowest score in the specified grid, without rounding intermediate values. These Gaussian-bias, sample-reuse, fractional-guessing and success-test assumptions are heuristic; in particular the nominal score is not a certified operation count.

Parameters and reproducibility. For ML-KEM, $d = 256$, $q = 3329$, $n = T = 256k$, and $v = \eta_1/2$, using the parameter sets of FIPS 203, Table 2 [Nat24b]. For ML-DSA, use $d = 256$, $q = 8380417$, $n = 256\ell$, $T = 256k$, and $v = \eta(\eta + 1)/3$ from FIPS 204, Table 1 [Nat24a]. The latter calculation assumes access to the full unrounded vector $\mathbf{t}$, not just the published high bits $\mathbf{t}_1$. It is an easier-input model and is not a forgery estimate.

Table 2: Best modeled scores on the grids $20 \leq \lambda < 900$. Scores and list exponents are base-two logarithms. Values are rounded only for display; the JSON files retain full floating-point output. The last column is the classical gate count of the NIST security category [Nat16].

Parameter set	n	T	score, step 0.5	score, step 0.05	target
ML-KEM-512	512	512	184.00	184.00	2^{143}
ML-KEM-768	768	768	249.00	248.63	2^{207}
ML-KEM-1024	1024	1024	332.50	332.10	2^{272}
ML-DSA-44 (full $\mathbf{t}$)	1024	1024	271.03	271.03	2^{146}
ML-DSA-65 (full $\mathbf{t}$)	1280	1536	368.50	368.25	2^{207}
ML-DSA-87 (full $\mathbf{t}$)	1792	2048	450.00	449.90	2^{272}

Run `python heuristic_grid.py -step .5` and repeat with `-step .05`. Large modeled lists illustrate the gap between an asymptotic theorem and a plausible implementation, but do not establish minimum attack costs. Even within the model, a finite grid cannot exclude a better parameter choice between grid points or outside the scanned interval.

A.2 Homogeneous MSIS instances associated with Dilithium

The Dilithium specification distinguishes an MSIS instance used in its strong-unforgeability analysis from the SelfTargetMSIS problem used in its other security statements [BDK⁺21, Sect. 6.2, Eqs. (6)–(8)]. They have different dimensions and norm bounds. The following calculation treats the former instance and an ordinary homogeneous relaxation of the latter; it is not a reduction from either solver to an ML-DSA forgery.

Use ring degree $d = 256$ and $q = 8380417$. To avoid confusing the ring degree with the rounding parameter of [BDK⁺21], denote that parameter by $D_{\text{round}} = 13$. Write τ_D for the challenge weight and $\beta_D = \tau_D \eta$, and define

$$\begin{aligned}\zeta &= \max\{\gamma_1 - \beta_D, 2\gamma_2 + 1 + 2^{D_{\text{round}}-1}\tau_D\}, \\ \zeta' &= \max\{2(\gamma_1 - \beta_D), 4\gamma_2 + 2\}.\end{aligned}$$

These are the bounds in the cited specification. The parameter triples (k, ℓ, η) are $(4, 4, 2)$, $(6, 5, 4)$, $(8, 7, 2)$; the challenge weights are 39, 49, 60, and the corresponding (γ_1, γ_2) are $(2^{17}, (q-1)/88)$, $(2^{19}, (q-1)/32)$, $(2^{19}, (q-1)/32)$. These values also appear in FIPS 204, Table 1 [Nat24a]. The common parameter values do not make all aspects of the older Dilithium specification and ML-DSA identical.

For the SelfTarget-shaped proxy, set $n = 256k$, $m = 256(k + \ell + 1)$ and the infinity-norm bound to ζ . For the strong-unforgeability MSIS instance, set $n = 256k$, $m = 256(k + \ell)$ and the bound to ζ' . The former proxy ignores the hash-dependent condition in SelfTargetMSIS, and does not impose any target-column coefficient constraint needed to turn a homogeneous

relation into an inhomogeneous solution. The latter is the homogeneous problem appearing in a security reduction, not a demonstrated forgery algorithm. These distinctions prevent interpreting either calculation as an attack on the signature scheme.

Model and finite search. For a candidate λ , initialize the sparse weight and standard deviation as above, but now use $T = m - n$ for the number of free relation coordinates. Apply the same round update, with n_{left} clamped to $[0, n]$. If β is the chosen norm bound, approximate the log-probability of a short output by

$$\log_2 p_{\text{short}} = (m - n_{\text{left}}) \log_2 \left(\text{erf} \frac{\beta}{\sqrt{2}s} \right) + n_{\text{left}} \log_2 (2\beta/q).$$

The model accepts a round when $\lambda + \log_2 p_{\text{short}} > -1$. This expected-count criterion assumes Gaussian reduced coordinates and uniform unreduced coordinates; it proves neither a success probability nor the nonzero condition for an actual output. The script `sis_grid.py` scans $5 \leq \lambda < 900$ with step 0.1 and records the first passing grid value, using full binary64 precision for the objective and rounding only the reported table.

Table 3: First accepted list exponents on the stated homogeneous-SIS grid. The SelfTarget column is an ordinary homogeneous proxy, not a SelfTargetMSIS solver. Neither column is a lower bound on attack cost.

Parameter label	ζ	proxy λ	ζ'	SUF-MSIS λ	target
ML-DSA-44	350209	269.9	380930	265.0	2^{146}
ML-DSA-65	724481	343.0	1048184	311.4	2^{207}
ML-DSA-87	769537	450.2	1048336	414.5	2^{272}

The clamp matters: applying the same formulas after n_{left} becomes negative no longer describes coordinate counts or a probability model. The code includes an explicitly invalid unclamped diagnostic; on the last SUF-MSIS row it reports 211.3 with negative remaining dimension. It is not an alternative attack estimate. The valid entries above are outputs of this model and this grid. They omit certification costs, list dependence, integer block constraints and alternative algorithmic strategies, including module-specific optimizations. Large modeled lists do not establish a concrete-security lower bound.

A.3 Comparison with the NIST security targets

The NIST security categories are defined by the cost of key search on AES-128, AES-192 and AES-256 and of collision search on a 256-bit hash function. These costs are about 2^{143} , 2^{207} , 2^{272} and 2^{146} classical gates [Nat16]. ML-KEM-512, ML-KEM-768 and ML-KEM-1024 target categories 1, 3 and 5 [Nat24b], and ML-DSA-44, ML-DSA-65 and ML-DSA-87 target categories 2, 3 and 5 [Nat24a]. The list exponent λ lower-bounds both the time and the memory of the modelled attack, counted in vectors rather than gates.

Every entry of Tables 2 and 3 exceeds the corresponding target. For ML-KEM the smallest margin is about 2^{41} , for ML-KEM-512 and ML-KEM-768. For ML-DSA, the cheapest of the three problems is the strong-unforgeability MSIS instance, and the smallest margin is about 2^{104} , for ML-DSA-65. For ML-KEM the binding constraint is the small modulus $q = 3329$. Only about 18 levels fit before the bias $\exp(-2\pi^2\sigma_r^2\|x\|^2/q^2)$ becomes too small to detect with N vectors. Reducing all secret coordinates in so few levels forces blocks of $\log N / \log p_i$ coordinates with small precisions p_i , hence $N \approx 2^{184}$ for ML-KEM-512.

We therefore conclude, as [DEL25] did for Dilithium, that Wagner-style attacks do not reach the NIST security targets of ML-KEM or ML-DSA in this model. The conclusion is limited to the model. It does not cover refinements such as coded-BKW with

sieving [GJMS17], FFT-based guessing of larger secret blocks, quantum speed-ups of the collision search, attacks that exploit the ring structure, or MLWE instances derived from ML-KEM ciphertexts. The SIS-type entries concern hardness assumptions in the security analysis of Dilithium rather than forgeries. All estimates also lie far above the published costs of the lattice-reduction attacks on which the parameter sets are based [ABD⁺21, BDK⁺21]. Wagner-style algorithms are therefore not the relevant threat to these standards at present.

B Reproducible Checks and Their Scope

The accompanying code is available at <https://github.com/wjyoumans/module-wagner>. It contains `sanity_checks.py` with its JSON output, the two heuristic grid scripts of Appendix A, and a release-regression script. These executed finite regression checks do not prove the asymptotic theorems.

Exact finite algebra. Using integer modular arithmetic, exact Gaussian elimination and polynomial remainders, the script checks 640 adjoint identities, 1898 CRT image-rank and support instances, 15044 nonzero integer lifts against the uncertainty inequality, and 4800 scalar-search, narrow-search and affine identities. The rings have $(d, q) = (1, 3), (2, 3), (2, 5), (4, 3), (4, 5), (4, 17), (8, 3), (8, 17)$. The image test verifies both containment in V_S^ℓ and rank $\delta\ell|S|$. Random examples use seed 20261005; exhaustive domains are recorded in the script.

Five small shear cases were counted exhaustively at one CRT factor and combined across factors with exact rational probabilities. For $(Q, g, H, k) = (5, 2, 3, 2)$, the exact all-shears-fail probability is 58849/390625, below the bound 32/135. Another 800 categorical-rounding instances use exact rational total variation. No floating-point SVP result is claimed as an exact minimum.

Numerical diagnostics. For $d = 2, q = 5$, the affine theta check averages over all 625 matrix-perturbation pairs and all 25 labels, at four nonzero shifts and three widths. The truncated theta sums use 80 decimal digits and an analytic truncation-tail bound; roundoff is not interval-certified. Tail constants are tested on 4096 (b, K) pairs. Four high-precision analytic schedules use nominal nonprime moduli and do not execute prime search, SVP certification, the large lists, or an attack.

Reproduction. Run `python sanity_checks.py` in the top-level directory of the code; the recorded environment is Python 3.13.5, SymPy 1.14.0 and mpmath 1.3.0. Both grids in Table 2 and the clamped grid of Table 3 were rerun. The intentionally invalid unclamped variant is archived as a regression diagnostic, not an estimate. The release checks cover the norm corollary, accuracy-budget arithmetic, slack choices and grid constraints.

None of these finite tests constitutes formal verification of the theorems.