

A Provable Subexponential-Time Algorithm for LWE from $n + o(n)$ Samples via Wagner-Style Gaussian Sampling

Xuyuan Han¹, Yiming Gao¹, and Honggang Hu^{1,2}

¹ School of Cyber Science and Technology, University of Science and Technology of
China, Hefei, China

hanxuyuan@mail.ustc.edu.cn, qw1234567@mail.ustc.edu.cn

² Hefei National Laboratory, Hefei, China
hghu2005@ustc.edu.cn

Abstract. We give a randomized subexponential-time algorithm for decision Learning with Errors (LWE) in two complementary parameter regimes. For polynomial prime modulus $q = n^{\kappa+o(1)}$ with any fixed $\kappa > 1/2$, the algorithm handles independent discrete Gaussian errors drawn from $D_{\mathbb{Z}, s_e}$, with Gaussian parameter s_e satisfying $\ln(2+s_e) = (\ln n)^{a+o(1)}$ for any fixed $0 \leq a < 1$. Conversely, polynomial error $s_e = n^{\gamma+o(1)}$, for any fixed $\gamma > 0$, can be handled with a quasipolynomial prime modulus $q = \exp((\ln n)^{1+\eta+o(1)})$ for any fixed $\eta > 0$. In both regimes, for an arbitrary secret distribution, the algorithm succeeds with probability $1 - o(1)$ using $m = n + \Theta(n/\ln \ln n)$ original samples, with expected time and space $2^{\Theta(n/\ln \ln n)}$.

Our starting point is the Wagner-style Gaussian sampler of Ducas, Engelberts, and Loyer (CRYPTO 2025). We instantiate its auxiliary extensions with random P -ary lattices whose improved smoothing bound supports the accuracy required for LWE without increasing the leading list size exponent. At the smaller Gaussian width enabled by this geometry, the basis-dependent condition of the exact sampler used in prior work is no longer available. We therefore use the statistically approximate shifted discrete Gaussian sampler of Aggarwal, Dadush, and Stephens-Davidowitz (FOCS 2015) and control its error for adaptively chosen shifts. Finally, a martingale analysis applies the Aharonov–Regev Fourier distinguisher (JACM 2005) directly to the dependent output, avoiding the loss incurred by comparison with an independent product distribution.

Our analysis applies to unstructured LWE with independent uniform public vectors. In particular, it does not directly apply to ML-KEM, whose security is based on the Module-LWE problem: the algebraic dependencies arising from its module structure are not covered by our analysis.

Keywords: Learning with Errors · Wagner’s algorithm · BKW algorithm · discrete Gaussian sampling · cryptanalysis

1 Introduction

Lattice-based cryptography derives much of its security from computational problems on lattices. A central average-case problem is *Learning with Errors* (LWE) [Reg05]. An LWE sample is a pair $(\mathbf{a}, b) \in \mathbb{Z}_q^n \times \mathbb{Z}_q$, where $\mathbf{a}$ is uniform and $b = \langle \mathbf{a}, \mathbf{s} \rangle + e \bmod q$ for a secret $\mathbf{s}$ and a small random error e . Given m independent samples, search-LWE asks to recover $\mathbf{s}$, whereas decision-LWE asks to distinguish the samples from uniform pairs.

BKW Algorithm. An algorithmic approach to LWE originates from the Blum–Kalai–Wasserman (BKW) algorithm [BKW00]. It partitions the secret coordinates into blocks and subtracts samples whose current blocks collide, eliminating one block at the cost of increasing the effective error. For polynomial modulus the direct method is exponential. Modulus switching and quantization reduce the table sizes [AFFP14, ACF⁺15], and subsequent work introduced further optimizations to BKW-based algorithms [DTV15, GJS15, GJMS17, GJMW19, WBLW25].

Kirchner and Fouque [KF15] gave a refined BKW algorithm with subexponential running time for sufficiently small secrets and errors, and further claimed that only $\Theta(n)$ original samples suffice in the binary setting. This claim relies on sample amplification to generate the larger number of samples required by the BKW procedure. Herold, Kirshanova, and May [HKM18] later identified a gap in this analysis and obtained a rigorous amplification guarantee using $\Theta(n \ln n)$ original samples. Thus their analysis does not establish the claimed subexponential algorithm from only $\Theta(n)$ original samples.

Wagner-Style Gaussian Sampler. The discussion above identifies sample amplification as a central obstacle to obtaining rigorous BKW algorithms from few original samples. Ducas, Engelberts, and Loyer [DEL25] took a different approach. They isolated the Wagner step underlying BKW and realized it through discrete Gaussian sampling (DGS) on a chain of lattices. This avoids sample amplification and yields a provable Wagner-style Gaussian sampler for orthogonal q -ary lattices.

Let $\mathbf{A} = [\mathbf{A}' \mid \mathbf{I}_n] \in \mathbb{Z}_q^{n \times m}$. Partition the n equations into blocks of sizes $b_1, \dots, b_r$ and let $n_i = \sum_{j \leq i} b_j$. For each $i \in [r]$, let $\mathbf{A}'_i$ be the first n_i rows of $\mathbf{A}'$ and define $\mathbf{A}_i := [\mathbf{A}'_i \mid \mathbf{I}_{n_i}] \in \mathbb{Z}_q^{n_i \times (m - n + n_i)}$. Write $\Lambda_q^\perp(\mathbf{A}_i) := \{\mathbf{x} \in \mathbb{Z}^{m - n + n_i} : \mathbf{A}_i \mathbf{x} = \mathbf{0} \bmod q\}$. Set $\Lambda_0 := \mathbb{Z}^{m - n}$ and $\Lambda_i := \Lambda_q^\perp(\mathbf{A}_i)$ for $i \in [r]$. Then $\Lambda_r = \Lambda_q^\perp(\mathbf{A})$. Ducas et al. formulate the Wagner-style Gaussian sampler as a walk through this lattice chain, using an auxiliary superlattice $\Lambda'_i \supseteq \Lambda_i$ at each level. Figure 1 shows the walk in the direction of the algorithm.

At each level, the sampler uses a lattice extension. Let $\pi_i : \mathbb{R}^{m - n + n_i} \rightarrow \mathbb{R}^{m - n + n_{i-1}}$ denote the projection that removes the last b_i coordinates. By systematic form, $\pi_i(\Lambda_i) = \Lambda_{i-1}$ and $\ker(\pi_i|_{\Lambda_i}) = q\mathbb{Z}^{b_i}$. Identifying the kernel with its last b_i coordinates gives the exact sequence $0 \rightarrow q\mathbb{Z}^{b_i} \rightarrow \Lambda_i \xrightarrow{\pi_i} \Lambda_{i-1} \rightarrow 0$. The Wagner step introduces an auxiliary superlattice $\Lambda'_i \supseteq \Lambda_i$ with the same projection onto Λ_{i-1} . More precisely, let $\Gamma_i \subseteq \mathbb{R}^{b_i}$ be a full-rank lattice satisfying

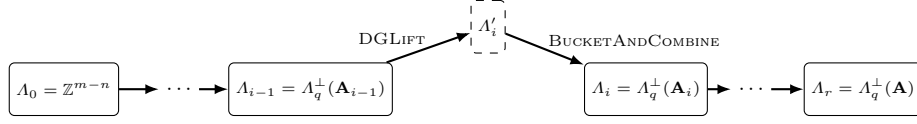


Fig. 1. The lattice walk underlying the Wagner-style Gaussian Sampler. At level i , samples are lifted from Λ_{i-1} to the auxiliary superlattice Λ'_i and then combined to obtain samples in Λ_i .

$q\mathbb{Z}^{b_i} \subseteq \Gamma_i$, and define

$$\Lambda'_i := \Lambda_i + (\mathbf{0}^{m-n+n_{i-1}} \times \Gamma_i).$$

Then $\pi_i(\Lambda'_i) = \Lambda_{i-1}$ and $\ker(\pi_i|_{\Lambda'_i}) = \mathbf{0}^{m-n+n_{i-1}} \times \Gamma_i$. Hence $0 \rightarrow \Gamma_i \rightarrow \Lambda'_i \xrightarrow{\pi_i} \Lambda_{i-1} \rightarrow 0$. Since $\Lambda_i \subseteq \Lambda'_i$ and the two maps have the same image, $\Lambda'_i/\Lambda_i \cong \Gamma_i/q\mathbb{Z}^{b_i}$. We call Λ'_i/Λ_i the *collision quotient* because BUCKETANDCOMBINE groups lifted vectors by its cosets. Thus the auxiliary lattice Γ_i determines both the collision quotient and the geometry of the Gaussian lifting step.

Ducas et al. [DEL25] instantiate this extension with the scaled integer lattice $\Gamma_i^{\text{DEL}} := (q/p_i)\mathbb{Z}^{b_i}$. Hence $\Lambda'_i/\Lambda_i \cong \Gamma_i^{\text{DEL}}/q\mathbb{Z}^{b_i} \cong (\mathbb{Z}/p_i\mathbb{Z})^{b_i}$, so the quotient consists of $p_i^{b_i}$ classes. This is the lattice realization of lazy modulus switching: the residue of the new block modulo q is replaced by a coarser label modulo p_i .

The lifting step requires shifted DGS over Γ_i^{DEL} . In their instantiation, [DEL25] use the polynomial-time exact discrete Gaussian sampler of [BLP⁺13, Lemma 2.3]. The parameter choices ensure that the Gaussian width satisfies the condition required by this sampler. The subsequent BUCKETANDCOMBINE step groups lifted vectors by their classes in Λ'_i/Λ_i and subtracts vectors from the same class. This returns to Λ_i and increases the Gaussian width by a factor $\sqrt{2}$.

Repeating these two steps walks from $\Lambda_0 = \mathbb{Z}^{m-n}$ to $\Lambda_r = \Lambda_q^\perp(\mathbf{A})$. The collision and subtraction steps introduce dependencies among the output samples, so Ducas et al. maintain an invariant of *conditional similarity* to an independent discrete Gaussian list rather than independence itself. For $m = n + \omega(n/\ln \ln n)$ and polynomial q , this yields a sampler of width $q/\text{polylog}(n)$ in time $\exp(O(n/\ln \ln n))$.

The LWE Barrier in Prior Work. Ducas et al. [DEL25, Section 4.3] observe that their Wagner-style Gaussian sampler does not directly yield an LWE algorithm running in subexponential time. There are two obstacles. A direct transfer of an independent sample analysis of the Aharonov–Regev distinguisher [AR05] through statistical distance would require the conditionally similar output distribution to be extremely close to an independent product distribution. Second, achieving such accuracy with their sampler makes the smoothing requirement of Γ_i^{DEL} incompatible with the block sizes needed for a subexponential recursion.

To make the first obstacle explicit, suppose that the final list $L = (\mathbf{X}_1, \dots, \mathbf{X}_N)$ is conditionally δ_r -similar to independent samples from $D_{\Lambda_q^\perp(\mathbf{A}), q/f}$. The sampler

analysis gives $\delta_r = O(4^r \varepsilon)$, where ε is the smoothing accuracy. The closure property of conditional similarity and the probability chain rule give

$$\Pr[L = (\mathbf{x}_1, \dots, \mathbf{x}_N)] = e^{\pm N\delta_r} \prod_{j=1}^N D_{\Lambda_q^\perp(\mathbf{A}), q/f}(\mathbf{x}_j).$$

Hence, whenever $N\delta_r = o(1)$, $\text{SD}(\text{Law}(L), D_{\Lambda_q^\perp(\mathbf{A}), q/f}^N) = O(N\delta_r)$. This direct comparison therefore requires $N\delta_r = o(1)$. Since $\delta_r = O(4^r \varepsilon)$, it suffices to take $\varepsilon = o(1/(N4^r))$. For $\ln N = \Theta(n/\ln \ln n)$ and $r = O(\ln n)$, we have $\ln 4^r = O(\ln n) = o(\ln N)$ and therefore $\ln(1/\varepsilon) = \Omega(n/\ln \ln n)$. Thus a direct product-distribution transfer forces a subexponentially small smoothing error, far beyond the inverse polynomial accuracy used in the SIS^∞ application.

The lattice Γ_i^{DEL} then becomes the geometric bottleneck. Its smoothing parameter has the asymptotic form³ $\eta_\varepsilon((q/p_i)\mathbb{Z}^{b_i}) = \Theta\left((q/p_i)\sqrt{\ln(b_i/\varepsilon)}\right)$. Since the procedure DGLIFT requires $s_{i-1} \geq \eta_\varepsilon((q/p_i)\mathbb{Z}^{b_i})$, this accuracy forces $p_i = \Omega((q/s_{i-1})\sqrt{\ln(1/\varepsilon)})$. On the other hand, Λ'_i/Λ_i has $p_i^{b_i}$ cosets, and keeping the list size within $N^{1+o(1)}$ requires $b_i \ln p_i = O(\ln N)$. Combining the two requirements gives $b_i = O(\ln N / (\ln(q/s_{i-1}) + \frac{1}{2} \ln \ln(1/\varepsilon)))$. In the accuracy regime above, $\ln(1/\varepsilon) = \Omega(\frac{n}{\ln \ln n})$, and therefore $\ln \ln(1/\varepsilon) = \Omega(\ln n)$. Since $s_{i-1} < q$, it follows that $b_i = O\left(\frac{\ln N}{\ln n}\right)$. With $r = O(\ln n)$ levels, the total number of equations that can be processed is therefore $\sum_{i=1}^r b_i = O(\ln N) = O\left(\frac{n}{\ln \ln n}\right) = o(n)$. Thus, within their instantiation and the above product-distribution analysis, the sampler cannot simultaneously guarantee the required accuracy, keep the lists subexponential, and process all n equations. Ducas et al. [DEL25, Section 4.3] explicitly suggest replacing the scaled integer auxiliary lattice by a lattice with better smoothing behavior.

1.1 Contributions

We realize the direction suggested in [DEL25] and obtain a subexponential-time decision-LWE algorithm from their Wagner-style Gaussian sampler. The proof addresses three obstacles. First, the auxiliary lattice must remain smooth at the accuracy required for LWE without forcing the block dimensions to become too small. Second, the resulting narrower Gaussian width must still permit an efficient implementation of the lifting step. Third, the sampler output must be analyzed without comparing its full law with an independent product distribution.

Our main result gives a general tradeoff between the logarithmic growth scales of the modulus and the Gaussian error width.

³ Lemma 2.8 of [DEL25] gives the corresponding upper bound. The reverse dependence follows already from the $2b_i$ dual vectors $\pm(p_i/q)\mathbf{e}_j$. Writing $\eta := \eta_\varepsilon((q/p_i)\mathbb{Z}^{b_i})$, we have $2b_i \exp(-\pi\eta^2 p_i^2/q^2) \leq \varepsilon$, and hence $\eta \geq (q/p_i)\sqrt{\ln(2b_i/\varepsilon)/\pi}$.

Theorem 1.1 (Informal Main Theorem). Fix constants $0 \leq a < b$. Let q be prime and let $s_e \geq 1$ satisfy

$$\ln(2 + s_e) = (\ln n)^{a+o(1)} \quad \text{and} \quad \ln\left(\frac{q}{\sqrt{n}}\right) = (\ln n)^{b+o(1)}.$$

For every secret distribution over $\mathbb{Z}_q^n$, there is a randomized distinguisher for decision-LWE with independent errors from $D_{\mathbb{Z}, s_e}$ that succeeds with probability $1 - o(1)$ using $m = n + \Theta\left(\frac{n}{(b-a) \ln \ln n}\right)$ original samples. Its expected time and space complexities are $2^{\left(\frac{1}{2(b-a)} + o(1)\right)n / \ln \ln n}$.

For polynomial prime modulus $q = n^{\kappa+o(1)}$ with any fixed $\kappa > 1/2$, we have $b = 1$, and the theorem applies whenever $\ln(2 + s_e) = (\ln n)^{a+o(1)}$ for a fixed $0 \leq a < 1$. Conversely, polynomial error $s_e = n^{\gamma+o(1)}$ corresponds to $a = 1$, and is covered by a quasipolynomial prime modulus $q = \exp\left((\ln n)^{1+\eta+o(1)}\right)$ for any fixed $\gamma, \eta > 0$. In particular, in the polynomial-modulus regime with $\ln(2 + s_e) = (\ln n)^{o(1)}$, the expected time and space become $2^{(1/2+o(1))n / \ln \ln n}$ using $n + \Theta(n / \ln \ln n)$ original samples.

Random P -ary Auxiliary Lattices. The first ingredient is a new choice of auxiliary lattice in the framework of Section 1. The abstract quotient-based sampler of [DEL25] remains unchanged. Their concrete instantiation uses the scaled integer lattice $\Gamma_i^{\text{DEL}} = (q/p_i)\mathbb{Z}^{b_i}$. We instead choose a prime P and use

$$\Gamma_i := \frac{q}{P} \Lambda_P(\mathbf{a}_i^T) = q\mathbb{Z}^{b_i} + \frac{q}{P} \mathbf{a}_i \mathbb{Z}, \quad \mathbf{a}_i \leftarrow \mathbb{F}_P^{b_i} \setminus \{\mathbf{0}\}.$$

It satisfies $\Gamma_i/q\mathbb{Z}^{b_i} \cong \mathbb{Z}/P\mathbb{Z}$, so the collision quotient has cardinality P , independently of block dimension b_i . At the same time, Lemma 3.2 shows that, with high probability,

$$\eta_\varepsilon(\Gamma_i) = O\left(qP^{-1/b_i} \sqrt{\max\left\{1, \frac{\ln(1/\varepsilon)}{b_i}\right\}}\right).$$

For the determinant-matched scaled integer benchmark introduced in Section 3.1, the corresponding dual-minimum smoothing bound is larger by a factor of order $\sqrt{b_i}$ in the high-accuracy regime $\ln(1/\varepsilon) = \omega(b_i)$. At the accuracy required for LWE, this gain removes the block-size obstruction described above: the block lengths can sum to n while $\ln N = \Theta(n / \ln \ln n)$.

Implementing DGLIFT at Narrow Gaussian Width. The random P -ary auxiliary lattices above permit substantially smaller Gaussian widths, but this raises an implementation issue. DGLIFT requires DGS over the lattice coset. Ducas et al. implement this step using the exact sampler of [BLP⁺13, Lemma 2.3].

Although that sampler applies to general lattice bases, its polynomial-time guarantee is basis dependent. The basis of Γ_i^{DEL} satisfies the required condition in their instantiation, whereas for our random P -ary lattices we do not have an efficiently computable basis satisfying the required condition at the narrower widths obtained above.

We therefore implement the shifted-DGS oracle using the statistically approximate algorithm of Aggarwal, Dadush, and Stephens-Davidowitz [ADS15, Theorem 22]. The shifted-DGS condition of [ADS15] is controlled by the distance from the shift to the auxiliary lattice rather than by a basis-dependent Gram-Schmidt bound. In block dimension b_i , one call takes time $2^{b_i + o(b_i)}$ and incurs statistical error $2^{-\Omega(b_i^2)}$. Since the shift used in a later call may depend on all preceding outputs, these approximation errors cannot simply be composed as independent errors. Lemma 3.3 gives a coupling argument that controls their accumulation throughout the adaptive execution.

Separately, BUCKETANDCOMBINE requires smoothing of the orthogonal q -ary lattices. We establish this smoothing directly from the Gaussian mass of their dual lattices. Writing $b_{\min} := \min_i b_i$ and $b_{\max} := \max_i b_i$, we choose the block parameters so that $b_{\max} = o(\ln N)$ and $b_{\min}^2 = \omega(\ln N)$. With these choices, the shifted-DGS overhead does not change the leading list-size exponent and the accumulated implementation error is $o(1)$. Consequently, Theorem 3.2 gives an implementable Wagner-style Gaussian sampler.

Fourier Distinguishing under Conditional Similarity. The final ingredient is a direct analysis of the Aharonov–Regev periodic-Gaussian distinguisher [AR05] on the vectors produced by the sampler. For a vector $\mathbf{x} \leftarrow D_{\Lambda_q^\perp(\mathbf{A}), q/f}$, the distinguisher uses the Fourier score $h_{\mathbf{b}}(\mathbf{x}) := \cos\left(\frac{2\pi}{q}\langle \mathbf{x}, \mathbf{b} \rangle\right)$. The analysis in Lemma 4.1 gives a separation of order $\tau := \exp(-s_e^2 m / f^2)$ between LWE and uniform inputs.

The actual sampler output is dependent rather than independent. However, conditional similarity controls the conditional expectation of each Fourier score given all preceding vectors. Centering the scores by these conditional expectations therefore yields a bounded martingale-difference sequence. Martingale concentration shows that the empirical Fourier score differs from its ideal mean by $O(\delta + N^{-1/2})$, where δ is the conditional-similarity parameter. Hence it suffices that δ be a sufficiently small constant fraction of τ and that $N\tau^2 \rightarrow \infty$.

By contrast, transferring the analysis through statistical distance to the full independent product distribution would require the much stronger condition $N\delta = o(1)$. Thus the required sampler accuracy is governed by the Fourier bias τ , rather than by the inverse list size $1/N$.

Finally, we choose the smoothing accuracy so that the sampler guarantee satisfies $q^4 \varepsilon \leq \tau/64$, and select the remaining parameters so that the sampler and Fourier conditions hold simultaneously. Combining these choices with the block-capacity bound above yields the modulus–error tradeoff of Theorem 1.1.

1.2 Related Work

Algebraic approaches provide another line of subexponential algorithms for LWE [AG11,ACFP14,STA20,Ste24]. We defer a detailed comparison, together with the corresponding BKW-based results, to Section C.

A related geometric construction appears in the recent exact-SVP algorithm of Gao, Feng, and Hu [GFH26], which uses a random prime-index superlattice analogous to our random P -ary auxiliary lattice. Its role is different: their construction is used to make the honest discrete Gaussian sampler available at the shortest vector scale, whereas ours improves the smoothing behavior of the Wagner-style Gaussian sampler while preserving a cyclic collision quotient of size P .

2 Preliminaries

Throughout the paper, $\ln$ denotes the natural logarithm and $\log$ denotes the logarithm to base two. For an integer $M \geq 1$, let $[M] := \{1, \dots, M\}$. We write $\mathbb{Z}_q := \mathbb{Z}/q\mathbb{Z}$. When q is prime, we identify $\mathbb{Z}_q$ with the finite field $\mathbb{F}_q$. We use $\mathbf{I}_d$ for the $d \times d$ identity matrix and $\mathbf{0}$ for a zero vector or matrix. We write $\|\cdot\|_2$ for the Euclidean norm and $\langle \cdot, \cdot \rangle$ for the standard inner product.

We write $\mathbf{x} \leftarrow D$ to denote sampling $\mathbf{x}$ from a distribution D , and $\text{Law}(X)$ for the distribution of a random variable X . For any distribution D , we write D^N for the joint distribution of N independent samples from D . For a function h of a random variable $\mathbf{X} \leftarrow D$, we write $\mathbb{E}_{\mathbf{X} \leftarrow D}[h(\mathbf{X})]$ for its expectation. For $u, v > 0$ and $\delta \geq 0$, the notation $u = e^{\pm\delta}v$ means $e^{-\delta}v \leq u \leq e^{\delta}v$.

2.1 Statistical Distance and Similarity of Distributions

Throughout this subsection, all distributions are discrete and supported on a countable set $\mathcal{X}$. For distributions P and Q on $\mathcal{X}$, their statistical distance is $\text{SD}(P, Q) := \frac{1}{2} \sum_{x \in \mathcal{X}} |P(x) - Q(x)|$. For a tuple $\mathbf{X} = (X_1, \dots, X_N)$, we write $\mathbf{X}_{-i}$ for the tuple obtained by deleting X_i , $\mathbf{X}_{<i} := (X_1, \dots, X_{i-1})$, and $\mathbf{X}_{>i} := (X_{i+1}, \dots, X_N)$.

We use the notions of similarity and conditional similarity from [DEL25, Definitions 2.1 and 2.2].

Definition 2.1 (Similarity, [DEL25, Definition 2.1]). *Let D be a distribution on $\mathcal{X}$ and let $\delta \geq 0$. A random variable X taking values in $\mathcal{X}$ is δ -similar to D if, for every $x \in \mathcal{X}$, $\Pr[X = x] = e^{\pm\delta}D(x)$.*

Definition 2.2 (Conditional Similarity, [DEL25, Definition 2.2]). *Let D be a distribution on $\mathcal{X}$ and let $\delta \geq 0$. Random variables $X_1, \dots, X_N$ taking values in $\mathcal{X}$ are conditionally δ -similar to independent samples from D if, for every $i \in [N]$ and every $\mathbf{x} = (x_1, \dots, x_N) \in \mathcal{X}^N$ for which the conditional probability is defined, $\Pr[X_i = x_i \mid \mathbf{X}_{-i} = \mathbf{x}_{-i}] = e^{\pm\delta}D(x_i)$.*

Conditional δ -similarity implies δ -similarity of every marginal. More generally, conditioning may be restricted to any subset of the remaining variables by the same averaging argument, equivalently by the closure property of [DEL25, Lemma 2.1]. In particular, for every $i \in [N]$ and every prefix in the support,

$$\Pr[X_i = x_i \mid \mathbf{X}_{<i} = \mathbf{x}_{<i}] = e^{\pm\delta} D(x_i).$$

2.2 Lattices and Quotients

A rank- d lattice $\Lambda \subseteq \mathbb{R}^k$ is the integer span $\Lambda = \mathbf{B}\mathbb{Z}^d$ of linearly independent columns of $\mathbf{B} \in \mathbb{R}^{k \times d}$. We write $\text{span}(\Lambda)$ for its linear span and $\det(\Lambda) = \sqrt{\det(\mathbf{B}^T \mathbf{B})}$ for its determinant. Its dual lattice is $\Lambda^* := \{\mathbf{y} \in \text{span}(\Lambda) : \langle \mathbf{y}, \mathbf{x} \rangle \in \mathbb{Z} \text{ for every } \mathbf{x} \in \Lambda\}$, and its first minimum is $\lambda_1(\Lambda) := \min_{\mathbf{0} \neq \mathbf{x} \in \Lambda} \|\mathbf{x}\|_2$. For $\mathbf{c} \in \text{span}(\Lambda)$, we write $\text{dist}(\mathbf{c}, \Lambda) := \min_{\mathbf{x} \in \Lambda} \|\mathbf{c} - \mathbf{x}\|_2$.

A sublattice $\Gamma \subseteq \Lambda$ is *primitive* if Λ/Γ is torsion free, or equivalently $\Gamma = \Lambda \cap \text{span}(\Gamma)$. In this case, orthogonal projection onto $\text{span}(\Gamma)^\perp$ maps Λ onto a lattice, and there exists a lattice complement $\Lambda_c \subseteq \Lambda$ such that $\Lambda = \Gamma \oplus \Lambda_c$. The projection restricts to an isomorphism from Λ_c onto the projected lattice. We also use that if $\Lambda_0 \subseteq \Lambda_1$ have the same rank, then Λ_1/Λ_0 is finite and $|\Lambda_1/\Lambda_0| = [\Lambda_1 : \Lambda_0]$.

2.3 LWE and q -ary Lattices

We regard the public LWE vectors as the columns of a matrix.

Definition 2.3 (Learning with Errors (LWE)). *Let $n, m, q \in \mathbb{Z}_{>0}$, let $D_{\mathbf{s}}$ be a distribution over $\mathbb{Z}_q^n$, and let $D_{\mathbf{e}}$ be a distribution over $\mathbb{Z}^m$. Sample $\mathbf{A} \leftarrow \mathbb{Z}_q^{n \times m}$ uniformly, $\mathbf{s} \leftarrow D_{\mathbf{s}}$, and $\mathbf{e} \leftarrow D_{\mathbf{e}}$ independently, and set $\mathbf{b} = \mathbf{A}^T \mathbf{s} + \mathbf{e} \bmod q$. In decision-LWE, the goal is to distinguish $(\mathbf{A}, \mathbf{b})$ from $(\mathbf{A}, \mathbf{u})$, where $\mathbf{u} \leftarrow \mathbb{Z}_q^m$ is uniform and independent of $\mathbf{A}$.*

A full-rank lattice $\Lambda \subseteq \mathbb{Z}^m$ is called a q -ary lattice if $q\mathbb{Z}^m \subseteq \Lambda \subseteq \mathbb{Z}^m$. For $\mathbf{A} \in \mathbb{Z}_q^{n \times m}$, we define the *primal q -ary lattice* and the *orthogonal q -ary lattice* by

$$\begin{aligned} \Lambda_q(\mathbf{A}) &:= \{\mathbf{x} \in \mathbb{Z}^m : \mathbf{x} = \mathbf{A}^T \mathbf{s} \bmod q \text{ for some } \mathbf{s} \in \mathbb{Z}_q^n\} = q\mathbb{Z}^m + \mathbf{A}^T \mathbb{Z}_q^n, \\ \Lambda_q^\perp(\mathbf{A}) &:= \{\mathbf{x} \in \mathbb{Z}^m : \mathbf{A}\mathbf{x} = \mathbf{0} \bmod q\}. \end{aligned}$$

They satisfy the duality relation $\Lambda_q(\mathbf{A})^* = q^{-1} \Lambda_q^\perp(\mathbf{A})$.

The following standard first-moment bound controls the Euclidean minimum of orthogonal q -ary lattice. Its proof counts nonzero integer vectors in a Euclidean ball and bounds the probability that any such vector lies in the random lattice. The proof is given in Section B.1.

Lemma 2.1 (Minimum of a Random Orthogonal q -ary Lattice, [Mic11, Exercise 12]). *Let q be prime and let $\mathbf{A} \leftarrow \mathbb{F}_q^{n \times m}$ be uniform. Assume that $q^{n/m} \geq 16$ and that $\sqrt{m} q^{n/m} \leq q$. Then*

$$\Pr \left[\lambda_1(\Lambda_q^\perp(\mathbf{A})) < \frac{1}{16} \sqrt{m} q^{n/m} \right] \leq 2^{-m}.$$

2.4 Discrete Gaussian Distribution

For any real $s > 0$ and $\mathbf{c} \in \mathbb{R}^d$, we define the Gaussian function on $\mathbb{R}^d$ by

$$\forall \mathbf{x} \in \mathbb{R}^d, \rho_{s,\mathbf{c}}(\mathbf{x}) := e^{-\pi \frac{\|\mathbf{x}-\mathbf{c}\|_2^2}{s^2}}.$$

For any countable set X , we define $\rho_{s,\mathbf{c}}(X) := \sum_{\mathbf{x} \in X} \rho_{s,\mathbf{c}}(\mathbf{x})$. We write $\rho_s := \rho_{s,\mathbf{0}}$. For a full rank lattice $\Lambda \subseteq \mathbb{R}^n$, center $\mathbf{c} \in \text{span}(\Lambda)$, and parameter $s > 0$, the discrete Gaussian distribution $D_{\Lambda,s,\mathbf{c}}$ is the probability distribution supported on Λ defined by

$$\forall \mathbf{x} \in \Lambda, D_{\Lambda,s,\mathbf{c}}(\mathbf{x}) := \frac{\rho_{s,\mathbf{c}}(\mathbf{x})}{\rho_{s,\mathbf{c}}(\Lambda)}.$$

We abbreviate the centered distribution $D_{\Lambda,s,\mathbf{0}}$ as $D_{\Lambda,s}$. For $\mathbf{c} \in \text{span}(\Lambda)$, define the lattice shift $\Lambda - \mathbf{c} := \{\mathbf{x} - \mathbf{c} : \mathbf{x} \in \Lambda\}$, and let $D_{\Lambda-\mathbf{c},s}$ denote the centered discrete Gaussian on this lattice shift. The two conventions are related by translation. If $\mathbf{X} \leftarrow D_{\Lambda,s,\mathbf{c}}$, then $\mathbf{X} - \mathbf{c}$ is distributed as $D_{\Lambda-\mathbf{c},s}$. Conversely, adding $\mathbf{c}$ to a sample from $D_{\Lambda-\mathbf{c},s}$ gives a sample from $D_{\Lambda,s,\mathbf{c}}$.

For a full-rank lattice $\Lambda \subset \mathbb{R}^d$ and $s > 0$, the *periodic Gaussian* associated with Λ at width s is the Λ -periodic function that maps $\mathbf{c} \in \mathbb{R}^d$ to $\rho_s(\Lambda - \mathbf{c})/\rho_s(\Lambda)$. The following form of Poisson summation gives its Fourier representation.

Lemma 2.2 (Poisson Summation, [MR04, Lemmas 2.8 and 2.9]). *Let $\Lambda \subset \mathbb{R}^d$ be full rank, let $s > 0$, and let $\mathbf{c} \in \mathbb{R}^d$. Then*

$$\rho_s(\Lambda + \mathbf{c}) = \frac{s^d}{\det(\Lambda)} \sum_{\mathbf{w} \in \Lambda^*} e^{-\pi s^2 \|\mathbf{w}\|_2^2} e^{2\pi i \langle \mathbf{w}, \mathbf{c} \rangle}.$$

In particular, $\rho_s(\Lambda + \mathbf{c}) \leq \rho_s(\Lambda)$.

We use the following lower bound on the periodic Gaussian. It appears in the proof of [AR05, Lemma 3.2]. The same statement with an explicit Gaussian width appears in [PS24, Lemma 1].

Lemma 2.3 (Periodic-Gaussian Lower Bound). *Let $\Lambda \subset \mathbb{R}^d$ be a full-rank lattice, let $s > 0$, and let $\mathbf{c} \in \mathbb{R}^d$. Then*

$$\frac{\rho_s(\Lambda - \mathbf{c})}{\rho_s(\Lambda)} \geq \rho_s(\mathbf{c}) = \exp\left(-\pi \frac{\|\mathbf{c}\|_2^2}{s^2}\right).$$

We use the following standard tail estimate for a product discrete Gaussian.

Lemma 2.4 (Discrete Gaussian Error Norm). *If $\mathbf{e} \leftarrow D_{\mathbb{Z},s_e}^m$, then*

$$\Pr\left[\|\mathbf{e}\|_2^2 > s_e^2 m / \pi\right] \leq (2/e)^{m/2}.$$

2.5 Smoothing Parameter

For $0 < \varepsilon < 1$ the smoothing parameter of a lattice Λ is

$$\eta_\varepsilon(\Lambda) := \inf \left\{ s > 0 : \rho_{1/s}(\Lambda^* \setminus \{\mathbf{0}\}) \leq \varepsilon \right\}.$$

The smoothing parameter measures when the Gaussian mass of the dual lattice away from the origin becomes negligible. Above this threshold, Gaussian mass is nearly invariant under shifts, a property used in our Gaussian lifting analysis.

We also use a bound on the smoothing parameter in terms of the Euclidean minimum of the dual lattice. For $\varepsilon \leq e^{-d}$, the following bound follows from Banaszczyk's Gaussian tail estimate [Ban93, Lemma 1.5] and appears in [PRS17, Lemma 2.6] and [KNSW20, Lemma 6]. Monotonicity extends it to all $\varepsilon \in (0, 1)$.

Lemma 2.5 (Dual-Minimum Smoothing Bound, [KNSW20, Lemma 6]).

For every rank- d lattice Λ and every $\varepsilon \in (0, 1)$,

$$\eta_\varepsilon(\Lambda) \leq \frac{\sqrt{\max\{d, \ln(1/\varepsilon)\}}}{\lambda_1(\Lambda^*)}.$$

Smoothing of a Random Orthogonal q -ary Lattice. By directly bounding the Gaussian mass of the dual lattice, we obtain the following smoothing bound for a random orthogonal q -ary lattice. The proof is given in Section B.2.

Lemma 2.6. *Let q be prime, let $1 \leq n < m$, let $\mathbf{A} \leftarrow \mathbb{F}_q^{n \times m}$ be uniform, and let $0 < \varepsilon \leq \frac{1}{2}$. Whenever $8 \max \left\{ \sqrt{\ln\left(\frac{2m}{\varepsilon}\right)}, q^{n/m} \varepsilon^{-1/m} \right\} \leq q$, we have, except with probability 2^{-m} over $\mathbf{A}$,*

$$\eta_\varepsilon(\Lambda_q^\perp(\mathbf{A})) \leq 8 \max \left\{ \sqrt{\ln\left(\frac{2m}{\varepsilon}\right)}, q^{n/m} \varepsilon^{-1/m} \right\}.$$

A natural alternative is to bound the smoothing parameter through a minimum of the dual lattice, as in Lemma 2.5. For $\Lambda_q^\perp(\mathbf{A})$, applying Lemma 2.1 to the corresponding primal q -ary lattice would require $\sqrt{m} q^{(m-n)/m} \leq q$, equivalently $q^{n/m} \geq \sqrt{m}$. This condition need not hold in the early levels of our recursion, where the number of imposed equations is still small. This obstruction is intrinsic: since $\mathbb{Z}^m \subseteq \Lambda_q^\perp(\mathbf{A})^*$, one always has $\lambda_1(\Lambda_q^\perp(\mathbf{A})^*) \leq 1$. We therefore bound the Gaussian mass of the integer part and the random nontrivial cosets separately.

[DEL25, Lemma 2.10–2.11] follow the analogous route using an ℓ_∞ minimum and obtain $\eta_\varepsilon(\Lambda_q^\perp(\mathbf{A})) = O\left(q^{n/m} \sqrt{\ln(1/\varepsilon)}\right)$ under their stated conditions. For the smaller smoothing errors required here, directly bounding the Gaussian mass of the dual yields $q^{n/m} \varepsilon^{-1/m}$ for the random-coset contribution. Since $\varepsilon^{-1/m} = \exp(\ln(1/\varepsilon)/m)$, this factor is $1 + o(1)$ in our high-accuracy regime $m = (1 + o(1))n$ and $\ln(1/\varepsilon) = \Theta(n/\ln n)$.

2.6 Discrete Gaussian Sampling Algorithms

We use two discrete Gaussian sampling results. An *exact* sampler has output law exactly equal to the target discrete Gaussian. A *statistically approximate* sampler has output law within a stated statistical distance of the target distribution.

The exact sampler below applies to an arbitrary lattice basis. Its width condition depends on the longest Gram–Schmidt vector of that basis.

Lemma 2.7 (Exact Shifted DGS, [BLP⁺13, Lemma 2.3]). *Let $\mathbf{B}$ be a basis of a d -dimensional lattice Λ , let $\tilde{\mathbf{B}}$ be its Gram–Schmidt orthogonalization, and let $\mathbf{c} \in \text{span}(\Lambda)$. If $s \geq \|\tilde{\mathbf{B}}\| \sqrt{\ln(2d+4)/\pi}$, where $\|\tilde{\mathbf{B}}\|$ is the largest Euclidean norm of its columns, then there is a probabilistic polynomial-time algorithm that samples exactly from $D_{\Lambda,s,\mathbf{c}}$.*

We also use the approximate shifted DGS algorithm. [ADS15, Theorem 22] is stated for sampling on the lattice shift $\Lambda - \mathbf{c}$ and returns at least one sample. Retaining one output and translating it by $\mathbf{c}$ gives the following result.

Lemma 2.8 (Statistically Approximate Shifted DGS, [ADS15, Theorem 22]). *There is an absolute constant $c_{\text{DGS}} > 0$ for which the following holds. Let $\Lambda \subset \mathbb{R}^d$ be a d -dimensional lattice, let $\mathbf{c} \in \text{span}(\Lambda)$, and let $g(d) = d^{\omega(1)}$ be efficiently computable. If $s > \text{dist}(\mathbf{c}, \Lambda)/g(d)$, then one can sample from a distribution within statistical distance at most $2^{-c_{\text{DGS}}d^2}$ of $D_{\Lambda,s,\mathbf{c}}$ in time $2^{d+O(\log d \log g(d))}$.*

We instantiate Lemma 2.8 with $g(d) = e^{\sqrt{d}}$. Its width condition then becomes $s > \text{dist}(\mathbf{c}, \Lambda)e^{-\sqrt{d}}$, and its running time becomes $2^{d+o(d)}$.

3 A Wagner-Style Gaussian Sampler with Random P -ary Auxiliary Lattices

This section constructs a Wagner-style Gaussian sampler for orthogonal q -ary lattices $\Lambda_q^\perp(\mathbf{A})$ using random P -ary auxiliary lattices. We first establish the algebraic structure and smoothing bounds required by the recursion. We then instantiate DGLIFT and BUCKETANDCOMBINE steps and analyze the resulting sampler at target Gaussian width q/f and smoothing-error parameter ε .

As in previous BKW-style analyses [KF15, DEL25], we place the input matrix in systematic form. Write $\mathbf{A} = [\mathbf{C} \mid \mathbf{B}]$, where $\mathbf{B} \in \mathbb{Z}_q^{n \times n}$. Conditioned on $\mathbf{B}$ being invertible, replace $\mathbf{A}$ by $\tilde{\mathbf{A}} := \mathbf{B}^{-1}\mathbf{A} = [\mathbf{A}' \mid \mathbf{I}_n]$. This transformation leaves $\Lambda_q^\perp(\mathbf{A})$ unchanged. We write $\mathbf{A} = [\mathbf{A}' \mid \mathbf{I}_n]$ for the resulting systematic matrix.⁴

⁴ Conditioned on $\mathbf{B}$ being invertible, $\mathbf{B}^{-1}\mathbf{C}$ is uniform, and the same holds for every prefix. Moreover, $\Pr[\mathbf{B} \text{ is singular}] = O(q^{-1})$.

3.1 Random P -ary Auxiliary Lattices

We first define the lattice chain underlying the Wagner-style Gaussian sampler [DEL25].

Let $b_1, \dots, b_r$ be positive integers with $\sum_{i=1}^r b_i = n$, and put $n_i := \sum_{j=1}^i b_j$ and $n_0 := 0$. For $i \in [r]$, let $\mathbf{A}'_i$ be the submatrix consisting of the first n_i rows of $\mathbf{A}'$ and define $\mathbf{A}_i := [\mathbf{A}'_i \mid \mathbf{I}_{n_i}]$, $\Lambda_i := \Lambda_q^\perp(\mathbf{A}_i)$. Thus $\Lambda_0 = \mathbb{Z}^{m-n}$ and $\Lambda_r = \Lambda_q^\perp(\mathbf{A})$. Deleting the last b_i coordinates defines a surjective projection $\pi_i : \Lambda_i \rightarrow \Lambda_{i-1}$ with $\ker(\pi_i) = q\mathbb{Z}^{b_i}$, where $q\mathbb{Z}^{b_i}$ is embedded in the newest coordinate block.

At level i , the sampler introduces an auxiliary superlattice $\Lambda'_i \supseteq \Lambda_i$. The lifting and collision steps require the following two properties.

1. **Primitive extension.** The kernel Γ_i of the projection $\pi_i : \Lambda'_i \rightarrow \Lambda_{i-1}$ is primitive in Λ'_i , and the projection induces the exact sequence $0 \rightarrow \Gamma_i \rightarrow \Lambda'_i \xrightarrow{\pi_i} \Lambda_{i-1} \rightarrow 0$. This is the lattice extension used by DGLIFT.
2. **Finite collision quotient.** The inclusion $\Lambda_i \subseteq \Lambda'_i$ has finite index. The procedure BUCKETANDCOMBINE groups lifted vectors by their classes in Λ'_i/Λ_i . Two vectors in the same class differ by an element of the lattice Λ_i .

We realize both properties using a random P -ary auxiliary lattice. Let P be prime. For each $i \in [r]$, sample $\mathbf{a}_i \leftarrow \mathbb{F}_P^{b_i} \setminus \{\mathbf{0}\}$ uniformly and independently, and define the random P -ary auxiliary lattice by

$$\Gamma_i := \frac{q}{P} \Lambda_P(\mathbf{a}_i^T) = q\mathbb{Z}^{b_i} + \frac{q}{P} \mathbf{a}_i \mathbb{Z}.$$

We identify Γ_i with its embedding $\mathbf{0} \oplus \Gamma_i$ in the newest b_i coordinates and set $\Lambda'_i := \Lambda_i + (\mathbf{0} \oplus \Gamma_i)$. Lemma 3.1 verifies the two required algebraic properties. The proof is given in Section B.3.

Lemma 3.1 (Algebraic Properties of the Random P -ary Extension).

For every $i \in [r]$, the coordinate projection induces the exact sequence $0 \rightarrow \Gamma_i \rightarrow \Lambda'_i \xrightarrow{\pi_i} \Lambda_{i-1} \rightarrow 0$. Consequently, Γ_i is primitive in Λ'_i . Moreover, $\Lambda'_i/\Lambda_i \cong \Gamma_i/q\mathbb{Z}^{b_i} \cong \mathbb{Z}/P\mathbb{Z}$.

The lemma shows that the random P -ary auxiliary lattice provides both the primitive extension required by DGLIFT and the finite cyclic quotient required by BUCKETANDCOMBINE. Its geometry determines the Gaussian width needed for smoothing and the cost of implementing the lifting step.

Geometry and Smoothing. Standard P -ary duality gives $\Gamma_i^* = q^{-1} \Lambda_P^\perp(\mathbf{a}_i^T)$. Thus the smoothing parameter of Γ_i is controlled by the Euclidean minimum of the random orthogonal P -ary lattice $\Lambda_P^\perp(\mathbf{a}_i^T)$. Combining Lemmas 2.1 and 2.5 gives the following result.

Lemma 3.2 (Smoothing of the Random P -ary Auxiliary Lattice).

Suppose that $P^{1/b_i} \geq 16$ and $\sqrt{b_i} P^{1/b_i} \leq P$. Then, with probability at least $1 - 2^{1-b_i}$ over $\mathbf{a}_i$, simultaneously for every $\varepsilon \in (0, 1)$,

$$\eta_\varepsilon(\Gamma_i) \leq \frac{16q}{P^{1/b_i}} \sqrt{\max \left\{ 1, \frac{\ln(1/\varepsilon)}{b_i} \right\}}.$$

Proof. Apply Lemma 2.1 with $n = 1$, $m = b_i$, and $q = P$. Before conditioning on $\mathbf{a}_i \neq \mathbf{0}$, the failure probability is at most 2^{-b_i} . Since $(1 - P^{-b_i})^{-1} \leq 2$, conditioning on $\mathbf{a}_i \neq \mathbf{0}$ increases this probability to at most 2^{1-b_i} . Hence, except with probability 2^{1-b_i} , $\lambda_1(A_P^\perp(\mathbf{a}_i^T)) \geq \frac{\sqrt{b_i} P^{1/b_i}}{16}$, and therefore $\lambda_1(\Gamma_i^*) \geq \frac{\sqrt{b_i} P^{1/b_i}}{16q}$. On this event, Lemma 2.5 gives the claim. $\square$

Remark 3.1. The use of the dual Euclidean minimum is essential for the dependence on the smoothing accuracy in Lemma 3.2. A bound based on an ℓ_∞ minimum, as in [DEL25, Lemmas 2.10 and 2.11], would lose the factor $\sqrt{b_i}$ obtained above. Alternatively, a direct first-moment bound on the Gaussian mass of the random nontrivial cosets introduces a term of the form $qP^{-1/b_i}\varepsilon^{-1/b_i}$. By contrast, Lemma 3.2 depends only on $\sqrt{\ln(1/\varepsilon)/b_i}$. This distinction is crucial in the high-accuracy regime used in the LWE application.

We next quantify the geometric improvement over a scaled integer lattice of the same determinant. Define $\Gamma_i^{\text{DEL}} := \frac{q}{P^{1/b_i}} \mathbb{Z}^{b_i}$. This lattice is used only as a determinant-matched geometric benchmark.⁵

Both auxiliary lattices have determinant $\det(\Gamma_i^{\text{DEL}}) = \det(\Gamma_i) = \frac{q^{b_i}}{P}$. The dual of the scaled integer lattice is $(\Gamma_i^{\text{DEL}})^* = \frac{P^{1/b_i}}{q} \mathbb{Z}^{b_i}$, and hence $\lambda_1((\Gamma_i^{\text{DEL}})^*) = \frac{P^{1/b_i}}{q}$. On the event established in the proof of Lemma 3.2, which has probability at least $1 - 2^{1-b_i}$, the random P -ary auxiliary lattice satisfies $\lambda_1(\Gamma_i^*) \geq \frac{\sqrt{b_i} P^{1/b_i}}{16q}$. Thus the random construction achieves an $\Omega(\sqrt{b_i})$ improvement in the dual minimum over the determinant-matched scaled integer lattice.

This improvement is optimal up to an absolute constant for the dual minimum among determinant-matched lattices. Indeed, for any b_i -dimensional lattice $\mathcal{L}$ with $\det(\mathcal{L}) = P/q^{b_i}$, Minkowski's first theorem gives $\lambda_1(\mathcal{L}) = O\left(\frac{\sqrt{b_i} P^{1/b_i}}{q}\right)$.

Together with $\lambda_1(\Gamma_i^*) \geq \frac{\sqrt{b_i} P^{1/b_i}}{16q}$, this shows that the random P -ary construction achieves the optimal $\Theta(\sqrt{b_i})$ improvement over the determinant-matched scaled integer lattice at the level of the dual minimum. The same gain appears in the corresponding smoothing bounds. We stress that this optimality concerns only the dual-minimum analysis and does not imply optimality of the smoothing parameter among all determinant-matched lattices.

3.2 Instantiating the Wagner-Style Gaussian Sampler

We now instantiate the Wagner-style Gaussian sampler of [DEL25, Algorithm 4] with the random P -ary auxiliary extensions constructed above. Each level uses the same two operations: DGLIFT lifts the current list from A_{i-1} to A'_i , and BUCKETANDCOMBINE combines vectors with the same class modulo A_i . The abstract lifting and collision transformations are unchanged. What requires a new implementation is the shifted DGS oracle inside DGLIFT; the cyclic quotient labels used by BUCKETANDCOMBINE are described in Section 3.2.

⁵ When P^{1/b_i} is not integral, it need not contain $q\mathbb{Z}^{b_i}$ and therefore does not define the collision quotient required by the sampler.

DGLIFT over Random P -ary Auxiliary Lattices. We follow the abstract lifting procedure of [DEL25, Algorithm 2 and Lemma 3.1]. Its distributional analysis depends only on the primitive lattice extension and the smoothing condition on its kernel, both of which were established above. The remaining issue is to implement its shifted DGS oracle at the Gaussian width available here.

For the scaled integer auxiliary lattice in [DEL25], the natural basis satisfies the basis-dependent width condition of Lemma 2.7. For the random P -ary auxiliary lattice, the efficiently computable Hermite Normal Form (HNF) basis constructed below is not guaranteed to satisfy that condition at our target width. We therefore use Lemma 2.8, whose width condition depends on the distance from the center to the lattice.

Concrete Lift. Write $\mathbf{A}'_i = \begin{pmatrix} \mathbf{A}'_{i-1} \\ \Delta \mathbf{A}'_i \end{pmatrix}$, $\Delta \mathbf{A}'_i \in \mathbb{Z}^{b_i \times (m-n)}$. The orthogonal q -ary lattice A_i has basis

$$\mathbf{B}_i = \begin{pmatrix} \mathbf{I}_{m-n} & \mathbf{0} \\ -\mathbf{A}'_i & q\mathbf{I}_{n_i} \end{pmatrix}.$$

Let $\mathcal{C}_i$ be the lattice generated by the first $m-n+n_{i-1}$ columns of $\mathbf{B}_i$. The remaining b_i columns generate the kernel $q\mathbb{Z}^{b_i}$ of $\pi_i : A_i \rightarrow A_{i-1}$. Consequently, $A_i = \mathcal{C}_i \oplus q\mathbb{Z}^{b_i}$ and $A'_i = \mathcal{C}_i \oplus \Gamma_i$, where the last components are embedded in the newest coordinate block.

Let $\mathbf{x} = (\mathbf{u}, \mathbf{v}) \in A_{i-1}$, where $\mathbf{u} \in \mathbb{Z}^{m-n}$ and $\mathbf{v} \in \mathbb{Z}^{n_{i-1}}$. Its unique lift in $\mathcal{C}_i$ is $\mathbf{y} = (\mathbf{u}, \mathbf{v}, -\Delta \mathbf{A}'_i \mathbf{u})$. Writing $\bar{\mathbf{x}} = (\mathbf{u}, \mathbf{v}, \mathbf{0})$, $\mathbf{c}_i := \Delta \mathbf{A}'_i \mathbf{u} \in \mathbb{R}^{b_i}$, we have, for every $\mathbf{z} \in \Gamma_i$, $\mathbf{y} + (\mathbf{0}, \mathbf{0}, \mathbf{z}) = \bar{\mathbf{x}} + (\mathbf{0}, \mathbf{0}, \mathbf{z} - \mathbf{c}_i)$. The two components on the right are orthogonal. Hence $\rho_s(\mathbf{y} + (\mathbf{0}, \mathbf{0}, \mathbf{z})) = \rho_s(\bar{\mathbf{x}}) \rho_{s, \mathbf{c}_i}(\mathbf{z})$. Therefore, conditioned on the vector $\mathbf{x}$, the ideal lifting step samples $\mathbf{z} \leftarrow D_{\Gamma_i, s, \mathbf{c}_i}$ and returns $\mathbf{y} + (\mathbf{0}, \mathbf{0}, \mathbf{z})$.

An exact rational basis of Γ_i can be precomputed once at every level. Specifically, compute an HNF basis $\mathbf{H}_i$ of the column lattice generated by $[P\mathbf{I}_{b_i} \mid \mathbf{a}_i]$ and set $\mathbf{G}_i := \frac{q}{P}\mathbf{H}_i$. Then $\mathbf{G}_i$ is an exact rational basis of $\Gamma_i = (q/P)\Lambda_P(\mathbf{a}_i^T)$, and its computation takes polynomial time in the input bit length.

Algorithm 1 summarizes this concrete lift. It remains to verify that Lemma 2.8 applies to every shift and to quantify the resulting statistical error.

Algorithm 1: Implemented DGLIFT over a Random P -ary Extension

Input : Level data $(\Delta \mathbf{A}'_i, \Gamma_i, \mathbf{G}_i)$, a Gaussian width $s > 0$, and $\mathbf{x} \in A_{i-1}$.
Output : A vector $\mathbf{x}' \in A'_i$ satisfying $\pi_i(\mathbf{x}') = \mathbf{x}$.

- 1 Write $\mathbf{x} = (\mathbf{u}, \mathbf{v})$, where $\mathbf{u} \in \mathbb{Z}^{m-n}$ and $\mathbf{v} \in \mathbb{Z}^{n_{i-1}}$
- 2 Set $\mathbf{y} \leftarrow (\mathbf{u}, \mathbf{v}, -\Delta \mathbf{A}'_i \mathbf{u})$
- 3 Set $\mathbf{c}_i \leftarrow \Delta \mathbf{A}'_i \mathbf{u}$
- 4 Sample $\mathbf{z} \leftarrow D_{\Gamma_i, s, \mathbf{c}_i}$ using Lemma 2.8 with basis $\mathbf{G}_i$
- 5 Set $\mathbf{x}' \leftarrow \mathbf{y} + (\mathbf{0}, \mathbf{0}, \mathbf{z})$
- 6 **return** $\mathbf{x}'$

Shifted DGS Implementation. Because Γ_i contains $q\mathbb{Z}^{b_i}$, its covering radius satisfies $\mu(\Gamma_i) \leq \mu(q\mathbb{Z}^{b_i}) = \frac{q\sqrt{b_i}}{2}$. Therefore every center $\mathbf{c}_i \in \mathbb{R}^{b_i}$ arising in Algorithm 1 satisfies $\text{dist}(\mathbf{c}_i, \Gamma_i) \leq \frac{q\sqrt{b_i}}{2}$. Applying Lemma 2.8 with $d = b_i$ and $g(b_i) = e^{\sqrt{b_i}}$, we obtain a valid shifted DGS implementation whenever $s > \frac{q\sqrt{b_i}}{2e^{\sqrt{b_i}}}$. Each call then takes time $2^{b_i+o(b_i)}$ and has statistical error at most $2^{-c_{\text{DGS}}b_i^2}$ relative to the target shifted discrete Gaussian.

The centers used in the shifted DGS calls may depend adaptively on previous sampler outputs throughout the recursion. All invocations of the shifted DGS sampler use fresh independent randomness. We therefore bound the accumulated statistical error in Lemma 3.3. The proof is the standard adaptive hybrid argument and is given in Section B.4.

Lemma 3.3 (Adaptive Composition of Shifted DGS Errors). *Consider an adaptive computation making Q shifted DGS calls. Immediately before the j -th call, let $\mathbf{H}_{j-1}$ denote the complete algorithmic state. Conditioned on $\mathbf{H}_{j-1} = h$, let $D_j(\cdot \mid h)$ be the target distribution and let $\tilde{D}_j(\cdot \mid h)$ be the distribution produced by the implemented sampler. Suppose that $\text{SD}(\tilde{D}_j(\cdot \mid h), D_j(\cdot \mid h)) \leq \xi_j$ for every $j \in [Q]$ and every state h at which the j -th call is made. Then the joint distributions of the implemented and ideal computations satisfy $\text{SD}(\mathbf{Law}_{\text{impl}}, \mathbf{Law}_{\text{ideal}}) \leq \sum_{j=1}^Q \xi_j$.*

We now apply the adaptive composition bound to one complete lifting stage. In the ideal stage, the primitive extension controls conditional similarity. In the implemented stage, one shifted DGS call is made for each of the M input vectors.

Lemma 3.4 (Implemented DGLIFT). *Fix a level i , a width $s > 0$, and $0 < \varepsilon \leq \frac{1}{2}$. Suppose that*

$$s \geq \eta_\varepsilon(\Gamma_i), \quad s > \frac{q\sqrt{b_i}}{2e^{\sqrt{b_i}}}.$$

Let the input list consist of M vectors in Λ_{i-1} , with joint distribution conditionally δ -similar to $D_{\Lambda_{i-1}, s}^M$. Then Algorithm 1 outputs M vectors in Λ'_i , whose joint distribution is within statistical distance $M2^{-c_{\text{DGS}}b_i^2}$ of a distribution conditionally $(\delta+3\varepsilon)$ -similar to $D_{\Lambda'_i, s}^M$. It makes M shifted DGS calls and runs, up to polynomial factors, in time $M2^{b_i+o(b_i)}$.

Proof. With target shifted discrete Gaussian samples, Lemma 3.1 and [DEL25, Lemma 3.1] give conditional $(\delta + 3\varepsilon)$ -similarity to $D_{\Lambda'_i, s}^M$.

Under the second width condition, every implemented shifted-DGS call is within statistical distance $2^{-c_{\text{DGS}}b_i^2}$ of its target distribution for every center. Applying Lemma 3.3 to the M calls gives total statistical distance at most $M2^{-c_{\text{DGS}}b_i^2}$ between the implemented and ideal output laws. The running-time bound follows by summing the cost of the M shifted DGS calls. $\square$

Algorithm 2: BUCKETANDCOMBINE at Level i

Input : An ordered list $L' = (\mathbf{y}_1, \dots, \mathbf{y}_M) \subseteq \Lambda'_i$ with $M \geq 3P$.
Output : An ordered list of $\lfloor M/3 \rfloor$ vectors in Λ_i .
1 Initialize an empty ordered bucket $\mathcal{B}(c)$ for each $c \in \mathbb{Z}/P\mathbb{Z}$
2 **for** $j \leftarrow 1$ **to** M **do**
3 $c_j \leftarrow \ell_i(\mathbf{y}_j)$
4 Append $\mathbf{y}_j$ to $\mathcal{B}(c_j)$
5 $L \leftarrow ()$
6 **for** $j \leftarrow 1$ **to** M **do**
7 **if** $|L| < \lfloor M/3 \rfloor$ **and** $|\mathcal{B}(c_j)| \geq 2$ **then**
8 Let $\mathbf{y}, \mathbf{y}'$ be the first two vectors in $\mathcal{B}(c_j)$
9 Append $\mathbf{y} - \mathbf{y}'$ to L
10 Remove $\mathbf{y}, \mathbf{y}'$ from $\mathcal{B}(c_j)$
11 **return** L

BUCKETANDCOMBINE over the Cyclic Quotient. After lifting, BUCKETANDCOMBINE returns the list from Λ'_i to Λ_i by subtracting pairs that belong to the same coset modulo Λ_i . We use the collision procedure of [DEL25, Algorithm 3]. It remains only to give an efficient representation of the quotient classes.

By Lemma 3.1, $\Lambda'_i/\Lambda_i \cong \mathbb{Z}/P\mathbb{Z}$. Using the decomposition $\Lambda'_i = \mathcal{C}_i \oplus \Gamma_i$, write $\mathbf{y} = \mathbf{y}_c + \mathbf{z}$ with $\mathbf{y}_c \in \mathcal{C}_i$ and $\mathbf{z} \in \Gamma_i$. This decomposition is efficiently recoverable from $\mathbf{y}$: the vector $\mathbf{y}_c$ is the unique lift in $\mathcal{C}_i$ of $\pi_i(\mathbf{y})$, and then $\mathbf{z} = \mathbf{y} - \mathbf{y}_c$. Choose $r_i \in [b_i]$ such that $a_{i,r_i} \neq 0 \pmod{P}$, and define the quotient label by

$$\ell_i(\mathbf{y}) := a_{i,r_i}^{-1} P z_{r_i} / q \pmod{P}.$$

To see that the expression is integral, write $\mathbf{z} = q\mathbf{w} + (q/P)k\mathbf{a}_i$ for $\mathbf{w} \in \mathbb{Z}^{b_i}$ and $k \in \mathbb{Z}$. Then $\ell_i(\mathbf{y}) = k \pmod{P}$. It follows that $\ell_i(\mathbf{y}) = \ell_i(\mathbf{y}') \iff \mathbf{y} - \mathbf{y}' \in \Lambda_i$. Thus ℓ_i represents the quotient classes in Λ'_i/Λ_i .

We now apply BUCKETANDCOMBINE from [DEL25, Algorithm 3] using the labels ℓ_i . For $c \in \mathbb{Z}/P\mathbb{Z}$, let $N_c := |\mathcal{B}(c)|$ denote the size of bucket c . If its input list has length $M \geq 3P$, then the P buckets contain at least

$$\sum_{c \in \mathbb{Z}/P\mathbb{Z}} \left\lfloor \frac{N_c}{2} \right\rfloor \geq \frac{M - P}{2} \geq \frac{M}{3},$$

disjoint pairs. We retain the first $\lfloor M/3 \rfloor$ differences in the order prescribed by the collision procedure.

We use the collision analysis of [DEL25, Lemmas 3.2–3.4] with the quotient labels ℓ_i , under the stronger smoothing condition stated below.⁶

⁶ In [DEL25, Lemma 3.3], the stated smoothing condition is on the superlattice, while the convolution step in its proof uses smoothing of the underlying sublattice. Our assumption $s \geq \sqrt{2}\eta_\varepsilon(\Lambda_i)$ is exactly the condition required for that step, so the same argument gives the stated $(4\delta + 3\varepsilon)$ -similarity bound.

Algorithm 3: Wagner-Style Gaussian Sampler with Random P -ary Auxiliary Lattices

Input : A systematic matrix $\mathbf{A} = [\mathbf{A}' \mid \mathbf{I}_n] \in \mathbb{Z}_q^{n \times m}$, block lengths $b_1, \dots, b_r$ with $\sum_{i=1}^r b_i = n$, a prime P , an initial Gaussian width $s_0 > 0$, and a final list size N .

Output : An ordered list $L_r = (\mathbf{X}_1, \dots, \mathbf{X}_N) \subseteq \Lambda_q^\perp(\mathbf{A})$.

- 1 Set $n_0 \leftarrow 0$, $\Lambda_0 \leftarrow \mathbb{Z}^{m-n}$
- 2 Initialize $L_0 \leftarrow 3^r N$ independent samples from D_{Λ_0, s_0}
- 3 **for** $i \leftarrow 1$ **to** r **do**
- 4 $n_i \leftarrow n_{i-1} + b_i$ and $s_i \leftarrow \sqrt{2} s_{i-1}$
- 5 Let $\mathbf{A}'_i$ be the first n_i rows of $\mathbf{A}'$
- 6 Set $\mathbf{A}_i \leftarrow [\mathbf{A}'_i \mid \mathbf{I}_{n_i}]$ and $\Lambda_i \leftarrow \Lambda_q^\perp(\mathbf{A}_i)$
- 7 Sample $\mathbf{a}_i \leftarrow \mathbb{F}_P^{b_i} \setminus \{\mathbf{0}\}$
- 8 Set $\Gamma_i \leftarrow \frac{q}{P} \Lambda_P(\mathbf{a}_i^T)$, $\Lambda'_i \leftarrow \Lambda_i + (\mathbf{0} \oplus \Gamma_i)$
- 9 Compute a rational basis $\mathbf{G}_i$ of Γ_i , and choose $r_i \in [b_i]$ such that $a_{i, r_i} \neq 0 \pmod{P}$
- 10 Initialize $L'_i \leftarrow ()$
- 11 **foreach** $\mathbf{x} \in L_{i-1}$ *in list order* **do**
- 12 $\mathbf{x}' \leftarrow \text{DGLIFT}(i, s_{i-1}, \mathbf{x}; \Delta \mathbf{A}'_i, \Gamma_i, \mathbf{G}_i)$
- 13 Append $\mathbf{x}'$ to L'_i
- 14 $L_i \leftarrow \text{BUCKETANDCOMBINE}(L'_i)$
- 15 **return** L_r

Lemma 3.5 (Distribution of BUCKETANDCOMBINE). *Let $M \geq 3P$ and $0 < \varepsilon \leq \frac{1}{2}$. Suppose that the input list is conditionally δ -similar to $D_{\Lambda'_i, s}^M$. If $s \geq \sqrt{2} \eta_\varepsilon(\Lambda_i)$, then Algorithm 2 returns $\left\lfloor \frac{M}{3} \right\rfloor$ vectors in Λ_i whose joint distribution is conditionally $(4\delta + 3\varepsilon)$ -similar to $D_{\Lambda_i, \sqrt{2}s}^{\left\lfloor \frac{M}{3} \right\rfloor}$.*

Lemmas 3.4 and 3.5 give the two transformations used at each level. We now combine them into the full Wagner-style Gaussian sampler.

Correctness of the Wagner-Style Gaussian Sampler. The preceding lifting and collision lemmas determine the conditional similarity of the ideal sampler, in which every shifted DGS call returns an exact sample from its target distribution. The implemented sampler replaces these calls by Lemma 2.8. The next theorem accounts for the accumulated implementation error.

Theorem 3.1 (Correctness of the Sampler). *Let $0 < \varepsilon \leq \frac{1}{2}$, assume $s_0 \geq \sqrt{\ln 2(m-n) + 4/\pi}$, and suppose that, for every $i \in [r]$,*

$$s_{i-1} > \max \left\{ \eta_\varepsilon(\Gamma_i), \sqrt{2} \eta_\varepsilon(\Lambda_i), \frac{q\sqrt{b_i}}{2 \exp(\sqrt{b_i})} \right\}.$$

Assume also that $P \leq N$, and let $b_{\min} := \min_{i \in [r]} b_i$.

The ideal version of Algorithm 3 outputs exactly N vectors in $\Lambda_q^\perp(\mathbf{A})$ whose joint distribution is conditionally $5(4^r - 1)\varepsilon$ -similar to $D_{\Lambda_q^\perp(\mathbf{A}), 2^{r/2}s_0}^N$.

The output law of the implemented sampler using Lemma 2.8 is within statistical distance $((3^{r+1} - 3)/2)N2^{-c_{\text{DGS}}b_{\min}^2}$ of this ideal output law. The implemented sampler makes exactly $((3^{r+1} - 3)/2)N$ shifted DGS calls.

Proof. We first analyze the ideal procedure. Let δ_i denote the conditional similarity parameter after level i . Initially, $M_0 = 3^r N$ and $\delta_0 = 0$. Suppose that L_{i-1} is conditionally δ_{i-1} -similar to $D_{\Lambda_{i-1}, s_{i-1}}^{M_{i-1}}$. Lemma 3.4 gives conditional $(\delta_{i-1} + 3\varepsilon)$ -similarity on Λ'_i , and Lemma 3.5 gives $M_i = M_{i-1}/3$, $s_i = \sqrt{2}s_{i-1}$, and $\delta_i = 4\delta_{i-1} + 15\varepsilon$. Since $P \leq N$, we have $M_{i-1} \geq 3P$ at every level. Solving these recurrences yields $M_i = 3^{r-i}N$, $s_i = 2^{i/2}s_0$, $\delta_i = 5(4^i - 1)\varepsilon$. At $i = r$, $\Lambda_r = \Lambda_q^\perp(\mathbf{A})$, which proves the claim for the ideal procedure.

At level i , the implementation makes $M_{i-1} = 3^{r-i+1}N$ DGS calls, each within statistical distance $2^{-c_{\text{DGS}}b_i^2}$ of its target distribution. Hence Lemma 3.3 gives

$$\text{SD}(\text{Law}_{\text{impl}}, \text{Law}_{\text{ideal}}) \leq \sum_{i=1}^r 3^{r-i+1}N 2^{-c_{\text{DGS}}b_i^2} \leq \frac{3^{r+1} - 3}{2}N 2^{-c_{\text{DGS}}b_{\min}^2}.$$

Finally, $\sum_{i=1}^r M_{i-1} = ((3^{r+1} - 3)/2)N$, which is the claimed number of shifted DGS calls. $\square$

3.3 Sampler Tradeoff and Complexity

We now combine the requirements established above to derive explicit parameters and complexity bounds. We separate the parameter analysis into three levels. Theorem 3.2 treats the output-list size N as a prescribed parameter. Corollary 3.1 then chooses the smallest asymptotically admissible N , and Corollary 3.2 specializes the resulting bound to the parameter regime used for LWE.

For compactness, throughout this subsection we write

$$L := \ln(2m/\varepsilon), \quad W := \ln \ln \left(\frac{q}{\sqrt{L}} \right) - \ln(1 + \ln f). \quad (1)$$

Theorem 3.2 gives sufficient conditions, in terms of N , under which the sampler produces N Gaussian lattice vectors at target width q/f , together with the corresponding time and space bounds.

Theorem 3.2 (Wagner-Style Gaussian Sampler with a Prescribed List Size). *Consider an asymptotic parameter family with $m > n$, prime q , $f > 1$, and $0 < \varepsilon \leq 1/2$. Define L and W as in Equation (1), and suppose that*

$$L = O\left(\frac{n}{\ln \ln q}\right), \quad \frac{q}{f} \geq 64\sqrt{L}.$$

There exist constants $C_{\text{sam}}, C > 0$ such that the following holds. Let $N \geq 2$ be a power of two satisfying

$$W > C_{\text{sam}}, \quad \log_2 N \geq \frac{n}{2(W - C_{\text{sam}})}, \quad (2)$$

and

$$(\ln q)^3 = o(\ln N), \quad \ln N = o(n).$$

Suppose further that

$$m - n \geq C \left(\log_2 N + \frac{L}{\ln(2 + L)} \right).$$

Then, for all sufficiently large n , with probability $1 - o(1)$ over $\mathbf{A} \leftarrow \mathbb{Z}_q^{n \times m}$ and the random auxiliary lattices, the sampler outputs N vectors in $\Lambda_q^\perp(\mathbf{A})$. Its output distribution is within statistical distance $o(1)$ of a distribution conditionally $q^4 \varepsilon$ -similar to $D_{\Lambda_q^\perp(\mathbf{A}), q/f}^N$. The expected time and space complexities are $N^{1+O(1/(1+\ln f))+o(1)} \text{poly}(m, \ln q, \ln(1/\varepsilon))$.

Proof. The proof constructs feasible internal parameters for the sampler from the instance parameters n, m, q , the target parameters f, ε , and the list-size parameter N . Starting from the target width q/f , we choose the number r of recursion levels and the Gaussian widths $s_0, \dots, s_r$ backward so that $s_r = q/f$ while the initial width s_0 is large enough for exact sampling. At level i , we write $t_i := s_{i-1}/\sqrt{2}$ for the auxiliary smoothing width. The prescribed list size N determines a prime P of comparable size and, through the smoothing bound for the random P -ary auxiliary lattices, an upper bound for each block length b_i . We call this upper bound the *block capacity* at level i .

The central step is to choose b_i that stay below the per-level smoothing capacities, remain large enough for the shifted-DGS implementation, and satisfy $\sum_{i=1}^r b_i = n$. This is provided by Lemma B.1. Once these parameters have been fixed, we verify smoothing of every auxiliary lattice Γ_i , smoothing of every intermediate orthogonal q -ary lattice Λ_i , and the shifted-DGS width condition. We then apply Theorem 3.1, bound the accumulated shifted-DGS error, and estimate the running time and space.

Choice of Parameters. Set $\varepsilon_0 := \varepsilon/5$, $L_0 := \ln(2m/\varepsilon_0) = L + \ln 5$, and $r := \lfloor 2 \log_2(q/f) - \log_2(512L_0) \rfloor$. For $i \in [r]$, define

$$s_0 := \frac{q}{f 2^{r/2}}, \quad s_i := 2^{i/2} s_0, \quad t_i := \frac{s_{i-1}}{\sqrt{2}}, \quad x_i := \ln(q/t_i). \quad (3)$$

The assumed lower bound on q/f ensures that $r \geq 1$ for all sufficiently large n . From Equation (3), we have

$$s_r = \frac{q}{f}, \quad t_1 \geq 16\sqrt{L_0}, \quad t_r = \frac{q}{2f}, \quad x_r = \ln(2f).$$

Moreover, the x_i form a decreasing arithmetic progression and $r = O(\ln q)$.

Apply Lemma B.1 to Equation (3). It gives a prime $P \in (N/2, N)$ and positive integers $b_1, \dots, b_r$ such that $\sum_{i=1}^r b_i = n$ and, for every $i \in [r]$,

$$\frac{16q}{P^{1/b_i}} \sqrt{\max \left\{ 1, \frac{\ln(1/\varepsilon_0)}{b_i} \right\}} \leq t_i.$$

Write $b_{\min} := \min_{i \in [r]} b_i$, $b_{\max} := \max_{i \in [r]} b_i$, and $n_i := \sum_{j \leq i} b_j$. The same lemma gives, for every $i \in [r]$,

$$b_{\min} = \Omega \left(\frac{\ln N}{\ln q} \right), \quad b_{\max} = O \left(\frac{\ln N}{1 + \ln f} \right), \quad n_i x_i = O(\ln N \ln t_i).$$

Auxiliary Lattice Smoothing. By the choice of b_i , Lemma 3.2 gives $\eta_{\varepsilon_0}(\Gamma_i) \leq t_i$ simultaneously for all $i \in [r]$, except with probability at most $\sum_{i=1}^r 2^{1-b_i} \leq 2r 2^{-b_{\min}}$. The lower bound on $b_{\min}$ and the assumption $(\ln q)^3 = o(\ln N)$ give $\frac{b_{\min}}{(\ln q)^2} = \Omega \left(\frac{\ln N}{(\ln q)^3} \right) = \omega(1)$. Since $r = O(\ln q)$, the preceding failure probability is $o(1)$. Since $\eta_{\varepsilon_0}(\Gamma_i) \leq t_i < s_{i-1}$, the auxiliary-lattice smoothing condition in Theorem 3.1 holds at every level.

Orthogonal Lattice Smoothing. Fix $i \in [r]$ and put $m_i := m - n + n_i$. From Lemma 2.6, it is enough to prove

$$\frac{t_i}{2} \geq 8 \max \left\{ \sqrt{\ln \left(\frac{2m_i}{\varepsilon_0} \right)}, q^{n_i/m_i} \varepsilon_0^{-1/m_i} \right\}. \quad (4)$$

Indeed, $t_i/2 \leq t_r/2 = q/(4f) < q$, so Equation (4) also verifies the range condition in Lemma 2.6.

Since $m_i \leq m$, $t_i \geq t_1$, and $t_1 \geq 16\sqrt{L_0}$, the first term in Equation (4) is bounded as required. For the second term, taking logarithms shows that it is enough to prove

$$m_i \ln t_i - n_i \ln q \geq \ln(1/\varepsilon_0) + m_i \ln 16. \quad (5)$$

Using $m_i = m - n + n_i$ and $\ln q = \ln t_i + x_i$, the left-hand side of Equation (5), after subtracting $m_i \ln 16$, becomes

$$(m - n)(\ln t_i - \ln 16) - n_i(x_i + \ln 16).$$

Because $x_i \geq x_r = \ln(2f) > \ln 2$, the prefix bound implies that this expression is at least $\left(\frac{m-n}{2} - C_1 \ln N \right) \ln t_i$ for some absolute constant $C_1 > 0$ and all sufficiently large n . Using the assumed lower bound on $m - n$, this is at least

$$\left(\frac{C}{2 \ln 2} - C_1 \right) \ln N \ln t_i + \frac{C}{2} \frac{L}{\ln(2 + L)} \ln t_i.$$

Choose C sufficiently large that the first term is nonnegative. Since $t_i \geq t_1 \geq 16\sqrt{L_0}$, we have $\ln t_i \geq \frac{1}{4} \ln(2 + L)$ for all sufficiently large n . Increasing C if necessary, the second term is therefore at least L , which is at least $\ln(1/\varepsilon_0)$. This proves Equations (4) and (5).

We next bound the failure probability for the prefix-lattice smoothing conditions. By the systematic-form reduction in Footnote 4, the failure bound of Lemma 2.6 increases by at most a factor two for all sufficiently large n . Hence $\Pr[\eta_{\varepsilon_0}(\Lambda_i) > t_i/2] \leq 2^{1-m_i}$. Since $m_i \geq m - n = \Omega(\log_2 N)$ and $r = O(\ln q) = o(\ln N)$, the union bound over all prefixes is $o(1)$. Consequently, $\sqrt{2} \eta_{\varepsilon_0}(\Lambda_i) \leq t_i/\sqrt{2} < s_{i-1}$, as required by Theorem 3.1.

Shifted-DGS Implementation. It remains to verify the shifted-DGS width condition $s_{i-1} > \frac{q\sqrt{b_i}}{2e^{\sqrt{b_i}}}$ in Theorem 3.1. Since $s_{i-1} = \sqrt{2}t_i$ and $t_i = qe^{-x_i}$, this inequality is equivalent to

$$\sqrt{b_i} - x_i - \frac{1}{2} \ln b_i + \ln(2\sqrt{2}) > 0.$$

The assumption $(\ln q)^3 = o(\ln N)$ gives $\ln q = o(\sqrt{\ln N / \ln q})$. Moreover, $b_{\max} = O(\ln N)$ gives $\ln b_{\max} = O(\ln \ln N) = o(\sqrt{\ln N / \ln q})$. Together with the lower bound on $b_{\min}$, this yields $\sqrt{b_{\min}} - \ln q - \frac{1}{2} \ln b_{\max} + \ln(2\sqrt{2}) = \omega(1)$. Since $b_i \geq b_{\min}$, $b_i \leq b_{\max}$, and $x_i < \ln q$, it follows that, for every $i \in [r]$,

$$\sqrt{b_i} - x_i - \frac{1}{2} \ln b_i + \ln(2\sqrt{2}) \geq \sqrt{b_{\min}} - \ln q - \frac{1}{2} \ln b_{\max} + \ln(2\sqrt{2}) > 0$$

for all sufficiently large n . The shifted-DGS condition therefore holds at every level for all sufficiently large n .

Distributional Guarantee and Complexity. The initial list can be sampled exactly using Lemma 2.7. Indeed, $s_0 \geq \sqrt{512L_0} \geq \sqrt{\ln(2(m-n)+4)/\pi}$, which is the required width for the standard basis of $\Lambda_0 = \mathbb{Z}^{m-n}$.

Moreover, $P < N$, so every call to BUCKETANDCOMBINE receives at least $3P$ inputs. We may therefore apply Theorem 3.1 with smoothing error ε_0 . The final Gaussian width is $s_r = q/f$, and the conditional-similarity parameter is $5(4^r - 1)\varepsilon_0 = (4^r - 1)\varepsilon < q^4\varepsilon$. Indeed, $4^r = \left(2^{r/2}\right)^4 = \left(\frac{q}{fs_0}\right)^4 < q^4$.

The total number of shifted-DGS calls is $\frac{3^{r+1}-3}{2}N$. Since $r = O(\ln q) = o(\ln N)$, this number is $N^{1+o(1)}$. Furthermore, $\frac{b_{\min}^2}{\ln N} = \Omega\left(\frac{\ln N}{(\ln q)^2}\right) = \omega(1)$. It follows that the accumulated implementation error $\frac{3^{r+1}-3}{2}N2^{-c_{\text{DGS}}b_{\min}^2}$ is $o(1)$.

A shifted-DGS call at level i costs $2^{b_i+o(b_i)}$ up to polynomial bit factors. Since $b_{\max} = O(\ln N/(1 + \ln f))$, the expected running time, and likewise the space, is

$$N^{1+O(1/(1+\ln f))+o(1)} \text{poly}(m, \ln q, \ln(1/\varepsilon)).$$

A union bound over the auxiliary-lattice, prefix-lattice, and systematic-form failure events gives total failure probability $o(1)$. Conditioned on the complementary good event, the distributional guarantee above holds over the remaining sampler randomness, as required later in Theorem 4.1. $\square$

We next choose N at the smallest asymptotic scale allowed by Equation (2), thereby obtaining a complexity bound directly in terms of n , q , f , and ε .

Corollary 3.1 (Asymptotic Sampler Tradeoff). *Let $m > n$, q be prime, $f > 1$, and $0 < \varepsilon \leq 1/2$. Define L and W as in Equation (1). Suppose that*

$$L = O\left(\frac{n}{\ln \ln q}\right), \quad \frac{q}{f} \geq 64\sqrt{L}, \quad W = \omega(1), \quad \frac{W(\ln q)^3}{n} = o(1).$$

If $m - n \geq C_0 \left(\frac{n}{W} + \frac{L}{\ln(2+L)}\right)$ for a sufficiently large constant $C_0 > 0$, then there exists a power of two N satisfying $\log_2 N = \left(\frac{1}{2} + o(1)\right) \frac{n}{W}$ for which the conclusions of Theorem 3.2 hold. In particular, the expected time and space are $2^{\left(\frac{1}{2} + O\left(\frac{1}{1+\ln f}\right) + o(1)\right) n/W} \text{poly}(m, \ln q, \ln(1/\varepsilon))$.

Proof. Set $\log_2 N = \left\lceil \frac{n}{2(W - C_{\text{sam}})} \right\rceil = \left(\frac{1}{2} + o(1)\right) \frac{n}{W}$, since $W = \omega(1)$. The assumptions on W , q , and $m - n$ imply all remaining conditions of Theorem 3.2, and substituting this choice of N gives the claimed bounds. $\square$

The following specialization evaluates the tradeoff under explicit logarithmic growth assumptions on q and f .

Corollary 3.2 (Logarithmic-Scale Sampler Specialization). *Fix constants $0 \leq a < b$. Let $m > n$, q be prime, $f > 1$, and $0 < \varepsilon \leq 1/2$. Define L as in Equation (1). Suppose that*

$$\ln f = (\ln n)^{a+o(1)}, \quad \ln\left(\frac{q}{\sqrt{L}}\right) = (\ln n)^{b+o(1)}, \quad L = O\left(\frac{n}{\ln \ln n}\right).$$

If, for a sufficiently large constant $C_{\log} > 0$, $m - n \geq C_{\log} \frac{n}{(b-a) \ln \ln n}$, then there exists a power of two N satisfying $\log_2 N = \left((2(b-a))^{-1} + o(1)\right) \frac{n}{\ln \ln n}$ for which the conclusions of Theorem 3.2 hold. The expected time and space are

$$2^{\left(\frac{1}{2(b-a)} + O\left(\frac{1}{1+\ln f}\right) + o(1)\right) n / \ln \ln n} \text{poly}(m, \ln q, \ln(1/\varepsilon)).$$

Proof. The assumptions give $\ln(q/\sqrt{L}) = (b + o(1)) \ln \ln n$ and $\ln(1 + \ln f) = (a + o(1)) \ln \ln n$. The second equality also covers $a = 0$, when $\ln(1 + \ln f) = o(\ln \ln n)$. Hence Equation (1) gives $W = (b - a + o(1)) \ln \ln n$. The claim follows by substituting $W = (b - a + o(1)) \ln \ln n$ into Corollary 3.1. The remaining conditions of Corollary 3.1 follow directly from the assumptions. $\square$

Remark 3.2 (Comparison with the Scaled Integer Auxiliary Lattice). The gain over the scaled integer construction of [DEL25] comes from the dependence of the auxiliary-lattice smoothing bound on the required accuracy. For the scaled integer lattice, the global factor $\sqrt{\ln(1/\varepsilon)}$ contributes $\ln(\ln f + \frac{1}{2} \ln \ln(1/\varepsilon))$ to the second term in the logarithmic width W . In contrast, Lemma 3.2 replaces this loss by $\sqrt{\max\{1, \ln(1/\varepsilon)/b_i\}}$, leaving the second term as $\ln(1 + \ln f)$.

This difference is decisive when $\ln(1/\varepsilon) = \Theta(n/\ln \ln n)$. For polynomial q and $\ln f = (\ln n)^{a+o(1)}$ with $a < 1$, our construction gives $W = (1 - a + o(1)) \ln \ln n$ and hence $\log_2 N = \Theta(n/\ln \ln n)$. With the scaled integer smoothing bound, both terms in the corresponding logarithmic difference are $\ln \ln n + O(1)$, leaving only an $O(1)$ gap and therefore requiring $\log_2 N = \Theta(n)$. The price for exploiting the narrower widths enabled by the random P -ary geometry is the $N^{O(1/(1+\ln f))}$ overhead of the statistically approximate shifted-DGS implementation, which is $N^{o(1)}$ whenever $f = \omega(1)$.

4 A Subexponential-Time Decision-LWE Algorithm

The preceding section constructed a Gaussian sampler that, given $\mathbf{A} \in \mathbb{Z}_q^{n \times m}$, produces vectors in $\Lambda_q^\perp(\mathbf{A})$ at target width q/f . A natural next step is to use these vectors as the dual vectors in a dual attack.

We briefly recall the general principle of dual attack. Given $\mathbf{A} \in \mathbb{Z}_q^{n \times m}$, the input vector $\mathbf{b}$ is drawn either from $\mathbf{b} = \mathbf{A}^T \mathbf{s} + \mathbf{e} \bmod q$ or uniformly from $\mathbb{Z}_q^m$. For $\mathbf{x} \in \Lambda_q^\perp(\mathbf{A})$, consider the residue $\langle \mathbf{x}, \mathbf{b} \rangle \bmod q$. For an LWE input, the defining relation $\mathbf{A}\mathbf{x} = \mathbf{0} \bmod q$ gives

$$\langle \mathbf{x}, \mathbf{b} \rangle = \langle \mathbf{x}, \mathbf{A}^T \mathbf{s} + \mathbf{e} \rangle = \langle \mathbf{A}\mathbf{x}, \mathbf{s} \rangle + \langle \mathbf{x}, \mathbf{e} \rangle = \langle \mathbf{x}, \mathbf{e} \rangle \bmod q.$$

Thus the dual vector eliminates the secret and leaves an observation determined by the error. On the other hand, if $\mathbf{b} \leftarrow \mathbb{Z}_q^m$ is uniform, then, for every fixed $\mathbf{x} \notin q\mathbb{Z}^m$, $\langle \mathbf{x}, \mathbf{b} \rangle \bmod q$ is uniform. Hence, for a fixed dual vector $\mathbf{x}$, the original decision-LWE problem is reduced to distinguishing $\langle \mathbf{x}, \mathbf{e} \rangle \bmod q$ from $U(\mathbb{Z}_q)$.

More generally, given dual vectors $\mathbf{X}_1, \dots, \mathbf{X}_N$, the attack forms

$$z_j := \langle \mathbf{X}_j, \mathbf{b} \rangle \bmod q, \quad j \in [N],$$

and applies a statistical test to $(z_1, \dots, z_N)$. The sampler of the preceding section supplies the first step. It remains to choose and analyze the distinguisher.

Several dual distinguishers can be used in this framework. Laarhoven and Walter [LW21] compare the simple dual, Aharonov–Regev, and Neyman–Pearson decoders. The Aharonov–Regev distinguisher [AR05] is particularly well suited to our setting because its periodic-Gaussian score is naturally represented using Gaussian-distributed dual vectors. The Gaussian weights are already encoded in the sampling distribution, while the resulting empirical statistic is an average of bounded cosine values. This makes its ideal bias explicit and allows concentration to be proved under conditional similarity.

Algorithm 4: Decision-LWE Algorithm

Input : $(\mathbf{A}, \mathbf{b}) \in \mathbb{Z}_q^{n \times m} \times \mathbb{Z}_q^m$, a target bias $0 < \tau \leq 1$,
a target-width factor f , a list size N ,
and admissible parameters for Algorithm 3 with final width q/f .

Output : LWE or UNIFORM.

- 1 Write $\mathbf{A} = [\mathbf{C} \mid \mathbf{B}]$ with square $\mathbf{B}$
- 2 **if** $\mathbf{B}$ is singular **then**
- 3 **return** an independent uniform element of $\{\text{LWE}, \text{UNIFORM}\}$
- 4 $\mathbf{A} \leftarrow \mathbf{B}^{-1} \mathbf{A}$
- 5 Run Algorithm 3 on $\mathbf{A}$ with the supplied sampler parameters to obtain
 $L = (\mathbf{X}_1, \dots, \mathbf{X}_N)$
- 6 $\hat{\mu} \leftarrow N^{-1} \sum_{j=1}^N \cos(2\pi \frac{\langle \mathbf{X}_j, \mathbf{b} \rangle}{q})$
- 7 **if** $\hat{\mu} \geq \frac{\tau}{2}$ **then**
- 8 **return** LWE
- 9 **return** UNIFORM

4.1 The Aharonov–Regev Dual Distinguisher

We now formalize the Aharonov–Regev distinguisher selected above. Let $\Lambda := \Lambda_q(\mathbf{A})$. The duality relation $\Lambda^* = q^{-1} \Lambda_q^\perp(\mathbf{A})$ implies that $\mathbf{X} \leftarrow D_{\Lambda_q^\perp(\mathbf{A}), q/f} \iff \frac{\mathbf{X}}{q} \leftarrow D_{\Lambda^*, 1/f}$. For $\mathbf{b} \in \mathbb{Z}_q^m$, define its Fourier mean by

$$\mu_{\mathbf{A}}(\mathbf{b}) := \mathbb{E}_{\mathbf{X} \leftarrow D_{\Lambda_q^\perp(\mathbf{A}), q/f}} \cos\left(\frac{2\pi}{q} \langle \mathbf{X}, \mathbf{b} \rangle\right) = \frac{\rho_f(\Lambda_q(\mathbf{A}) - \mathbf{b})}{\rho_f(\Lambda_q(\mathbf{A}))}. \quad (6)$$

The equality follows from the Fourier representation of the periodic Gaussian [AR05, Lemma 1.3]. Hence $\mu_{\mathbf{A}}(\mathbf{b})$ is precisely the Aharonov–Regev periodic-Gaussian score.

For a list $L = (\mathbf{X}_1, \dots, \mathbf{X}_N) \subseteq \Lambda_q^\perp(\mathbf{A})$, define the empirical Fourier score

$$\hat{\mu}_L(\mathbf{b}) := \frac{1}{N} \sum_{j=1}^N \cos\left(\frac{2\pi}{q} \langle \mathbf{X}_j, \mathbf{b} \rangle\right).$$

For independent $\mathbf{X}_j \leftarrow D_{\Lambda_q^\perp(\mathbf{A}), q/f}$, this is an unbiased estimator of $\mu_{\mathbf{A}}(\mathbf{b})$. The Aharonov–Regev distinguishing step compares this empirical score with a threshold determined by a lower bound τ on the LWE value of $\mu_{\mathbf{A}}(\mathbf{b})$.

Our full decision-LWE algorithm combines this Fourier test with the Wagner-style Gaussian sampler developed in Section 3. It first generates a list $L \subseteq \Lambda_q^\perp(\mathbf{A})$ at width q/f , and then returns LWE if $\hat{\mu}_L(\mathbf{b}) \geq \tau/2$, and UNIFORM otherwise; see Algorithm 4. The sampled vectors need not be independent, and the effect of their conditional similarity on this test is analyzed below.

4.2 Analysis of the Dual Distinguisher

We analyze the dual test for a fixed public matrix $\mathbf{A}$. The analysis separates the two sources of distinguishing error. Define

$$\nu_\chi(\tau; f, m) := \Pr_{\mathbf{e} \leftarrow \chi^m} [\rho_f(\mathbf{e}) < \tau], \quad \Delta := (m - n) \ln q - m \ln(2f), \quad (7)$$

where χ is the error distribution over $\mathbb{Z}$, $0 < \tau \leq 1$ is a target bias, and $f \geq 1$. Thus $\nu_\chi(\tau; f, m)$ bounds the probability that the ideal LWE score falls below τ , while $e^{-\Delta}$ will bound the expected ideal score of a uniform input.

Ideal Fourier Gap. We now compare the ideal mean $\mu_{\mathbf{A}}(\mathbf{b})$ defined in Equation (6) under LWE and uniform inputs.

Lemma 4.1 (Ideal Fourier Gap). *Let $m > n$, q be prime. Fix $\mathbf{A} \in \mathbb{Z}_q^{n \times m}$. Let $f \geq 1$, let χ be an error distribution over $\mathbb{Z}$, and let $0 < \tau \leq 1$. Then the following hold.*

- (i) *For every secret $\mathbf{s} \in \mathbb{Z}_q^n$, if $\mathbf{b} = \mathbf{A}^T \mathbf{s} + \mathbf{e} \bmod q$ with $\mathbf{e} \leftarrow \chi^m$, then $\mu_{\mathbf{A}}(\mathbf{b}) \geq \tau$ except with probability at most $\nu_\chi(\tau; f, m)$ over $\mathbf{e}$.*
- (ii) *If $\mathbf{A}$ has rank n and $\mathbf{b} \leftarrow \mathbb{Z}_q^m$ is uniform, then $\mathbb{E}_{\mathbf{b}} \mu_{\mathbf{A}}(\mathbf{b}) \leq e^{-\Delta}$. Consequently, $\Pr_{\mathbf{b}} [\mu_{\mathbf{A}}(\mathbf{b}) > \frac{\tau}{8}] \leq \frac{8e^{-\Delta}}{\tau}$.*

Proof. For an LWE input, $\mathbf{b} - \mathbf{e} \in \Lambda_q(\mathbf{A})$. Hence Lemma 2.3 gives $\mu_{\mathbf{A}}(\mathbf{b}) \geq \rho_f(\mathbf{e})$. The first claim follows directly from the definition of $\nu_\chi(\tau; f, m)$.

For uniform $\mathbf{b}$, character orthogonality over $\mathbb{Z}_q^m$ gives, for every fixed $\mathbf{X} \in \mathbb{Z}^m$, $\mathbb{E}_{\mathbf{b}} e^{2\pi i \langle \mathbf{X}, \mathbf{b} \rangle / q} = \mathbf{1}_{\{\mathbf{X} \in q\mathbb{Z}^m\}}$. Because $\mathbf{A}$ is fixed, the distribution $D_{\Lambda_q^\perp(\mathbf{A}), q/f}$ is also fixed. Taking real parts, averaging over the reference vector $\mathbf{X} \leftarrow D_{\Lambda_q^\perp(\mathbf{A}), q/f}$, and exchanging the two expectations, which is justified because the cosine is bounded, gives

$$\begin{aligned} \mathbb{E}_{\mathbf{b}} \mu_{\mathbf{A}}(\mathbf{b}) &= \mathbb{E}_{\mathbf{b}} \mathbb{E}_{\mathbf{X}} \cos\left(\frac{2\pi}{q} \langle \mathbf{X}, \mathbf{b} \rangle\right) = \mathbb{E}_{\mathbf{X}} \mathbb{E}_{\mathbf{b}} \cos\left(\frac{2\pi}{q} \langle \mathbf{X}, \mathbf{b} \rangle\right) \\ &= \mathbb{E}_{\mathbf{X} \leftarrow D_{\Lambda_q^\perp(\mathbf{A}), q/f}} \Pr[\mathbf{X} \in q\mathbb{Z}^m] \\ &= \frac{\rho_{q/f}(q\mathbb{Z}^m)}{\rho_{q/f}(\Lambda_q^\perp(\mathbf{A}))}. \end{aligned}$$

By scaling and Poisson summation, $\rho_{q/f}(q\mathbb{Z}^m) = f^{-m} \rho_f(\mathbb{Z}^m)$ and $\rho_{q/f}(\Lambda_q^\perp(\mathbf{A})) = q^{m-n} f^{-m} \rho_f(\Lambda_q(\mathbf{A}))$, where the second identity uses $\Lambda_q(\mathbf{A})^* = q^{-1} \Lambda_q^\perp(\mathbf{A})$ and $\det(\Lambda_q(\mathbf{A})) = q^{m-n}$. Therefore $\mathbb{E}_{\mathbf{b}} \mu_{\mathbf{A}}(\mathbf{b}) = \rho_f(\mathbb{Z}^m) / (q^{m-n} \rho_f(\Lambda_q(\mathbf{A})))$. Since $x \mapsto e^{-\pi x^2 / f^2}$ is decreasing on $\mathbb{R}_{\geq 0}$, $2 \sum_{k \geq 1} e^{-\pi k^2 / f^2} \leq 2 \int_0^\infty e^{-\pi x^2 / f^2} dx = f$. Hence $\rho_f(\mathbb{Z}) \leq 1 + f \leq 2f$. Thus $\mathbb{E}_{\mathbf{b}} \mu_{\mathbf{A}}(\mathbf{b}) \leq q^{-(m-n)} (2f)^m = e^{-\Delta}$. Finally, $\mu_{\mathbf{A}}(\mathbf{b}) \geq 0$, so Markov's inequality gives $\Pr_{\mathbf{b}} [\mu_{\mathbf{A}}(\mathbf{b}) > \frac{\tau}{8}] \leq \frac{8e^{-\Delta}}{\tau}$. $\square$

For $\chi = D_{\mathbb{Z}, s_e}$, set $\tau := \exp(-s_e^2 m / f^2)$. By Lemma 2.4, $\nu_\chi(\tau; f, m) \leq (2/e)^{m/2}$. Thus τ is a valid target bias for the discrete Gaussian specialization. It remains to show that the empirical score concentrates around $\mu_{\mathbf{A}}(\mathbf{b})$ when the Gaussian dual vectors are conditionally similar but not independent.

Concentration Under Conditional Similarity. The sampler output need not consist of independent vectors, so Hoeffding's inequality does not apply directly. Instead, we center each summand by its conditional mean given the preceding samples. The resulting partial sums form a martingale, while conditional similarity bounds the discrepancy between each conditional mean and the target mean.

We recall the relevant definitions from [MU05, Chapter 12]. A sequence $M_0, \dots, M_N$ is a martingale with respect to $\mathbf{X}_1, \dots, \mathbf{X}_N$ if, for every $k \in [N]$, M_k is determined by $\mathbf{X}_1, \dots, \mathbf{X}_k$, $\mathbb{E}[|M_k|] < \infty$, and $\mathbb{E}[M_k \mid \mathbf{X}_1, \dots, \mathbf{X}_{k-1}] = M_{k-1}$. The Azuma–Hoeffding inequality states that if, in addition, $|M_k - M_{k-1}| \leq c_k$ almost surely for every $k \in [N]$, then, for every $u > 0$,

$$\Pr[|M_N - M_0| \geq u] \leq 2 \exp \left(-u^2 / (2 \sum_{k=1}^N c_k^2) \right).$$

Lemma 4.2 (Conditional Similarity Implies Martingale Concentration).

Suppose $\mathbf{X}_1, \dots, \mathbf{X}_N$ are conditionally δ -similar to independent samples from a distribution D , where $0 \leq \delta \leq 1$. Then, for every $h : \mathcal{X} \rightarrow [-1, 1]$ and every $t > 0$,

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

Proof. For each $k \in [N]$, define the conditional mean $a_k := \mathbb{E}[h(\mathbf{X}_k) \mid \mathbf{X}_1, \dots, \mathbf{X}_{k-1}]$. Fix any prefix $\mathbf{x}_{<k}$ in the support of $\mathbf{X}_{<k}$. By conditional δ -similarity, the conditional distribution $P_k(\mathbf{x} \mid \mathbf{x}_{<k}) := \Pr[\mathbf{X}_k = \mathbf{x} \mid \mathbf{X}_{<k} = \mathbf{x}_{<k}]$ satisfies $P_k(\mathbf{x} \mid \mathbf{x}_{<k}) = e^{\pm\delta} D(\mathbf{x})$ for every $\mathbf{x} \in \mathcal{X}$. Consequently,

$$\begin{aligned} |a_k - \mathbb{E}_D h| &\leq \sum_{\mathbf{x} \in \mathcal{X}} |h(\mathbf{x})| |P_k(\mathbf{x} \mid \mathbf{x}_{<k}) - D(\mathbf{x})| \\ &\leq (e^\delta - 1) \sum_{\mathbf{x} \in \mathcal{X}} D(\mathbf{x}) = e^\delta - 1 \leq 2\delta, \end{aligned}$$

Now define $M_0 := 0$, $M_k := \sum_{j=1}^k (h(\mathbf{X}_j) - a_j)$ for $k \in [N]$. Then $(M_k)_{k=0}^N$ is a martingale with respect to $\mathbf{X}_1, \dots, \mathbf{X}_N$.⁷ We next verify the bounded-difference hypothesis required by the Azuma–Hoeffding inequality. Since $h(\mathbf{X}_k) \in [-1, 1]$ and $a_k \in [-1, 1]$, we have $|M_k - M_{k-1}| = |h(\mathbf{X}_k) - a_k| \leq 2$. Thus the Azuma–Hoeffding inequality gives

$$\Pr[|M_N| \geq Nt] \leq 2 \exp \left(-\frac{N^2 t^2}{2 \sum_{k=1}^N 2^2} \right) = 2 \exp \left(-\frac{Nt^2}{8} \right).$$

⁷ Indeed, M_k is determined by $\mathbf{X}_1, \dots, \mathbf{X}_k$; moreover, $|M_k| \leq 2k$, so $\mathbb{E}[|M_k|] < \infty$; and $\mathbb{E}[M_k \mid \mathbf{X}_1, \dots, \mathbf{X}_{k-1}] = M_{k-1} + \mathbb{E}[h(\mathbf{X}_k) - a_k \mid \mathbf{X}_1, \dots, \mathbf{X}_{k-1}] = M_{k-1}$.

Finally,

$$\left| \frac{1}{N} \sum_{k=1}^N h(\mathbf{X}_k) - \mathbb{E}_D h \right| = \left| \frac{M_N}{N} + \frac{1}{N} \sum_{k=1}^N (a_k - \mathbb{E}_D h) \right| \leq \frac{|M_N|}{N} + 2\delta.$$

Therefore, if the left-hand side exceeds $2\delta + t$, then $|M_N| > Nt$. The preceding Azuma–Hoeffding bound proves Equation (8). $\square$

Lemma 4.1 bounds the ideal Aharonov–Regev mean under LWE and uniform inputs, while Lemma 4.2 controls the deviation of the empirical score from this mean under conditional similarity. Combining these bounds and accounting for the statistical distance of the actual list distribution gives the following guarantee.

Theorem 4.1 (Aharonov–Regev Test under Conditional Similarity).

Let $m > n$, let q be prime, and fix a full-row-rank matrix $\mathbf{A} \in \mathbb{Z}_q^{n \times m}$. Let $f \geq 1$, let χ be an error distribution over $\mathbb{Z}$, let $N \geq 1$, and fix a target bias $0 < \tau \leq 1$. Define Δ as in Equation (7).

Let $L = (\mathbf{X}_1, \dots, \mathbf{X}_N) \subseteq \Lambda_q^\perp(\mathbf{A})$ be a random list generated independently of $\mathbf{b}$. Suppose that the law of L is within statistical distance ξ of a distribution conditionally δ -similar to $D_{\Lambda_q^\perp(\mathbf{A}), q/f}^N$, where $\delta \leq \tau/64$. If

$$\nu_\chi(\tau; f, m) = o(1), \quad e^{-\Delta} = o(\tau), \quad N^{-1/2} = o(\tau), \quad \xi = o(1),$$

then the distinguisher succeeds with probability $1 - o(1)$.

Proof. Let $D := D_{\Lambda_q^\perp(\mathbf{A}), q/f}$ and $h_{\mathbf{b}}(\mathbf{x}) := \cos(2\pi \langle \mathbf{x}, \mathbf{b} \rangle / q)$. Applying Lemma 4.2 to $h_{\mathbf{b}}$ and using Equation (6), we have $\mathbb{E}_D[h_{\mathbf{b}}] = \mu_{\mathbf{A}}(\mathbf{b})$. Taking $t = \tau/16$ and using $\delta \leq \tau/64$ gives $2\delta + t \leq 3\tau/32$. Therefore

$$\Pr \left[|\hat{\mu} - \mu_{\mathbf{A}}(\mathbf{b})| > \frac{3\tau}{32} \right] \leq 2 \exp \left(-\frac{N\tau^2}{2048} \right).$$

For an LWE input, Lemma 4.1 gives $\mu_{\mathbf{A}}(\mathbf{b}) \geq \tau$ except with probability at most $\nu_\chi(\tau; f, m)$. Outside these two events, $\hat{\mu} > \tau/2$. For a uniform input, Lemma 4.1 gives $\mu_{\mathbf{A}}(\mathbf{b}) \leq \tau/8$ except with probability at most $8e^{-\Delta}/\tau$. Outside these events, $\hat{\mu} < \tau/2$. Thus, under the conditionally similar reference law, the error probabilities on LWE and uniform inputs are at most, respectively, $\nu_\chi(\tau; f, m) + 2 \exp \left(-\frac{N\tau^2}{2048} \right)$ and $\frac{8e^{-\Delta}}{\tau} + 2 \exp \left(-\frac{N\tau^2}{2048} \right)$.

The statistical distance between the actual list law and the conditionally similar reference law is at most ξ . Therefore the probability of either decision event changes by at most ξ . Adding this term to the preceding bounds proves the two claims. Under the displayed asymptotic assumptions, every term in the two error bounds is $o(1)$. $\square$

4.3 Combining Sampler and Distinguisher

We now combine the sampler tradeoff from Corollary 3.2 with the Fourier distinguisher of Theorem 4.1. The following theorem states the resulting modulus-error tradeoff directly. Its two corollaries isolate the polynomial-modulus and polynomial-error regimes.

Theorem 4.2 (Logarithmic-Scale Modulus-Error Tradeoff). *Fix constants $0 \leq a < b$. Let q be prime and let $s_e \geq 1$ satisfy*

$$\ln(2 + s_e) = (\ln n)^{a+o(1)}, \quad \ln\left(\frac{q}{\sqrt{n}}\right) = (\ln n)^{b+o(1)}.$$

For every secret distribution over $\mathbb{Z}_q^n$, Algorithm 4 distinguishes decision-LWE with independent errors from $D_{\mathbb{Z}, s_e}$ from uniform using $m = n + \Theta\left(\frac{n}{(b-a)\ln \ln n}\right)$ original samples. Its success probability is $1 - o(1)$, and its expected time and space complexities are $2^{\left(\frac{1}{2(b-a)} + o(1)\right)n / \ln \ln n}$.

The proof is given in Section B.6.

Corollary 4.1 (Polynomial-Modulus Specialization). *Fix constants $\kappa > \frac{1}{2}$ and $0 \leq a < 1$. Let $q = n^{\kappa+o(1)}$ be prime and suppose that $\ln(2 + s_e) = (\ln n)^{a+o(1)}$, $s_e \geq 1$. For every secret distribution over $\mathbb{Z}_q^n$, Algorithm 4 distinguishes decision-LWE with independent errors from $D_{\mathbb{Z}, s_e}$ from uniform using $m = n + \Theta\left(\frac{n}{(1-a)\ln \ln n}\right)$ original samples, with success probability $1 - o(1)$ and expected time and space $2^{\left(\frac{1}{2(1-a)} + o(1)\right)n / \ln \ln n}$. In particular, if $\ln(2 + s_e) = (\ln n)^{o(1)}$, then the expected time and space are $2^{(1/2+o(1))n / \ln \ln n}$.*

Corollary 4.2 (Polynomial Error with Quasipolynomial Modulus). *Fix constants $\gamma > 0$ and $\eta > 0$. Let q be prime and suppose that $s_e = n^{\gamma+o(1)}$ and $q = \exp\left((\ln n)^{1+\eta+o(1)}\right)$. For every secret distribution over $\mathbb{Z}_q^n$, Algorithm 4 distinguishes decision-LWE with independent errors from $D_{\mathbb{Z}, s_e}$ from uniform using $m = n + \Theta\left(\frac{n}{\eta \ln \ln n}\right)$ original samples, with success probability $1 - o(1)$ and expected time and space $2^{\left(\frac{1}{2\eta} + o(1)\right)n / \ln \ln n}$.*

Use of Generative AI. During the preparation of this manuscript, the authors used OpenAI GPT-5.6 Sol for exploratory discussion concerning the choice of discrete Gaussian sampling algorithm and for clarifying the presentation of proof ideas. All claims, proofs, and technical arguments in the final manuscript were developed and independently verified by the authors, who take full responsibility for the content.

References

- ACF⁺15. Martin R. Albrecht, Carlos Cid, Jean-Charles Faugère, Robert Fitzpatrick, and Ludovic Perret. On the complexity of the BKW algorithm on LWE. *DCC*, 74(2):325–354, 2015. doi:10.1007/s10623-013-9864-x.
- ACFP14. Martin R. Albrecht, Carlos Cid, Jean-Charles Faugère, and Ludovic Perret. Algebraic algorithms for LWE. Cryptology ePrint Archive, Report 2014/1018, 2014. URL: <https://eprint.iacr.org/2014/1018>.
- ADS15. Divesh Aggarwal, Daniel Dadush, and Noah Stephens-Davidowitz. Solving the closest vector problem in 2^n time - the discrete Gaussian strikes again! In Venkatesan Guruswami, editor, *56th FOCS*, pages 563–582. IEEE Computer Society Press, October 2015. doi: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 Hugo Krawczyk, editor, *PKC 2014*, volume 8383 of *LNCS*, pages 429–445. Springer, Berlin, Heidelberg, March 2014. doi:10.1007/978-3-642-54631-0_25.
- AG11. Sanjeev Arora and Rong Ge. New algorithms for learning in presence of errors. In Luca Aceto, Monika Henzinger, and Jiri Sgall, editors, *ICALP 2011, Part I*, volume 6755 of *LNCS*, pages 403–415. Springer, Berlin, Heidelberg, July 2011. doi:10.1007/978-3-642-22006-7_34.
- AR05. Dorit Aharonov and Oded Regev. Lattice problems in $\text{np} \cap \text{conp}$. *J. ACM*, 52(5):749–765, September 2005. doi:10.1145/1089023.1089025.
- Ban93. Wojciech Banaszczyk. New bounds in some transference theorems in the geometry of numbers. *Mathematische Annalen*, 296(1):625–635, 1993.
- BGG⁺14. Dan Boneh, Craig Gentry, Sergey Gorbunov, Shai Halevi, Valeria Nikolaenko, Gil Segev, Vinod Vaikuntanathan, and Dhinakaran Vinayagamurthy. Fully key-homomorphic encryption, arithmetic circuit ABE and compact garbled circuits. In Phong Q. Nguyen and Elisabeth Oswald, editors, *EUROCRYPT 2014*, volume 8441 of *LNCS*, pages 533–556. Springer, Berlin, Heidelberg, May 2014. doi:10.1007/978-3-642-55220-5_30.
- BGV12. Zvika Brakerski, Craig Gentry, and Vinod Vaikuntanathan. (Leveled) fully homomorphic encryption without bootstrapping. In Shafi Goldwasser, editor, *ITCS 2012*, pages 309–325. ACM, January 2012. doi:10.1145/2090236.2090262.
- BKW00. Avrim Blum, Adam Kalai, and Hal Wasserman. Noise-tolerant learning, the parity problem, and the statistical query model. In *32nd ACM STOC*, pages 435–440. ACM Press, May 2000. doi:10.1145/335305.335355.
- BLP⁺13. Zvika Brakerski, Adeline Langlois, Chris Peikert, Oded Regev, and Damien Stehlé. Classical hardness of learning with errors. In Dan Boneh, Tim Roughgarden, and Joan Feigenbaum, editors, *45th ACM STOC*, pages 575–584. ACM Press, June 2013. doi:10.1145/2488608.2488680.
- CKKS17. Jung Hee Cheon, Andrey Kim, Miran Kim, and Yong Soo Song. Homomorphic encryption for arithmetic of approximate numbers. In Tsuyoshi Takagi and Thomas Peyrin, editors, *ASIACRYPT 2017, Part I*, volume 10624 of *LNCS*, pages 409–437. Springer, Cham, December 2017. doi:10.1007/978-3-319-70694-8_15.
- DEL25. Léo Ducas, Lynn Engelberts, and Johanna Loyer. Wagner’s algorithm provably runs in subexponential time for rmSIS^∞ . In Yael Tauman Kalai and Seny F. Kamara, editors, *CRYPTO 2025, Part I*, volume 16000 of *LNCS*, pages 353–384. Springer, Cham, August 2025. doi:10.1007/978-3-032-01855-7_12.

- DTV15. Alexandre Duc, Florian Tramèr, and Serge Vaudenay. Better algorithms for LWE and LWR. In Elisabeth Oswald and Marc Fischlin, editors, *EUROCRYPT 2015, Part I*, volume 9056 of *LNCS*, pages 173–202. Springer, Berlin, Heidelberg, April 2015. doi:10.1007/978-3-662-46800-5_8.
- FV12. Junfeng Fan and Frederik Vercauteren. Somewhat practical fully homomorphic encryption. Cryptology ePrint Archive, Report 2012/144, 2012. URL: <https://eprint.iacr.org/2012/144>.
- 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. URL: <https://eprint.iacr.org/2026/1587>.
- GJMS17. Qian Guo, Thomas Johansson, Erik Mårtensson, and Paul Stankovski. Coded-BKW with sieving. In Tsuyoshi Takagi and Thomas Peyrin, editors, *ASIACRYPT 2017, Part I*, volume 10624 of *LNCS*, pages 323–346. Springer, Cham, December 2017. doi:10.1007/978-3-319-70694-8_12.
- GJMW19. Qian Guo, Thomas Johansson, Erik Mårtensson, and Paul Stankovski Wagner. On the asymptotics of solving the lwe problem using coded-bkw with sieving. *IEEE Transactions on Information Theory*, 65(8):5243–5259, 2019.
- GJS15. Qian Guo, Thomas Johansson, and Paul Stankovski. Coded-BKW: Solving LWE using lattice codes. In Rosario Gennaro and Matthew J. B. Robshaw, editors, *CRYPTO 2015, Part I*, volume 9215 of *LNCS*, pages 23–42. Springer, Berlin, Heidelberg, August 2015. doi:10.1007/978-3-662-47989-6_2.
- GLN⁺25. Lewis Glabush, Patrick Longa, Michael Naehrig, Chris Peikert, Douglas Stebila, and Fernando Virdia. FrodoKEM: A CCA-secure learning with errors key encapsulation mechanism. *CiC*, 2(3):25, 2025. doi:10.62056/ayivom2hd.
- GSW13. Craig Gentry, Amit Sahai, and Brent Waters. Homomorphic encryption from learning with errors: Conceptually-simpler, asymptotically-faster, attribute-based. In Ran Canetti and Juan A. Garay, editors, *CRYPTO 2013, Part I*, volume 8042 of *LNCS*, pages 75–92. Springer, Berlin, Heidelberg, August 2013. doi:10.1007/978-3-642-40041-4_5.
- GVW13. Sergey Gorbunov, Vinod Vaikuntanathan, and Hoeteck Wee. Attribute-based encryption for circuits. In Dan Boneh, Tim Roughgarden, and Joan Feigenbaum, editors, *45th ACM STOC*, pages 545–554. ACM Press, June 2013. doi:10.1145/2488608.2488677.
- HKM18. Gottfried Herold, Elena Kirshanova, and Alexander May. On the asymptotic complexity of solving LWE. *DCC*, 86(1):55–83, 2018. doi: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 Rosario Gennaro and Matthew J. B. Robshaw, editors, *CRYPTO 2015, Part I*, volume 9215 of *LNCS*, pages 43–62. Springer, Berlin, Heidelberg, August 2015. doi:10.1007/978-3-662-47989-6_3.
- KNSW20. Elena Kirshanova, Huyen Nguyen, Damien Stehlé, and Alexandre Wallet. On the smoothing parameter and last minimum of random orthogonal lattices. *DCC*, 88(5):931–950, 2020. doi:10.1007/s10623-020-00719-w.
- LW21. Thijs Laarhoven and Michael Walter. Dual lattice attacks for closest vector problems (with preprocessing). In Kenneth G. Paterson, editor, *CT-RSA 2021*, volume 12704 of *LNCS*, pages 478–502. Springer, Cham, May 2021. doi:10.1007/978-3-030-75539-3_20.

- Mic11. Daniele Micciancio. The geometry of lattice cryptography. In Alessandro Aldini and Roberto Gorrieri, editors, *Foundations of Security Analysis and Design VI*, volume 6858 of *Lecture Notes in Computer Science*, pages 185–210. Springer, 2011. doi:[10.1007/978-3-642-23082-0_7](https://doi.org/10.1007/978-3-642-23082-0_7).
- MM11. Daniele Micciancio and Petros Mol. Pseudorandom knapsacks and the sample complexity of LWE search-to-decision reductions. In Phillip Rogaway, editor, *CRYPTO 2011*, volume 6841 of *LNCS*, pages 465–484. Springer, Berlin, Heidelberg, August 2011. doi:[10.1007/978-3-642-22792-9_26](https://doi.org/10.1007/978-3-642-22792-9_26).
- MR04. Daniele Micciancio and Oded Regev. Worst-case to average-case reductions based on Gaussian measures. In *45th FOCS*, pages 372–381. IEEE Computer Society Press, October 2004. doi:[10.1109/FOCS.2004.72](https://doi.org/10.1109/FOCS.2004.72).
- MU05. Michael Mitzenmacher and Eli Upfal. *Probability and Computing: Randomized Algorithms and Probabilistic Analysis*. Cambridge University Press, 2005. doi:[10.1017/CB09780511813603](https://doi.org/10.1017/CB09780511813603).
- Nat24. National Institute of Standards and Technology. Module-lattice-based key-encapsulation mechanism standard. Technical Report FIPS 203, August 2024. doi:[10.6028/NIST.FIPS.203](https://doi.org/10.6028/NIST.FIPS.203).
- PRS17. Chris Peikert, Oded Regev, and Noah Stephens-Davidowitz. Pseudorandomness of ring-LWE for any ring and modulus. In Hamed Hatami, Pierre McKenzie, and Valerie King, editors, *49th ACM STOC*, pages 461–473. ACM Press, June 2017. doi:[10.1145/3055399.3055489](https://doi.org/10.1145/3055399.3055489).
- PS24. Amaury Pouly and Yixin Shen. Provable dual attacks on learning with errors. In Marc Joye and Gregor Leander, editors, *EUROCRYPT 2024, Part VII*, volume 14657 of *LNCS*, pages 256–285. Springer, Cham, May 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 Harold N. Gabow and Ronald Fagin, editors, *37th ACM STOC*, pages 84–93. ACM Press, May 2005. doi:[10.1145/1060590.1060603](https://doi.org/10.1145/1060590.1060603).
- SAB⁺20. Peter Schwabe, Roberto Avanzi, Joppe Bos, Léo Ducas, Eike Kiltz, Tancrede Lepoint, Vadim Lyubashevsky, John M. Schanck, Gregor Seiler, and Damien Stehlé. CRYSTALS-KYBER. Technical report, National Institute of Standards and Technology, 2020. available at <https://csrc.nist.gov/projects/post-quantum-cryptography/post-quantum-cryptography-standardization/round-3-submissions>.
- STA20. Chao Sun, Mehdi Tibouchi, and Masayuki Abe. Revisiting the hardness of binary error LWE. In Joseph K. Liu and Hui Cui, editors, *ACISP 20*, volume 12248 of *LNCS*, pages 425–444. Springer, Cham, November / December 2020. doi:[10.1007/978-3-030-55304-3_22](https://doi.org/10.1007/978-3-030-55304-3_22).
- Ste24. Matthias Johann Steiner. The complexity of algebraic algorithms for LWE. In Marc Joye and Gregor Leander, editors, *EUROCRYPT 2024, Part III*, volume 14653 of *LNCS*, pages 375–403. Springer, Cham, May 2024. doi:[10.1007/978-3-031-58734-4_13](https://doi.org/10.1007/978-3-031-58734-4_13).
- WBLW25. Yu Wei, Lei Bi, Xianhui Lu, and Kunpeng Wang. Memory-efficient BKW algorithm for solving the LWE problem. In Tibor Jager and Jiaxin Pan, editors, *PKC 2025, Part II*, volume 15675 of *LNCS*, pages 331–362. Springer, Cham, May 2025. doi:[10.1007/978-3-031-91823-0_11](https://doi.org/10.1007/978-3-031-91823-0_11).

A Asymptotic Implications for LWE-Based Cryptosystems

This section gives an asymptotic interpretation of our decision-LWE algorithm, rather than a concrete security assessment.

Recovering a secret corresponds to search-LWE. Sample-preserving search-to-decision reductions are known in several parameter regimes [MM11]. We therefore conditionally assume that such a reduction applies to the relevant secret and error distributions with polynomial overhead and without changing the asymptotic modulus, error width, or sample count. Under this assumption, Theorem 4.2 also gives a search-LWE algorithm with the same leading subexponential exponent.

Unstructured LWE. We denote the secret dimension and sample count by n and m . In the polynomial-modulus regime, our algorithm requires $m = n + \Theta(\frac{n}{\ln \ln n})$.

Regev’s encryption scheme supplies $\Theta(n \log q)$ samples [Reg05], so its sample count is sufficient. Its polynomial modulus and $\sqrt{n}$ -scale error, however, lie outside the range of Corollary 4.1. The original GSW construction and several related ABE schemes likewise use $\Theta(n \log q)$ LWE samples [GSW13, GVW13, BGG⁺14]; their sample counts are therefore compatible with our algorithm whenever the remaining modulus and error conditions hold.

FrodoKEM lies on the opposite side of the sample requirement. A public-key LWE instance has $m = n$, while the largest ciphertext-defined instance has $m = n + \bar{n}$ samples [GLN⁺25, Section 7.2]. Since all current parameter sets use $\bar{n} = 8$, $m - n$ is constant rather than $\Theta(n / \ln \ln n)$.

Structured LWE: Heuristic Implications and Limitations. Our theorem does not directly apply to Module-LWE or Ring-LWE. We nevertheless consider whether its sample and modulus–error requirements are compatible with representative structured schemes under the heuristic that coefficient-embedded relations can be treated as unstructured LWE samples with comparable error magnitude. This heuristic ignores the algebraic dependencies induced by module or ring multiplication.

For ML-KEM [Nat24], let k be the module rank and n the ring degree. Viewing the underlying Module-LWE instance as LWE gives scalar dimension $n' = kn$ and at most $m' = (k + 1)n$ scalar relations [SAB⁺20, Section 5.1.1]. Hence $m' - n' = n$, which for fixed k exceeds the $\Theta(n' / \ln \ln n')$ surplus required by our theorem. The error coefficients are also very small. The modulus scale is more subtle. ML-KEM chooses q subject to the NTT condition $n \mid (q - 1)$ and the required decryption-failure probability. Thus neither the sample count nor the modulus–error scale alone clearly rules out an asymptotic extension of our approach.

The main obstacle for ML-KEM is structural. After coefficient embedding, the scalar equations remain strongly correlated, whereas our sampler assumes independent uniform public vectors. The additional ciphertext relation also contains mixed error terms, message encoding, and compression effects. Consequently, the preceding parameter comparison does not imply an attack on ML-KEM; a

rigorous application would require a separate analysis of these dependencies and of the resulting error distribution.

RLWE-based homomorphic encryption gives another potentially relevant regime. In BGV, BFV, and particularly CKKS, fresh RLWE samples use small coefficient errors, while homomorphic depth is supported by comparatively large ciphertext moduli or modulus chains [BGV12,FV12,CKKS17]. This is qualitatively the direction favored by our modulus–error tradeoff: a large modulus can accommodate a substantially larger error scale, while very small initial errors fall on the favorable side of Theorem 4.2.

The basic public key contains only a constant number of Ring-LWE relations and therefore does not provide the required sample surplus. Evaluation, relinearization, rotation, and key-switching keys, however, may contain many additional RLWE-style encryptions. These are potentially more relevant to our sample requirement, but they involve ring dependencies, gadget decompositions, related or transformed secrets, and typically composite RNS moduli. Determining whether such evaluation material can be converted into sufficiently many usable relations is therefore a natural question for extending our analysis to structured LWE.

B Deferred Proofs and Parameter Calculations

This appendix collects the proofs and parameter calculations deferred from the main text, in the order in which they first arise.

B.1 Minimum of a Random Orthogonal q -ary Lattice

Proof (of Lemma 2.1). Set $R := \sqrt{m} q^{n/m}/16$. We show that, with probability at least $1 - 2^{-m}$, no nonzero integer vector of Euclidean norm smaller than R belongs to $\Lambda_q^\perp(\mathbf{A})$.

The assumption $\sqrt{m} q^{n/m} \leq q$ gives $R < q$. Hence every nonzero $\mathbf{x} \in \mathbb{Z}^m$ with $\|\mathbf{x}\|_2 < R$ remains nonzero modulo q . For such a fixed $\mathbf{x}$, each row of the uniform matrix $\mathbf{A} \leftarrow \mathbb{F}_q^{n \times m}$ has inner product zero with $\mathbf{x}$ modulo q with probability $1/q$, independently across the n rows. Therefore $\Pr_{\mathbf{A}}[\mathbf{x} \in \Lambda_q^\perp(\mathbf{A})] = q^{-n}$.

It remains to bound the number of candidate vectors. The unit cubes centered at the points of $\mathbb{Z}^m \cap RB_2^m$ have disjoint interiors and are contained in $(R + \sqrt{m}/2)B_2^m$. Consequently,

$$|\mathbb{Z}^m \cap RB_2^m| \leq V_m \left(R + \frac{\sqrt{m}}{2} \right)^m \leq \left[\sqrt{2\pi e} \left(\frac{q^{n/m}}{16} + \frac{1}{2} \right) \right]^m,$$

where the second inequality uses $V_m \leq (2\pi e/m)^{m/2}$.

A union bound over all nonzero candidate vectors now gives

$$\begin{aligned} \Pr\left[\lambda_1\left(\Lambda_q^\perp(\mathbf{A})\right) < R\right] &\leq q^{-n}|\mathbb{Z}^m \cap RB_2^m| \\ &\leq \left[\sqrt{2\pi e}\left(\frac{1}{16} + \frac{1}{2q^{n/m}}\right)\right]^m. \end{aligned}$$

Since $q^{n/m} \geq 16$, the quantity in brackets is at most $3\sqrt{2\pi e}/32 < 1/2$. The failure probability is therefore smaller than 2^{-m} , as claimed. $\square$

B.2 Smoothing of a Random Orthogonal q -ary Lattice

Proof (Proof of Lemma 2.6). Set

$$t := 8 \max \left\{ \sqrt{\ln\left(\frac{2m}{\varepsilon}\right)}, q^{n/m} \varepsilon^{-1/m} \right\}.$$

By assumption, $t \leq q$. To prove $\eta_\varepsilon(\Lambda_q^\perp(\mathbf{A})) \leq t$, it is enough, by the definition of the smoothing parameter, to show that $\rho_{1/t}(\Lambda_q^\perp(\mathbf{A})^* \setminus \{\mathbf{0}\}) \leq \varepsilon$.

The dual lattice has the representation $\Lambda_q^\perp(\mathbf{A})^* = \mathbb{Z}^m + \frac{1}{q}\mathbf{A}^T \mathbb{F}_q^n$. Its nonzero Gaussian mass has two qualitatively different contributions. The coset $\mathbb{Z}^m$ is present for every $\mathbf{A}$, whereas all other cosets depend on the random matrix. We bound the fixed integer contribution and the random nontrivial-coset contribution separately. The two terms in the definition of t are chosen precisely to control these two contributions.

The Integer Contribution. We first bound $\rho_{1/t}(\mathbb{Z}^m \setminus \{\mathbf{0}\})$.

Since $t \geq 8\sqrt{\ln(2m/\varepsilon)} > 1$,

$$\rho_{1/t}(\mathbb{Z}) - 1 = 2 \sum_{k \geq 1} e^{-\pi t^2 k^2} \leq 4e^{-\pi t^2}.$$

Moreover, $t^2 \geq 64 \ln(2m/\varepsilon)$, and hence

$$4me^{-\pi t^2} \leq 4m \left(\frac{\varepsilon}{2m}\right)^2 = \frac{\varepsilon^2}{m} \leq \frac{\varepsilon}{4},$$

where the last inequality uses $m \geq 2$ and $\varepsilon \leq 1/2$. Therefore

$$\rho_{1/t}(\mathbb{Z}^m \setminus \{\mathbf{0}\}) = \rho_{1/t}(\mathbb{Z})^m - 1 \leq \exp(4me^{-\pi t^2}) - 1 \leq \frac{\varepsilon}{2},$$

where the last step uses $e^v - 1 \leq 2v$ for $0 \leq v \leq 1$. Thus the fixed integer part contributes at most half of the allowed smoothing error.

The Random Coset Contribution. We next control the Gaussian mass outside $\mathbb{Z}^m$. For $\mathbf{z} \in \mathbb{F}_q^n \setminus \{\mathbf{0}\}$, define the coset $\mathcal{C}_{\mathbf{z}} := \mathbb{Z}^m + \frac{1}{q}\mathbf{A}^T\mathbf{z}$.

Because q is prime and $\mathbf{z} \neq \mathbf{0}$, $\mathbf{A}^T\mathbf{z}$ is uniform over $\mathbb{F}_q^m$. Therefore

$$\begin{aligned}\mathbb{E}_{\mathbf{A}}[\rho_{1/t}(\mathcal{C}_{\mathbf{z}})] &= \frac{1}{q^m} \sum_{\mathbf{v} \in \mathbb{F}_q^m} \rho_{1/t}\left(\mathbb{Z}^m + \frac{\mathbf{v}}{q}\right) \\ &= q^{-m} \rho_{q/t}(\mathbb{Z})^m.\end{aligned}$$

The last equality follows because the q^m cosets $\mathbb{Z}^m + \mathbf{v}/q$ partition $q^{-1}\mathbb{Z}^m$.

The nontrivial cosets of $\Lambda_q^\perp(\mathbf{A})^*$ form a subset of the cosets represented by $\mathcal{C}_{\mathbf{z}}$ for $\mathbf{z} \neq \mathbf{0}$. Different values of $\mathbf{z}$ may represent the same coset, so summing over all $\mathbf{z} \neq \mathbf{0}$ can only overcount their Gaussian mass. Hence

$$\mathbb{E}_{\mathbf{A}} \rho_{1/t}\left(\Lambda_q^\perp(\mathbf{A})^* \setminus \mathbb{Z}^m\right) \leq q^{-(m-n)} \rho_{q/t}(\mathbb{Z})^m.$$

Since $t \leq q$, we have $q/t \geq 1$. For every $s \geq 1$, $\rho_s(\mathbb{Z}) \leq 1 + s \leq 2s$, and therefore

$$\mathbb{E}_{\mathbf{A}} \rho_{1/t}\left(\Lambda_q^\perp(\mathbf{A})^* \setminus \mathbb{Z}^m\right) \leq 2^m q^n t^{-m}.$$

The second term in the definition of t gives $t \geq 8q^{n/m}\varepsilon^{-1/m}$, so the right-hand side is at most $\varepsilon 4^{-m}$. By Markov's inequality,

$$\Pr_{\mathbf{A}}\left[\rho_{1/t}\left(\Lambda_q^\perp(\mathbf{A})^* \setminus \mathbb{Z}^m\right) > \frac{\varepsilon}{2}\right] \leq 2 \cdot 4^{-m} \leq 2^{-m}.$$

Outside this event, the integer contribution and the nontrivial-coset contribution are each at most $\varepsilon/2$. Hence $\rho_{1/t}\left(\Lambda_q^\perp(\mathbf{A})^* \setminus \{\mathbf{0}\}\right) \leq \varepsilon$. By the definition of the smoothing parameter, $\eta_\varepsilon(\Lambda_q^\perp(\mathbf{A})) \leq t$, which proves the claim. $\square$

B.3 Algebraic Properties of the Random P -ary Extension

Proof (Proof of Lemma 3.1). We first determine the kernel of the projection. Since Γ_i is supported on the newest b_i coordinates, $\pi_i(\Gamma_i) = \{\mathbf{0}\}$. Hence $\pi_i(\Lambda'_i) = \pi_i(\Lambda_i) = \Lambda_{i-1}$, so the projection from Λ'_i is surjective.

Let $\mathbf{x} \in \Lambda'_i$ satisfy $\pi_i(\mathbf{x}) = \mathbf{0}$. By the definition of Λ'_i , write $\mathbf{x} = \mathbf{v} + \mathbf{z}$, where $\mathbf{v} \in \Lambda_i$ and $\mathbf{z} \in \Gamma_i$. Since $\pi_i(\mathbf{z}) = \mathbf{0}$, we have $\pi_i(\mathbf{v}) = \mathbf{0}$. Therefore $\mathbf{v} \in \ker(\pi_i|_{\Lambda_i}) = q\mathbb{Z}^{b_i} \subseteq \Gamma_i$. It follows that $\mathbf{x} \in \Gamma_i$, which proves $\ker(\pi_i) = \Gamma_i$ and gives

$$0 \longrightarrow \Gamma_i \longrightarrow \Lambda'_i \xrightarrow{\pi_i} \Lambda_{i-1} \longrightarrow 0.$$

Since the quotient $\Lambda'_i/\Gamma_i \cong \Lambda_{i-1}$ is torsion free, Γ_i is primitive in Λ'_i .

We next identify the finite quotient. The second isomorphism theorem gives

$$\Lambda'_i/\Lambda_i = (\Lambda_i + \Gamma_i)/\Lambda_i \cong \Gamma_i/(\Gamma_i \cap \Lambda_i).$$

A vector in $\Gamma_i \cap \Lambda_i$ is supported on the newest coordinate block and belongs to Λ_i . It therefore lies in $\ker(\pi_i|_{\Lambda_i}) = q\mathbb{Z}^{b_i}$. The reverse inclusion is immediate, so $\Gamma_i \cap \Lambda_i = q\mathbb{Z}^{b_i}$. Consequently, $\Lambda'_i/\Lambda_i \cong \Gamma_i/q\mathbb{Z}^{b_i}$. The quotient on the right is generated by the class of $\frac{q}{P}\mathbf{a}_i$. This class has order P : multiplication by P gives $q\mathbf{a}_i \in q\mathbb{Z}^{b_i}$, whereas an integer $1 \leq k < P$ cannot annihilate it because $k\mathbf{a}_i \neq \mathbf{0}$ in $\mathbb{F}_P^{b_i}$. Thus $\Gamma_i/q\mathbb{Z}^{b_i} \cong \mathbb{Z}/P\mathbb{Z}$ which proves the claim. $\square$

B.4 Adaptive Composition of Shifted DGS Errors

Proof (Proof of Lemma 3.3). For $j = 0, \dots, Q$, let $\mathcal{E}_j$ denote the law of the hybrid computation in which the first j calls use their target distributions and the remaining calls use the implemented sampler. Thus $\mathcal{E}_0 = \text{Law}_{\text{impl}}$ and $\mathcal{E}_Q = \text{Law}_{\text{ideal}}$. The adjacent hybrids $\mathcal{E}_{j-1}$ and $\mathcal{E}_j$ induce the same distribution on $\mathbf{H}_{j-1}$. Conditioned on $\mathbf{H}_{j-1} = h$, they differ only in the j th call, after which they apply the same adaptive randomized computation. Statistical distance does not increase under randomized post-processing, so $\text{SD}(\mathcal{E}_{j-1}, \mathcal{E}_j) \leq \xi_j$. The triangle inequality now gives

$$\text{SD}(\mathcal{E}_0, \mathcal{E}_Q) \leq \sum_{j=1}^Q \text{SD}(\mathcal{E}_{j-1}, \mathcal{E}_j) \leq \sum_{j=1}^Q \xi_j.$$

$\square$

B.5 Block Capacities for a Prescribed List Size

This subsection proves the block-allocation result used in Theorem 3.2. The list size N is prescribed externally and remains fixed throughout the argument.

At each level i , the smoothing requirement for the random P -ary auxiliary lattice imposes an upper bound on the admissible block size b_i . We refer to the largest block size guaranteed by this requirement as the *block capacity* of that level. The proof first derives these per-level capacities, then shows that their sum is at least n , and finally chooses actual block sizes that cover all n rows while retaining the lower, upper, and prefix bounds needed later in the sampler analysis.

Lemma B.1 (Block Allocation for a Prescribed List Size). *Fix the asymptotic parameter family and use $\varepsilon_0, r, t_i, x_i$ from the proof of Theorem 3.2. Let $N \geq 2$ be a power of two satisfying Equation (2) and*

$$\frac{(\ln q)^3}{\ln N} = o(1), \quad \ln N = o(n).$$

If the constant C_{sam} in Equation (2) is sufficiently large, then, for all sufficiently large n , there exist a prime $P \in (N/2, N)$ and positive integers $b_1, \dots, b_r$ with $\sum_{i=1}^r b_i = n$ such that, for every $i \in [r]$,

$$P^{1/b_i} \geq 16, \quad \sqrt{b_i} P^{1/b_i} \leq P, \quad \frac{16q}{P^{1/b_i}} \sqrt{\max \left\{ 1, \frac{\ln(1/\varepsilon_0)}{b_i} \right\}} \leq t_i.$$

Consequently, the hypotheses of Lemma 3.2 hold at every level, and the resulting smoothing upper bound is at most t_i .

Writing $n_i := \sum_{j \leq i} b_j$, $b_{\min} := \min_{i \in [r]} b_i$, and $b_{\max} := \max_{i \in [r]} b_i$, the same allocation satisfies

$$b_{\min} = \Omega\left(\frac{\ln N}{\ln q}\right), \quad b_{\max} = O\left(\frac{\ln N}{1 + \ln f}\right), \quad n_i x_i = O(\ln N \ln t_i)$$

for all $i \in [r]$.

Proof. The proof has three steps. We first derive a per-level block capacity from the auxiliary-lattice smoothing condition. We then show that the sum of these capacities is at least n . Finally, we choose the prime P and allocate the n rows while ensuring the required block-size bounds hold for every block.

Put $M := N/2$, let $L_0 := \ln(2m/\varepsilon_0)$, and write $\ell := \ln(1/\varepsilon_0)$. The quantity M will serve as a deterministic lower bound on the prime P , which will be chosen in $(M, 2M)$ after the per-level capacities have been determined.

We first record the scale of the smoothing-accuracy term. Since $W \leq \ln \ln q$, the block-capacity condition Equation (2) gives $\ln N = \Omega(n/\ln \ln q)$. Moreover, $L_0 = \ln(2m/\varepsilon) + \ln 5 = O(n/\ln \ln q)$, and hence $\beta_M := \ell/\ln M = O(1)$ along the fixed parameter family. The parameter β_M measures the smoothing-accuracy cost $\ln(1/\varepsilon_0)$ relative to the logarithmic list size.

From the Smoothing Constraint to Per-Level Capacities. Fix a level i . Our first goal is to determine a range of block sizes that automatically satisfies the auxiliary-lattice smoothing condition. For a candidate block size b , introduce the normalized inverse block size $y := \ln M/b$. This normalization is useful because the two b -dependent terms in the smoothing bound become $M^{-1/b} = e^{-y}$ and $\ell/b = \beta_M y$.

For every prime $P \geq M$, we therefore have

$$\frac{16q}{P^{1/b}} \sqrt{\max\left\{1, \frac{\ell}{b}\right\}} \leq 16qe^{-y} \sqrt{1 + \beta_M y}.$$

Since $x_i = \ln(q/t_i)$, equivalently $t_i = qe^{-x_i}$, it is sufficient that

$$y - \frac{1}{2} \ln(1 + \beta_M y) \geq x_i + \ln 16.$$

We define the *threshold inverse block size* y_i to be the unique solution

$$y_i = x_i + \frac{1}{2} \ln(1 + \beta_M y_i) + \ln 16.$$

To justify this definition, the function $y \mapsto y - \frac{1}{2} \ln(1 + \beta_M y)$ is strictly increasing for $y > 1/2$. Since $x_i \geq x_r = \ln(2f) > \ln 2$, the defining equation has a unique solution, and in fact $y_i > x_i + \ln 16$.

The corresponding *block capacity* is $B_i := \lfloor \ln M / y_i \rfloor$. Thus B_i is the largest block size, up to integer rounding, for which the preceding sufficient smoothing condition is guaranteed by the threshold y_i . Indeed, if $1 \leq b \leq B_i$, then $y = \ln M / b \geq y_i$, and monotonicity of $y - \frac{1}{2} \ln(1 + \beta_M y)$ gives $16qe^{-y} \sqrt{1 + \beta_M y} \leq qe^{-x_i} = t_i$. Hence the main smoothing inequality holds for every $b \leq B_i$.

The other two hypotheses of Lemma 3.2 also hold throughout this range. Since $P \geq M$ and $y \geq y_i > \ln 16$, we have $P^{1/b} \geq M^{1/b} = e^y > 16$. For $b = 1$, $\sqrt{b} P^{1/b} \leq P$ is immediate. If $b \geq 2$, then $P^{1/b} \leq \sqrt{P}$. Moreover, $b \leq B_i \leq \ln M / \ln 16 < M \leq P$ for all sufficiently large n , and therefore $\sqrt{b} P^{1/b} \leq \sqrt{b} \sqrt{P} \leq P$. Thus every positive integer $b \leq B_i$ is admissible at level i .

For later use, we also record the asymptotic size of the threshold. Because $\beta_M = O(1)$, its defining equation first gives $y_i = O(1 + x_i)$, and substitution back into the logarithmic term yields $y_i = x_i + O(\ln(1 + x_i))$ for all $i \in [r]$.

Total Block Capacity. We now show that the sum of the per-level capacities is large enough to accommodate all n rows. Let $S := \sum_{i=1}^r B_i$. By the definition of B_i ,

$$S = \ln M \sum_{i=1}^r \frac{1}{y_i} + O(r).$$

The preceding estimate for y_i , together with $y_i \geq x_i > \ln 2$, gives

$$\frac{1}{y_i} = \frac{1}{x_i} + O\left(\frac{\ln(1 + x_i)}{x_i^2}\right).$$

The values x_i form a decreasing arithmetic progression with spacing $\ln 2/2$. Since $x \mapsto \ln(1 + x)/x^2$ is integrable on $[\ln 2, \infty)$, summing the error term contributes only $O(1)$. A sum–integral comparison for $1/x$ therefore gives

$$\sum_{i=1}^r \frac{1}{y_i} = \sum_{i=1}^r \frac{1}{x_i} + O(1) = \frac{2}{\ln 2} \ln \frac{x_1}{x_r} + O(1).$$

It remains to identify the two endpoints. From Equation (3), $t_1 = \Theta(\sqrt{L_0})$ and $t_r = q/(2f)$. Hence $x_1 = \ln\left(q/\sqrt{\ln(2m/\varepsilon)}\right) + O(1)$ and $x_r = \ln(2f)$. Here replacing $L_0 = \ln(2m/\varepsilon_0)$ by $\ln(2m/\varepsilon)$ changes only the bounded term. Since $\ln x_r = \ln(1 + \ln f) + O(1)$, the quantity W from Equation (1) satisfies $\ln \frac{x_1}{x_r} = W + O(1)$. Thus W measures, up to an additive constant, the logarithmic width of the range of inverse block sizes available across the recursion.

Substituting the endpoint estimate into the capacity sum gives

$$S = 2 \log_2 M (W + O(1)) + O(r).$$

The growth condition $(\ln q)^3 / \ln N = o(1)$ implies $r = O(\ln q) = o(\ln N)$ and also $\ln \ln q = o(\ln N)$. Since $W \leq \ln \ln q$ and $\log_2 M = \log_2 N - 1$, replacing M

by N changes the preceding estimate by $o(\ln N)$. Hence, after choosing C_{sam} sufficiently large to absorb the bounded and lower-order terms,

$$S \geq 2 \log_2 N (W - C_{\text{sam}})$$

for all sufficiently large n . The block-capacity condition Equation (2) therefore gives $S \geq n$.

Block Allocation and Per-Level Size Bounds. The inequality $S \geq n$ shows that the per-level capacities are sufficient in total, but an arbitrary allocation could leave some blocks too small for the shifted-DGS analysis. We therefore impose a common baseline block size before distributing the remaining rows.

Since $N = \omega(1)$, Bertrand's postulate gives, for all sufficiently large n , a prime $P \in (M, 2M) = (N/2, N)$.⁸ This prime satisfies $P \geq M$, so every block size bounded by B_i is admissible by the preceding analysis.

We next lower-bound every B_i . Since $x_i \leq x_1 < \ln q$ and $y_i = x_i + O(\ln(1+x_i))$, we have $y_i = O(\ln q)$ for all i . Together with $\ln N / \ln q = \omega(1)$, which follows from $(\ln q)^3 / \ln N = o(1)$, this gives a constant $c_0 > 0$ such that $B_i \geq c_0 \ln N / \ln q$ for every level and all sufficiently large n .

Define the baseline block size $d := \lfloor c_0 \ln N / (2 \ln q) \rfloor$. Then $d = \Omega(\ln N / \ln q)$ and $1 \leq d \leq B_i$ at every level. Since $r = O(\ln q)$, the total number of rows used by these baseline blocks is $rd = O(\ln N) = o(n)$. Assign d rows to every level. The number of rows still to be assigned is $n - rd$, while the remaining capacity satisfies

$$\sum_{i=1}^r (B_i - d) = S - rd \geq n - rd.$$

We may therefore distribute the remaining rows among levels with unused capacity. This produces positive integers $b_1, \dots, b_r$ satisfying $\sum_{i=1}^r b_i = n$ and $d \leq b_i \leq B_i$ for every i . Since every block size at most B_i is admissible, the three auxiliary-smoothing conditions in the lemma follow immediately. Moreover,

$$b_{\min} = \Omega\left(\frac{\ln N}{\ln q}\right).$$

The upper bound follows from the upper capacity endpoint. Since $y_i \geq x_i \geq x_r = \ln(2f)$, we have $b_i \leq B_i \leq \ln M / x_r$. As $\ln(2f) = \Theta(1 + \ln f)$ for $f > 1$, it follows that

$$b_{\max} = O\left(\frac{\ln N}{1 + \ln f}\right).$$

⁸ The prime can be found deterministically by scanning the integers in $(M, 2M)$ and applying a polynomial-time primality test. Even a full scan costs $N \text{poly}(\ln N)$, which is absorbed by the sampler running time.

Prefix Bound. It remains to establish the prefix estimate used later to control the smoothing of the intermediate orthogonal q -ary lattices Λ_i . For every $j \leq i$, the capacity bound gives $b_j \leq B_j \leq \ln M/x_j$, and therefore

$$n_i \leq \ln M \sum_{j \leq i} \frac{1}{x_j}.$$

Because the x_j form an arithmetic progression with fixed spacing $\ln 2/2$, a sum–integral comparison gives

$$\sum_{j \leq i} \frac{1}{x_j} = O\left(\frac{1}{x_i} + \ln \frac{x_1}{x_i}\right).$$

Multiplying by x_i yields

$$n_i x_i = O\left(\ln N \left(1 + x_i \ln \frac{x_1}{x_i}\right)\right).$$

For $0 < x_i \leq x_1$, the inequality $x_i \ln(x_1/x_i) \leq x_1 - x_i$ follows from $\ln u \leq u - 1$. Moreover, $x_1 - x_i = \ln(t_i/t_1) \leq \ln t_i$. Finally, $t_1 = \Theta(\sqrt{L_0}) = \omega(1)$, so the additive constant is also absorbed by $O(\ln t_i)$. Consequently,

$$n_i x_i = O(\ln N \ln t_i)$$

for all $i \in [r]$, which completes the proof. $\square$

B.6 Proof of the Logarithmic-Scale Modulus–Error Tradeoff

Proof (Proof of Theorem 4.2). We combine the logarithmic-scale sampler of Corollary 3.2 with the Fourier test of Theorem 4.1. The parameter choices below are determined by two requirements. The conditional-similarity error must be a sufficiently small fraction of the target Fourier bias, and the sampler must output enough vectors to resolve that bias.

Matching the Sampler and Fourier Parameters. Let $C > C_{\log}$ be a sufficiently large constant, where $C_{\log}$ is the constant from Corollary 3.2. Set

$$m := n + \left\lceil C \frac{n}{(b-a) \ln \ln n} \right\rceil.$$

Thus $m = (1 + o(1))n$ and $m - n = \Theta(n/((b-a) \ln \ln n))$.

For Gaussian errors, Lemmas 2.4 and 4.1 identify $\exp(-s_e^2 m/f^2)$ as a valid target Fourier bias. We choose f so that this bias remains asymptotically larger than the sampling fluctuation at the list size supplied by Corollary 3.2. Fix a constant $K > 4(b-a)/\ln 2$ and define

$$f_\star^2 := K s_e^2 \ln \ln n, \quad \tau_\star := \exp\left(-\frac{s_e^2 m}{f_\star^2}\right).$$

Then

$$\ln(1/\tau_\star) = \left(\frac{1}{K} + o(1) \right) \frac{n}{\ln \ln n}.$$

The sampler guarantee in Theorem 3.2 has conditional-similarity parameter $q^4 \varepsilon$, while Theorem 4.1 requires this parameter to be at most $\tau_\star/64$. We therefore set $\varepsilon_\star := \tau_\star/(64q^4)$ and $L_\star := \ln(2m/\varepsilon_\star)$. For all sufficiently large n , we have $f_\star > 1$ and $0 < \varepsilon_\star \leq 1/2$.

Applying the Logarithmic-Scale Sampler. We verify the explicit hypotheses of Corollary 3.2. First, $\ln f_\star = (\ln n)^{a+o(1)}$. For $a > 0$, this follows from $\ln(2+s_e) = (\ln n)^{a+o(1)}$, which implies $\ln s_e = (\ln n)^{a+o(1)}$, while the factor $\sqrt{K} \ln \ln n$ is negligible on this scale. When $a = 0$, we have $\ln s_e \leq \ln(2+s_e) = (\ln n)^{o(1)}$, and the same conclusion follows because $\ln \sqrt{K} \ln \ln n = (\ln n)^{o(1)}$. In this case the additional factor also ensures $f_\star = \omega(1)$.

Next, the definition of $\varepsilon_\star$ gives

$$L_\star = \ln(2m) + \ln 64 + 4 \ln q + \ln(1/\tau_\star).$$

The modulus assumption implies $\ln q = (\ln n)^{\max\{1,b\}+o(1)} = o(n/\ln \ln n)$. Hence

$$L_\star = \left(\frac{1}{K} + o(1) \right) \frac{n}{\ln \ln n} = \Theta \left(\frac{n}{\ln \ln n} \right).$$

It follows that

$$\ln \left(\frac{q}{\sqrt{L_\star}} \right) = \ln \left(\frac{q}{\sqrt{n}} \right) + \frac{1}{2} \ln \left(\frac{n}{L_\star} \right) = (\ln n)^{b+o(1)},$$

because $b > 0$ and $\ln(n/L_\star) = O(\ln \ln \ln n)$. Finally, the choice $C > C_{\log}$ gives the required sample surplus. Thus Corollary 3.2 applies and provides a power of two N satisfying

$$\log_2 N = \left(\frac{1}{2(b-a)} + o(1) \right) \frac{n}{\ln \ln n}.$$

By Theorem 3.2, there is a good event of probability $1 - o(1)$, determined by the public matrix and the auxiliary-lattice choices. We condition on this event and on the realized lattices. The systematic-form reduction has succeeded, so the public matrix has full row rank. The remaining sampler randomness produces N vectors in $\Lambda_q^\perp(\mathbf{A})$ whose joint law is within statistical distance $o(1)$ of a distribution conditionally $\tau_\star/64$ -similar to $D_{\Lambda_q^\perp(\mathbf{A}), q/f_\star}^N$. The exceptional event will contribute only $o(1)$ to the final distinguishing error.

Applying the Fourier Test. We now check the remaining hypotheses of Theorem 4.1. The sampler uses only $\mathbf{A}$ and its internal randomness, so its output list is independent of $\mathbf{b}$. The statistical-distance and conditional-similarity requirements were established above.

For the sampling fluctuation, the list-size estimate gives

$$\frac{1}{2} \ln N - \ln(1/\tau_\star) = \left(\frac{\ln 2}{4(b-a)} - \frac{1}{K} + o(1) \right) \frac{n}{\ln \ln n}.$$

The coefficient is positive by the choice of K . Hence $N^{-1/2} = o(\tau_\star)$.

For the uniform-input condition, let $\Delta_\star := (m-n) \ln q - m \ln(2f_\star)$, as in Equation (7). We have

$$(m-n) \ln q = \frac{n(\ln n)^{\max\{1,b\}+o(1)}}{\ln \ln n},$$

whereas $m \ln(2f_\star) = n(\ln n)^{a+o(1)}$ and $\ln(1/\tau_\star) = \Theta(n/\ln \ln n)$. Since $a < b$,

$$m \ln(2f_\star) + \ln(1/\tau_\star) = o((m-n) \ln q).$$

Therefore

$$\Delta_\star - \ln(1/\tau_\star) = (1 - o(1))(m-n) \ln q = \omega(1),$$

which is equivalent to $e^{-\Delta_\star} = o(\tau_\star)$.

Finally, for $\chi = D_{\mathbb{Z}, s_e}$, the definition of $\tau_\star$ and Lemma 2.4 give

$$\nu_\chi(\tau_\star; f_\star, m) \leq (2/e)^{m/2} = o(1).$$

All hypotheses of Theorem 4.1 now hold on the good sampler event. The distinguisher therefore succeeds with conditional probability $1 - o(1)$, and adding the exceptional sampler event preserves overall success probability $1 - o(1)$. The argument is independent of the secret value, so the conclusion holds for every secret distribution over $\mathbb{Z}_q^n$.

Complexity. The logarithmic-scale sampler corollary gives expected time and space

$$2^{\left(\frac{1}{2(b-a)} + O\left(\frac{1}{1+\ln f_\star}\right) + o(1)\right)n/\ln \ln n} \text{poly}(m, \ln q, \ln(1/\varepsilon_\star)).$$

Here $f_\star = \omega(1)$, so the additional exponent is $o(1)$. Moreover, $m = O(n)$, $\ln q = (\ln n)^{O(1)}$, and $\ln(1/\varepsilon_\star) = O(n/\ln \ln n)$. Thus the polynomial factor is $2^{o(n/\ln \ln n)}$. The expected time and space are therefore

$$2^{\left(\frac{1}{2(b-a)} + o(1)\right)n/\ln \ln n}.$$

Together with the chosen value of m , this proves Theorem 4.2. $\square$

C Comparison of Subexponential-Time LWE Algorithms

Table 1 records the parameter regimes that are most directly relevant to the comparison with few original samples and subexponential running time. Several of the cited works contain additional tradeoffs that are not shown separately. Albrecht et

al. [ACFP14] give a more general algebraic family with $m = n(\log n)^{1/\varepsilon}$ samples for fixed $\varepsilon > 0$, while their semi-regularity analysis with only Cn samples, for fixed $C > 1$, remains exponential. Kirchner and Fouque [KF15] likewise give more general BKW tradeoffs for small-secret LWE. Our table records the linear-sample BinaryLWE specialization because it is the regime most directly comparable to our sample complexity.

Steiner [Ste24] derives subexponential complexity estimates not only for BinaryError-LWE, but also for general LWE and binary-secret LWE. These additional regimes require substantially more samples, whereas the BinaryError-LWE specialization is the one yielding a linear-sample subexponential estimate, and is therefore the row reported in the table. For the rigorous sparse-error result of Sun, Tibouchi, and Abe [STA20], the algorithm uses $3n$ original samples; we record this exact linear sample count rather than only $O(n)$. Finally, Herold, Kirshanova, and May [HKM18] also tabulate the classical BKW and BKW_2 baselines that require exponentially many samples. Since our comparison focuses on the regimes of few sample, these baselines with exponential sample are omitted.

Table 1. Comparison of selected asymptotic algorithms for unstructured LWE. Unless stated otherwise, the error column gives the discrete-Gaussian width s_e . Samples count original LWE samples, and the time column gives the leading exponent $\log_2 T$.

Reference	Error	Modulus	Samples	$\log_2 T$	Status
<i>Algebraic approaches</i>					
[AG11]	n^{c_s}	$\text{poly}(n)$	$2^{\tilde{O}(n^{2c_s})}$	$\omega(1 - 2c_s)n^{2c_s}(\log_2 n)^2$	rigorous
[ACFP14]	binary	$\text{poly}(n)$	$O(n \ln \ln n)$	$\left(\frac{\omega}{8} + o(1)\right) n \ln \ln \ln n (\ln \ln n)^{-1}$	heuristic
[STA20]	binary	$\text{poly}(n)$	$n^{1+\gamma}$	$\tilde{O}(n^{1-\gamma})$	heuristic
[STA20]	$\Pr[e = 1] = n^{-\beta}$	$\text{poly}(n)$	$O(n)$	$\begin{cases} O(n^{1-\beta} \ln n), & 0 < \beta < 1, \\ O(\ln n), & \beta \geq 1 \end{cases}$	rigorous
[Ste24]	binary	$\text{poly}(n)$	Cn	$\tilde{O}(\sqrt{n})$	estimate
<i>BKW and sample amplification</i>					
[KF15]	binary	$\text{poly}(n)$	$\Theta(n)$	$\left(\frac{\ln 2}{2} + o(1)\right) n (\ln \ln n)^{-1}$	claimed
[HKM18]	n^{c_s}	n^{c_q}	$\Theta(n \ln n)$	BKW : $\left(\frac{1}{2}c_q(c_q - c_s)^{-1} + o(1)\right) n$	rigorous
				BKW ₂ : $\left((2 \ln(c_q c_s^{-1}))^{-1} + o(1)\right) n$	
			$(c_m + o(1))n$	BKW : $\left(\frac{1}{2}c_q(c_q - c_s - c_q c_m^{-1})^{-1} + o(1)\right) n$ BKW ₂ : $\left(\left(2 \ln(c_q c_s^{-1}(1 - (c_m - 1)^{-1}))\right)^{-1} + o(1)\right) n$	
<i>This work</i>					
Corollary 4.1	$\ln(2 + s_e) = (\ln n)^{a+o(1)}$	$n^{\kappa+o(1)}$	$n + \Theta\left(n((1 - a) \ln \ln n)^{-1}\right)$	$\left((2(1 - a))^{-1} + o(1)\right) n (\ln \ln n)^{-1}$	rigorous
Corollary 4.2	$n^{\gamma+o(1)}$	$\exp\left((\ln n)^{1+\eta+o(1)}\right)$	$n + \Theta\left(n(\eta \ln \ln n)^{-1}\right)$	$\left((2\eta)^{-1} + o(1)\right) n (\ln \ln n)^{-1}$	rigorous

Conventions. For Gaussian rows, s_e denotes the width in $\rho_{s_e}(x) = \exp(-\pi x^2/s_e^2)$. When a cited work uses standard deviation σ , we convert it via $s_e = \sqrt{2\pi}\sigma$; this constant-factor conversion does not change the displayed asymptotics. Here ω denotes the matrix-multiplication exponent.

Parameter ranges. For Arora-Ge, $0 < c_s < 1/2$. For [STA20], $0 < \gamma < 1$ in the balanced-binary result and $\beta > 0$ in the sparse-error result. For [Ste24], $C > 1$ is constant. In [HKM18], c_q, c_s are fixed constants with $c_q > c_s + 1/2$; the linear-sample regime has constant $c_m > 1$ and satisfies the additional conditions of their Theorem 11. For this work, $0 \leq a < 1$, $\kappa > 1/2$, and $\gamma, \eta > 0$.

Qualifications. The bounds of [ACFP14] and the balanced-binary result of [STA20] rely on semi-regularity assumptions. The [Ste24] row uses the lowest achievable degree of regularity as a complexity estimate. The linear-sample BinaryLWE claim of [KF15], which additionally assumes a binary secret, relies on sample amplification; its limited-sample analysis contains the gap identified in [HKM18].