LEARNING WITH QUANTIZATION AND WORST CASE LATTICE PROBLEMS

ZIWEI HONG AND ZHONGQIU FANG

Abstract. Motivated by the MAMBA proposals, we study the worst case hardness of learning with quantization (LWQ) with public dither. We allow any efficiently computable dither function and efficiently samplable seed distribution; the matrix is uniform and independent of the secret and the seed.

Using the randomized polar lattice quantizer of Lyu et al., we construct explicit families with polynomial power of two moduli. For the search family, polynomial time recovery of the uniformly chosen secret with inverse polynomial probability yields quantum polynomial time algorithms for the gap shortest vector problem (GapSVP) and the shortest independent vectors problem (SIVP) on arbitrary n dimensional lattices, with approximation factors $O(n \log n)$ and $O(n(\log n)^2)$, respectively. Under stronger parameter conditions, a second family admits a decision reduction: a polynomial time algorithm with inverse polynomial distinguishing advantage yields quantum polynomial time algorithms for the same lattice problems with factors $O(n(\log n)^3)$ and $O(n(\log n)^4)$, respectively. Conditional on the public seed, the comparison output is uniform in the prescribed dither coset and independent of the matrix.

We also examine extensions to learning with errors (LWE) with interval uniform errors, as arising in MAMBA Frost. We identify obstacles caused by deterministic quantization boundaries and give a local perturbation bound with explicit dependence on the number of observations. This analysis does not establish a complete worst case reduction for these errors or hardness at the concrete MAMBA parameters.

School of Mathematics, Renmin University of China, Beijing, P.R. China

School of Mathematics and Statistics, UNSW Sydney, Sydney, Australia
Email address: hongziwei@live.com

China Post Securities Co., Ltd.
Email address: fangzhongqiu2024@outlook.com

1. Introduction

Learning with errors (LWE) is a central problem in lattice based cryptography. Its security foundation comes from reductions that relate solving LWE to solving worst case problems on lattices [**?**]. An LWE sample hides a linear expression As by adding a small error. Learning with rounding (LWR) replaces this addition by deterministic rounding, which reduces the size of the output [**?**]. The resulting error, however, depends on the input, and transferring hardness guarantees from LWE requires additional analysis.

Lyu, Liu, and Ling introduced learning with quantization (LWQ) to combine quantization with a reduction from LWE [**?**]. Instead of rounding each coordinate separately, LWQ quantizes As using a lattice. This allows the quantization rule to be chosen together with the desired error distribution. In particular, their polar lattice construction uses randomized successive cancellation (SC) encoding to obtain a distribution close to LWE with discrete Gaussian error. Dithering, which shifts the quantization lattice, is an essential part of this construction.

The practical interest in LWQ is reflected in the MAMBA family of first round candidates in China's Next Generation Commercial Cryptographic Algorithms (NGCC) program [**?**, **?**, **?**, **?**]. MAMBA Frost uses LWQ over matrices without ring or module structure, whereas MAMBA Viper and MAMBA NIKE use its module and ring variants, respectively. These proposals connect

Key words and phrases. Lattice based cryptography, learning with quantization, learning with errors, worst case reductions, polar lattices.

Corresponding author: Ziwei Hong (hongziwei@live.com).

the study of LWQ to the evaluation of postquantum algorithms for possible standardization. They also motivate a basic question: which choices of quantizer and public randomness admit a reduction from worst case lattice problems?

This paper addresses that question for LWQ with a prescribed public dither seed. The statistical distance bound of Lyu et al. averages over uniform dither. When a public seed determines the dither, a reduction must additionally produce the correct joint distribution of the seed, matrix, and quantized output. Closeness after averaging over the dither does not by itself establish this joint statement. We study the setting in which the matrix A is sampled uniformly and independently of both the secret and the seed. Within this setting, we construct a direct reduction in Regev's form and give explicit families for both search LWQ and decision LWQ, with different parameter requirements.

1.1. **Our contributions.** For search LWQ, our main result reduces worst case lattice problems to recovering the secret from LWQ samples. The reduction applies to randomized quantizers satisfying two conditions: a statistical distance bound to LWE with discrete Gaussian error, and an identity relating the output distributions under simultaneous translation of the input and the output coset. These conditions are stated in (**??**) and (**??**). We verify them for the polar construction and randomized SC encoding rule of Lyu et al. [**?**, Section 4], and give an efficient choice of nested information sets with explicit Bhattacharyya bounds. This yields the following result.

Main theorem (informal). *For all sufficiently large n, there is a polynomial time constructible LWQ family with quantization dimension $m = \Theta(n^2)$, modulus $q = 2^{\ell_q} = \Theta(\sqrt{n}\log n)$, and secret dimension $n+1$. Fix any efficiently computable public dither function and efficiently samplable seed distribution. The matrix A is uniform and independent of the secret and the seed. If a classical or quantum polynomial time algorithm recovers the uniformly chosen LWQ secret with inverse polynomial probability, then there are quantum polynomial time algorithms for*

$$\mathsf{GapSVP}_{O(n\log n)} \qquad \textit{and} \qquad \mathsf{SIVP}_{O(n(\log n)^2)}$$

on arbitrary n dimensional lattices.

Theorem **??** gives the general reduction to discrete Gaussian sampling, from which Regev's reductions yield the two lattice problems above. Corollary **??** specializes the result to the explicit polar family. A central ingredient is Lemma **??**: it accommodates any efficient public dither function by adding one secret coordinate, while using one source sample for each sample supplied to the LWQ algorithm.

Decision LWQ. We also establish hardness for distinguishing LWQ samples from uniform outputs in their prescribed cosets. Conditional on the public seed ρ, the comparison output is uniform in $H(\rho) + \Lambda/q\mathbb{Z}^m$ and independent of the matrix. Both distributions use the prescribed seed distribution and independently uniform matrices; the LWQ samples share one uniformly chosen secret, independent of the matrices and seeds. The reduction combines our transformation with Regev's LWE search to decision reduction for polynomial size prime moduli, or the reduction of Micciancio and Peikert [**?**] for prime power moduli with a polynomially bounded prime factor. The latter requires the additional noise conditions in Corollary **??**. In particular, a second explicit polar family has $m = \Theta(n^2)$, secret dimension $n+1$, and $q = 2^{\ell_q} = \Theta(\sqrt{n}(\log n)^3)$. A polynomial time algorithm with inverse polynomial distinguishing advantage for this family yields quantum polynomial time algorithms for

$$\mathsf{GapSVP}_{O(n(\log n)^3)} \qquad \text{and} \qquad \mathsf{SIVP}_{O(n(\log n)^4)}.$$

Thus decision LWQ also admits a worst case guarantee, but at larger parameters than the search family. Section **??** gives the precise conditions; this does not establish decision hardness at the search parameters above.

Connection to uniform error LWE. The practical proposals raise a further question: can the same route establish hardness for the interval uniform errors arising from deterministic integer lattice quantization? MAMBA Frost gives an exact equivalence with this form of LWE, but that equivalence does not supply the worst case reduction studied here. We explain why an

integer lattice subsequence of our construction still has a different output distribution, and analyze the effect of perturbations near deterministic quantization boundaries. A uniform addition gives a local error bound that includes the public dither and tracks the number of observations. Section **??** and Appendices **??** and **??** describe this possible extension and the remaining obstacles. These results do not establish a tight reduction for the concrete Frost parameters, or extend our theorem to the module and ring variants used by Viper and NIKE.

1.2. **Technical overview.**

Handling the public dither. The main difficulty is to generate the required samples without knowing the secret. Write $K_d(y)$ for the distribution of a dithered quantizer with output in the coset $d + \Lambda/q\mathbb{Z}^m$, where $q\mathbb{Z}^m \subseteq \Lambda \subseteq \mathbb{Z}^m$. The algorithm receives

$$(\rho, A, b), \qquad b \leftarrow K_{H(\rho)}(AS),$$

with public seed ρ and prescribed dither function H. Simply replacing $H(\rho)$ by independent uniform dither would change a relation that is visible from the sample. Conversely, given a sample with uniform dither, the reduction may be unable to find a seed that produces that dither. The construction must therefore work using only forward evaluation of H.

Our construction uses the prescribed output coset and the translation identity (**??**). Let (A, b) be an LWQ sample with uniform dither and unknown secret $s \in \mathbb{Z}_q^n$. Choose a private invertible matrix $M \in \mathrm{GL}_{n+1}(\mathbb{Z}_q)$ once for an invocation of the LWQ algorithm. For each sample, draw ρ from its prescribed distribution and λ uniformly from $\Lambda/q\mathbb{Z}^m$, and put

$$b' = H(\rho) + \lambda, \qquad t = b' - b, \qquad A' = [A \mid t]M.$$

Uniform source dither makes $[A \mid t]$ uniform, and (**??**) gives the required conditional distribution of b'. All samples have the same target secret $S = M^{-1}(s, 1)$. The change of variables makes this secret negligibly close to uniform, and a correct answer recovers s through MS. Lemma **??** establishes the required joint distribution of the secret and all triples (ρ, A', b').

From LWQ to lattice problems. The reduction follows Regev's proof through bounded distance decoding and discrete Gaussian sampling. Given Gaussian samples from a lattice and a point close to its dual, the calculation in Regev's Lemma 3.11 produces linear observations with Gaussian errors. The secret consists of the coefficients of the closest dual lattice point modulo q. Gaussian randomized rounding [**?**] converts these observations to discrete Gaussian samples. The statistical distance bound (**??**) and the transformation above then allow the LWQ algorithm to recover the secret. The closest point remains unknown throughout the construction.

Verification and repetition amplify the success probability, while Regev's noise adjustment handles the unknown Gaussian width. The resulting closest vector algorithm is used in Regev's quantum sampling procedure. Each iteration reduces the Gaussian width by a factor greater than two, yielding the discrete Gaussian sampling algorithm of Theorem **??**. The appendices account for finite precision, statistical distance, and sample complexity.

An efficient quantizer. For the polar construction, we use the multilevel lattice structure and randomized SC encoding rule of Lyu et al. Information bits are drawn from their conditional distributions. Prefixes of the integer lattice coefficients specify each dither coset; accounting for the carries between levels relates these prefixes to the frozen bits and proves (**??**). Averaging over uniform dither gives the statistical distance bound (**??**). To make the construction explicit, we compute nested information sets satisfying the required Bhattacharyya bounds and give a uniform polynomial time implementation of the randomized quantizer.

The decision reduction. The same transformation sends independent uniform pairs (A, b) to the comparison distribution specified above. Lemma **??** therefore converts an LWQ distinguisher using polynomially many samples into an LWE distinguisher, with negligible loss of advantage. Gaussian randomized rounding also preserves uniformity modulo q. We can therefore apply an LWE search to decision reduction and then Regev's lattice reduction. For prime power moduli, this step reduces the Gaussian width available to the search problem. This accounts for the additional conditions and larger approximation factors in Section **??**.

1.3. **Related work.**

LWQ and lattice quantization. Lyu et al. establish the statistical distance bounds underlying the polar quantization approach [**?**, Theorems 4 and 7]. Their use of a public seed to generate the matrix and dither motivates our question, although our theorem samples the matrix independently of that seed. We use their multilevel randomized SC encoding rule [**?**, equations (32)–(36)], with integer coefficient labels for dither cosets. Lemma **??** relates these labels to the frozen bits, and Lemma **??** proves the translation identity needed by our reduction. Their equivalence for nearest neighbor quantization with uniform dither instead gives an error uniform on the integer points of the quantization region [**?**, Lemma 1 and Theorem 5]. This is a different distribution from that of randomized polar quantization.

Lyu, Liu, Li, and Song introduce Learning With Sampling (LWS), whose output is obtained by adding a discrete Gaussian on the coset $\Lambda - As$ to As [**?**]. Our sequential quantizer need not sample this distribution for arbitrary nested information sets, because a prescribed frozen bit may follow an information bit. On the subsequence in Proposition **??**, all frozen bits precede all information bits. Proposition **??** then identifies the output as As plus a discrete Gaussian on $d + \Lambda - As$. For zero dither, this is the LWS distribution modulo q; nonzero dither gives the corresponding shifted coset.

MAMBA and uniform error LWE. Lyu, Ma, and Lai's MAMBA Frost specializes LWQ to scaled integer lattices with aligned power of two moduli. Its public bijection relates LWQ samples with the dither written separately to samples with interval uniform additive errors, and preserves the corresponding uniform distributions [**?**, Theorem 1, Lemma 3, and Theorem 3]. The inverse map reconstructs a uniform dither value from the additive noise sample. Our transformation instead handles an arbitrary efficient dither function H and prescribed seed distribution using only forward evaluation. Frost's equivalence identifies its uniform error LWE assumption, but does not replace that error by a discrete Gaussian; its discussion of worst case hardness invokes results for a prescribed number of samples [**?**, Section 3.4].

MAMBA Viper extends integer lattice quantization to modules over $\mathbb{Z}_q[X]/(X^{256} + 1)$. Its specification relates the resulting distribution to Module LWE with interval uniform errors and gives conditional security proofs for encryption and key encapsulation [**?**, Appendices A and C]. MAMBA NIKE combines Ring LWQ with key reconciliation; its key security theorem concerns passive attacks under additional reconciliation assumptions and a random oracle model for key derivation [**?**, Theorem 6.1 and Appendix B]. Both use algebraic structure beyond the independently uniform matrices considered here.

For LWE with interval uniform errors, Döttling and Müller-Quade give a worst case reduction using lossy codes, with a sample bound fixed in advance and quantitative conditions on the error width [**?**]. Bai et al. use Rényi divergence to obtain reductions from LWE with Gaussian errors [**?**]. These approaches provide relevant starting points for the extension discussed in Section **??**. Their guarantees depend on the error width, sample count, modulus, and loss in success or distinguishing advantage, so an exact correspondence of samples alone does not settle the parameter comparison.

Transformations of LWE samples. Our use of an extra secret coordinate and an invertible change of variables builds on established techniques. Regev's secret shifts randomize a fixed secret while preserving the additive error [**?**, Lemma 4.1]. Applebaum, Cash, Peikert, and Sahai give a normal form transformation in which the secret follows the error distribution [**?**, Lemma 2]. Brakerski, Langlois, Peikert, Regev, and Stehlé add one coordinate and apply an invertible change of variables in a reduction to LWE with one error free sample [**?**, Lemma 4.3]. The additional task in Lemma **??** is to obtain the prescribed joint distribution of public seeds, matrices, and quantized outputs across samples sharing one secret. Drawing a uniform extra column and rejecting until the output lies in the required coset would succeed with probability $1/\det\Lambda$. We instead draw the output directly in that coset and derive the extra column from it.

Rounding and polar codes. For deterministic rounding, Banerjee, Peikert, and Rosen control boundary events with a super polynomial parameter margin [**?**, Theorem 1]. Alwen, Krenn, Pietrzak, and Wichs use lossy matrices and entropy to allow polynomial moduli with a prescribed number of samples [**?**, Theorem 4.1]. Our main reduction uses randomized quantization and an

exact change of public dither; the obstacle for deterministic rounding is considered separately in Appendix **??**.

The polar construction relies on Arıkan's polarization identities [**?**]. Tal and Vardy give efficient constructions of polar codes approaching capacity for fixed binary memoryless symmetric channels [**?**]. Here we compute explicit lower bounds on the Bhattacharyya parameters of the channels arising in the LWQ reduction. These bounds establish the distributional approximation and efficient sampling required by our theorem. They also admit scaled integer lattices, including the subsequence identified in Appendix **??**.

2. Preliminaries

2.1. **Notation.** We follow Regev [**?**, Section 2] for standard lattice notions. The asymptotic parameter n is the worst case lattice dimension and the source secret dimension; the target LWQ secret has dimension $n+1$. The quantization dimension m, parameter bit lengths, and rational input encodings are polynomial in n. Statements concern sufficiently large n. We write $\mathbb{Z}_q = \mathbb{Z}/q\mathbb{Z}$, $\mathbb{T} = \mathbb{R}/\mathbb{Z}$, and $\mathsf{U}(S)$ for the uniform distribution on a finite set S. In the polar construction $\ell_q = \log_2 q$ counts levels and β_{pol} is its polarization exponent; r and β denote Gaussian widths.

Lattices and decoding. For a basis B, write $L = \mathcal{L}(B) = B\mathbb{Z}^n$, $B^* = B^{-T}$, and $L^* = B^*\mathbb{Z}^n$. Thus $B^{-1}v$ is the integer coefficient vector of $v \in L$. The successive minima are $\lambda_i(L)$. For $d < \lambda_1(L)/2$, Regev's $\mathsf{CVP}_{L,d}$ is bounded distance decoding: given x with $\operatorname{dist}(x, L) \le d$, return its closest point $\kappa_L(x)$. Its modular form $\mathsf{CVP}^{(q)}_{L,d}$ returns $B^{-1}\kappa_L(x) \bmod q$. We use the standard problems SIVP_γ, requiring n independent vectors of length at most $\gamma\lambda_n(L)$, and GapSVP_γ, distinguishing $\lambda_1(L) \le d$ from $\lambda_1(L) > \gamma d$.

Gaussian conventions. For $r > 0$, put

$$\rho_r(x) = e^{-\pi\|x\|^2/r^2}, \qquad D_{L,r}(v) = \rho_r(v)/\rho_r(L),$$
$$\eta_\varepsilon(L) = \inf\{r > 0 : \rho_{1/r}(L^* \setminus \{0\}) \le \varepsilon\}.$$

The continuous distribution $\nu_r^{(j)}$ on $\mathbb{R}^j$ has density $r^{-j}\rho_r$; omit the superscript for $j = 1$. Its coordinate variance is $r^2/(2\pi)$. The centered distribution $D_{\mathbb{Z}^m,h,y}$ has mass proportional to $\rho_h(z-y)$. An algorithm for DGS_φ, given L and $r > \varphi(L)$, samples within negligible statistical distance of $D_{L,r}$. We use total variation $\Delta(P,Q) = \frac{1}{2}\sum_x |P(x) - Q(x)|$, and the corresponding integral; this is one half of the convention in [**?**]. We use its usual coupling, product, and data processing bounds. When a random variable or tuple is an argument of Δ, its probability distribution is meant.

Algorithms. Running times are bounded polynomially even outside input promises; failed lattice samplers return a fixed lattice vector, such as zero. Quantum algorithms are uniform polynomial size circuits whose purified forms and inverses may be used by the reduction. A fixed deterministic polynomial time algorithm, given 1^n, generates the public integer parameters, an integer basis B_Λ, and the finite descriptions of the quotient encodings. The seed sampler, the map H, and the quotient operations are fixed uniform polynomial time algorithms for these parameters; the seed length is polynomial in n. Every real parameter used by an algorithm has a finite description from which an approximation of absolute error at most 2^{-k} can be computed in time polynomial in n and k. Any finite implementation of the quantizer is likewise uniform in its public parameters, input, dither, and requested precision. Implementations use finite encodings, Gaussian truncation, and bounded sampling procedures. Finite implementation errors are measured on the induced discrete output distributions.

2.2. **LWQ distributions.**

LWQ samples. Let $q = p^a$, where p is prime and $a \ge 1$, and let

$$q\mathbb{Z}^m \subseteq \Lambda \subseteq \mathbb{Z}^m, \qquad 1 < \det\Lambda < q^m.$$

The symbol $Q_{\Lambda+d}$ denotes the specified quantization process for the instance under consideration. If this process is randomized, its internal randomness is part of its definition.

Definition 2.1 (LWQ and Gaussian distributions)**.** For a fixed secret $s \in \mathbb{Z}_q^n$, let $\mathcal{Q}_s$ be the distribution of

$$A \leftarrow \mathsf{U}(\mathbb{Z}_q^{m\times n}), \qquad d \leftarrow \mathsf{U}(\mathbb{Z}^m/\Lambda), \qquad (A,b) = \big(A, Q_{\Lambda+d}(As) \bmod q\big).$$

A complete pair (A,b) is one sample. Repeated samples are independent conditional on the common secret, with fresh dither and fresh internal quantization randomness.

For $\sigma > 0$, let $\mathcal{G}_{s,\sigma}$ be the LWE distribution with discrete Gaussian error, generated by

$$(A, As + E \bmod q), \qquad E \leftarrow D_{\mathbb{Z}^m,\sigma},$$

where A is uniform and independent of E. For $\beta > 0$, let $\mathcal{C}_{s,\beta}$ be the scalar LWE distribution with continuous Gaussian error on $\mathbb{Z}_q^n \times \mathbb{T}$, where $\mathbb{T} = \mathbb{R}/\mathbb{Z}$, generated by

$$\left(a, \frac{\langle a, s\rangle}{q} + e \bmod 1\right), \qquad a \leftarrow \mathsf{U}(\mathbb{Z}_q^n), \quad e \leftarrow \nu_\beta.$$

The LWQ input format follows [**?**, Definitions 9–10]. The Gaussian distributions are used in the reduction, with their recovery algorithms constructed from the assumed LWQ algorithm.

Definition 2.2 (Statistical distance between LWQ and LWE)**.** Put

$$\mathcal{S}_q = \{s \in \mathbb{Z}_q^n : \gcd(s_1, \ldots, s_n, q) = 1\}.$$

For $q = p^a$, the condition $\gcd(s_1, \ldots, s_n, q) = 1$ is equivalent to requiring at least one coordinate to be a unit modulo q. The reduction requires the following bound at width σ: there is a negligible function $\mu(n)$, uniform in the secret, such that

$$\sup_{s\in\mathcal{S}_q} \Delta(\mathcal{Q}_s, \mathcal{G}_{s,\sigma}) \le \mu(n). \tag{2.1}$$

The bounds for randomized SC encoding in [**?**, Theorems 4 and 7] yield this type of estimate when the linear input is uniform. Section **??** specifies the quantizer and proves (**??**) with $\mu(n)$ negligible in n, uniformly over the stated secrets.

Definition 2.3 (Translation of the quantizer)**.** Let $\pi_\Lambda : \mathbb{Z}_q^m \to \mathbb{Z}^m/\Lambda$ be the quotient map and $\mathcal{K} = \Lambda/q\mathbb{Z}^m$. We use a fixed, efficiently computable representative for every coset. Thus $|\mathcal{K}| = q^m/\det\Lambda$. Write $K_d(y)$ for the output distribution of the specified quantizer on input $y \in \mathbb{Z}_q^m$ with dither d; it is supported on $d + \mathcal{K}$. Besides (**??**), the reduction uses the identity

$$K_{d+\pi_\Lambda(t)}(y + t) \overset{d}{=} K_d(y) + t \qquad (d \in \mathbb{Z}^m/\Lambda,\ y, t \in \mathbb{Z}_q^m). \tag{2.2}$$

The input and the output coset are translated together, and $\overset{d}{=}$ denotes equality of probability distributions. All additions of outputs are modulo q.

The reduction in Section **??** uses only the fact that the output lies in $d + \mathcal{K}$ and the identity (**??**). The statistical distance bound (**??**) is an additional hypothesis for the lattice reduction; Section **??** verifies these properties for the specified randomized polar quantizer.

Definition 2.4 (LWQ with a public dither seed)**.** Fix a seed distribution sampled by a uniform polynomial time algorithm and a uniform polynomial time function

$$H : \{0,1\}^{\ell(n)} \longrightarrow \mathbb{Z}^m/\Lambda,$$

where $\ell(n)$ is polynomially bounded and the public parameters and finite encodings are generated as specified above. For a fixed $S \in \mathbb{Z}_q^{n+1}$, one sample is

$$\rho \leftarrow \mathsf{Seed}, \qquad A' \leftarrow \mathsf{U}(\mathbb{Z}_q^{m\times(n+1)}), \qquad b' \leftarrow K_{H(\rho)}(A'S); \qquad (\rho, A', b'). \tag{2.3}$$

The seed is independent of the secret, the matrix, and the quantizer's internal coins. Before the quantizer is evaluated, the matrix and these coins are also independent. Each sample uses a fresh seed and fresh coins, so samples are independent conditional on the common secret. The solver receives the seed itself. Equivalently, its last input may be $c' = b' - H(\rho) \in \mathcal{K}$, under a specified invertible encoding of $\mathcal{K}$.

The function H may be the output of a pseudorandom generator followed by a map to $\mathbb{Z}^m/\Lambda$, but no pseudorandomness property of H is assumed. Constant H and nonuniform dither distributions are also allowed. A seed sampled once for one invocation of the solver and reused across that invocation's samples is treated in Remark **??**. This sampling rule must be declared in the solver's promise. Independent repeated invocations draw fresh seeds.

3. Worst case lattice problems

Theorem 3.1 (Reduction to LWQ)**.** *Let an LWQ family satisfy* (**??**) *and* (**??**)*, with efficiently computable quotient operations on Λ. Fix H and the distribution of Definition* **??**. *Suppose there is a classical or quantum polynomial time algorithm W and a fixed constant $c_0 > 0$ with the following property. Choose S uniformly from $\mathbb{Z}_q^{n+1}$, once, and give W access to the classical samples* (**??**)*. Then*

$$\mathbb{P}[W \text{ returns } S] \geq \varepsilon_W(n) \geq n^{-c_0}.$$

The probability is over S, all public seeds and samples, and the internal randomness and measurements of W. Thus the success guarantee is averaged over the declared seed sampling rule; it need not hold for each fixed seed. The same result holds if S is uniform over the vectors in $\mathbb{Z}_q^{n+1}$ satisfying $\gcd(S_1, \ldots, S_{n+1}, q) = 1$. A uniform polynomial time finite implementation of K is allowed when its output distribution has a uniform negligible error over all inputs and dithers.

Let $\zeta(n) > 0$ be negligible. Choose an efficiently usable rounding width $h \geq \eta_\zeta(\mathbb{Z}^m)$, and put

$$g = \sqrt{\sigma^2 - h^2}, \qquad \alpha = \frac{g}{q}, \qquad 2\sqrt{n} < g < q. \tag{3.1}$$

For every negligible function $\varepsilon(n) > 0$, there is a quantum polynomial time algorithm $\mathcal{R}^W$ for DGS_φ, where

$$\varphi(L) = \frac{\sqrt{2n}\, q}{g} \eta_\varepsilon(L). \tag{3.2}$$

More precisely, on input a basis of L, a width $r > \varphi(L)$, and a prescribed polynomial output count J, the algorithm $\mathcal{R}^W$ returns J lattice vectors whose joint distribution is negligibly close to $D_{L,r}^{\otimes J}$.

The functions ε, ζ, and μ bound, respectively, the smoothing error on L, the rounding error on $\mathbb{Z}^m$, and the statistical distance between LWQ and LWE with discrete Gaussian error. Section **??** gives parameters satisfying the hypotheses. The proof is completed in Section **??**.

Corollary 3.2 (Standard worst case problems)**.** *Under the hypotheses of Theorem* **??**, *there are quantum polynomial time algorithms for*

$$\mathsf{SIVP}_{\widetilde{O}(nq/g)} \quad \text{and} \quad \mathsf{GapSVP}_{\widetilde{O}(nq/g)}$$

on arbitrary n dimensional input lattices.

Proof. We apply Regev's reductions from standard lattice problems to DGS, with their hypotheses unchanged. For SIVP choose $\varepsilon_S = \exp(-(\log n)^2)$. Lemma 3.17 and Lemma 2.12 of [**?**] give the sufficient approximation factor

$$\gamma_S(n) = \frac{2\sqrt{2}\, nq}{g} \sqrt{\frac{\ln(2n(1 + 1/\varepsilon_S))}{\pi}} = \widetilde{O}(nq/g).$$

Indeed, that reduction returns independent vectors of length at most $2\sqrt{n}\, \varphi(L)$.

For GapSVP choose $\varepsilon_G = 2^{-n}$. Lemma 2.11 of [**?**] gives

$$\varphi(L) \leq \frac{\sqrt{2}\, nq}{g\, \lambda_1(L^*)} = \frac{\sqrt{n}\, \gamma'(n)}{\lambda_1(L^*)}, \qquad \gamma'(n) = \frac{\sqrt{2n}\, q}{g}.$$

Lemma 3.20 of [**?**], together with the reduction from GapSVP to the restricted GapCVP problem used there, gives the sufficient factor $100\sqrt{2}\, nq/g$. The required polynomial numbers of samples are covered by the joint output guarantee of Theorem **??**. □

4. LWQ with public dither seeds

We use the source distributions of Definition **??**, the quantizer in Definition **??**, and the distribution of Definition **??**. The source dimension is n, and the target secret has dimension $n+1$. Put $G = \mathbb{Z}_q^m$, $\mathcal{D} = \mathbb{Z}^m/\Lambda$, and retain $\mathcal{K} = \Lambda/q\mathbb{Z}^m$ and $\pi_\Lambda : G \to \mathcal{D}$ from the preliminaries. The following lemma shows that uniform dither makes the quantization error independent of the input.

Lemma 4.1 (The distribution of the quantization error)**.** *Let $d \leftarrow \mathsf{U}(\mathcal{D})$, and let χ_Q be the distribution of $K_d(0)$ on G. Under* (**??**)*,*

$$K_d(y) - y \sim \chi_Q \qquad \text{for every } y \in G,$$

where the dither is resampled uniformly on the left. Consequently, for every fixed $s \in \mathbb{Z}_q^n$, $\mathcal{Q}_s$ is the distribution generated by

$$(A, As + E_Q), \qquad A \leftarrow \mathsf{U}(\mathbb{Z}_q^{m\times n}), \quad E_Q \sim \chi_Q, \quad E_Q \perp A.$$

If χ_σ is the distribution of $E \bmod q$ for $E \leftarrow D_{\mathbb{Z}^m,\sigma}$, then

$$\Delta(\mathcal{Q}_s, \mathcal{G}_{s,\sigma}) = \Delta(\chi_Q, \chi_\sigma) \qquad (s \in \mathbb{Z}_q^n). \tag{4.1}$$

In particular, (**??**) *and* (**??**) *imply the same statistical distance bound $\mu(n)$ for every source secret, including those with* $\gcd(s_1, \ldots, s_n, q) > 1$.

Proof. Equation (**??**) gives $K_d(y) - y \overset{\mathrm{d}}{=} K_{d-\pi_\Lambda(y)}(0)$. Translation of a uniform coset is uniform, proving the first assertion. Condition on each value of A and apply it to $y = As$; the resulting error distribution is independent of A. Finally, the bijection $(A, b) \mapsto (A, b - As)$ sends the two distributions in (**??**) to the product distributions with the common uniform matrix and respective errors χ_Q, χ_σ. Their total variation is exactly that of the error distributions. This identity holds for every s, so a bound for secrets satisfying $\gcd(s_1, \ldots, s_n, q) = 1$ also bounds all other secrets. □

For the target, write

$$\mathcal{P}_{n+1} = \{S \in \mathbb{Z}_q^{n+1} : S \bmod p \neq 0\}.$$

Since $q = p^a$, this is precisely the set of vectors satisfying $\gcd(S_1, \ldots, S_{n+1}, q) = 1$. The uniform distribution used for decision LWQ draws ρ from the prescribed seed distribution, then draws A' uniformly and b' uniformly in $H(\rho) + \mathcal{K}$, independently conditional on ρ. It is equivalently uniform when represented by the vector $b' - H(\rho) \in \mathcal{K}$.

Algorithm 1 (Reduction to LWQ with public dither)**. Input.** Public $n, q, B_\Lambda, H, \mathsf{Seed}$; a source of pairs (A, b) for one unknown $s \in \mathbb{Z}_q^n$; and the target search algorithm W. **Output.** A candidate $\widehat{s} \in \mathbb{Z}_q^n$.

1. Sample $M \leftarrow \mathsf{U}(\mathrm{GL}_{n+1}(\mathbb{Z}_q))$ once. Keep it private and fixed throughout this invocation of W. Its random choices are independent of the source secret, source samples, and seed generation randomness. Independent invocations use fresh M.
2. Run W. For each sample it requests, obtain a fresh source pair (A, b), draw fresh independent $\rho \leftarrow \mathsf{Seed}$ and $u \leftarrow \mathsf{U}(\mathbb{Z}_q^m)$, and put $\lambda = B_\Lambda u \bmod q$. Compute

$$b' = H(\rho) + \lambda, \qquad t = b' - b, \qquad A' = [A \mid t]M. \tag{4.2}$$

   Supply (ρ, A', b') to W.
3. If W returns $\widehat{S} \in \mathbb{Z}_q^{n+1}$, output the first n coordinates of $M\widehat{S}$; otherwise output $0 \in \mathbb{Z}_q^n$.

Lemma 4.2 (One additional secret coordinate)**.** *Assume that the quantizer's output lies in $d + \mathcal{K}$ and that* (**??**) *holds, as in Definition* **??**. *With ideal exact random choices, Algorithm* **??** *is an expected polynomial time transformation taking one source sample to one target sample. For every fixed source secret s, put $S = M^{-1}(s, 1)^T$. The joint distribution of S and the N triples supplied to W is exactly the following: choose S uniformly from $\mathcal{P}_{n+1}$, and generate N samples according to Definition* **??**, *all with this same secret. In particular, the equality includes every seed ρ, matrix A', and output b' jointly. A single hidden matrix is used to transform all*

samples within one invocation. From a correct target search answer the algorithm recovers s, using no further source samples.

Consequently, if a target search algorithm W uses at most N samples and succeeds with probability ϵ, the derived source algorithm $\widehat{W}$ invokes W once, uses at most N source samples, and recovers every fixed source secret with probability at least $\epsilon - \eta_{\mathrm{sec}}$, where

$$\eta_{\mathrm{sec}} = 0 \quad \text{for } S \leftarrow \mathsf{U}(\mathcal{P}_{n+1}), \qquad \eta_{\mathrm{sec}} = p^{-(n+1)} \quad \text{for uniform target secrets.} \tag{4.3}$$

The same transformation maps independent uniform source pairs (A, b) to the following distribution: draw the prescribed seeds and, conditional on them, draw each matrix A' uniformly and each output b' uniformly from $H(\rho) + \mathcal{K}$, with all matrices and outputs independent. After choosing the matrix, each output uses $O(m(n+1)^2 + m^2)$ ring operations, in addition to seed generation and evaluation of H; extraction uses $O(n^2)$ ring operations.

Proof. Fix s, M, ρ, and put $d_{\mathrm{t}} = H(\rho)$. Conditional on a source pair (A, b), the vector t in (**??**) is uniform on $d_{\mathrm{t}} - \pi_\Lambda(b) + \mathcal{K}$. By the support property, $\pi_\Lambda(b) = d$. Hence this sampling can equivalently be ordered as follows: first draw A, d as in Definition **??**; then draw $t \leftarrow \mathsf{U}(d_{\mathrm{t}} - d + \mathcal{K})$ independently of the quantizer's random choices; finally draw $b \sim K_d(As)$ and put $b' = b + t$.

For every $t_0 \in G$, exactly one dither coset, $d = d_{\mathrm{t}} - \pi_\Lambda(t_0)$, can produce t_0. Since $|\mathcal{D}|\,|\mathcal{K}| = |G|$,

$$\Pr[t = t_0 \mid s, M, \rho] = \frac{1}{|\mathcal{D}|}\frac{1}{|\mathcal{K}|} = \frac{1}{|G|}.$$

Moreover, A is independent of (d, t). Thus $[A \mid t]$ is uniform, and its multiplication by M leaves A' exactly uniform. Conditional on A, t, ρ, the dither equals $d_{\mathrm{t}} - \pi_\Lambda(t)$; the equivalent sampling order and (**??**) give

$$b + t \stackrel{\mathrm{d}}{=} K_{d_{\mathrm{t}}}(As + t).$$

For the analysis only, define

$$S = M^{-1}\begin{pmatrix} s \\ 1 \end{pmatrix}. \tag{4.4}$$

Then $A'S = As + t$, proving the complete conditional distribution in (**??**).

Conditioning on s, M and all seeds, different transformed samples are independent with the specified output distributions. Their joint distribution depends on M, s only through S, so it remains the same product distribution when conditioned on S and the seeds alone. The vector $(s, 1)$ belongs to $\mathcal{P}_{n+1}$ for every s, since its last coordinate is one. Over $\mathbb{Z}_q$, an invertible matrix can send any vector in $\mathcal{P}_{n+1}$ to a standard basis vector: move a unit coordinate to a pivot, scale it by a unit, and eliminate the other coordinates. Thus the general linear group acts transitively on $\mathcal{P}_{n+1}$. Uniform M makes (**??**) exactly uniform on this set, independently of the seeds. This establishes the asserted joint distribution for each fixed source secret. A correct answer gives $MS = (s, 1)$, proving extraction.

A uniform vector of $\mathbb{Z}_q^{n+1}$ has all coordinates divisible by p with probability $\eta_0 = p^{-(n+1)}$. If the success guarantee of W concerns uniform secrets, its success conditioned on $S \in \mathcal{P}_{n+1}$ is at least $\max\{0, (\epsilon - \eta_0)/(1 - \eta_0)\} \geq \epsilon - \eta_0$.

For the uniform distribution used in decision LWQ, let the source A, b be independent and uniform. Conditional on ρ, the new b' is uniform on $H(\rho) + \mathcal{K}$, independently of the source. Subtraction of uniform b makes $t = b' - b$ uniform on G and independent of b'. Therefore $A' = [A \mid t]M$ is uniform and independent of b', even conditional on M. This holds jointly for all samples and proves the stated distribution of the seeds, matrices, and outputs for independent uniform source pairs.

All steps are efficient. If B_Λ is an integer basis of Λ, sampling $u \leftarrow \mathsf{U}(\mathbb{Z}_q^m)$ and setting $\lambda = B_\Lambda u \bmod q$ gives uniform $\mathcal{K}$, by the equal fiber property of finite group homomorphisms. Uniform invertible matrices are sampled by rejection: the acceptance probability is $\prod_{i=1}^{n+1}(1 - p^{-i})$, bounded below by an absolute positive constant. Matrix inversion and all remaining operations take polynomial time. Truncating the rejection loops gives bounded polynomial running time and negligible additional error, as specified in the preliminaries. □

Statistical distance. The algorithm $\widehat{W}$ of Lemma **??** can replace a uniform dither source solver at the same dimension and for the same quantizer. Suppose, for each fixed source secret s, an available N sample distribution $P_s^{(N)}$ obeys

$$\Delta(P_s^{(N)}, \mathcal{Q}_s^{\otimes N}) \le \delta_N, \tag{4.5}$$

where $\mathcal{Q}_s$ is the distribution in Definition **??**. Apply the same randomized map (**??**) to either distribution, retaining the private M in the analysis. Data processing shows that the resulting joint distributions are still within δ_N. Since the recovery event uses only this joint distribution, the success probability of $\widehat{W}$ is at least

$$\epsilon - \eta_{\mathrm{sec}} - \delta_N.$$

This applies in particular to LWE samples with Gaussian error whenever (**??**) has been proved. The map is defined on every pair (A, b); for such samples, $\pi_\Lambda(b)$ is only a derived label and need not be an independently uniform dither. A per sample bound δ gives $\delta_N \le N\delta$ for independent samples. For a random source secret, the comparison must include that secret jointly, or follow from a pointwise bound as above. In particular, Lemma **??** extends the bound (**??**) from source secrets satisfying $\gcd(s_1, \dots, s_n, q) = 1$ to every source secret.

Remark 4.3 (The choice of the public seed). The same proof permits one seed to be drawn at the beginning of each invocation of W and reused across all samples in that invocation. Draw this seed independently of the source secret, the source samples, the private matrix M, and the other random choices of the reduction. Conditioning on its value proves the equality of the distributions of the supplied triples (ρ, A', b'). The success guarantee for W may be averaged over this seed; the reduction uses precisely the same sampling rule.

More generally, a polynomial length seed sequence may follow any prescribed efficiently samplable joint distribution independent of the source secret, source sample randomness, M, and the other random choices of the reduction. Conditional on the sequence, the matrices and quantizer coins have the independence specified in Definition **??**. This allows correlated or repeated seeds within one invocation, provided the solver's promise uses the same distribution.

Independent trials in the amplification draw a fresh seed sequence, including a fresh seed for each invocation when one seed is reused within that invocation. Reusing a single seed across all trials would require a success guarantee conditional on that seed; an average over seed success guarantee alone does not suffice.

5. Polar lattice quantization

We specify the randomized polar quantizer used in the reduction, including the carries between binary digits and its public coset label. The parameter n remains the dimension of the source lattice; the target secret in Lemma **??** has dimension $n + 1$. The quantization dimension is m. The quantizer samples bits according to the conditional probabilities in [**?**, Section 4.2, equations (32)–(36)].

5.1. **Conditional distributions and frozen bits.** Let $m = 2^{\ell_m}$, $q = 2^{\ell_q}$, and let G_m be a standard binary polar matrix with its specified SC ordering. Use its integer lift and put $B_{\mathrm{pol}} = G_m^T$. This matrix is unimodular: the usual base matrix $\left(\begin{smallmatrix} 1 & 0 \\ 1 & 1 \end{smallmatrix}\right)$, its Kronecker powers, and the bit reversal permutation all have determinant 1 or -1. We retain the original SC order when using this basis.

Write $\chi_{\sigma,q}$ for the distribution of $D_{\mathbb{Z}^m,\sigma} \bmod q$, with the Gaussian width convention of the preliminaries. Define the joint distribution P of the test channel by

$$X \leftarrow \mathsf{U}(\mathbb{Z}_q^m), \qquad E \leftarrow D_{\mathbb{Z}^m,\sigma}, \qquad Y = X + E \bmod q.$$

Both X and Y are uniform. Symmetry of the Gaussian gives $X \mid Y = y \overset{d}{=} y + E \bmod q$. For $Z = B_{\mathrm{pol}}^{-1} X \bmod q$, its conditional distribution given $Y = y$ is

$$P_y(z) := \mathbb{P}_P[Z = z \mid Y = y] = \chi_{\sigma,q}(y - B_{\mathrm{pol}} z). \tag{5.1}$$

Every value has positive mass.

Use the ordinary binary planes $X_\ell \in \{0,1\}^m$ of X, starting with its least significant plane, and define the raw polar variables by

$$X_\ell = B_{\mathrm{pol}} U_\ell \bmod 2, \qquad X = \sum_{\ell=1}^{\ell_q} 2^{\ell-1} X_\ell \bmod q.$$

The history for U_ℓ^j consists of all earlier planes, the preceding coordinates $U_\ell^1, \dots, U_\ell^{j-1}$, and Y. Let $\mathcal{Z}_{\ell j}$ be the Bhattacharyya parameter of this conditional bit under P.

Fix $0 < \beta_{\mathrm{pol}} < 1/2$, set $\delta = 2^{-m^{\beta_{\mathrm{pol}}}}$, and choose nested information sets I_ℓ and frozen sets F_ℓ satisfying

$$I_1 \subseteq \cdots \subseteq I_{\ell_q}, \qquad F_\ell = [m] \setminus I_\ell, \qquad \mathcal{Z}_{\ell j} \geq 1 - \delta \quad (j \in F_\ell). \tag{5.2}$$

The randomized SC encoding procedure samples U_ℓ^j for $j \in I_\ell$ from its conditional distribution under P, and replaces U_ℓ^j for $j \in F_\ell$ by an independent fair bit. At each step, this conditional distribution uses the bits already generated.

The exact threshold sets in [**?**, Section 4.2] are $I_\ell^0 = \{j : \mathcal{Z}_{\ell j} \leq 1 - \delta\}$. If their nesting has not separately been established, the choice $I_\ell = \bigcup_{a \leq \ell} I_a^0$ satisfies (**??**). This operation only removes indices from the frozen sets and preserves the statistical distance bound below.

5.2. **Integer coefficients and dither.** For each coordinate let

$$t_j = \min\big(\{\ell - 1 : j \in I_\ell\} \cup \{\ell_q\}\big).$$

Nestedness means that precisely the lowest t_j coefficient bits are frozen. Define the Construction D lattice and its quotient by

$$\Lambda = B_{\mathrm{pol}} \operatorname{diag}(2^{t_1}, \dots, 2^{t_m}) \mathbb{Z}^m, \qquad \mathcal{D} = \mathbb{Z}^m / \Lambda, \qquad \pi_\Lambda : \mathbb{Z}_q^m \longrightarrow \mathcal{D}. \tag{5.3}$$

This construction uses the same integer basis at every level, as in [**?**, Definition 13, equation (29)]. In particular,

$$q\mathbb{Z}^m \subseteq \Lambda \subseteq \mathbb{Z}^m, \qquad |\mathcal{D}| = \det \Lambda = 2^{\sum_j t_j} = 2^{\sum_\ell |F_\ell|}.$$

A coset label $d \in \mathcal{D}$ specifies $z_j \bmod 2^{t_j}$ for every j; no reordering of SC decisions is needed.

Definition 5.1 (The randomized quantizer)**.** Given $y \in \mathbb{Z}_q^m$ and $d \in \mathcal{D}$, process the binary digits z_ℓ^j in increasing ℓ, then in the original SC coordinate order. For $j \in F_\ell$, use the digit specified by d. For $j \in I_\ell$, sample from the conditional distribution P_y in (**??**), given the digits already generated. Output $B_{\mathrm{pol}} z \bmod q$. Its distribution is denoted by $K_d(y)$.

Lemma 5.2 (The carry correction)**.** *For every choice of raw polar variables there are deterministic functions c_ℓ, depending only on earlier planes, such that the binary planes of $z = B_{\mathrm{pol}}^{-1} X \bmod q$ satisfy*

$$z_\ell = U_\ell + c_\ell(U_{<\ell}) \bmod 2. \tag{5.4}$$

Under the randomized SC encoding procedure, $D = \pi_\Lambda(X)$ is uniform on $\mathcal{D}$ and independent of Y, even for a prescribed fixed $Y = y$. Conditional on $Y = y, D = d$, the output distribution is $K_d(y)$.

Proof. Modulo 2^ℓ, the terms from higher planes vanish. The contribution of the current plane to $B_{\mathrm{pol}}^{-1} X$ is $2^{\ell-1} U_\ell$, because $B_{\mathrm{pol}}^{-1} X_\ell = U_\ell \bmod 2$. The contribution of lower planes is already determined. Its carry into the current plane gives (**??**). This is a triangular bijection, with an inverse obtained one plane at a time. Within a plane, it only flips known individual bits. Consequently the histories and conditional distributions of the information bits in the raw and integer coefficient descriptions correspond exactly.

For fixed y, the whole raw frozen vector U_{ℓ, F_ℓ}, conditional on all earlier planes, is uniform and independent of those planes. One may draw this vector before processing the plane; the information decisions use only the frozen bits that have already been reached. Adding the known carry preserves this conditional uniformity. Thus the vectors z_{ℓ, F_ℓ} are jointly uniform over all

their possible labels and independent of y. These are exactly the coefficient prefixes specifying $\mathcal{D}$, which proves the assertion about D.

More explicitly, the probability of a sequence of raw bits, conditional on y, is 2^{-f} times the product of its information bit conditional probabilities, where $f = \sum_\ell |F_\ell|$. Fixing $D = d$ determines each raw frozen bit from d and the already determined lower plane carries. The remaining product is precisely the sequential probability in Definition **??** and sums to one. Conditioning cancels the constant 2^{-f}. □

This lemma supplies an integer interpretation of the dither statement in [**?**, equation (31)]. In multiple levels, an uncorrected sum of the raw frozen bits is not in general the coset label specified by the coefficient residues $z_j \bmod 2^{t_j}$. Accounting for the carries preserves the joint distribution of (Y, X); it specifies how to generate the quantizer output for a public coset chosen in advance.

Lemma 5.3 (Translation of the polar quantizer)**.** *For every $d \in \mathcal{D}$ and $y, t \in \mathbb{Z}_q^m$,*

$$K_{d+\pi_\Lambda(t)}(y+t) \overset{d}{=} K_d(y) + t \pmod q. \tag{5.5}$$

Proof. Put $c = B_{\mathrm{pol}}^{-1} t \bmod q$. The conditional distribution in (**??**) obeys

$$P_{y+t}(z+c) = P_y(z).$$

Coordinatewise addition by c, in increasing binary levels, has the form

$$(z+c)_\ell^j = z_\ell^j \oplus a_{\ell j}(z_{<\ell}, c),$$

where the flip is determined before the current level is processed. It therefore maps every SC history bijectively to a new history, and maps all possible completions of that history bijectively to their translates. The displayed identity implies equality of the corresponding conditional information bit probabilities after this known flip.

For $j \in F_\ell$, all lower bits of coordinate j are also frozen. Hence the translated prefix is determined solely by the prefix prescribed by d and by c_j. It is exactly the prefix prescribed by $d + \pi_\Lambda(t)$; no information bit is used to define this new public label. We can therefore couple the two sequential executions, for both information bits and frozen bits, so that their coefficients satisfy $z' = z + c \bmod q$ throughout. Their outputs differ by $B_{\mathrm{pol}} c = t \bmod q$.

At each step, the conditional distribution uses the digits already generated. The order of levels and the nested information sets in (**??**) therefore give the required coupling. □

A two dimensional example is given in Appendix **??**.

5.3. **Statistical distance between LWQ and LWE.**

Lemma 5.4 (Statistical distance from a discrete Gaussian)**.** *Let Y be uniform on $\mathbb{Z}_q^m$, choose $D \leftarrow \mathsf{U}(\mathcal{D})$ independently, and sample $X \leftarrow K_D(Y)$. Let $E \leftarrow D_{\mathbb{Z}^m,\sigma}$ be independent. Then*

$$\Delta\big((Y, X), (Y, Y + E \bmod q)\big) \le \mu_{\mathrm{pol}} := \ell_q m \sqrt{\ln 2}\, 2^{-m^{\beta_{\mathrm{pol}}}/2}. \tag{5.6}$$

For any secret dimension k and any $s \in \mathbb{Z}_q^k$ satisfying $\gcd(s_1, \dots, s_k, q) = 1$, the same bound holds between $(A, K_D(As))$ and $(A, As + E \bmod q)$ when A is uniform in $\mathbb{Z}_q^{m\times k}$.

Proof. By Lemma **??**, averaging the output distribution of Definition **??** over D gives exactly the randomized SC encoding procedure for the raw bits. Apply the successive triangle inequality bounds of [**?**, Lemmas 3–4 and Theorem 7, equation (39)]. The conditional distributions of information bits agree with those under P. For a frozen bit with $j \in F_\ell$, Pinsker's inequality and $H \ge Z^2$ bound the average conditional statistical distance by $\sqrt{(\ln 2)\delta}$. Summing over at most $\ell_q m$ frozen bits, as in that proof, gives (**??**). The argument only requires the inequalities in (**??**), so removing indices from the frozen sets is allowed. The estimate concerns the discrete Gaussian reduced modulo q.

For such s, As is uniform. Both conditional output distributions depend on A only through As. Averaging their statistical distance over A therefore equals averaging it over $Y = As$, exactly

as in [**?**, Theorem 4, equation (14)]. The bound is uniform in such s and applies to $k = n$ as well as $k = n + 1$. □

The bound (**??**) averages over uniform random dither. Lemma **??** uses the equality of distributions in (**??**) to obtain samples with the prescribed public dither.

5.4. **Parameters.**

Proposition 5.5. *For sufficiently large n, choose*

$$m = 2^{\lceil 2\log_2 n\rceil}, \qquad \beta_{\mathrm{pol}} = \tfrac{1}{4}, \qquad \sigma = 3\sqrt{n}, \qquad 8\sigma \ln n \le q < 20\sigma \ln n, \qquad q = 2^{\ell_q}, \tag{5.7}$$
$$h = \ln n, \qquad \zeta = e^{-(\ln n)^2}, \qquad g = \sqrt{\sigma^2 - h^2}.$$

The modulus can be chosen in polynomial time by the rule in Appendix **??**. *The exact threshold sets have $I_1^0 = \varnothing$ and $I_{\ell_q}^0 = [m]$. The sets $I_\ell = \bigcup_{a\le\ell} I_a^0$, and the construction in Theorem* **??**, *give $1 < \det\Lambda < q^m$. Moreover,*

$$\mu_{\mathrm{pol}} = \mathrm{poly}(n)2^{-\Omega(\sqrt{n})}, \qquad h \ge \eta_\zeta(\mathbb{Z}^m), \qquad 2\sqrt{n} < g < q, \qquad q/g = \Theta(\ln n).$$

Proof. The first partition channel is the binary additive channel with noise $E \bmod 2$. Poisson summation gives its parity bias

$$\gamma = \mathbb{E}[(-1)^E] = \frac{\sum_{a\in\mathbb{Z}} e^{-\pi\sigma^2(a-1/2)^2}}{\sum_{a\in\mathbb{Z}} e^{-\pi\sigma^2 a^2}} \le 4e^{-\pi\sigma^2/4}.$$

Its Bhattacharyya parameter is $Z_1 = \sqrt{1-\gamma^2}$. The recursions $Z(W^+) = Z(W)^2$ and $Z(W^-) \ge Z(W)$ [**?**, Proposition 5] show that every length m subchannel has parameter at least $Z_1^m \ge 1 - m\gamma^2$. Here $m\gamma^2 = e^{-\Omega(n)} < 2^{-m^{1/4}}$. Thus the first level frozen set is $[m]$.

At the last level, subtracting the known lower bits gives the channel with inputs $0, q/2$ and noise $E \bmod q$. Expanding its aliased Gaussian masses and using $\sqrt{(\sum a_i)(\sum b_j)} \le \sum_{i,j}\sqrt{a_i b_j}$ yields

$$Z_{\ell_q} \le \sum_{a\in\mathbb{Z}} e^{-\pi(q/2+aq)^2/(4\sigma^2)} \le 4e^{-\pi q^2/(16\sigma^2)} \le 4e^{-4\pi(\ln n)^2}.$$

For the first inequality, completing the square leaves a Gaussian mass centered at $(q/2 + aq)/2$, which is an integer once $4 \mid q$, so its normalizing ratio is one. The remaining geometric tail gives the factor four for sufficiently large n. The upper polar recursion $Z(W^-) \le 2Z(W)$ gives $\mathcal{Z}_{\ell_q j} \le mZ_{\ell_q} = \mathrm{negl}(n) < 1 - 2^{-m^{1/4}}$. Thus the last level information set is $[m]$.

Taking $I_\ell = \bigcup_{a\le\ell} I_a^0$ preserves these endpoints. Hence

$$m \le \sum_{\ell=1}^{\ell_q} |F_\ell| \le (\ell_q - 1)m, \qquad 2^m \le \det\Lambda \le 2^{(\ell_q-1)m}.$$

The stated bound on μ_{pol} follows from $n^2 \le m < 2n^2$ and $\ell_q = O(\log n)$. Finally, the standard smoothing bound [**?**, Lemma 2.12] gives

$$\eta_\zeta(\mathbb{Z}^m) \le \sqrt{\frac{\ln(2m(1+1/\zeta))}{\pi}} < \ln n.$$

The remaining width inequalities and $q/g = \Theta(\ln n)$ follow directly from (**??**). □

5.5. **Efficient construction.** The conditional sampler has a polynomial time implementation for every nested choice of information sets. We also give a deterministic construction of sets satisfying the Bhattacharyya bounds. The proof is in Appendix **??**.

Theorem 5.6 (Efficient construction and sampling)**.** *For the parameters of Proposition* **??**, *with q chosen by* (**??**), *there are a deterministic algorithm* $\mathsf{Build}(1^n)$, *running in* $\mathrm{poly}(n)$ *bit operations, and a randomized algorithm* $\mathsf{Sample}(1^n, \mathcal{I}, y, d, 1^k)$, *running in* $\mathrm{poly}(n, k)$ *bit operations, with the following properties for all sufficiently large n.*

The construction returns nested sets $I_1 \subseteq \cdots \subseteq I_{\ell_q}$ satisfying (??), with $I_1 = \varnothing$ and $I_{\ell_q} = [m]$, and the integer basis

$$B_\Lambda = B_{\mathrm{pol}} \operatorname{diag}(2^{t_1}, \ldots, 2^{t_m}).$$

In particular, $1 < \det \Lambda < q^m$.

The sampler accepts any nested tuple $\mathcal{I} = (I_1, \ldots, I_{\ell_q})$ of subsets of $[m]$, encoded by its incidence bits. For these supplied sets, form the lattice Λ and quotient $\mathcal{D}$ in (??), and let $K_d(y)$ denote the output distribution of Definition ?? at input y. For every $k \geq 1$, $y \in \mathbb{Z}_q^m$, and $d \in \mathcal{D}$, encoded by its coefficient residues $d_j \bmod 2^{t_j}$, the sampler's distribution $\widetilde{K}_d^{(k)}(y)$ is supported on $\pi_\Lambda^{-1}(d)$ and satisfies

$$\sup_{y,d} \Delta\big(\widetilde{K}_d^{(k)}(y), K_d(y)\big) \leq 2^{-k}. \tag{5.8}$$

The running time and the bound are uniform in the supplied nested tuple $\mathcal{I}$.

Corollary 5.7 (LWQ on polar lattices)**.** *For the parameters of Proposition ??, with q chosen by (??), fix any deterministic polynomial time algorithm constructing nested information sets satisfying (??) and*

$$0 < \sum_{\ell=1}^{\ell_q} |F_\ell| < \ell_q m.$$

Define Λ by (??), set $k(n) = n$, and choose a seed distribution and map H as in Definition ??. Use the implementation $\widetilde{K}_{H(\rho)}^{(k(n))}$ of the randomized quantizer from Definition ?? and Theorem ?? as the LWQ quantizer. If a polynomial time algorithm W recovers its uniformly chosen secret with probability at least n^{-c_0}, for a fixed constant $c_0 > 0$, the conclusions of Theorem ?? and Corollary ?? hold for this family. The algorithm Build *supplies one such construction. For every prescribed polynomial output count, the DGS width threshold is*

$$\varphi(L) = \sqrt{2n}\frac{q}{g}\eta_\varepsilon(L) = O(\sqrt{n}\log n)\eta_\varepsilon(L).$$

With the smoothing parameters of Corollary ??, sufficient approximation factors are

$$\gamma_{\mathsf{GapSVP}} = O(n \log n), \qquad \gamma_{\mathsf{SIVP}} = O\big(n(\log n)^2\big). \tag{5.9}$$

The input lattice has dimension n, and the LWQ secret has dimension $n + 1$.

Proof. The frozen bit count gives $1 < \det \Lambda < q^m$. Let $N = \mathrm{poly}(n)$ bound the number of samples used by W. The uniform estimate (??) bounds the statistical distance between its ideal and implemented public inputs by $N2^{-n} = \mathrm{negl}(n)$, including adaptive sample requests. Lemmas ?? and ?? apply to the constructed information sets satisfying (??). Together with Theorem ??, they therefore give the stated DGS algorithm. Proposition ?? gives $q/g = \Theta(\log n)$. Substitution into the factors of Corollary ?? proves (??). □

Example 5.8 (A family with congruent coordinates)**.** Let $\mathbf{1} = (1, \ldots, 1)^T \in \mathbb{Z}^m$, and choose the index j_* whose column in B_{pol} is $\mathbf{1}$. For the usual Kronecker polar transform, with or without bit reversal, this is $j_* = m$. The original SC order is retained. For each sufficiently large n, run $\mathsf{Build}(1^n)$ and replace its information sets by

$$I'_\ell = I_\ell \cup \{j_*\}, \qquad F'_\ell = [m] \setminus I'_\ell \quad (1 \leq \ell \leq \ell_q).$$

These sets define a deterministic polynomial time constructible family to which Corollary ?? applies, with the same parameters and approximation factors.

In particular, for $n = 2^{4r}$ and all sufficiently large integers r, put $a = 8n^{1/4} = 2^{r+3}$. The resulting lattice is

$$\Lambda' = a\mathbb{Z}^m + \mathbb{Z}\mathbf{1} = \{x \in \mathbb{Z}^m : x_i \equiv x_j \pmod a \text{ for all } i, j\}, \qquad \det \Lambda' = a^{m-1}. \tag{5.10}$$

This lattice is not a Cartesian product of one dimensional lattices in the given coordinates.

Proof. The new sets are nested and satisfy $F'_\ell \subseteq F_\ell$. Consequently every index $j \in F'_\ell$ satisfies the same inequality $\mathcal{Z}_{\ell j} \geq 1 - 2^{-m^{1/4}}$ as before. The endpoints of Build give $I'_1 = \{j_*\}$ and $I'_{\ell_q} = [m]$, so

$$0 < m - 1 \leq \sum_{\ell=1}^{\ell_q} |F'_\ell| \leq (\ell_q - 1)(m-1) < \ell_q m.$$

These are precisely the additional frozen bit count conditions in Corollary **??**; that corollary does not require the first information set to be empty. The uniform sampler of Theorem **??** accepts the new sets. In particular, its error is at most 2^{-k}, and the bound in (**??**) remains $\mu_{\mathrm{pol}} + 2^{-k}$. The values of q, σ, h, g are unchanged, so all width inequalities and the factors in (**??**) remain valid.

On the stated subsequence, Proposition **??** gives

$$I'_\ell = \{j_*\} \quad (\ell \leq r+3), \qquad I'_\ell = [m] \quad (\ell \geq r+4).$$

Thus the coefficient exponent of column j_* is zero, and every other coefficient exponent is $r+3$. Writing b_j for the columns of B_{pol}, the lattice in (**??**) becomes

$$\Lambda' = \mathbb{Z} b_{j_*} + \sum_{j \neq j_*} a\mathbb{Z} b_j = a B_{\mathrm{pol}} \mathbb{Z}^m + \mathbb{Z} b_{j_*} = a\mathbb{Z}^m + \mathbb{Z}\mathbf{1}.$$

Unimodularity of B_{pol} also gives $\det \Lambda' = a^{m-1}$. The congruence description in (**??**) follows by subtracting $x_1\mathbf{1}$ from any vector whose coordinates are congruent modulo a. Since a and q are powers of two and $q > a$, one has $q\mathbb{Z}^m \subseteq \Lambda' \subseteq \mathbb{Z}^m$, as required.

Finally, every coordinate projection of Λ' is $\mathbb{Z}$, because $\mathbf{1} \in \Lambda'$. If Λ' were a Cartesian product in these coordinates, its factors would therefore all be $\mathbb{Z}$. This would give $\Lambda' = \mathbb{Z}^m$, whereas $e_1 \notin \Lambda'$ for $m \geq 2$ and $a > 1$, by the congruence description above. □

Information sets and compressed coordinates. The inequalities in (**??**) suffice for the statistical distance bound (**??**) and allow a range of nested sets, including choices that give scaled integer grids. Proposition **??** identifies such a subsequence of Build. Example **??** gives an explicit family that is not a Cartesian product in the given coordinates. The exact output distribution on the grid subsequence and its relation to LWS are stated in Proposition **??**. Once the public coset d is specified, the remaining coefficient bits give a bijective encoding of the quantized vector using exactly

$$\sum_{j=1}^{m} (\ell_q - t_j) = \ell_q m - \sum_{\ell=1}^{\ell_q} |F_\ell|$$

bits.

6. From LWQ to bounded distance decoding

6.1. **Gaussian randomized rounding.** For $y \in \mathbb{R}^m$, define the integer valued random variable

$$\mathsf{RR}_h(y) \leftarrow D_{\mathbb{Z}^m, h, y} \quad \text{or equivalently} \quad \mathsf{RR}_h(y) \stackrel{d}{=} y + D_{\mathbb{Z}^m - y, h}. \tag{6.1}$$

The second expression denotes a shifted lattice discrete Gaussian followed by addition of y, so its result is in $\mathbb{Z}^m$.

Lemma 6.1 (Continuous to discrete Gaussian rounding)**.** *For $u \in \mathbb{Z}^m$,*

$$\mathsf{RR}_h(y + u) \stackrel{d}{=} \mathsf{RR}_h(y) + u.$$

Consequently $\mathsf{RR}_h(y) \bmod q$ is well defined in distribution when only $y \bmod q\mathbb{Z}^m$ is supplied. If $h \geq \eta_\zeta(\mathbb{Z}^m)$, $0 < \zeta \leq 1/2$, $E \leftarrow \nu_g^{(m)}$, and $\sigma^2 = g^2 + h^2$, then

$$\Delta\big(\mathsf{RR}_h(E), D_{\mathbb{Z}^m, \sigma}\big) \leq 8\zeta. \tag{6.2}$$

The rounding operation is efficiently implementable to negligible error.

Proof. The identity under integer translation follows by translating the summation index in the normalizing mass $\rho_h(\mathbb{Z}^m - y)$. For (**??**), apply the continuous Gaussian case of [**?**, Theorem 1] with rounding lattice $\mathbb{Z}^m$, rounding covariance parameter $h^2 I$, continuous covariance parameter $g^2 I$, and zero shift. That case requires the smoothing condition on the rounding lattice only. The coordinatewise implementation is described in [**?**, Section 4.1]. □

The reduction uses this randomized rounding with $\sigma^2 = g^2 + h^2$.

6.2. **Recovery of the secret.** We first construct a recovery algorithm from W, and then use it to solve the closest vector problem.

Lemma 6.2 (A fixed noise width)**.** *Under the hypotheses of Theorem* **??***, there is a polynomial time procedure $\mathcal{F}_\alpha^W$ that, for every fixed $s \in \mathbb{Z}_q^n$, recovers s from independent observations with distribution $\mathcal{C}_{s,\alpha}$, with probability $1 - 2^{-\Omega(n)}$. It has the same classical or quantum computational type as W.*

Proof. In one trial choose $M \leftarrow \mathsf{U}(\mathrm{GL}_{n+1}(\mathbb{Z}_q))$, once for the entire call to W, and keep it hidden. Whenever W requests one sample, collect m fresh scalar observations, write their first components as the rows of A and their second components as a vector $c \in \mathbb{T}^m$, and compute

$$b_0 = \mathsf{RR}_h(qc) \bmod q. \tag{6.3}$$

Use any representatives for c; Lemma **??** makes the result independent of that choice in distribution. Apply the algorithm of Lemma **??**:

$$\begin{gathered} \rho \leftarrow \mathsf{Seed}, \qquad \lambda \leftarrow \mathsf{U}(\mathcal{K}), \qquad b' = H(\rho) + \lambda, \qquad t = b' - b_0, \\ A' = [\, A \mid t \,]M, \qquad \text{give } (\rho, A', b') \text{ to } W. \end{gathered} \tag{6.4}$$

If W returns $\widehat{S} \in \mathbb{Z}_q^{n+1}$, the candidate for the original secret is

$$\widehat{s} = (M\widehat{S})_{1:n} \bmod q.$$

An invalid output is assigned a default candidate.

The source secret s remains arbitrary and fixed. The inputs before rounding have the Gaussian distribution with mean As modulo q. Lemma **??** shows that each rounded sample is within 8ζ of $\mathcal{G}_{s,\sigma}$. Lemma **??** extends (**??**) to this fixed secret, whether or not $s \in \mathcal{S}_q$. Let $N = N(n)$ bound the number of complete samples supplied in one invocation of W. Replacing the rounded samples by LWQ samples with uniform dither changes the success probability by at most $N(8\zeta + \mu)$. This comparison is made before applying Lemma **??**; data processing applies to the joint distribution of all inputs supplied to W, including their seeds. For LWQ source samples with uniform dither, Lemma **??** gives the secret $S = M^{-1}(s, 1)$, uniform among the vectors in $\mathbb{Z}_q^{n+1}$ satisfying $\gcd(S_1, \ldots, S_{n+1}, q) = 1$. Changing that target secret distribution to uniform on all $\mathbb{Z}_q^{n+1}$ costs at most $p^{-(n+1)}$. Consequently the success probability of this trial is at least $\varepsilon_W - e_W$, where

$$e_W = N(8\zeta + \mu) + p^{-(n+1)} + \nu_{\mathrm{impl}} = \mathrm{negl}(n). \tag{6.5}$$

Here ν_{impl} includes numerical errors and, when applicable, the error of the finite precision quantizer over all N samples.

If W is promised a uniformly chosen S satisfying $\gcd(S_1, \ldots, S_{n+1}, q) = 1$, the term $p^{-(n+1)}$ is omitted. For adaptive sample requests, replace their conditional distributions one at a time and apply the triangle inequality, using the polynomial running time bound on the number of requests. Thus this trial returns the correct original secret with probability at least $n^{-c_0}/2$, for all sufficiently large n.

We validate each candidate on fresh, unrounded observations from $\mathcal{C}_{s,\alpha}$, using [**?**, Lemma 3.6]. The statistic is the empirical mean of

$$\cos\left(2\pi\left(c - \frac{\langle a, \widehat{s}\rangle}{q}\right)\right).$$

Its expectation is $e^{-\pi\alpha^2} > 0.04$ for the correct candidate and zero for any wrong candidate. The latter statement holds also for composite q, since the mean of a nontrivial character of $\mathbb{Z}_q^n$ is

zero. A fixed threshold between zero and $e^{-\pi}$, evaluated to sufficient constant precision, gives validation error $2^{-\Omega(n)}$ from $O(n)$ fresh observations. Concrete choices are given in Appendix **??**.

Make $O(n^{c_0+1})$ independent trials, with fresh invertible matrices, seeds, observations, rounding randomness, and work registers, and accept the first verified candidate. The probability that no trial produces the correct candidate is $2^{-\Omega(n)}$. A union bound over the polynomially many validation tests bounds erroneous acceptance by $2^{-\Omega(n)}$ as well. Finite sampling and arithmetic errors can be made exponentially small. The negligible error in (**??**) preserves the inverse polynomial success probability of each trial. □

Lemma 6.3 (An unknown noise width)**.** *There is a polynomial time procedure $\mathcal{N}_\alpha^W$ which, for every fixed $s \in \mathbb{Z}_q^n$ and every unknown $0 < \beta \le \alpha$, recovers s from independent observations with distribution $\mathcal{C}_{s,\beta}$, with probability $1 - 2^{-\Omega(n)}$.*

Proof. We use the algorithmic noise adjustment of [**?**, Lemma 3.7], with $\mathcal{F}_\alpha^W$ as its fixed width procedure. The procedure is as follows.

Choose an integer constant $c \ge 1$ so that n^c bounds the total number of scalar observations used by one execution of $\mathcal{F}_\alpha^W$, including its internal candidate verification. Consider the added variance grid

$$\mathcal{V} = \{j\alpha^2 n^{-2c} : j = 0, \dots, n^{2c}\}. \tag{6.6}$$

For each $t \in \mathcal{V}$, add independent $\nu_{\sqrt{t}}$ noise modulo one to each observation supplied to $\mathcal{F}_\alpha^W$; for $t = 0$ add zero. Verify the resulting candidate with fresh observations from the original, unmodified $\mathcal{C}_{s,\beta}$ source. Repeat each grid trial $O(n)$ times and accept a verified candidate.

There is a grid value with

$$\alpha^2 - \beta^2 \le t \le \alpha^2 - \beta^2 + \alpha^2 n^{-2c}.$$

At that value the modified width $\alpha' = \sqrt{\beta^2 + t}$ satisfies

$$\alpha \le \alpha' \le (1 + n^{-2c})\alpha.$$

By [**?**, Claim 2.2], in our convention the distance per scalar observation from width α is at most $(9/2)n^{-2c}$. The full input used by $\mathcal{F}_\alpha^W$ is therefore within $(9/2)n^{-c}$ of its fixed width input. Subsequent randomized rounding and computation cannot increase this distance. Lemma **??** consequently gives success at least $1/2$ for this grid value, for sufficiently large n. Fresh original observations validate candidates with exponentially small error, by the same cosine test since $\beta \le \alpha < 1$. Repetition proves the assertion.

The parameters g, h, σ remain fixed throughout the procedure. Candidates from every grid value are checked using fresh observations from the original distribution. □

6.3. **Solving the closest vector problem.**

Lemma 6.4 (Bounded distance decoding)**.** *Under the hypotheses of Theorem* **??**, *let $L = \mathcal{L}(B)$ and*

$$R > \sqrt{2}\, q\, \eta_\varepsilon(L), \qquad d_R = \frac{g}{\sqrt{2}\, R}. \tag{6.7}$$

Given a polynomial number of independent samples from $D_{L,R}$, there is a polynomial time algorithm using W that, for every fixed target x with $\operatorname{dist}(x, L^*) \le d_R$, returns

$$(B^*)^{-1} \kappa_{L^*}(x) \bmod q \tag{6.8}$$

with probability $1 - \operatorname{negl}(n)$. Probability is over the Gaussian input samples and the algorithm's randomness and measurements. The algorithm has the same classical or quantum computational type as W.

Proof. The decoding radius is below $\lambda_1(L^*)/2$ for sufficiently large n. Indeed, $g < q$ and (**??**) imply $d_R < 1/(2\eta_\varepsilon(L))$, and [**?**, Claim 2.13] implies $\eta_\varepsilon(L)\lambda_1(L^*) > 1$.

For the analysis let

$$u = \kappa_{L^*}(x), \qquad z = (B^*)^{-1} u \in \mathbb{Z}^n, \qquad s = z \bmod q.$$

The algorithm does not know u, z, s. For each fresh vector $v \leftarrow D_{L,R}$, compute

$$a = B^{-1}v \bmod q, \qquad c = \frac{\langle x, v\rangle}{q} + e_0 \bmod 1, \qquad e_0 \leftarrow \nu_{\alpha/\sqrt{2}}. \tag{6.9}$$

These operations use only known quantities.

By the distributional calculation in [**?**, Lemma 3.11], the following two properties hold. First, the coset mass estimate makes a negligibly close to uniform. Second, the identity

$$\langle u, v\rangle = \langle z, B^{-1}v\rangle$$

gives the required modular linear term. Conditional on a, the remaining projection of v, together with e_0, has negligible distance from a continuous Gaussian of width β, where

$$\beta^2 = \left(\frac{R\|x-u\|}{q}\right)^2 + \frac{\alpha^2}{2} \le \alpha^2. \tag{6.10}$$

This is exactly the smoothing calculation of that lemma and its Corollary 3.10, with modulus q and base Gaussian width $\alpha/\sqrt{2}$. Its hypothesis follows from $R/\sqrt{2} > q\eta_\varepsilon(L) = \eta_\varepsilon(qL)$. Consequently the distribution of (**??**) is negligibly close to $\mathcal{C}_{s,\beta}$, for the same unknown β on every row. More explicitly, the conditional projection error is at most 2ε in our convention, and the residue distribution contributes at most $\varepsilon/(1-\varepsilon)$. For $\varepsilon \le 1/2$, the resulting row error is at most 4ε. Using fresh vectors for all requested rows preserves negligible distance over the full polynomial computation.

Run $\mathcal{N}_\alpha^W$ from Lemma **??** on these rows. To specify its calls to W, consider one value $v_{\mathrm{add}} \in \mathcal{V}$ and choose one $M \in \mathrm{GL}_{n+1}(\mathbb{Z}_q)$ for that call. Group m rows to form A. First compute the auxiliary vector b_0 coordinatewise:

$$(b_0)_i = \mathsf{RR}_h\left(\langle x, v_i\rangle + q(e_{0,i} + e_{\mathrm{add},i})\right) \bmod q, \quad e_{\mathrm{add},i} \leftarrow \nu_{\sqrt{v_{\mathrm{add}}}}. \tag{6.11}$$

Here the scalar version of RR_h is applied independently to each coordinate, and the added noise is zero when $v_{\mathrm{add}} = 0$. The actual public input supplied to W is

$$\begin{aligned} \rho \leftarrow \mathsf{Seed}, \quad \lambda \leftarrow \mathsf{U}(\mathcal{K}), \quad b' = H(\rho) + \lambda, \quad t = b' - b_0, \\ (\rho, A', b'), \qquad A' = [\,A \mid t\,]M. \end{aligned} \tag{6.12}$$

The matrix M is fixed and hidden throughout this call. If W returns $\widehat{S}$, its candidate for the modular closest vector coefficients is $(M\widehat{S})_{1:n}$. All input vectors, noises, and rounding choices are fresh within this execution of the decoding algorithm. The verification in $\mathcal{N}_\alpha^W$ uses fresh original rows (**??**), before the added grid noise and rounding. The verification in $\mathcal{F}_\alpha^W$ uses the continuous observations after adding the noise for that grid value.

Every quantity in (**??**)–(**??**) is computed from x, the lattice samples, and the reduction's randomness. At a suitable grid value, randomized rounding, (**??**), Lemma **??**, and Lemma **??** give the input distribution used in Lemmas **??**–**??**. Those lemmas recover s with exponentially high probability from independent observations with distribution $\mathcal{C}_{s,\beta}$. The negligible statistical distance between those observations and (**??**) then gives success $1 - \mathrm{negl}(n)$, proving (**??**). □

To use this algorithm in Regev's quantum sampling procedure, we amplify it uniformly over the finite set of queries used by that procedure.

7. Discrete Gaussian sampling

Lemma 7.1 (Amplification on a finite query set)**.** *Let $\mathcal{T}$ be a finite set of finite input encodings, fixed independently of sampled data, and let f be specified on a promised subset $\mathcal{T}_{\mathrm{good}}$. Suppose an algorithm retaining its input uses a fresh list of samples whose distribution does not depend on the query and, for each fixed promised x, returns $f(x)$ with probability at least $2/3$. The algorithm may be quantum. Take K independent copies, each with its own list of samples and classical preprocessing randomness, where*

$$K \ge 18\big(\ln|\mathcal{T}| + \ln(1/\delta) + \ln(1/\eta)\big). \tag{7.1}$$

With probability at least $1-\eta$ over this joint draw of the samples and preprocessing randomness, all of which are then fixed for subsequent reuse, the majority circuit has conditional error at most δ on every promised x. Deferring measurements, copying the answer by XOR, and uncomputing gives a circuit within state norm error $2\sqrt{\delta}$ of the oracle for f that restores the work registers to zero, uniformly on superpositions of promised inputs, including entanglement with other registers.

Proof. Run K independent copies using independent lists of samples and fresh work registers. Return the strict majority of the complete output strings, or a fixed default if no strict majority exists. The majority can be computed reversibly.

Let P comprise the lists of samples and any classical preprocessing randomness fixed for subsequent reuse, and let $e_P(x)$ be the conditional measurement error of the majority circuit. For each fixed promised x, independence before conditioning on P and Hoeffding's inequality imply

$$\mathbb{E}_P e_P(x) \le e^{-K/18}.$$

Markov's inequality and a union bound give

$$\mathbb{P}_P[\exists x \in \mathcal{T}_{\mathrm{good}} : e_P(x) > \delta] \le |\mathcal{T}| e^{-K/18}/\delta \le \eta.$$

The randomness included in P is classical preprocessing randomness; measurements are deferred in the circuit below.

Fix a P for which $e_P(x) \le \delta$ for every promised x, defer measurements, and let U_P be the resulting circuit. Compute with U_P, copy the answer by XOR into a designated output register, and uncompute with $U_P^\dagger$. For fixed x, the incorrect answer component of the computed state has norm at most $\sqrt{\delta}$. Copying agrees with ideal copying on the correct answer component; the difference between the two copying operators has norm at most two on the remaining component. Thus this computation differs from

$$|x, y, 0\rangle \longmapsto |x, y \oplus f(x), 0\rangle$$

by at most $2\sqrt{\delta}$. The query register is preserved, so the sectors for different x are orthogonal. The same bound holds for any superposition of promised inputs and any entangled outside register. □

Lemma 7.2 (The iterative step)**.** *Under the hypotheses of Theorem* **??**, *let $R > \sqrt{2}\, q\, \eta_\varepsilon(L)$ and put*

$$\theta = \frac{g}{\sqrt{n}} > 2.$$

A polynomial size list of independent $D_{L,R}$ samples, together with W, suffices for a quantum polynomial time algorithm to produce any prescribed polynomial number J of samples whose joint distribution is negligibly close to

$$D_{L,R/\theta}^{\otimes J}.$$

The same input samples can be reused for all J outputs.

Proof. Set $d = g/(\sqrt{2}R)$. Lemma **??** supplies an algorithm for recovering the closest vector coefficients modulo q, with success at least $2/3$ for each fixed promised target. Before drawing its samples, pad the loops over $\mathcal{V}$, repetitions, calls to W, and candidate verification to their polynomial bounds. Allocate disjoint lists of Gaussian samples for every potential observation request, including validation requests. A success flag may record early acceptance while all remaining work follows the fixed schedule. Let N_0 bound the required number of Gaussian vectors. Once the finite query encodings and the running time bound of W are specified, N_0 is fixed independently of the query and the sampled vectors.

We identify the finite queries needed for Regev's quantum sampling step. In original coordinates, the grid in [**?**, Lemma 3.14] can be chosen as L^*/M_{grid}, where

$$M_{\mathrm{grid}} \ge 2^{3n} \frac{\sqrt{n}}{d} \lambda_n(L^*).$$

Select M_{grid} using a computable upper bound for $\lambda_n(L^*)$ from the input basis. Then $\log M_{\mathrm{grid}}$ is polynomial in the allowed input size. Initial queries are B^*a/M_{grid}, $a \in \{0,\dots,M_{\mathrm{grid}}-1\}^n$. The reduction from CVP to its modular version in [**?**, Lemma 3.5] uses the recurrence

$$x_{i+1} = (x_i - B^*c_i)/q, \qquad c_i \in \{0,\dots,q-1\}^n.$$

All branches, including erroneous output branches, stay within the predetermined set

$$\mathcal{T} = \bigcup_{i=0}^{n-1} \left\{ B^* \frac{a}{M_{\mathrm{grid}} q^i} : a \in \mathbb{Z}^n,\ |a_j| \le M_{\mathrm{grid}} q^i \right\}.$$

Indeed, the basis coefficients remain in $[-1,1]^n$. Moreover,

$$|\mathcal{T}| \le n(2M_{\mathrm{grid}}q^n + 1)^n, \qquad \ln|\mathcal{T}| = O(n\log M_{\mathrm{grid}} + n^2\log q + \log n). \tag{7.2}$$

Only a computable upper bound for this logarithm is needed; the set is not enumerated.

First fix the truncation bounds and the finite implementation of the decoding algorithm of Lemma **??**. Its negligible sampling errors and the verification accuracy described above still leave success at least $2/3$ for every fixed promised query. Apply Lemma **??** to this actual finite circuit, with $\delta = \eta = 2^{-n^2}$. The total number $N = KN_0$ of samples is polynomial. Their joint distribution and the set $\mathcal{T}$ are fixed before the samples are drawn. All invertible matrices, seeds, and noises can either be generated inside the circuit with deferred measurements or, when independent of the query, included in its classical preprocessing. The circuit implements the input dependent randomized rounding of (**??**).

For samples and preprocessing randomness satisfying Lemma **??**, the circuit approximates the oracle that computes the closest vector coefficients modulo q and restores the work registers to zero, with error $2\sqrt{\delta}$. The reversible implementation of [**?**, Lemma 3.5] uses $O(n)$ such calls and a Babai nearest plane computation to obtain the full closest point. Along an ideal promised computation the distance shrinks by a factor q at each step of that recurrence. Replacing the oracle calls one at a time and applying the triangle inequality gives full decoding error $O(n\sqrt{\delta})$.

Apply Regev's quantum sampling procedure [**?**, Lemma 3.14], using the corrected proof of Claim 3.13.[1] Decoding radius d gives output Gaussian width

$$\frac{\sqrt{n}}{\sqrt{2}\,d} = \frac{R\sqrt{n}}{g} = \frac{R}{\theta}. \tag{7.3}$$

Its ideal Gaussian tail and finite grid errors are exponentially small. The error of the reversible decoding circuit contributes $O(n\sqrt{\delta})$. Consequently, conditional on any P satisfying Lemma **??**, the output distribution $\mathcal{D}_P$ satisfies

$$\Delta(\mathcal{D}_P, D_{L,R/\theta}) \le 2^{-\Omega(n)} + O(n\sqrt{\delta}) + \xi_{\mathrm{impl}} =: \xi, \tag{7.4}$$

where ξ_{impl} is a negligible implementation error.

All operations admit a uniform finite implementation. For $v = Bz$ and a query $x = B^*a/(M_{\mathrm{grid}}q^i)$,

$$\langle x, v\rangle = \frac{a^T z}{M_{\mathrm{grid}} q^i},$$

which is an exactly computable rational number. Discrete Gaussian tails permit polynomial bit bounds on the sample coefficients, with total overflow probability at most 2^{-n^2}; one may use a norm tail bound and the input basis inverse to obtain those bounds. Overflow produces default lattice outputs and is charged to the probability of an unsuccessful choice of samples. Gaussian noise generation, the randomized rounding in (**??**), and the cosine tests admit arbitrary negligible accuracy in their finitely encoded outputs. Regev's finite Fourier and Gaussian state procedures supply the remaining approximations. All register sizes and circuit lengths are bounded before sampling.

Finally, retain only the samples and classical preprocessing randomness, and restart all work registers for each output. The J outputs are independent conditional on P; hence their

[1] `https://arxiv.org/abs/2401.03703v1`.

conditional distance from the desired product is at most $J\xi$. Averaging over P, and including overflow, gives total distance at most $2\eta + J\xi$, which is negligible. □

Proof of Theorem **??**. Let the DGS input be $L = \mathcal{L}(B)$ and $r > \varphi(L)$. Set

$$r_i = r\theta^i, \qquad i = 0, \ldots, 3n, \qquad \theta = g/\sqrt{n} > 2.$$

For every $i \geq 1$,

$$r_i \geq r_1 = r\theta > \sqrt{2}\, q\, \eta_\varepsilon(L).$$

The iterative step therefore applies at each such width. Since $q/g > 1$, [**?**, Claim 2.13] gives

$$r > \sqrt{2n}\, \eta_\varepsilon(L) \geq \sqrt{\frac{2\ln(1/\varepsilon)}{\pi n}}\lambda_n(L).$$

Together with $\theta > 2$, this implies $r_{3n} > 2^{2n}\lambda_n(L)$ for all sufficiently large n. Lemma 3.2 of [**?**] supplies a polynomial size initial list of samples at width r_{3n}, to exponentially small joint error.

Choose a common polynomial sample count large enough for all $3n$ instances of Lemma **??**; their encoding and resource bounds are fixed before drawing samples. Starting with the initial samples, repeatedly produce new lists of samples at widths $r_{3n-1}, \ldots, r_1$, and finally produce the requested J samples at $r_0 = r$. Each new list is generated using independent fresh work registers, with the input samples retained only as read only data.

If the joint distribution of the input samples has distance a_i from the distribution of independent D_{L,r_i} samples, and the sampling step on independent D_{L,r_i} inputs has joint statistical distance b_i from the required output distribution, data processing gives

$$a_{i-1} \leq a_i + b_i.$$

Every b_i is negligible by Lemma **??**, including its polynomial output count. Summing over $3n$ transitions proves the claimed joint output guarantee. □

8. Decision LWQ

The search theorem concerns a solver that returns the secret. For decision, compare the samples (ρ, A', b') from Definition **??** with the following distribution:

$$\rho \leftarrow \mathsf{Seed}, \quad A' \leftarrow \mathsf{U}(\mathbb{Z}_q^{m\times(n+1)}), \quad b' \leftarrow \mathsf{U}(H(\rho) + \mathcal{K}). \tag{8.1}$$

Conditional on the seed, the matrix and the output are independent. All samples follow the same declared rule for drawing seeds in both distributions. The decision advantage is the absolute difference between acceptance probabilities for samples from Definition **??**, with one common secret drawn as in Theorem **??**, and samples from (**??**). Taking b' uniform on all of $\mathbb{Z}_q^m$ would yield a different problem: the test $\pi_\Lambda(b') = H(\rho)$ already distinguishes it.

Lemma 8.1 (From decision LWQ to decision LWE)**.** *Suppose W_{dec} distinguishes the distribution in Definition* **??** *from* (**??**) *with advantage $\varepsilon_{\mathrm{dec}}$ and uses at most N complete samples. Under the distribution hypotheses of Theorem* **??**, *there is a distinguisher for LWE with discrete Gaussian error, $\mathcal{G}_{s,\sigma}$ versus independent uniform (A, b), with advantage at least*

$$\varepsilon_{\mathrm{dec}} - N\mu - p^{-(n+1)} - \nu_{\mathrm{impl}}.$$

There is also a distinguisher for m row collections of $\mathcal{C}_{s,g/q}$ versus independent uniform observations on $\mathbb{Z}_q^n \times \mathbb{T}$, grouped into m row collections, with advantage at least

$$\varepsilon_{\mathrm{dec}} - N(\mu + 8\zeta) - p^{-(n+1)} - \nu_{\mathrm{impl}}.$$

The term $p^{-(n+1)}$ is omitted if the target secret is uniform over the vectors satisfying $\gcd(S_1, \ldots, S_{n+1}, q) = 1$. In this lemma, ν_{impl} bounds the sum of the implementation errors for the two distributions being distinguished.

Proof. Run the map in Lemma **??** and then W_{dec}. For samples from $\mathcal{G}_{s,\sigma}$, Lemma **??** and data processing bound the distance from the samples generated with quantizer K by $N\mu$; replacing a secret uniform over $\{S \in \mathbb{Z}_q^{n+1} : \gcd(S_1, \dots, S_{n+1}, q) = 1\}$ by one uniform over $\mathbb{Z}_q^{n+1}$ adds at most $p^{-(n+1)}$. The uniform source maps exactly to (**??**), as proved in Lemma **??**. For continuous Gaussian samples, first apply $\mathsf{RR}_h(qc) \bmod q$ to each m row collection. Lemma **??** gives the additional $8N\zeta$ loss. On uniform $c \in \mathbb{T}^m$, this rounded output is exactly uniform modulo q: integer translations preserve the uniform input distribution, and Lemma **??** shows that rounding commutes with every such translation in distribution. The output distribution is therefore invariant under all translations of $\mathbb{Z}_q^m$. The matrix remains independent of it. □

Prime moduli. For a prime $q \le \mathrm{poly}(n)$, apply [**?**, Lemmas 4.1–4.2] to decision LWE with error $\chi = D_{\mathbb{Z},\sigma} \bmod q$, grouping the scalar observations into the required complete samples. These lemmas give a search procedure for that discrete error distribution. Gaussian randomized rounding then supplies its inputs from the continuous width g/q, and Regev's decoding and DGS construction gives (**??**) and Corollary **??**.

Regev's coordinate test uses the invertibility of every nonzero coefficient modulo q, and enumerates q guesses. This accounts for the prime modulus and polynomial size conditions. For the power of two moduli used in polar quantization, we use the following reduction with additional noise.

Corollary 8.2 (Prime power moduli)**.** *Retain the hypotheses on the quantizer, seeds, and negligible error bounds above. Let $q = p^a$, with $p \le \mathrm{poly}(n)$, and fix an efficiently usable $t(n) = \omega(\sqrt{\log n})$, $t \ge 1$. In addition to $\sigma^2 = g^2 + h^2$ and $g < q$, assume*

$$g > 2\sqrt{n}\, t^2, \qquad g \ge t, \qquad 0 < \frac{g}{qt^2} \le \frac{1}{t}. \tag{8.2}$$

*A polynomial time algorithm that distinguishes the distribution in Definition **??** from (**??**) with inverse polynomial advantage gives a quantum polynomial time algorithm for $\mathsf{DGS}_{\varphi_{\mathrm{dec}}}$, where*

$$\varphi_{\mathrm{dec}}(L) = \frac{\sqrt{2n}\, q\, t^2}{g}\eta_\varepsilon(L).$$

*It also gives the SIVP factor in Corollary **??** multiplied by t^2, and the sufficient GapSVP factor $100\sqrt{2}\, nqt^2/g$.*

Proof. Use the continuous Gaussian distributions in Lemma **??** and the search to decision reduction of Micciancio and Peikert [**?**], in the form of Theorem 3.1 of the full version,[2] with

$$\alpha_{\mathrm{src}} = \frac{g}{qt^2}, \qquad \alpha_{\mathrm{dec}} = \frac{g}{q}.$$

For the single prime factor, the theorem's integer ℓ is either zero or one. Its conditions are $\alpha_{\mathrm{src}} \le 1/t$, $\alpha_{\mathrm{dec}} \ge \alpha_{\mathrm{src}}$, $\alpha_{\mathrm{dec}} \ge t/q$, and $\alpha_{\mathrm{dec}}^{\ell} \ge \alpha_{\mathrm{src}} t^{1+\ell}$. All follow from (**??**) and $t \ge 1$. The theorem yields continuous Gaussian search at width α_{src}, with polynomial overhead. Regev's lattice reduction applies because $\alpha_{\mathrm{src}} q > 2\sqrt{n}$. Its DGS threshold is $\sqrt{2n}\eta_\varepsilon(L)/\alpha_{\mathrm{src}}$; the stated lattice factors follow from the same worst case reductions as before. □

An explicit polar decision family. Keep m, β_{pol}, and ζ as in (**??**), and take

$$t = \ln n, \qquad \sigma = 3\sqrt{n}(\ln n)^2, \qquad h = \ln n, \qquad g = \sqrt{\sigma^2 - h^2}.$$

Choose q by (**??**) with this value of σ. Then

$$q = \Theta\big(\sqrt{n}(\log n)^3\big), \qquad g = \Theta\big(\sqrt{n}(\log n)^2\big), \qquad \frac{qt^2}{g} = \Theta\big((\log n)^3\big),$$

and (**??**) holds for all sufficiently large n. The proof of Theorem **??** extends to these parameters: the first level channel deficit is $e^{-\Omega(n(\log n)^4)}$, while the last level overlap is $e^{-\Omega((\log n)^2)}$. Moreover, $q/\sigma = \Theta(\log n)$ and $\log \sigma = O(\log n)$ give a uniform channel mass lower bound of $2^{-O((\log n)^2)}$. The width σ^2 is efficiently computable. Using a dyadic upper bound on σ^2 with a guaranteed

[2] `https://eprint.iacr.org/2011/501`.

error bound to select the integer truncation radius, interval evaluation of the Gaussian sums provides the same polynomial time construction and uniform finite precision sampling guarantees for this scaled family.

Use the resulting quantizer with precision $k(n) = n$, and choose a seed distribution and map H as in Definition **??**. If a polynomial time algorithm distinguishes the samples in Definition **??**, generated with this quantizer and one uniformly chosen $S \in \mathbb{Z}_q^{n+1}$ common to all samples, from (**??**) with inverse polynomial advantage, Corollary **??** gives quantum algorithms with

$$\varphi_{\mathrm{dec}}(L) = O(\sqrt{n}(\log n)^3)\eta_\varepsilon(L),$$
$$\gamma_{\mathsf{GapSVP}} = O(n(\log n)^3), \qquad \gamma_{\mathsf{SIVP}} = O(n(\log n)^4).$$

Here the SIVP and GapSVP bounds use the respective smoothing error choices in Corollary **??**.

9. Discussion

The reduction applies to the samples in Definition **??**: a seed specifies the dither, and the matrix is uniform and independent of the secret and that seed. The polar construction uses the structure in Section 4 of Lyu et al., integer coefficient prefixes to specify cosets, and sampling from the conditional distributions at indices in the information sets. The Bhattacharyya bounds give a security reduction for every efficiently constructed family satisfying Corollary **??**. The quantization dimension is explicit; achieving capacity approaching rates and quantization goodness requires stronger information set constructions.

Section **??** gives decision consequences for the same matrix and dither sampling, with uniform outputs in the prescribed coset as the comparison distribution. For prime power moduli, these consequences require the additional noise conditions of Corollary **??**; the explicit polar decision family has larger parameters than the search family.

In the implementation described in [**?**, Remark 3], a single seed generates both (A, d). Extending our result to those samples requires simulating the additional public relation $A = G_A(\rho)$. Seeds specifying raw frozen bits must also account for the lower plane carries in Appendix **??**. Replacing randomized information bit sampling by the maximum a posteriori rule in [**?**, Algorithm 1] changes the output distribution and requires a separate analysis.

9.1. Integer lattice quantization and uniform errors.

The integer lattice specialization in MAMBA Frost provides a concrete target for extending the present results. Let q and p_c be powers of two with $1 < p_c < q$, put $k = q/p_c$, and take $\Lambda = k\mathbb{Z}^m$. Here p_c is a compression modulus, not the prime factor in our notation $q = p^a$. In Frost's convention, the dither is added before nearest neighbor quantization and subtracted afterwards:

$$d \leftarrow \mathsf{U}(\{0, \ldots, k-1\}^m), \qquad b = Q_\Lambda(As + d) - d \pmod q.$$

The change $d \mapsto -d$ in $\mathbb{Z}^m/\Lambda$ relates this convention to our notation for the output coset. For (A, s) independent of d, and using rounding ties towards the larger integer, the resulting distribution is exactly

$$(A, As + e \bmod q), \qquad e_i \overset{\text{iid}}{\leftarrow} \mathsf{U}(\{-k/2 + 1, \ldots, k/2\}). \tag{9.1}$$

The errors are independent of (A, s). In the separate representation (A, c, d), put $c = Q_\Lambda(As + d) \bmod q$; the compact label is $c/k \in \mathbb{Z}_{p_c}^m$. The public map is $b = c - d \bmod q$, with inverse $d = (-b) \bmod k$, $c = (b + d) \bmod q$. These maps preserve the real and uniform distributions. A further public translation $b \mapsto b - \mathbf{1}$ gives uniform errors on $\{-k/2, \ldots, k/2 - 1\}$, with the same number of samples and distinguishing advantage [**?**, Theorems 1–3 and Lemma 4]. This is the precise uniform error LWE problem, often denoted uLWE, associated with this quantization rule.

The equivalence preserves dimensions, sample counts, and distinguishing advantage. It does not itself provide a reduction from worst case lattice problems, or an approximation factor for such a reduction. Thus the question is whether one can produce this particular uniform error distribution from worst case lattice instances with controlled parameter and advantage losses. It cannot be answered by identifying the quantization lattice alone.

Existing reductions. The lossy code result of Döttling and Müller-Quade provides one baseline [**?**, Theorem 1]. For search LWE with M scalar samples fixed in advance and errors uniform on an integer interval of radius ρq, its main theorem requires

$$M = \mathrm{poly}(n), \qquad M \geq 3n, \qquad 0 < \rho < 1/10, \qquad \rho q \geq 2n^{1/2+c}M$$

for a fixed constant $c \in (0, 1)$, and gives a quantum reduction from worst case GapSVP and SIVP with approximation factor $\widetilde{O}(n^{1+c}M/\rho)$. Here M is a scalar sample count, not our quantization dimension m. These conditions exhibit a dependence on the sample budget that cannot be dropped when applying the theorem to a narrow error interval. Rényi divergence methods offer another route from Gaussian to uniform errors [**?**]; they too require checking the quantitative hypotheses of the particular reduction. Neither approach can be invoked merely by calling the error uniform. The precise interval endpoints, modulus, secret distribution, and search or decision formulation must agree with the theorem being applied.

What the integer lattice subsequence implies. Theorem **??** does not apply merely because the quantization lattice has the form $k\mathbb{Z}^m$. Its additional hypothesis (**??**) concerns the entire distribution and requires negligible distance from discrete Gaussian errors. On the subsequence $n = 2^{4r}$, Appendix **??** gives $\Lambda = 8n^{1/4}\mathbb{Z}^m$, but Proposition **??** still prescribes random sampling from a discrete Gaussian on a coset, not deterministic nearest neighbor quantization. Equality of lattices therefore does not identify the two LWQ problems, and equality of error variances would not supply the missing statistical distance bound.

Steps that can be reused and a local perturbation bound. Some parts of our proof can nevertheless be reused. Deterministic nearest neighbor quantization on $k\mathbb{Z}^m$, with the tie rule in Proposition **??**, satisfies the coset support condition and (**??**). Hence Lemma **??** applies to its public dither transformation without assuming Gaussian errors. This observation does not supply the preceding reduction from worst case lattice problems to the required uniform error distribution. Proposition **??** shows that an arbitrarily small continuous perturbation followed by deterministic quantization need not yield a negligible simulation error. It is not an impossibility result for all reductions to uLWE. Lemma **??** shows that adding independent uniform randomness in $[0, 1)$ before quantization preserves every integer input's output and bounds the perturbation error by its expected magnitude divided by the quantization step. The bound includes the public dither, but accumulates with the total number of scalar observations. It does not by itself give a worst case reduction or preserve our approximation factors.

More explicitly, consider one batch satisfying the independence hypotheses of Lemma **??**. Let $T \geq 1$ count its scalar observations and let $\tau \in (0, 1)$ be the statistical error budget allocated to this batch. The finite version of Lemma **??** gives the sufficient conditions

$$\max_{i \leq T} \mathbb{E}|\delta_i| \leq \frac{k\tau}{2T}, \qquad h \geq \max\left\{0, \left\lceil \log_2 \frac{2T}{k\tau} \right\rceil\right\}. \tag{9.2}$$

The perturbation contribution is then at most $\tau/2$, and the finite addition contributes at most $\tau/2$. This allocates a budget to one simulation step, not to the complete lattice reduction; other calls and other error terms require separate allocations. If the joint distribution of the secret and the solver's input changes by distance at most τ, its recovery success probability changes by at most τ. For a decision problem, the losses for the real and uniform branches must both be counted: if each is at distance at most τ, an advantage ε may decrease by 2τ. The local lemma does not supply the uniform branch of a complete decision reduction. For a fixed error budget, this sufficient bound allows a smaller mean perturbation as T increases. The uniform addition removes the boundary discontinuity, not the sample dependence in this bound; it does not prove a lower bound for all possible reductions.

Parameters and remaining conditions. For comparison, the Frost parameters in [**?**, Table 2] use $q = 2^{16}$, $\ell = 8$, and $n = m$ equal to 576, 968, or 1320. The public key and ciphertext component U have error widths 64, 32, and 8 in these three profiles, respectively, while component V has widths 256, 512, and 256. Both long term and ephemeral secrets have centered binomial distributions. The encryption proof also uses a joint ciphertext instance with different error widths in different rows, and an explicit term for the seed expansion [**?**, Theorem 4]. Our

allowance for power of two moduli does not resolve these differences: our target secret is uniform (or uniform over vectors whose coordinates have greatest common divisor one with q), and our decision theorem retains Gaussian distribution and width hypotheses. We do not obtain security estimates for any of these fixed parameter sets from the present asymptotic theorems.

Viper adds a separate structural requirement. It uses $\mathbb{Z}_q[X]/(X^{256}+1)$ with module rank $2, 3, 4, 7$, or 9 and $q = 2^{12}$ or 2^{13}, with centered binomial secrets and interval uniform errors [**?**, Table 3 and Appendix A]. Expanding polynomial multiplication into an integer matrix leaves blocks whose rows are related by signed cyclic shifts; the entries are not independent uniform elements of $\mathbb{Z}_q$. Thus this expansion does not meet the matrix distribution in Definition **??**. The general invertible change of variables in Lemma **??** is not shown to preserve this structure. The same distinction applies to Ring LWQ in MAMBA NIKE. An extension to ordinary uLWE would therefore not, by itself, establish the corresponding module or ring assumptions, nor replace the protocol analyses in those specifications.

What a complete extension still requires. A complete extension must track a prescribed sample budget, the exact error support and width, the secret distribution, and all public information jointly. It must also provide a decision reduction for the chosen modulus and, where relevant, handle different error widths within the same instance. Existing lossy code and Rényi divergence methods [**?**, **?**] offer possible approaches, but their quantitative hypotheses cannot be omitted. Along the direct route, one would first have to generate the uniform error samples from the lattice decoding instance, with the required joint independence and the error budget in (**??**); the public dither transformation can then be applied when its hypotheses hold. For decision hardness, the uniform branch and a suitable decision reduction must also be established. Reducing a perturbation enough to control the accumulated error can restrict the decoding radius and worsen the final lattice approximation factor. Whether this route, or a combination with existing methods, yields a reduction with small loss for the concrete integer lattice quantization parameters remains unresolved here. The present extension analysis identifies reusable steps and a way to control local boundary effects, but does not claim to resolve that question.

References

[1] Oded Regev. On lattices, learning with errors, random linear codes, and cryptography. *Journal of the ACM*, 56(6):34:1–34:40, 2009. https://doi.org/10.1145/1568318.1568324.

[2] Shanxiang Lyu, Ling Liu, and Cong Ling. Learning with quantization: A ciphertext efficient lattice problem with tight security reduction from LWE. *Cryptology ePrint Archive*, Paper 2024/714, 2024. https://eprint.iacr.org/2024/714.

[3] Chris Peikert. An efficient and parallel Gaussian sampler for lattices. In *Advances in Cryptology—CRYPTO 2010*, volume 6223 of *Lecture Notes in Computer Science*, pages 80–97. Springer, 2010. https://doi.org/10.1007/978-3-642-14623-7_5.

[4] Daniele Micciancio and Chris Peikert. Trapdoors for lattices: Simpler, tighter, faster, smaller. In *Advances in Cryptology—EUROCRYPT 2012*, volume 7237 of *Lecture Notes in Computer Science*, pages 700–718. Springer, 2012. https://doi.org/10.1007/978-3-642-29011-4_41.

[5] Erdal Arıkan. Channel polarization: A method for constructing capacity achieving codes for symmetric binary input memoryless channels. *IEEE Transactions on Information Theory*, 55(7):3051–3073, 2009. https://doi.org/10.1109/TIT.2009.2021379.

[6] Benny Applebaum, David Cash, Chris Peikert, and Amit Sahai. Fast cryptographic primitives and circular secure encryption based on hard learning problems. In *Advances in Cryptology—CRYPTO 2009*, volume 5677 of *Lecture Notes in Computer Science*, pages 595–618. Springer, 2009. https://doi.org/10.1007/978-3-642-03356-8_35.

[7] Zvika Brakerski, Adeline Langlois, Chris Peikert, Oded Regev, and Damien Stehlé. Classical hardness of learning with errors. In *Proceedings of the 45th Annual ACM Symposium on Theory of Computing (STOC 2013)*, pages 575–584. ACM, 2013. https://doi.org/10.1145/2488608.2488680.

[8] Abhishek Banerjee, Chris Peikert, and Alon Rosen. Pseudorandom functions and lattices. In *Advances in Cryptology—EUROCRYPT 2012*, volume 7237 of *Lecture Notes in Computer Science*, pages 719–737. Springer, 2012. https://doi.org/10.1007/978-3-642-29011-4_42.

[9] Joël Alwen, Stephan Krenn, Krzysztof Pietrzak, and Daniel Wichs. Learning with rounding, revisited: New reduction, properties and applications. In *Advances in Cryptology—CRYPTO 2013, Part I*, volume 8042 of *Lecture Notes in Computer Science*, pages 57–74. Springer, 2013. https://doi.org/10.1007/978-3-642-40041-4_4.

[10] Ido Tal and Alexander Vardy. How to construct polar codes. *IEEE Transactions on Information Theory*, 59(10):6562–6582, 2013. https://doi.org/10.1109/TIT.2013.2272694.
[11] Shanxiang Lyu, Ke Ma, and Junzuo Lai. MAMBA Frost: A lattice KEM from learning with quantization. *Cryptology ePrint Archive*, Paper 2026/2136, 2026. https://eprint.iacr.org/2026/2136.
[12] Institute of Commercial Cryptography Standards. Announcement on launching the Next Generation Commercial Cryptographic Algorithms Program (NGCC). February 5, 2025. https://www.niccs.org.cn/niccs/Notice/pc/content/content_1937428197396713472.html.
[13] Shanxiang Lyu, Ke Ma, Jiarui Deng, Junzuo Lai, Ling Liu, and Xianhui Lu. MAMBA Viper: Specification for a key encapsulation mechanism from module learning with quantization. NGCC first round candidate specification, April 28, 2026. https://www.niccs.org.cn/niccs/Proposal/Public-Key%20Cryptographic%20Algorithms/Round%201%20candidates/MAMBA-Viper.zip.
[14] Shanxiang Lyu, Hao Cheng, Tang Zhou, Ke Ma, Chengjing Tian, Junzuo Lai, Zheng Wang, Jinming Wen, Tao Yang, Zheng Qin, and Zhengchun Zhou. MAMBA NIKE: Specification for a noninteractive key exchange from ring learning with quantization. NGCC first round candidate specification, June 29, 2026. https://www.niccs.org.cn/niccs/Proposal/Public-Key%20Cryptographic%20Algorithms/Round%201%20candidates/MAMBA-NIKE.zip.
[15] Nico Döttling and Jörn Müller-Quade. Lossy codes and a new variant of the learning with errors problem. In *Advances in Cryptology—EUROCRYPT 2013*, volume 7881 of *Lecture Notes in Computer Science*, pages 18–34. Springer, 2013. https://doi.org/10.1007/978-3-642-38348-9_2.
[16] Shi Bai, Tancrède Lepoint, Adeline Roux-Langlois, Amin Sakzad, Damien Stehlé, and Ron Steinfeld. Improved security proofs in lattice based cryptography: Using the Rényi divergence rather than the statistical distance. *Journal of Cryptology*, 31(2):610–640, 2018. https://doi.org/10.1007/s00145-017-9265-9.
[17] Shanxiang Lyu, Ling Liu, Yong Li, and Linqi Song. Towards ciphertext efficient LWE: Learning with sampling over sublattices. In *2026 IEEE International Symposium on Information Theory (ISIT)*. IEEE, 2026. https://doi.org/10.1109/ISIT62367.2026.11653664.

Appendix A. Efficient implementation of polar quantization

The modulus is chosen by (**??**), and the Gaussian parameter is the integer $\sigma^2 = 9n$. We give the construction and sampling algorithms used in Theorem **??**. The sampling analysis is uniform over supplied nested information sets.

Proof of Theorem **??**. For an effective choice of q, compute a positive dyadic number v_n and set

$$\sigma \ln n \le v_n \le \tfrac{5}{4}\sigma \ln n, \qquad q = \min\{2^a : 2^a \ge 8v_n\}. \tag{A.1}$$

Approximations with guaranteed constant relative error to $\sqrt{n}$ and $\ln n$ suffice. The final comparison is between rational numbers, and $\sigma^2 = 9n$ is the exact Gaussian parameter.

Finite partition channels and Gaussian masses. Put $Q_\ell = 2^\ell$ and $\chi_{\sigma,Q}(a) = \mathbb{P}[E = a \bmod Q]$, where $E \leftarrow D_{\mathbb{Z},\sigma}$. Given the lower bits $a_i = X_i \bmod 2^{\ell-1}$, marginalizing the uniform higher bits in the distribution P gives

$$\mathbb{P}[Y_i = y_i \mid a_i, X_{\ell,i} = b] = \frac{Q_\ell}{q}\, \chi_{\sigma,Q_\ell}(y_i - a_i - 2^{\ell-1}b).$$

Hence $o_i = (y_i - a_i) \bmod Q_\ell$ is sufficient, and the layer uses m independent copies of the binary input memoryless channel

$$W_\ell(o \mid b) = \chi_{\sigma,Q_\ell}(o - 2^{\ell-1}b), \qquad o \in \mathbb{Z}_{Q_\ell}. \tag{A.2}$$

Its alphabet has size $Q_\ell \le q = \mathrm{poly}(n)$; the common factor above cancels in every conditional probability. These are the channels of [**?**, Section 4.2], conditional on the lower plane history.

All these masses can be enclosed in computed intervals of width 2^{-P} in $\mathrm{poly}(n, P)$ time. Indeed, truncate the integer Gaussian sum at

$$T = \left\lceil \sqrt{9n\big(P + \lceil \log_2(9n) \rceil + 10\big)} \right\rceil.$$

The omitted unnormalized mass is at most $(9n/(\pi T))e^{-\pi T^2/(9n)} < 2^{-P-8}$. Evaluate each retained exponential with absolute error at most $2^{-P-\lceil \log_2(2T+1) \rceil - 10}$, sum by residue class, and divide by the interval for the total mass. Its lower endpoint can be taken at least one, since the term at zero is exactly one. These operations give the claimed intervals, with room for final dyadic rounding. Exponentials and π are evaluated by convergent series with range reduction; their

arguments and precision are polynomially bounded. Square roots and the integer cutoff use rational or integer bisection.

Computing lower bounds for the Bhattacharyya parameters. Set

$$\tau = \lceil m^{1/4} \rceil, \quad \delta_0 = 2^{-\tau-1}, \quad C = \tau + \log_2 m + 8, \qquad \lfloor x \rfloor_C = 2^{-C} \lfloor 2^C x \rfloor.$$

For each channel compute a dyadic lower bound a_ℓ on $Z(W_\ell) = \sum_o \sqrt{W_\ell(o \mid 0) W_\ell(o \mid 1)}$, with $0 \le Z(W_\ell) - a_\ell \le 2^{1-C}$. Mass precision $2C + 2\ell_q + 12$, square root precision $C + \ell_q + 8$, and a final downward rounding to C bits suffice: use $|\sqrt{u} - \sqrt{v}| \le \sqrt{|u - v|}$ and sum over at most q terms. Propagate the lower bounds through the polar tree by

$$a^+ = \lfloor a^2 \rfloor_C, \qquad a^- = a.$$

These are valid by [**?**, Proposition 5]. Include an index j in the frozen set at level ℓ only when its resulting lower bound is at least $1 - \delta_0$. Finally replace each preliminary information set I'_ℓ by $I_\ell = \bigcup_{a \le \ell} I'_a$. This only removes indices from the frozen sets, and $\delta_0 \le 2^{-m^{1/4}}/2$ proves $\mathcal{Z}_{\ell j} \ge 1 - 2^{-m^{1/4}}$ for every $j \in F_\ell$.

Relative to the scalar tree rooted at the exact $Z(W_\ell)$, with maps $x \mapsto x^2$ and $x \mapsto x$ and no rounding, the error at a leaf is less than $3m2^{-C} < \delta_0/4$: the initial error is at most 2^{1-C}, squaring is Lipschitz with constant 2 on $[0, 1]$, and the depth is $\log_2 m$. The first level bound in Proposition **??** has deficit $e^{-\Omega(n)} < \delta_0/4$, so every computed first level lower bound is at least $1 - \delta_0$. Its last level bound makes every computed last level lower bound smaller than $1 - \delta_0$. Taking $I_\ell = \bigcup_{a \le \ell} I'_a$ preserves both endpoints. There are $O(\ell_q m)$ tree operations; all the preceding precisions, alphabet sizes, and Gaussian cutoffs are polynomial in n. The integer basis has m^2 entries of $O(\log q)$ bits. This proves the polynomial running time of Build.

Sampling conditional distributions. Fix any nested tuple $\mathcal{I} = (I_1, \ldots, I_{\ell_q})$ supplied as incidence bits to Sample. Form its coefficient prefixes and lattice by (**??**). The channel likelihoods depend on the test channel distribution and the generated history; bits with $j \in I_\ell$ are sampled, and bits with $j \in F_\ell$ receive their prescribed digits. Nestedness suffices for the digit relation in Lemma **??** and the equality of output distributions in (**??**). The following explicit lower bound holds for every entry of every channel:

$$W_\ell(o \mid b) \ge \frac{e^{-\pi q^2/(36n)}}{1 + 3\sqrt{n}} \ge 2^{-D}, \qquad D = \lceil q^2/n \rceil + \lceil \log_2(4n) \rceil + 2. \tag{A.3}$$

Choose a representative of absolute value at most $Q_\ell/2$ for the first inequality, and use $\rho_\sigma(\mathbb{Z}) \le 1 + \sigma$. Notice that $D = O((\ln n)^2)$. For the requested precision choose

$$P_{\mathrm{s}} = D + k + \lceil \log_2(64 \ell_q m^2) \rceil, \qquad P_{\mathrm{b}} = k + \lceil \log_2(4 \ell_q m) \rceil. \tag{A.4}$$

Compute intervals of width at most $2^{-P_{\mathrm{s}}-2}$ using the mass procedure above, take their midpoints, and round to the nearest P_{s} bit dyadic number in $[0, 1]$. The resulting positive weights $\widehat{W}_\ell$ have absolute entrywise error less than $2^{-P_{\mathrm{s}}}$. In particular

$$\left| \widehat{W}_\ell / W_\ell - 1 \right| \le \eta := 2^{D - P_{\mathrm{s}}} \le \frac{2^{-k}}{64 \ell_q m^2}.$$

The weights need not sum exactly to one; only their ratios in the conditional probabilities will be used.

At a given layer and raw SC prefix, the two candidate likelihoods L_0, L_1 are sums of products of m channel entries over all completions. They are strictly positive, and their unnormalized sum is at least 2^{-Dm}, since at least one completion exists. Compute the corresponding $\widehat{L}_0, \widehat{L}_1$ by the standard SC recursion [**?**, Theorem 5 and Section VIII], using likelihood pairs. The elementary maps on pairs $a = (a_0, a_1)$, $b = (b_0, b_1)$ are, up to common factors,

$$(a_0 b_0 + a_1 b_1,\ a_1 b_0 + a_0 b_1), \qquad (a_u b_0,\ a_{u \oplus 1} b_1),$$

where u is the already decided preceding bit. Use exact rational addition and multiplication throughout. Computing each requested pair separately uses $O(\ell_q m^2 \log m)$ operations. Every intermediate quantity is a positive polynomial of degree at most m, with at most 2^m monomials.

Its dyadic numerator and denominator have $O(m(P_{\mathrm{s}}+1))$ bits. The computation therefore uses polynomially many bit operations.

Positivity gives a uniform relative error bound. For either candidate, every monomial changes by a factor in $[(1-\eta)^m,(1+\eta)^m]$; the same holds for their sums. Since $m\eta \le 1/8$, normalization therefore gives

$$\left|\frac{\widehat{L}_1}{\widehat{L}_0+\widehat{L}_1}-\frac{L_1}{L_0+L_1}\right| \le 8m\eta.$$

Compute the first ratio as a rational number, round down to a multiple of $2^{-P_{\mathrm{b}}}$, and sample that Bernoulli distribution using exactly P_{b} fair bits and an integer comparison. Its conditional error is at most $8m\eta + 2^{-P_{\mathrm{b}}} \le 2^{-k}/(\ell_q m)$.

To sample the coefficient digits in Definition **??**, recover the already generated raw planes using the digit relation in Lemma **??**; compute a_i and o_i in (**??**). Within the current plane its carry is already known, so the preceding coefficient digits specify precisely one history of the raw bits. The current coefficient digit differs from the corresponding raw bit only by a known flip. Apply the preceding sampler to that raw bit and undo the flip. For $j \in F_\ell$, set the prescribed coefficient digit exactly. The information bit distributions are conditioned on the current history. Sequential coupling over at most $\ell_q m$ decisions proves (**??**), uniformly in all y, d and all supplied nested tuples. The precision choices and operation counts depend only on n, k, and reading and checking the incidence bits takes $O(\ell_q m)$ operations. Setting these frozen bits exactly and performing the final multiplication by $B_{\mathrm{pol}} \bmod q$ exactly keeps the output in the prescribed coset.

The conditional Bernoulli error is at most $2^{-k}/(\ell_q m)$ for every information bit with $j \in I_\ell$, uniformly over all admissible histories. For every $y, t \in \mathbb{Z}_q^m$ and $d \in \mathcal{D}$, the triangle inequality and Lemma **??** give

$$\Delta\big(\widetilde{K}^{(k)}_{d+\pi_\Lambda(t)}(y+t), \widetilde{K}^{(k)}_{d}(y)+t\big) \le 2^{1-k}.$$

When the supplied sets also satisfy (**??**), Lemma **??** bounds the corresponding joint statistical distances by

$$\mu_{\mathrm{pol}} + 2^{-k}, \qquad \mu_{\mathrm{pol}} = \ell_q m\sqrt{\ln 2}\, 2^{-m^{1/4}/2}, \tag{A.5}$$

for the implemented quantizer. This includes the estimate for every secret satisfying the condition in Lemma **??**. All computations and random bits are polynomial in n, k. □

A.1. **A grid subsequence.**

Proposition A.1. *For $n = 2^{4r}$ and all sufficiently large integers r, the algorithm* Build *above returns*

$$I_\ell = \varnothing \quad (\ell \le r+3), \qquad I_\ell = [m] \quad (\ell \ge r+4), \qquad \Lambda = 8n^{1/4}\mathbb{Z}^m.$$

Proof. For a power of two $Q \ge 2$, put

$$p_Q(o) = \chi_{\sigma,Q}(o), \qquad p'_Q(o) = p_Q(o - Q/2), \qquad Z_Q = \sum_{o\in\mathbb{Z}_Q}\sqrt{p_Q(o)p'_Q(o)}, \qquad \theta_Q = e^{-\pi\sigma^2/Q^2}.$$

Write $S_\sigma = \sum_{a\in\mathbb{Z}} e^{-\pi\sigma^2 a^2}$, so $1 \le S_\sigma \le 2$ for the parameters here. Poisson summation gives

$$p_Q(o) = \frac{1}{QS_\sigma}\sum_{k\in\mathbb{Z}} \theta_Q^{k^2} e^{2\pi i k o/Q}.$$

If $\theta_Q \le 1/10$, the bound $\sum_{k\ge1}\theta_Q^{k^2} \le \theta_Q/(1-\theta_Q)$ implies

$$p_Q(o), p'_Q(o) \ge \frac{1}{4Q}, \qquad |p_Q(o) - p'_Q(o)| \le \frac{5\theta_Q}{Q}.$$

Indeed, the difference contains only the odd Fourier terms. Consequently,

$$1 - Z_Q = \frac{1}{2}\sum_o \frac{(p_Q(o)-p'_Q(o))^2}{(\sqrt{p_Q(o)}+\sqrt{p'_Q(o)})^2} \le \frac{25}{2}\theta_Q^2 \le 16\theta_Q^2.$$

For a lower bound, summing the same Fourier series against $e^{2\pi io/Q}$ retains all indices $k \equiv -1 \pmod{Q}$, giving

$$\sum_o p_Q(o) e^{2\pi io/Q} = \frac{\sum_{a\in\mathbb{Z}} e^{-\pi\sigma^2(a-1/Q)^2}}{S_\sigma} \geq \frac{\theta_Q}{S_\sigma}.$$

The translation by $Q/2$ changes the sign of this Fourier coefficient, so $\Delta(p_Q, p'_Q) \geq \theta_Q/S_\sigma$. Cauchy Schwarz gives $\Delta(p_Q, p'_Q)^2 \leq 1 - Z_Q^2 \leq 2(1 - Z_Q)$, whence, for every Q,

$$1 - Z_Q \geq \theta_Q^2/8.$$

Now $m = n^2$, $\tau = \sqrt{n}$, and $\delta_0 = \frac{1}{2}e^{-(\ln 2)\sqrt{n}}$. Set $Q_0 = 8n^{1/4} = 2^{r+3}$. For $Q \leq Q_0$, the upper bound above gives

$$1 - Z_Q \leq 16e^{-(9\pi/32)\sqrt{n}}.$$

Every leaf value obtained by the updates $x \mapsto x^2$ and $x \mapsto x$, without rounding, is at least $Z_Q^m \geq 1 - m(1 - Z_Q)$. The preceding rounding analysis therefore bounds the deficit of every computed leaf by

$$16me^{-(9\pi/32)\sqrt{n}} + 3m2^{-C} < \delta_0$$

for sufficiently large n: the ratio of the first term to δ_0 tends to zero because $9\pi/32 > \ln 2$, while the ratio of the second is exactly $3/128$. Thus the frozen set is $[m]$ at each of these levels.

For $Q \geq 2Q_0$, the lower bound gives

$$1 - Z_Q \geq \tfrac{1}{8}e^{-(9\pi/128)\sqrt{n}} > \delta_0$$

for sufficiently large n, since $9\pi/128 < \ln 2$. The computed root lower bound is at most Z_Q, and each update $a^+ = \lfloor a^2 \rfloor_C$, $a^- = a$ produces a value at most its input. Every leaf at these levels therefore is smaller than $1 - \delta_0$, so the frozen sets are empty at these levels. Taking $I_\ell = \bigcup_{a\leq\ell} I'_a$ leaves the resulting sets unchanged, and $q > 2Q_0$ for sufficiently large n. Hence every $t_j = r + 3$; unimodularity of B_{pol} yields $\Lambda = 2^{r+3}\mathbb{Z}^m = 8n^{1/4}\mathbb{Z}^m$. □

Proposition A.2 (Output distribution on the grid subsequence)**.** *Assume the parameters and information sets of Proposition* **??***, and put $a = 8n^{1/4}$, so that $\Lambda = a\mathbb{Z}^m$. For $d \in \mathcal{D}$, choose an integer representative $d_0 \in \mathbb{Z}^m$. The ideal quantizer of Definition* **??** *satisfies, for every $y \in \mathbb{Z}_q^m$,*

$$\mathbb{P}[K_d(y) = x] = \frac{\chi_{\sigma,q}(y - x)}{\sum_{u\in(d_0+\Lambda)/q\mathbb{Z}^m} \chi_{\sigma,q}(y - u)} \quad (x \in (d_0 + \Lambda)/q\mathbb{Z}^m), \tag{A.6}$$

$$K_d(y) \overset{d}{=} y + E_d \pmod{q}, \qquad E_d \leftarrow D_{\Lambda+d_0-y,\sigma}.$$

The probability is zero outside the displayed coset. Here $D_{C,\sigma}$ assigns mass $\rho_\sigma(e)/\rho_\sigma(C)$ to each $e \in C$; any integer representative of y may be used.

Proof. Put $t = r + 3$, so $a = 2^t$. All coefficient bits at levels $1, \ldots, t$ are fixed by d, and every higher bit is sampled from its distribution under P_y, conditional on the preceding bits. The chain rule therefore gives exactly P_y conditioned on the entire fixed prefix. Since B_{pol} is unimodular, this prefix event is precisely $X \in (d_0 + \Lambda)/q\mathbb{Z}^m$. Equation (**??**) gives the first line of (**??**). Summing the weights $\rho_\sigma(e)$ over all integer lifts $e \equiv x - y \pmod{q}$ gives the numerator, up to the common normalization $\rho_\sigma(\mathbb{Z}^m)$. Summing over the whole output coset gives $\rho_\sigma(\Lambda + d_0 - y)$ with the same normalization. Their ratio proves the second line. Changes of representatives leave this coset unchanged because $q\mathbb{Z}^m \subseteq \Lambda$. □

For $d = 0$ and $y = As$, equation (**??**) is the LWS sampling prescription [**?**], reduced modulo q and expressed in our Gaussian width convention. More generally, on this same subsequence, the public seed distribution of Definition **??** uses $E \leftarrow D_{\Lambda+H(\rho)-A'S,\sigma}$, where an integer representative of $H(\rho)$ is understood. The seed and its distribution remain part of the public input. Identifying the zero dither output with LWS does not remove this public input or identify arbitrary dither distributions with zero dither. For the finite precision implementation, (**??**) bounds the distance from (**??**) by 2^{-k}, uniformly in y, d, and hence by $N2^{-k}$ for N samples.

Remark A.3 (The lattice does not determine the quantizer). Proposition **??** describes random sampling on the coset $a\mathbb{Z}^m + d_0 - y$. It does not describe nearest neighbor rounding onto $a\mathbb{Z}^m + d_0$. In particular, the integer lattice geometry in Proposition **??** does not change this distribution into the interval uniform error distribution in (**??**). The latter arises from a different, deterministic quantization rule with uniform dither. Consequently this subsequence alone does not extend our worst case theorem to the integer lattice instantiation of MAMBA Frost.

Appendix B. Sample complexity and statistical distance

We give concrete bounds for the algorithms in Sections **??** and **??**. Each sample supplied to W contains m rows.

B.1. **Sample, call, and success bounds.** Let N bound the complete m row LWQ samples consumed in one invocation of W, and retain e_W from (**??**). For a finite instance suppose

$$a_0 = \varepsilon_W - e_W > 0.$$

Given $0 < \eta < 1/2$, the procedure $\mathcal{F}_\alpha^W$ may use

$$K_{\text{trial}} = \left\lceil \frac{\ln(2/\eta)}{a_0} \right\rceil, \qquad V = \lceil 20000 \ln(2K_{\text{trial}}/\eta) \rceil \tag{B.1}$$

trials and fresh validation observations per trial. Indeed, the cosine statistic lies in $[-1, 1]$. The 0.02 threshold, the 0.001 evaluation error, and Hoeffding's inequality give validation error at most $e^{-V/20000}$. The failure probability is at most

$$e^{-a_0 K_{\text{trial}}} + K_{\text{trial}} e^{-V/20000} \le \eta,$$

in addition to the error of the validation arithmetic or source sampler. The following are padded upper bounds; early acceptance may use fewer calls and observations:

$$\#W \text{ calls} \le K_{\text{trial}}, \qquad H_F = K_{\text{trial}}(mN + V) \quad \text{scalar observations.}$$

Repeated instances use independent samples, seeds, matrices M, and work registers.

For the procedure $\mathcal{N}_\alpha^W$, choose $n^c \ge H_F$. There are $L_{\text{grid}} = n^{2c} + 1$ grid values. If each is repeated R_{grid} times and each resulting candidate uses V_{out} fresh original observations for verification, then the algorithm of Lemma **??** has the bounds

$$K_0 = L_{\text{grid}} R_{\text{grid}} K_{\text{trial}}, \qquad H_0 = L_{\text{grid}} R_{\text{grid}} (H_F + V_{\text{out}}). \tag{B.2}$$

Here $R_{\text{grid}} = O(n)$ and $V_{\text{out}} = O(n + \log L_{\text{grid}})$ suffice for exponentially small failure probability on independent $\mathcal{C}_{s,\beta}$ observations. At a grid value satisfying the inequalities in Lemma **??**, the complete input to $\mathcal{F}_\alpha^W$ has the additional distance $(9/2)n^{-c}$. Taking the fixed width error at most $1/4$ and $n^c \ge 18$ gives success at least $1/2$ at this grid value. The lattice construction contributes at most $4H_0\varepsilon$ to the statistical distance of the output of Lemma **??**.

Equivalently, at such a grid value the error bound for an individual N sample call to W, before validation, can be written as

$$\begin{aligned} &N(8\zeta + \mu) + p^{-(n+1)} + \nu_{\text{impl}} \\ &\quad + 4mN\varepsilon + \tfrac{9}{2} mN n^{-2c}. \end{aligned} \tag{B.3}$$

The last term is omitted when the input width is exactly α. The term $p^{-(n+1)}$ is omitted if W is promised a uniformly chosen S satisfying $\gcd(S_1, \ldots, S_{n+1}, q) = 1$. Lemma **??** gives the bound (**??**) for every fixed source secret; the term $p^{-(n+1)}$ concerns conditioning the target secret on $\gcd(S_1, \ldots, S_{n+1}, q) = 1$.

For the stated negligible error in (**??**) and negligible implementation errors, e_W is negligible for every polynomial N. Concrete parameters must satisfy $e_W < \varepsilon_W$.

Total resources of the DGS construction. For all $3n$ widths, choose common padded upper bounds K_0, H_0 as in (**??**), and a common bit bound $b_{\mathrm{enc}} \geq \log_2 |\mathcal{T}|$. With

$$C = \left\lceil 18(b_{\mathrm{enc}} + 2n^2) \ln 2 \right\rceil, \qquad P = CH_0,$$

Lemma **??** applies with $\delta = \eta = 2^{-n^2}$. Each modular decoding call, with the work registers restored to zero, uses at most $2CK_0$ executions of the circuit for W or its inverse. Regev's reduction from CVP to its modular version and subsequent sampling step use $O(n)$ such calls per Gaussian output. The total number of W or $W^\dagger$ circuit executions is therefore bounded by

$$O\big(nCK_0((3n-1)P + J)\big). \tag{B.4}$$

The initial wide Gaussian sampler does not use W.

B.2. **Errors for the polar quantizer.** Consider the quantizer of Corollary **??**, with its fixed precision $k(n) = n$. Let N bound the number of complete samples used in one invocation of W, and let ε_W be its success probability on the stated LWQ distribution. For every fixed source secret, the corresponding trial in Lemma **??** succeeds with probability at least

$$\varepsilon_W - N\big(8\zeta + \mu_{\mathrm{pol}} + 2^{-k(n)}\big) - 2^{-(n+1)} - \nu_{\mathrm{other}}. \tag{B.5}$$

Here $\nu_{\mathrm{other}} = \mathrm{negl}(n)$ is the sum of the implementation and stopping errors of the remaining operations. The term $N2^{-k(n)}$ accounts for the polar quantizer. The term $2^{-(n+1)}$ is omitted when W is promised a uniformly chosen S satisfying $\gcd(S_1, \ldots, S_{n+1}, q) = 1$.

To obtain (**??**), condition on the public seed and every preceding sample request. Theorem **??** gives statistical distance at most $2^{-k(n)}$ for the next output. Replacing the conditional distributions one at a time and applying the triangle inequality bounds the statistical distance of all inputs jointly by $N2^{-k(n)}$. Apply Lemma **??** to the ideal quantizer. Randomized rounding and Lemma **??** contribute $N(8\zeta + \mu_{\mathrm{pol}})$, and conditioning the target secret on $\gcd(S_1, \ldots, S_{n+1}, q) = 1$ contributes $2^{-(n+1)}$. The remaining implementation errors give ν_{other}.

For a finite choice of the parameters and a precision k, allocating at most $\epsilon_{\mathrm{tot}}/8$ to each of the two polar errors is ensured by

$$N\ell_q m\sqrt{\ln 2}\, 2^{-m^{1/4}/2} \leq \epsilon_{\mathrm{tot}}/8, \qquad N2^{-k} \leq \epsilon_{\mathrm{tot}}/8. \tag{B.6}$$

For the family of Corollary **??**, both errors are negligible for every polynomial N.

The numerical illustration in [**?**, Figure 3] uses $\sigma = 3$, eight polar levels, and $m = 2^{20}$. These parameters differ from the asymptotic family in Proposition **??**. In particular, the width condition $\sigma^2 > 4n + h^2$ requires a larger σ for $n \geq 3$. A finite parameter choice must satisfy this width condition and the error bounds above.

B.3. **Statistical distance of the output.** In Table **??**, N counts complete samples in one invocation of W, H_0 counts scalar observations used by the algorithm of Lemma **??**, and J_i is the output count of one sampling step. Statistical distances concern joint distributions of all inputs or outputs.

The first four rows establish a constant success probability for the algorithm of Lemma **??** before applying Lemma **??**; the candidate verification bounds for $\mathcal{F}_\alpha^W$ are given in (**??**). Write a_i for the joint error of the input samples at width r_i. Data processing includes it once per sampling step. Thus the final induction in the proof of Theorem **??** is summarized by

$$a_{i-1} \leq a_i + 2\eta_i + J_i\xi_i, \qquad a_0 \leq a_{3n} + \sum_{i=1}^{3n}(2\eta_i + J_i\xi_i) = \mathrm{negl}(n). \tag{B.7}$$

Here $a_{3n} = 2^{-\Omega(n)}$, the output counts are polynomial, and $\delta_i = \eta_i = 2^{-n^2}$. The statistical distance of the input samples is included through the complete sampling step.

Stage	Bound	Unit and role
Rounding and distance between LWQ and LWE	$N(8\zeta+\mu)+\nu_{\mathrm{impl}}$	One N sample call to W, including its public seeds; (**??**). The implemented polar quantizer's $N2^{-k(n)}$ error is included once in ν_{impl}.
Target secret distribution	$p^{-(n+1)}$	Once per invocation of W, as in (**??**); zero if S is uniform subject to $\gcd(S_1,\dots,S_{n+1},q)=1$.
Lattice generated observations	$4H_0\varepsilon$	One execution of Lemma **??**. For a single call to W, the corresponding term is $4mN\varepsilon$; (**??**) and Section **??**.
Noise adjustment	$\frac{9}{2}n^{-c}\le\frac{1}{4}$	One full $\mathcal{F}_\alpha^W$ execution at the grid value used in Lemma **??**, with $n^c\ge 18$, before candidate verification and repetition; (**??**) and Lemma **??**.
Choice of samples	η, plus η for overflow	One list of samples at one width, simultaneously for all promised finite queries; (**??**) and Lemma **??**. This probability concerns all outputs.
Reversible decoding	$2\sqrt{\delta}$, then $O(n\sqrt{\delta})$	State norm error of one modular call with work registers restored to zero, conditional on the event in Lemma **??**; $O(n)$ calls give the full decoding error. Lemma **??** and (**??**).
One sampling step	$2\eta_i+J_i\xi_i$	Joint distance of all J_i outputs, where $\xi_i=2^{-\Omega(n)}+O(n\sqrt{\delta_i})+\xi_{\mathrm{impl},i}$ is (**??**) at this width; Lemma **??**.

TABLE 1. Errors at their respective stages of the reduction. The bound for the choice of samples is a failure probability, and the reversible decoding bound is a state norm bound; the sampling bounds are statistical distances.

APPENDIX C. AN EXAMPLE OF THE CARRY CORRECTION

Remark C.1 (Two levels of quantization). Take $m=2$, $q=4$, and the standard column vector convention

$$B_{\mathrm{pol}}=G_2^T=\begin{pmatrix}1&1\\0&1\end{pmatrix}.$$

Write $U_1=(a,b)^T$, $U_2=(c,e)^T$, with all four entries binary; level 1 is processed before level 2, and coordinate 1 before coordinate 2. The ordinary bit planes give, modulo four,

$$\begin{aligned} X&=\big((a\oplus b)+2(c\oplus e),\ b+2e\big)^T,\\ z=B_{\mathrm{pol}}^{-1}X&=\big(a+2(c\oplus ab),\ b+2e\big)^T. \end{aligned}\tag{C.1}$$

Thus the coefficient planes are $z_{\mathrm{low}}=U_1$ and $z_{\mathrm{high}}=U_2\oplus(ab,0)^T$: the carry depends on both lower plane bits.

Choose the nested sets $I_1=I_2=\{2\}$. Then $t_1=2,t_2=0$, so

$$\Lambda=B_{\mathrm{pol}}\,\mathrm{diag}(4,1)\mathbb{Z}^2=\{(4h+j,j)^T:h,j\in\mathbb{Z}\},\qquad \pi_\Lambda(X)=X_1-X_2 \bmod 4=z_1\in\mathbb{Z}_4.$$

The frozen bits of U_1,U_2 are (a,c). Fix them to $(1,0)$; the following two outcomes lie in different cosets. Both have positive probability for every y, since P_y is positive everywhere. Pairs in the table denote column vectors.

U_1	U_2	X	z	$\pi_\Lambda(X)$
$(1,0)$	$(0,0)$	$(1,0)$	$(1,0)$	1
$(1,1)$	$(0,0)$	$(0,1)$	$(3,1)$	3

In particular, the vector formed from these frozen bits, $B_{\mathrm{pol}}(a+2c,0)^T=(1,0)^T$ does not specify the output coset: the information bit b in the first level changes the output coset.

To prescribe the coset determined by $\delta=\delta_0+2\delta_1\in\mathbb{Z}_4$ with binary digits δ_0,δ_1, instead set $a=\delta_0$ at level 1 and, after generating b, set $c=\delta_1\oplus\delta_0 b$ at level 2. Equation (**??**) then gives

$z_1 = \delta$ on every outcome. For $\delta = 1$, this rule is $c = b$; choosing $(b, e) = (1, 0)$ yields $X = (2, 1)^T$ and $z = (1, 1)^T$.

For any integer translation v, the residue $z_1 \bmod 4$ changes by $v_1 - v_2 \bmod 4$. In particular, $v = (1, 1)^T = B_{\mathrm{pol}}(0, 1)^T \in \Lambda$ sends $X = (1, 0)^T$ to $(2, 1)^T$ and preserves $z_1 = 1$. The pair of frozen bits (a, c) changes from $(1, 0)$ to $(1, 1)$, exactly as the correction $c = b$ requires. This is the translation in Lemma **??**: $K_1(y + v) \overset{d}{=} K_1(y) + v$, with the same output coset even though the pair of frozen bits changes. The statistical distance bound (**??**) additionally requires the inequalities for indices in the frozen sets in (**??**); the calculation here establishes the relation between the carries and the coset.

Appendix D. Quantization boundaries and perturbed inputs

A small perturbation before deterministic quantization can change the output distribution by a constant amount. The following example separates this boundary effect from the statistical distance assumption (**??**) for uniform dither.

Proposition D.1 (A boundary discontinuity)**.** *Let $k \geq 2$ be even, let $k \mid q$, and assume $q > k$. Take $\Lambda = k\mathbb{Z}$ and the deterministic nearest neighbor quantizer*

$$Q_\Lambda(y) = k \left\lfloor \frac{y}{k} + \frac{1}{2} \right\rfloor, \qquad Q_{\Lambda+d}(y) = Q_\Lambda(y - d) + d,$$

with d uniform in $\{0, \ldots, k-1\}$. Put

$$E = \{-k/2 + 1, \ldots, k/2\} \bmod q, \qquad P = \mathsf{U}(E).$$

The distribution of $Q_{\Lambda+d}(0) \bmod q$ is P. For every $-1 < u < 0$,

$$Q_{\Lambda+d}(u) \bmod q \sim \mathsf{U}(E - 1), \qquad \Delta(\mathsf{U}(E - 1), P) = \frac{1}{k}.$$

If $0 < t < 1$, H is uniform on $\{-t, t\}$, and F is the distribution of $Q_{\Lambda+d}(H) \bmod q$ with independent d, then for every $N \geq 1$,

$$\Delta(F, P) = \frac{1}{2k}, \qquad \Delta(F^{\otimes N}, P^{\otimes N}) = 1 - \left(1 - \frac{1}{2k}\right)^N. \tag{D.1}$$

Proof. At zero, the possible outputs are precisely the k elements of E, each with probability $1/k$. For $-1 < u < 0$, only the tie at $d = k/2$ changes side: the output $k/2$ becomes $-k/2$. The resulting set is $E - 1$. The endpoints are distinct modulo q because $q > k$, proving the first distance. For $0 < t < 1$ no output changes from its value at zero. Therefore

$$F = \tfrac{1}{2}\mathsf{U}(E) + \tfrac{1}{2}\mathsf{U}(E - 1).$$

On E, the mass of F is pointwise at most that of P, and $F(E) = 1 - 1/(2k)$. Thus $F^{\otimes N}$ is pointwise at most $P^{\otimes N}$ on E^N, and the remaining mass lies outside the support of $P^{\otimes N}$. Its mass is exactly the right hand side of (**??**). □

Even an arbitrarily small negative perturbation therefore has distance $1/k$, and symmetrizing the perturbation leaves distance $1/(2k)$. For N comparable with k, the joint distance is already a constant.

The same issue can occur inside a lattice Gaussian projection. Take $L = \mathbb{Z}^n$, $x = te_1$, $0 < t < 1/2$, so $\kappa_{L^*}(x) = 0$. Let $v \leftarrow D_{\mathbb{Z}^n, R}$, $a = v \bmod q$, and choose τ uniformly in $\mathbb{Z}_q^n$ as the secret for the simulated samples. Direct quantization would give

$$\widetilde{b} = Q_{\Lambda+d}(\langle a, \tau \rangle + tv_1) \bmod q.$$

In the joint distribution of secret and sample consider the event

$$b - \langle a, \tau \rangle \equiv -k/2 \pmod q.$$

Its probability in a true LWQ sample is zero. Conditional on any v with $-1 < tv_1 < 0$, its probability under direct quantization is $1/k$. To see this, translate the uniform dither by the

known integer $\langle a, \tau \rangle$ modulo k, and apply the preceding proposition. Consequently the joint distance is at least

$$\frac{1}{k}\mathbb{P}[-1 < tv_1 < 0].$$

Choose R above the relevant smoothing thresholds and $t = 1/(Rn)$. Symmetry and the discrete Gaussian tail bound of [**?**, Lemma 2.5] give

$$\Delta(\text{simulated distribution}, \text{LWQ distribution}) \geq \frac{1 - D_{\mathbb{Z},R}(0)}{2k} - \text{negl}(n).$$

For a wide Gaussian and polynomial k, this is inverse polynomial, not negligible. The target displacement can be very small while the boundary effect persists.

Thus a direct simulation that perturbs the linear input before deterministic quantization must account for boundary crossings in its joint output distribution. This observation concerns that simulation, not all possible reductions to LWE with interval uniform errors.

D.1. **Uniform addition before quantization.** The following elementary estimate shows that the boundary discontinuity can be removed locally without changing the quantized output at any integer input. The uniform addition is part of a possible simulator; it is not a change to the target integer input distribution.

Lemma D.2 (Uniform addition before scalar quantization)**.** *Let $k \geq 2$ be even, let $k \mid q$, and define*

$$Q_d(x) = k\left\lfloor \frac{x-d}{k} + \frac{1}{2} \right\rfloor + d \pmod q \qquad (x \in \mathbb{R},\ d \in \{0, \dots, k-1\}).$$

Let Z be integer valued, let δ be real valued, and let C be an auxiliary random variable. Their joint distribution may be arbitrary. Draw d uniformly from $\{0, \dots, k-1\}$ and V uniformly from $[0, 1)$, independently of one another and of (Z, C, δ). Then $Q_d(Z + V) = Q_d(Z)$ almost surely, and

$$\Delta\big((C, d, Q_d(Z + \delta + V)), (C, d, Q_d(Z))\big) \leq \mathbb{E}\min\left\{\frac{|\delta|}{k}, 1\right\}. \tag{D.2}$$

The bound also holds with Z and δ retained in both tuples.

For T scalar observations, let $(\mathbf{Z}, C, \boldsymbol{\delta})$ have any joint distribution, with $\mathbf{Z} \in \mathbb{Z}^T$ and $\boldsymbol{\delta} \in \mathbb{R}^T$. Draw all d_i, V_i independently as above, and independently of this tuple. Put

$$\mathbf{B} = (Q_{d_i}(Z_i))_{i=1}^T, \qquad \widetilde{\mathbf{B}} = (Q_{d_i}(Z_i + \delta_i + V_i))_{i=1}^T.$$

Then

$$\Delta\big((C, \mathbf{d}, \widetilde{\mathbf{B}}), (C, \mathbf{d}, \mathbf{B})\big) \leq \min\left\{1, \sum_{i=1}^T \mathbb{E}\min\left\{\frac{|\delta_i|}{k}, 1\right\}\right\}. \tag{D.3}$$

This bound also holds with $\mathbf{Z}, \boldsymbol{\delta}$ retained.

For an integer $h \geq 0$, replacing each V_i by a uniform choice from $\{j2^{-h} : j = 0, \dots, 2^h - 1\}$ preserves the identity at integer inputs. A sufficient scalar bound is then $\mathbb{E}\min\{(|\delta| + 2^{-h})/k, 1\}$, and a sufficient joint bound is

$$\min\left\{1, \sum_{i=1}^T \mathbb{E}\min\left\{\frac{|\delta_i|}{k}, 1\right\} + \frac{T}{k2^h}\right\}. \tag{D.4}$$

Proof. Condition on $Z = z, C = c, \delta = u$. For each d, write

$$j_d = \left\lfloor \frac{z-d}{k} + \frac{1}{2} \right\rfloor, \qquad r_d = z - d + \frac{k}{2} - kj_d.$$

Since k is even, r_d is an integer in $\{0, \dots, k-1\}$. For $0 \leq V < 1$, therefore,

$$Q_d(z + V) = kj_d + d = Q_d(z) \pmod q,$$

whereas

$$Q_d(z + u + V) = kj_d + d + k\left\lfloor \frac{r_d + V + u}{k} \right\rfloor \pmod q.$$

As d ranges uniformly over its domain, r_d ranges uniformly over the same integers. Thus $R = r_d + V$ is uniform on $[0, k)$, even conditional on (z, c, u). Generate the two outputs with the same d, V. They agree whenever $0 \le R + u < k$, so their probability of disagreement is at most

$$\mathbb{P}[R + u \notin [0, k)] = \min\{|u|/k, 1\}.$$

Reduction modulo q can only decrease this probability. The coupling bound for statistical distance and averaging prove (**??**), including the retained variables. For the vector statement, condition on the entire tuple $(\mathbf{Z}, C, \boldsymbol{\delta})$, use the same coupling at each coordinate, and apply a union bound. The perturbations need not be independent of one another.

For the finite choice of V, set $M = 2^h$. Conditional on (z, c, u), the variable R is uniform on $\{j/M : j = 0, \ldots, kM - 1\}$. An interval of length $|u|$ contains at most $M|u| + 1$ such grid points. For $|u| \le k$, this bounds the probability of crossing the appropriate endpoint by $(|u| + M^{-1})/k$; for larger $|u|$, use the bound one. The scalar coupling and the vector union bound give the asserted finite bounds, using $\min\{(x + M^{-1})/k, 1\} \le \min\{x/k, 1\} + 1/(kM)$. □

The independence conditions in this lemma are essential: the estimate averages over fresh uniform dither and does not assert the same bound for each fixed dither or for an arbitrary prescribed dither function. The variable C allows other public information, and even the secret, to be retained, provided the stated joint independence holds. The finite addition uses exactly h unbiased bits; evaluating the quantizer on an encoded real input still requires a separate precision analysis.

If $\mathbb{E}|\delta_i| \le \varepsilon$ for all i, the continuous bound is at most $T\varepsilon/k$. Here T counts scalar observations: N complete samples with m rows require $T = Nm$, before accounting for any additional calls in a reduction. Thus the lemma removes the discontinuity exhibited above but not the dependence on the sample count. A reduction would have to keep this accumulated error small relative to the solver's success or distinguishing advantage, establish the required distributions and independence, and track the effect of a smaller allowable perturbation on the decoding radius. The lemma does not establish such a reduction, its lattice approximation factor, or security for the parameters of Frost or Viper.