CERTIFIED CONDITIONAL PROFILE OPTIMA FOR BLOCK REDUCTION, AND THE KYBER ROUND-3 PRIMAL CHAIN RE-EVALUATED

SEAN WARD AND MAESTRO

Abstract. The parameter chains of lattice-based schemes model the Gram–Schmidt profile of a BKZ-reduced basis by the geometric series assumption, a straight line whose slope is fixed by the blocksize. We replace the line by the class of profiles satisfying, at every projected block, the linear inequality that the floor side of the Gaussian heuristic imposes, and show that over this class the head $\log \|b_1\|$ and every prefix volume have exact extremal values, given by a linear programme whose dual multipliers are positive rationals computed in $O(d)$ exact operations and whose value is enclosed by interval arithmetic; the extremal profile is the unique all-tight one. Inserted into our reconstruction of the round-3 Kyber specification's primal-attack condition in place of the GSA line, the extremal profile moves the crossing blocksize from 406/624/874 to 417/642/900, each certified least in its scanned interval; the crossings coincide with those of the converged Chen–Nguyen simulator, whose fixed points are the profiles obeying the reverse block inequalities and whose limit is, up to its tail convention, the class's extremal profile clipped at the q-ary ceiling, move only to 415/639/897 under a uniform slack of 0.01 in every inequality, and by at most one blocksize under a slack at the head alone eight times the largest deficit measured. This is a statement about a class of profiles, not about any attack: whether the outputs of block reduction belong to the class is open, and at the sizes where full enumeration is feasible they leave it at the head, by a certified amount that the paper prices exactly. The paper proves no lower bound on the cost of any attack.

Contents

1. Introduction 2
2. The profile class and the linear programme 3
2.1. Setting 3
2.2. The head floor 4
2.3. Positivity of the certificate 4
2.4. The value 5
3. The prefix-volume floor 5
4. Per-basis identities 6
5. Certificates 8
6. The round-3 primal chain re-evaluated 8
6.1. The specification's chain and its reconstruction 8
6.2. Three forms of the re-evaluation 9
6.3. Simulators, the q-ary ceiling and the dual route 11
6.4. Provenance of every blocksize triple 14
7. What block reduction actually does: an applicability audit 15
8. Related work 18
9. Escapes and limitations 19
Appendix A. The identities 19

Date: 1 September 2026.

Appendix B. The two-sided floor 20
Appendix C. The idealised recurrence 21
Appendix D. Reproducibility 21
References 21

1. Introduction

The security estimates of the lattice schemes whose parameters were later standardised in FIPS 203 [6] were argued, in the round-3 Kyber specification [10], through a chain of conventions: an attack is modelled as a block reduction of blocksize β on a lattice of dimension d; the reduced basis's Gram–Schmidt profile is modelled by the geometric series assumption (GSA) – a straight line in $\ell_i := \log \|b_i^*\|$ whose slope is fixed by β through the Gaussian heuristic; the attack succeeds when a planted short vector is detected against that line; and the reduction is assigned the conventional core-SVP exponent 0.292β [1]. Each link is a heuristic, and the second is known to be inaccurate in recognised ways: the profile of a BKZ-reduced basis has a concave head and an upturned tail [4, 3]. The parameter chains absorb these effects into the blocksize by simulation. This paper studies that one historical primal-estimation chain; it does not audit the current estimators, whose attack models and conventions differ.

The paper concerns the second link only, and asks what the Gaussian heuristic, taken as an inequality at every block, forces on the profile. Four statements must be kept apart. (a) A theorem about a restricted class of profiles: among all profiles satisfying the block inequalities, the head and every prefix volume have exact extremal values, characterised by a linear programme. (b) Its certified optimum at given (d, β, ε): an exact rational dual certificate and an interval enclosure, so that a comparison with a target is a rigorous decision. (c) An observation about real block reduction at small sizes: whether the bases that tours of unpruned enumeration produce belong to the class, measured where full enumeration is feasible. (d) A statement about the security of a deployed scheme. This paper proves (a), computes (b), reports (c), and makes no statement of kind (d). Its headline is of kind (b): in the round-3 specification's primal-attack condition, with the class's extremal profile in place of the GSA line, the crossing blocksize on the scanned range moves from 406/624/874 to 417/642/900 – each certified least in its scanned interval, a difference of 3.2, 5.3 and 7.6 in the conventional core-SVP exponent 0.292β; the crossings move only to 415/639/897 under a uniform slack of 0.01 in every inequality (a double-precision scan, certified at Kyber512 at both slacks and at Kyber768 at 0.01) and by at most one blocksize under a head slack of 1. The same crossings 417/642/900 are what the converged Chen–Nguyen simulator returns (Section 6.3): its fixed points are exactly the profiles obeying the reverse block inequalities, and the profile it reaches from the standard start is, up to its tail convention, the class's extremal profile clipped at the q-ary ceiling; so the shift over the GSA chain is the GSA-versus-simulator discrepancy, here with a certificate that no profile obeying the inequalities does better. Its principal limitation is that membership of block-reduction outputs in the class is open, and fails at the head, by a certified amount, at the small sizes where it can be examined.

The results are as follows. Section 2: the class, the linear programme, the head floor (Theorem 2.2), the exact positive dual (Theorem 2.3) and the identities that express the extremal head as the GSA head plus a tail correction $\kappa_{d,\beta}$. Section 3: every prefix volume is bounded below by the extremal profile's (Theorem 3.1), while no single entry beyond the head is bounded at all. Section 4: exact identities for every basis expressing the head, and every prefix sum, as the extremal value minus a weighted sum of the blocks' deficits. Section 5: what is certified and how. Section 6: the round-3 chain reproduced, re-evaluated with the floor in three forms, and a provenance table. Section 7: the applicability audit. Sections 8 and 9: related work and the operations the model forbids. Appendices give the identities' proofs, the two-sided (primal and dual) floor as a diagnostic, the idealised recurrence whose fixed point the extremal profile is, and the reproducibility statement.

2. The profile class and the linear programme

2.1. **Setting.** Let $L \subset \mathbb{R}^d$ be a lattice with basis $b_1, \dots, b_d$, Gram–Schmidt vectors b_i^*, and $\ell_i := \log \|b_i^*\|$; $S := \sum_i \ell_i = \log \operatorname{vol} L$. For a lattice M write $\lambda_1(M)$ for the length of its shortest nonzero vector. Let π_i be the orthogonal projection onto $\operatorname{span}(b_1, \dots, b_{i-1})^\perp$, fix $2 \le \beta \le d$, put $\beta_i := \min(\beta, d - i + 1)$, and let

$$B_i := \sum_{j=i}^{i+\beta_i-1} \mathbb{Z}\,\pi_i(b_j) \qquad (1 \le i \le d-1)$$

be the projected block at i; for $i > d - \beta + 1$ it is the tail projection $\pi_i(L)$, and we call the positions $i > d-\beta+1$ the *tail*. Then $b_i^* = \pi_i(b_i) \in B_i \setminus \{0\}$ and $\operatorname{vol} B_i = \prod_{j=i}^{i+\beta_i-1} \|b_j^*\|$. Write V_n for the volume of the unit n-ball, $\hat{c}_n := V_n^{-1/n}$ for the dimensionless Gaussian-heuristic constant ($\hat{c}_1 = 1/2$), so that the Gaussian-heuristic length of an n-dimensional lattice M is $\mathrm{GH}(M) := \hat{c}_n \operatorname{vol}(M)^{1/n}$, and $L(n) := \log \hat{c}_n$. Note that $\hat{c}_n < 1$ for $n \le 12$, and that $f(n) := L(n)/(n-1)$ attains its maximum over the integers at $n = 36$ ($f(35) = 0.0125299$, $f(36) = 0.0125320$, $f(37) = 0.0125252$, computed exactly to the digits shown; the often-quoted value near 46 arises from the leading term of $L(n)$ only). Let e_1 be the first standard basis vector of $\mathbb{R}^d$, $\mathbf{1}$ the all-ones vector, and $\mathbf{1}_{\le m}$ the indicator of the first m coordinates.

Definition 2.1 (admissibility)**.** For $\varepsilon \ge 0$ the basis is *block-*(β, ε)*-admissible* if

$$\lambda_1(B_i) \;\ge\; e^{-\varepsilon}\, \hat{c}_{\beta_i}\, \operatorname{vol}(B_i)^{1/\beta_i} \qquad (1 \le i \le d-1).$$

Since b_i^* is a nonzero vector of B_i, block-admissibility implies the linear constraints on the profile

$$(A_i) \qquad \Big(1 - \frac{1}{\beta_i}\Big)\ell_i - \frac{1}{\beta_i} \sum_{j=i+1}^{i+\beta_i-1} \ell_j \;\ge\; c_i := L(\beta_i) - \varepsilon.$$

A vector $\ell \in \mathbb{R}^d$ satisfying $(A_1), \dots, (A_{d-1})$ is a *linearly* (β, ε)*-admissible profile*; the set of such profiles of volume S is the *profile class*. The profiles of block-admissible bases lie in the profile class; the converse is not asserted – a profile does not determine the block minima – and every extremal statement below is a statement about the profile class, the linear relaxation.

The inequality is the floor side of the Gaussian heuristic, imposed at every block and every tail, with a uniform slack ε. It is not the Gaussian heuristic, which is a distributional statement about random lattices: for a random n-dimensional lattice the number of pairs $\pm v$ of length at most $t \cdot \mathrm{GH}$ is asymptotically Poisson of mean $t^n/2$ [3, §3], so a block's minimum is *shorter* than its GH length with probability about $1 - e^{-1/2}$. Zero slack is an extremal convention, and the class is to be read with the ε-sensitivity reported below. With indices increasing from head to tail, the GSA line of blocksize β has slope $-2L(\beta)/(\beta-1)$; it satisfies every full-size constraint with equality, and a shrinking tail constraint of size $n < \beta$ with left side $(n-1)f(\beta)$ against $(n-1)f(n)$, i.e. iff $f(\beta) \ge f(n)$. Hence for $\beta \le 36$ the GSA line lies in the zero-slack class, and for $\beta > 36$ it violates the tail constraints of the sizes n with $f(n) > f(\beta)$; the class is in any case much larger than the line.

2.2. **The head floor.** Let $a_i \in \mathbb{R}^d$ be the coefficient vector of the left side of (A_i).

Theorem 2.2 (head floor)**.** *There is a unique solution $(y, z) \in \mathbb{R}^{d-1} \times \mathbb{R}$ of $e_1 = \sum_{i=1}^{d-1} y_i a_i + z\mathbf{1}$, and $z = 1/d$. If $y_i \geq 0$ for every i, then every linearly (β, ε)-admissible profile of volume S satisfies*

$$\ell_1 \;\geq\; \frac{S}{d} + h_{d,\beta}(\varepsilon), \qquad h_{d,\beta}(\varepsilon) := \sum_{i=1}^{d-1} y_i\, c_i = \sum_i y_i L(\beta_i) - \varepsilon \sum_i y_i.$$

The profile with every (A_i) tight and volume S exists, is unique, lies in the class, and attains the bound; if every $y_i > 0$ it is the unique minimiser of ℓ_1 over the class. Consequently every block-(β, ε)-admissible basis has

$$\delta_0 := \Big(\frac{\|b_1\|}{\mathrm{vol}(L)^{1/d}}\Big)^{1/(d-1)} \;\geq\; \delta_0^{\mathrm{floor}}(d, \beta, \varepsilon) := \exp\Big(\frac{h_{d,\beta}(\varepsilon)}{d-1}\Big).$$

Proof. The coordinate equations of $e_1 = \sum_i y_i a_i + z\mathbf{1}$ are triangular: at position 1, $(1-1/\beta)y_1 + z = 1$; at $2 \leq k \leq d-1$, $(1 - 1/\beta_k)y_k - \sum_{i \in W_k} y_i/\beta_i + z = 0$ with $W_k := \{i : \max(1, k - \beta + 1) \leq i \leq k-1\}$ the blocks covering k; and at position d, $z - \sum_{i \in W_d} y_i/\beta_i = 0$. Summing all d equations, and using that every a_i has coordinate sum 0, gives $dz = 1$; the first $d-1$ equations then determine $y_1, \ldots, y_{d-1}$ successively (each has the positive diagonal coefficient $1 - 1/\beta_k$), and the d-th holds automatically. Taking the inner product of the identity with ℓ gives $\ell_1 = \sum_i y_i \langle a_i, \ell\rangle + zS \geq \sum_i y_i c_i + S/d$ by (A_i) and $y_i \geq 0$, with equality iff $\langle a_i, \ell\rangle = c_i$ wherever $y_i > 0$. The tight system $\langle a_i, \ell\rangle = c_i$ $(1 \leq i \leq d-1)$, $\sum \ell = S$ is uniquely solvable: (A_i) tight reads $\ell_i = (\beta_i c_i + \sum_{j=i+1}^{i+\beta_i-1} \ell_j)/(\beta_i - 1)$, which determines $\ell_{d-1}, \ldots, \ell_1$ from ℓ_d as $\ell_i = \ell_d + q_i$ with constants q_i (the coefficients of the later ℓ_j sum to one), and the volume fixes ℓ_d. The all-tight profile satisfies every (A_i) with equality, hence lies in the class and attains the bound; under strict positivity every minimiser is tight everywhere, hence equals it. $\square$

We call the all-tight profile $\ell^{\mathrm{tight}}(d, \beta, \varepsilon; S)$. It is the fixed point of the idealised block recurrence in which each block's first entry is set to its Gaussian-heuristic value (Appendix C); its tail turns up, $\ell_d^{\mathrm{tight}} - \ell_{d-1}^{\mathrm{tight}} = -2L(2) + 2\varepsilon$, which is 1.1447 at $\varepsilon = 0$.

2.3. **Positivity of the certificate.** From the coordinate equations, $y_1 = \beta(d-1)/(d(\beta-1))$ and $y_2 = \beta(d-\beta)/(d(\beta-1)^2)$; in the head $y_{k+1} - y_k = y_k/(\beta - 1)$ for $2 \leq k \leq \min(\beta - 1, d - \beta)$, and in the constant-size body the differences $\Delta_k := y_k - y_{k-1}$ obey the moving average $\Delta_k = \frac{1}{\beta-1}\sum_{j=k-\beta+1}^{k-1} \Delta_j$. These recurrences do not by themselves prove positivity and give no control in the shrinking tail; the following argument does.

Theorem 2.3 (positivity)**.** *For $2 \leq \beta < d$, every $y_k > 0$. Hence the head objective is bounded below on the class, Theorem 2.2 applies for every (d, β, ε), and the all-tight profile is the unique minimiser. For $\beta = d$, $y_1 = 1$ and all remaining multipliers are zero.*

Proof. Put $\beta_d := 1$ and $y_d := 0$. For $2 \leq k \leq d-2$ let $S_k := \sum_{i \in W_k} y_i/\beta_i$; then $S_k - S_{k+1} = [k \geq \beta]\, y_{k-\beta+1}/\beta_{k-\beta+1} - y_k/\beta_k$, since the block starting at $k - \beta + 1$ covers k but not $k+1$, and the block starting at k covers $k+1$ but not k. Subtracting the coordinate-$(k+1)$ equation from the coordinate-k equation gives

$$y_k \;=\; \frac{\beta_{k+1} - 1}{\beta_{k+1}}\, y_{k+1} \;+\; [k \geq \beta]\, \frac{y_{k-\beta+1}}{\beta},$$

the last coefficient being $1/\beta$ because $k \geq \beta$ implies $k - \beta + 1 \leq d - \beta$, so that block has full size. The same formula at $k = d-1$, under $\beta_d = 1$, $y_d = 0$, is the position-d equation, $y_{d-1} = y_{d-\beta}/\beta$; and subtracting the coordinate-2 equation from the coordinate-1 equation gives $y_1 = 1 + \frac{\beta-1}{\beta} y_2$.

This is $y = Ty + e_1$ with $T \geq 0$: row 1 has weight $(\beta-1)/\beta$ on column 2; for $2 \leq k \leq d-1$ row k has weight $(\beta_{k+1}-1)/\beta_{k+1}$ on column $k+1$ and, when $k \geq \beta$, weight $1/\beta$ on column $k-\beta+1$. The row sums are $r_1 = (\beta-1)/\beta$ and $r_k = (\beta_{k+1}-1)/\beta_{k+1} + [k \geq \beta]/\beta$: for $k < \beta$, $r_k \leq (\beta-1)/\beta < 1$; for $\beta \leq k \leq d-\beta$, $r_k = 1$; for $k \geq \beta$ and $k > d-\beta$, $\beta_{k+1} = d-k < \beta$ and $r_k = 1 - 1/(d-k) + 1/\beta < 1$. So T is substochastic. Every state $k \leq d-2$ has a positive forward edge $k \to k+1$ (as $\beta_{k+1} \geq 2$), and $d-1$ is a deficient row (sum $1/\beta$), so from every state there is a path of length at most $d-1$ to a deficient row; since $T\mathbf{1} \leq \mathbf{1}$ the vectors $T^n\mathbf{1}$ are componentwise non-increasing in n, and along such a path the mass leaks, so $T^{d-1}\mathbf{1} < \mathbf{1}$ componentwise, $\rho(T) < 1$, and the Neumann series gives $y = (I-T)^{-1}e_1 = \sum_{n\geq 0} T^n e_1 \geq 0$. For strict positivity, $(T^n e_1)_k > 0$ exactly when the directed graph of T has a length-n path from k to 1: state β has the edge $\beta \to 1$; every $k < \beta$ reaches β by forward edges; if $k > \beta$, repeated backward edges replace k by $k-(\beta-1)$ until the value lies in $[1,\beta]$, whence it is 1 or reaches 1 directly or via β. For $\beta = d$ the identity reads $e_1 = y_1 a_1 + z\mathbf{1}$ with $a_1 = e_1 - \mathbf{1}/d$, so $y_1 = 1$, $z = 1/d$ and $y_i = 0$ for $i \geq 2$. □

Equivalently $y_k = \partial \ell_1^{\text{tight}}/\partial c_k$ at fixed volume: raising any block's constant raises the minimal head. The body multipliers decrease slowly (drift $-2/(d(\beta-1))$), $y_1 \approx 1$ dominates, $y_2 \approx \beta(d-\beta)/(d\beta^2)$ is small, and the least multiplier is $y_{d-1} = y_{d-\beta}/\beta$.

2.4. **The value.** Pairing the certificate identity with the profile $\ell_j = j$ gives the exact identity

$$\sum_{i=1}^{d-1} (\beta_i - 1)\, y_i = d-1,$$

so $\sum_{i \leq d-\beta+1} y_i = (d-1-\sum_{i>d-\beta+1}(\beta_i-1)y_i)/(\beta-1)$ and $\sum_i y_i$ exceeds $(d-1)/(\beta-1)$ only by a tail term. With $g_{d,\beta} := (d-1)L(\beta)/(\beta-1)$ the head of the pure GSA line of volume 0, the *tail correction* is exactly

$$\kappa_{d,\beta} := g_{d,\beta} - h_{d,\beta}(0) = -\sum_{i>d-\beta+1} y_i(\beta_i-1)\big[f(\beta_i) - f(\beta)\big], \qquad f(n) = \frac{L(n)}{n-1},$$

a weighted deviation of the tail constants from linear interpolation (Appendix A); and

$$\delta_0^{\text{floor}}(d,\beta,\varepsilon) = \delta_{\text{GSA}}(\beta)\exp\Big(-\frac{\kappa_{d,\beta}}{d-1} - \frac{\varepsilon\sum_i y_i}{d-1}\Big), \qquad \delta_{\text{GSA}}(\beta) := \hat{c}_\beta^{\,1/(\beta-1)},$$

the ε-dependence being exactly linear. Since f increases to its integer maximum at 36 and decreases beyond, for $\beta > 36$ the tail dimensions in $[36,\beta)$ push κ negative and those with $\hat{c}_{\beta_i} < 1$ push it positive. At the printed round-3 tuples $(d,\beta) = (1003,403), (1424,625), (1885,877)$ one has $\kappa = -0.059, -0.086, -0.106$: the extremal head is *above* the GSA head by that amount in $\log\|b_1\|$, and since its last entries also lie above the line's and the volumes agree, its body lies below the line (a numerical observation at these tuples; the entries at the detection positions are quantified in Section 6). The floor is not monotone in β: $\hat{c}_n < 1$ for $n \leq 12$ makes the constraints nearly vacuous at small blocksizes, and $\delta_{\text{GSA}}(\beta)$ has its maximum at $\beta = 36$; a "least blocksize" is therefore always relative to a stated range, and every such statement below names its range.

3. The prefix-volume floor

Let $F_m := \sum_{i\leq m} \mathbb{Z} b_i$ be the members of the basis's flag, with volume taken in the real span, so that $\operatorname{vol} F_m = \exp P_m(\ell)$ with $P_m(\ell) := \sum_{i \leq m} \ell_i = \langle \mathbf{1}_{\leq m}, \ell\rangle$.

Theorem 3.1 (prefix-volume floor). *Let ℓ be linearly (β,ε)-admissible of volume S and ℓ^{tight} the all-tight profile of the same volume. Then for every $m = 1,\dots,d-1$,*

$$P_m(\ell) \ \ge\ P_m(\ell^{\text{tight}}).$$

The case $m=1$ is Theorem 2.2. Each prefix functional has an exact nonnegative rational dual certificate $(w^{(m)}, z^{(m)})$, $\mathbf{1}_{\le m} = \sum_i w_i^{(m)} a_i + z^{(m)}\mathbf{1}$ with $w^{(m)} \ge 0$ and $z^{(m)} = m/d$, computed by the same triangular recurrence as y.

Proof. Put $x := \ell - \ell^{\text{tight}}$, so that $\langle a_i, x\rangle \ge 0$ for all i and $\sum x = 0$, and let $T_j := \sum_{i\ge j} x_i$, so that $T_1 = T_{d+1} = 0$. The constraint at i says $x_i \ge (T_i - T_{i+\beta_i})/\beta_i$, i.e.

$$T_{i+1} = T_i - x_i \ \le\ \Big(1-\frac{1}{\beta_i}\Big)T_i + \frac{1}{\beta_i}T_{i+\beta_i}:$$

every suffix sum is at most a convex combination of the one before it and one after it. Let $M := \max_{1\le j\le d+1} T_j$ and suppose $M > 0$; let J be the least index with $T_J = M$, so $J \ge 2$, $T_{J-1} < M$ and $T_{J-1+\beta_{J-1}} \le M$. Then $M = T_J \le (1-1/\beta_{J-1})T_{J-1} + T_{J-1+\beta_{J-1}}/\beta_{J-1} < M$, a contradiction. Hence $T_j \le 0$ for all j and $P_m(x) = -T_{m+1} \ge 0$. For the certificate: the tight system is uniquely solvable, so $a_1,\dots,a_{d-1},\mathbf{1}$ form a basis of $\mathbb{R}^d$ and every rational vector, in particular $\mathbf{1}_{\le m}$, has a unique representation $\sum_i w_i a_i + z\mathbf{1}$ with rational coefficients, $z = m/d$ by summing coordinates; the minimum of P_m over the class is attained (at ℓ^{tight}), so by strong LP duality there is a nonnegative dual solution, and by uniqueness of the representation it is $(w^{(m)}, z^{(m)})$. □

Lemma 3.2 (what the class does not bound). *For $2 \le \beta < d$ and every $k \ge 2$, $\inf\{\ell_k : \ell$ in the class$\} = -\infty$.*

Proof. Put $r := d-k+1$ and let $v_j := 0$ for $j<k$, $v_j := -1$ for $j \ge k$, and $u := v + (r/d)\mathbf{1}$, so that $\sum_j u_j = 0$ and $\langle a_i, u\rangle = \langle a_i, v\rangle$ for every i, each a_i having coordinate sum 0. If block i lies entirely before k or entirely at or after k, the coordinates of v on it are constant and $\langle a_i, v\rangle = 0$; if it crosses k ($i < k \le i+\beta_i-1$), its first coordinate is $v_i = 0$ while the later ones are 0 or -1, so $\langle a_i, v\rangle = (1-1/\beta_i)(v_i - \operatorname{avg}_{[i+1,i+\beta_i-1]} v) \ge 0$. Hence $\ell^{\text{tight}} + \lambda u$ lies in the class for every $\lambda \ge 0$, and its k-th coordinate $\ell_k^{\text{tight}} + \lambda u_k$ with $u_k = -1 + r/d = -(k-1)/d < 0$ tends to $-\infty$. (For $k=1$ the direction has $u_1 = 0$, as it must: Theorem 2.2 bounds ℓ_1 below.) □

The following are computational statements, verified by linear programming at $(d,\beta) = (60,20)$, $(100,40)$ and $(200,40)$ and not proved in general: $\sup \ell_{d-\beta+1} = +\infty$ over the class; the same two unboundedness statements hold under the two-sided constraints of Appendix B; and $\inf P_{d-\beta} = -\infty$ once the constraint of the last full block, $(A_{d-\beta+1})$, is dropped. The second of these says that the prefix-volume floor needs the axiom at every block, including the block whose Gaussian-heuristic length it bounds; both families of constraints say only that the profile descends at least at the GSA rate, and bound no tail entry from above. Since the all-tight profile is the unique solution of the tight system, every rational linear functional $\langle g, \cdot\rangle$ has an exact rational representation $g = \sum_i w_i a_i + z\mathbf{1}$ (multipliers of either sign in general), and $\langle g, \ell^{\text{tight}}\rangle = \sum_i w_i c_i + zS$; this is how the entries ℓ_k^{tight} and the prefix volumes are enclosed rigorously in Section 6.

4. Per-basis identities

For *every* basis of every lattice and $2 \le \beta \le d$, put

$$\varepsilon_i(B) := \log\frac{\hat c_{\beta_i}\operatorname{vol}(B_i)^{1/\beta_i}}{\lambda_1(B_i)}, \qquad \nu_i(\ell) := L(\beta_i) - \langle a_i, \ell\rangle = \log \mathrm{GH}(B_i) - \ell_i \qquad (1 \le i \le d-1),$$

the signed *sub-Gaussian ratio* of block i (positive where the block's minimum is shorter than its Gaussian-heuristic length) and the *profile deficit* of position i (positive where b_i^* is shorter than the block's Gaussian-heuristic length). Since $b_i^* \in B_i \setminus \{0\}$, $\nu_i = \varepsilon_i(B) - \log(\|b_i^*\|/\lambda_1(B_i)) \le \varepsilon_i(B)$. Pairing the certificates with ℓ gives, at $\varepsilon = 0$,

$$\ell_1 - \frac{S}{d} \;=\; h_{d,\beta}(0) - \sum_{i=1}^{d-1} y_i\,\nu_i \;=\; h_{d,\beta}(0) - \sum_{i=1}^{d-1} y_i\,\varepsilon_i(B) + R(B), \qquad R(B) := \sum_{i=1}^{d-1} y_i \log \frac{\|b_i^*\|}{\lambda_1(B_i)} \;\ge\; 0,$$

and likewise, for every m,

$$P_m(\ell) \;=\; P_m(\ell^{\text{tight}}) - \sum_{i=1}^{d-1} w_i^{(m)}\,\nu_i.$$

We call $\sum_i y_i \nu_i$ the *weighted profile deficit* and $\sum_i y_i \varepsilon_i(B)$ the *weighted sub-Gaussian mass*; they differ by the non-tightness $R(B)$. For $\beta < d$, where every $y_i > 0$, $R(B) = 0$ exactly when $\|b_i^*\| = \lambda_1(B_i)$ for all $i \le d-1$ – every block SVP-tight at its first position, which is the defining property of strict BKZ-β reduction without the size-reduction clause (conventions differ on the shrinking tail); for $\beta = d$ only the first block enters. The floor of Theorem 2.2 is the case $\nu_i \le \varepsilon$ for all i. The identities turn the question of applicability into numbers: *a basis lies below the zero-slack head floor by exactly its weighted profile deficit, and below the prefix optimum at m by its $w^{(m)}$-weighted deficit*; the contrapositive of the floor is exact – a basis with ℓ_1 below the floor has, at some position, b_i^* shorter than $e^{-\varepsilon}$ times its block's Gaussian-heuristic length, hence a block minimum shorter still. Section 7 measures these deficits on the outputs of tours of unpruned enumeration.

The identities also price a deficit confined to the head, which is where Section 7 finds it.

Proposition 4.1 (the head-slack class)**.** *Let $2 \le \beta < d$ and $\varepsilon_1 \ge 0$, and let the* head-slack class *consist of the profiles of volume S with $\langle a_1, \ell\rangle \ge L(\beta) - \varepsilon_1$ and $\langle a_i, \ell\rangle \ge L(\beta_i)$ for $2 \le i \le d-1$. Then $\min \ell_1 = S/d + h_{d,\beta}(0) - y_1\varepsilon_1$, attained by the unique profile tight at every constraint of the modified system; for $1 \le m \le d-1$, $\min P_m = P_m(\ell^{\text{tight}}) - w_1^{(m)}\varepsilon_1$ with $w_1^{(m)} = (d-m)\beta/(d(\beta-1))$; and for $m = d-b$, $\beta = b$, every profile of the class has*

$$\log \mathrm{GH}_b(L/F_{d-b}) \;\le\; \ell^{\text{tight}}_{d-b+1} + \frac{b\,\varepsilon_1}{d(b-1)},$$

with equality at the modified tight profile.

Proof. Pairing $e_1 = \sum_i y_i a_i + \mathbf{1}/d$ with ℓ gives $\ell_1 \ge S/d + \sum_i y_i c_i' = S/d + h_{d,\beta}(0) - y_1\varepsilon_1$ for the modified constants $c_1' = L(\beta) - \varepsilon_1$, $c_i' = L(\beta_i)$, with equality iff every constraint is tight (all $y_i > 0$ by Theorem 2.3); the modified tight system has the same triangular structure as in Theorem 2.2 and is uniquely solvable. For the prefix functionals, the proof of Theorem 3.1 uses only $\langle a_i, \ell - \ell'\rangle \ge 0$ for the reference profile ℓ', which holds with ℓ' the modified tight profile, so $\min P_m$ is attained there and equals $\sum_i w_i^{(m)} c_i' + mS/d = P_m(\ell^{\text{tight}}) - w_1^{(m)}\varepsilon_1$. Coordinate 1 of $\mathbf{1}_{\le m} = \sum_i w_i^{(m)} a_i + (m/d)\mathbf{1}$ reads $(1 - 1/\beta)w_1^{(m)} + m/d = 1$, no earlier block covering position 1, which is the closed form. Finally $\log \mathrm{GH}_b(L/F_{d-b}) = L(b) + (S - P_{d-b}(\ell))/b \le L(b) + (S - P_{d-b}(\ell^{\text{tight}}))/b + w_1^{(d-b)}\varepsilon_1/b$, and $w_1^{(d-b)}/b = b/(d(b-1))$. □

The head-slack class is the class of Definition 2.1 with the head constraint relaxed and every other constraint kept at zero slack; the detection bound rises by exactly $b\varepsilon_1/(d(b-1))$, which is about ε_1/d.

5. Certificates

The word "certified" is used in this paper in four senses, stated wherever a numerical claim is made: (i) a *theorem* (Sections 2–4); (ii) a *finite LP optimum*: for given (d,β,ε), the exact rational multipliers are computed and the identity $e_1 = \sum y_i a_i + z\mathbf{1}$ and $y \geq 0$ re-verified exactly, and the value $h_{d,\beta}(\varepsilon)$, or an entry ℓ_k^{tight}, is enclosed in an interval by evaluating $L(n) = -\frac{1}{n}(\frac{n}{2}\log\pi - \log\Gamma(\frac{n}{2}+1))$, the dot product and the exponential in ball arithmetic (a real number is represented by a midpoint and a radius, every operation returning a ball that provably contains the result); a comparison is *decided* when the two balls are disjoint – every comparison reported here was decided at 256 bits of precision; a comparison whose balls overlap is reported undecided, and none was – and the decision is directed: an inequality $x \leq y$ is certified when the upper endpoint of x lies below the lower endpoint of y, and $x \geq y$ when the lower endpoint of x lies above the upper endpoint of y; (iii) a *range scan*: every parameter value in a stated set decided in the sense (ii), so that a least passing value is certified least in that set and in that set only; (iv) a *software computation* in double precision, used only to locate candidates, never as a certificate. The empirical statistics of Section 7 are measurements, labelled as such.

The multipliers are computed by the forward recurrence of Theorem 2.2: the incoming sum $\sum_{i\in W_k} y_i/\beta_i$ runs over the window $[\max(1,k-\beta+1), k-1]$, every block of which covers k, so it is a sliding sum maintained in $O(1)$ exact rational operations per position, and the whole certificate costs $O(d)$ such operations; the verification re-accumulates the same sliding sum independently. On the machine used (single wall-clock measurements, archived), at $d = 1003$ the exact certificate takes 0.12 s and a decision against a target 0.17 s; at $d = 1424$, 0.16 s and 0.28 s; at $d = 1885$, 0.36 s and 0.68 s. The prefix-volume and entry certificates use the same recurrence with a different right-hand side; the enclosure of the detection entry $\ell_{d-b+1}^{\text{tight}}$ at the Kyber512 crossing $(d,b) = (1030, 417)$, $S = 517\log q$, has radius below $5\cdot 10^{-74}$ at 256 bits in both the main computation and the checker. An independent checker, which recomputes the multipliers by a direct forward solve over all earlier blocks with an explicit coverage test (no sliding window; $\Theta(d^2)$ coverage tests, $O(d\beta)$ nonzero terms), re-verifies the identity and positivity exactly, re-encloses the value in ball arithmetic and re-decides the comparison, is part of the repository [5]; it is independent in its recurrence and verification logic and shares the ball-arithmetic kernel (FLINT/Arb) with the main computation. Its archived transcript agrees with the sliding-window computation on the head decisions at $\beta = 412, 413$, $644, 645$ and $907, 908$ (the first failing, the second passing in each pair) and on the detection crossings of Section 6.

6. The round-3 primal chain re-evaluated

6.1. **The specification's chain and its reconstruction.** The round-3 Kyber specification [10, §5.1.2] estimates the primal attack as follows. From m LWE samples one builds a lattice of dimension $d = m+kn+1$ and volume q^m containing the short vector $(s,e,1)$; BKZ-b is modelled by the GSA with

$$\delta(b) = \Big(\frac{(\pi b)^{1/b}\, b}{2\pi e}\Big)^{1/(2(b-1))},$$

and the attack is declared to succeed iff

$$\varsigma\sqrt{b} \;\leq\; \delta(b)^{2b-d-1}\, q^{m/d} \tag{9}$$

for some integer $m \in [0,(k+1)n]$, where $\varsigma^2 = \eta_1/2$ is the variance of the secret and error coefficients ($n = 256$, $q = 3329$; $k = 2, 3, 4$ and $\eta_1 = 3, 2, 2$ for the three parameter sets); the least such b is the reported blocksize, and its Table 4 prints $(d,b) = (1003, 403), (1424, 625), (1885, 877)$. The left side of (9) is the root-mean-square norm of the planted vector's projection onto the last b Gram–Schmidt directions; the right side is the GSA line's value of $\|b^*_{d-b+1}\|$, i.e. the

Gaussian-heuristic length of the last block read off the line. Since $\delta(b) < 1$ for small b, (9) holds trivially there (the exponent $2b-d-1$ is negative), so a blocksize range must be declared: we scan integers $b \in [300, 1100]$ and $m \in [0, (k+1)n]$, restricted to $2b < d+1$, the domain in which (9) is the attack's condition. Our reconstruction from the printed parameters returns $b_{\mathrm{GSA}} = 406, 624, 874$, within ± 3 of the printed values. At the least passing blocksize many sample counts pass – $m \in [486, 546]$, $[658, 740]$, $[842, 946]$, i.e. $d \in [999, 1059]$, $[1427, 1509]$, $[1867, 1971]$ – with margins of order 10^{-5} in the logarithm, so a witness convention is needed, and two are used. For the reconstructed GSA baseline the full interval of passing sample counts is reported and, where a single dimension is quoted, it is the smallest passing one ($d = 999, 1427, 1867$); the interior maximum-margin count is not used for it, since its margin is flat and its argmax moves by several units under different roundings of the formula. For the profile-relaxation crossings of Section 6.2 the certified witness is the maximum-margin sample count ($m = 517, 698, 890$, i.e. $d = 1030, 1467, 1915$), whose margin of 4–$5 \cdot 10^{-3}$ is well conditioned, and the interval of passing counts found by the double-precision scan is reported beside it: $m \in [481, 555]$, $[651, 747]$ and $[833, 949]$ at $b = 417, 642, 900$. The specification's own script (`Kyber.py` of the pq-crystals security-estimates repository [12], run at its current head, commit `75c2694` of 16 March 2021) returns for the primal attack $b = 406, 626, 878$ with $486, 650, 860$ samples ($d = 999, 1419, 1885$), so the printed 403/625/877 are themselves within ± 3 of the script's current output; the script reads the entry at position $d-b+1$ of a GSA profile clipped at $\log q$ (its `construct_BKZ_shape`), of which the unclipped condition (9) is the special case, which accounts for the remaining differences of 0, 2 and 4 blocksizes. All comparisons below are against this reconstruction, in double precision for the GSA line (sense (iv)) and rigorously for the floors, and they concern the round-3 chain as printed. Note that (9) uses the approximate formula $\delta(b)$, whereas the floors of Section 2 and the fixed-target comparison use the exact $\delta_{\mathrm{GSA}}(\beta) = \hat{c}_\beta^{1/(\beta-1)}$.

6.2. **Three forms of the re-evaluation.** There are three ways to put the floor into the chain, giving different numbers; the first two are reported because their discrepancy is instructive, the third is the form consistent with the linear programme.

Fixed target (quality comparison). Hold the target root-Hermite factor fixed at $\delta_{\mathrm{GSA}}(b_{\mathrm{spec}})$, b_{spec} the printed blocksize, and ask which β the floor needs to reach it: the least β with $\delta_0^{\mathrm{floor}}(d, \beta, 0) \le \delta_{\mathrm{GSA}}(b_{\mathrm{spec}})$ at the printed d. Table 1 reports the result. At $\varepsilon = 0$ every β from 2 to the crossing was decided rigorously, each β by its own certificate (sense (ii), all decisions archived): the passing set below the crossing is exactly $[2, 14]$, $[2, 14]$, $[2, 13]$ – the permissive small-blocksize segment, where $L(\beta) \le 0$ through $\beta = 12$ and remains small just above – so every $\beta \in [15, 412]$, $[15, 644]$, $[14, 907]$ fails rigorously and 413, 645, 908 are the least passing blocksizes among all $\beta \ge 15$, $\beta \ge 15$, $\beta \ge 14$ respectively (values above each crossing were not examined). This measures the change in the profile-relaxation optimum, i.e. how much more reduction quality the class demands than the GSA line at fixed target; it is not the attack's condition.

set	(d, b_{spec})	$\delta_0^{\mathrm{floor}}(b_{\mathrm{spec}})$	$\delta_{\mathrm{GSA}}(b_{\mathrm{spec}})$	$\kappa_{d,b}$	least β $(\varepsilon = 0)$	first β (0.01)	first β (0.03)
Kyber512	$(1003, 403)$	1.0040210	1.0039615	-0.059	413	409	401
Kyber768	$(1424, 625)$	1.0029592	1.0028984	-0.086	645	640	629
Kyber1024	$(1885, 877)$	1.0023122	1.0022558	-0.106	908	902	889

Table 1. The fixed-target comparison at the printed dimensions. Floors, GSA factors and κ are midpoints of balls of radius below 10^{-60}, rounded to the digits shown. At $\varepsilon = 0$ every β from 2 to the crossing is decided rigorously, and the entry is the least passing β beyond the permissive segment ($\beta \geq 15$, 15, 14; for Kyber1024 the segment is $[2, 13]$). At $\varepsilon = 0.01, 0.03$ the entries are the first passing β in the scanned intervals $[400, 415]$, $[622, 685]$, $[874, 937]$, every β of the interval up to the entry decided rigorously. The sensitivity to uniform slack is about 4, 5–5.5 and 6–6.5 blocksizes per 0.01, consistent with the exact loss $\varepsilon \sum_i y_i/(d-1)$.

Substitution of the floor into (9). Run the reconstructed selection with $\delta_0^{\mathrm{floor}}(d, b, 0)$ in place of $\delta(b)$ over the same domain: the least $b \in [300, 1100]$ for which some admissible m satisfies (9) is $408, 626, 876$ – only $+2$ blocksizes – with the crossings $(b, m, d) = (408, 509, 1022)$, $(626, 688, 1457)$, $(876, 866, 1891)$ decided rigorously and every (b, m) with $b \in [b_{\mathrm{GSA}} - 3, b^* - 1]$ and m admissible failing rigorously, so that these are certified least in $[b_{\mathrm{GSA}} - 3, b^*] \times [0, (k+1)n]$; below $b_{\mathrm{GSA}} - 3$ only the double-precision scan excludes, and that scan, run from $b = 300$ with the floor in place of $\delta(b)$, returns the same three values. The reason for the small shift is the sensitivity of (9): with

$$\log \delta_{\mathrm{req}}(b) := \frac{(m/d) \log q - \frac{1}{2} \log \varsigma^2 - \frac{1}{2} \log b}{d + 1 - 2b}, \qquad r_{\mathrm{req}} := \frac{d}{db} \log \delta_{\mathrm{req}} = \frac{2 \log \delta_{\mathrm{req}} - 1/(2b)}{d + 1 - 2b} > 0,$$

the required factor *rises* with b, by $r_{\mathrm{req}} = (3.1, 2.8, 3.4) \cdot 10^{-5}$ per blocksize at the three chain points, while $r_{\mathrm{model}} := \frac{d}{db} \log \delta(b)$ is only $-(6.7, 3.4, 1.9) \cdot 10^{-6}$; a shift Δ of the model factor at fixed b is converted into blocksize at the rate $\Delta/(r_{\mathrm{req}} + |r_{\mathrm{model}}|)$ in the attack's selection, against $\Delta/|r_{\mathrm{model}}|$ at fixed target, a ratio of 5.7, 9.3, 18. This substitution, however, keeps the GSA *shape*: it moves the line through the floor's head and reads the last block's Gaussian-heuristic length off that line, so it moves the tail entry by only $|\kappa|(d + 1 - 2b)/(d - 1) \approx 0.013$. It is not the form the linear programme supports.

The prefix-volume form. The specification's criterion declares the planted vector detected when its projection orthogonally to F_{d-b}, of root-mean-square norm $\varsigma\sqrt{b}$, is no longer than the Gaussian-heuristic length of the last block L/F_{d-b} – the only block whose numerator is the whole lattice:

$$\varsigma\sqrt{b} \;\leq\; \mathrm{GH}_b(L/F_{d-b}) = \hat{c}_b \Big(\frac{\operatorname{vol} L}{\operatorname{vol} F_{d-b}} \Big)^{1/b},$$

a function of the prefix volume $\operatorname{vol} F_{d-b}$ alone; condition (9) reads that volume off the GSA line. Over the profile class, Theorem 3.1 gives $P_{d-b}(\ell) \geq P_{d-b}(\ell^{\mathrm{tight}})$, hence the *upper bound*

$$\log \mathrm{GH}_b(L/F_{d-b}) \;=\; L(b) + \frac{S - P_{d-b}(\ell)}{b} \;\leq\; L(b) + \frac{S - P_{d-b}(\ell^{\mathrm{tight}})}{b} \;=\; \ell_{d-b+1}^{\mathrm{tight}}(d, b, \varepsilon; S) + \varepsilon$$

with $S = m \log q$, the last equality being the tight constraint (A_{d-b+1}) itself. So the criterion can hold for a profile of the class only if $\log(\varsigma\sqrt{b}) \leq \ell_{d-b+1}^{\mathrm{tight}} + \varepsilon$ for some admissible m, and it holds at the extremal profile when this inequality does: the all-tight profile maximises the modelled Gaussian-heuristic length of the detection block over the class. The result is therefore an extremal necessary condition inside the profile relaxation, and the blocksize at which it first holds is a conditional crossing, not a statement that any reduction algorithm detects the vector. The crossings below are at $\varepsilon = 0$; their sensitivity to a uniform slack is reported after them. The extremal profile's entry at the detection position lies below the GSA line's by about $|\kappa| \approx 0.06$, so the shift is several times the $+2$: at $\varepsilon = 0$ the GSA line reproduces the chain ($b = 406, 624, 874$, passing for $d \in [999, 1059]$, $[1427, 1509]$, $[1867, 1971]$ in this form), and the extremal profile gives the crossings

$$417,\ 642,\ 900 \quad \text{at } (m,d) = (517,1030),\ (698,1467),\ (890,1915): \quad +11,\ +18,\ +26 \text{ blocksizes},$$

which are differences of 3.2, 5.3, 7.6 in the conventional core-SVP exponent 0.292β – a conversion by a further cost convention, not security gained or lost. Certification (sense (iii)): for each set, the crossing (b^*, m^*) passes with the entry and $\log(\varsigma\sqrt{b})$ enclosed to 10^{-74} (at Kyber512, $3.22446\ldots$ against $3.21928\ldots$; at Kyber1024, $3.40516\ldots$ against $3.40120\ldots$); for each set every $b \in [b_{\mathrm{GSA}} - 3, b^* - 1]$ and every $m \in [0, (k+1)n]$ with $2b < d+1$ fails rigorously (every admissible m was decided; none was sampled), so 417, 642 and 900 are certified least in $[b_{\mathrm{GSA}} - 3, b^*] \times [0, (k+1)n]$, i.e. in $[403, 417]$, $[621, 642]$ and $[871, 900]$. The fixed-target comparison's $+10/+20/+31$ and the prefix-volume form's $+11/+18/+26$ agree in size; the $+2$ was an artefact of mixing a floor head with a GSA shape.

Uniform slack. With slack ε in every block inequality the bound becomes $\ell^{\mathrm{tight}}_{d-b+1}(d, b, \varepsilon; S) + \varepsilon$, and the least passing blocksize over the same domain is, in double precision (sense (iv)), 415, 639, 897 at $\varepsilon = 0.01$ and 411, 634, 891 at $\varepsilon = 0.03$ – about two blocksizes per 0.01 of slack, less than half the rate of the fixed-target comparison. At Kyber512 both crossings are certified (sense (iii)): every $b \in [403, 414]$ fails at every admissible m under $\varepsilon = 0.01$ and 415 passes, and every $b \in [403, 410]$ fails under $\varepsilon = 0.03$ and 411 passes; at Kyber768 the crossing at $\varepsilon = 0.01$ is certified likewise – every $b \in [621, 638]$ fails at every admissible m and 639 passes rigorously at 58 sample counts, $m \in [672, 729]$ – while the Kyber1024 values and the Kyber768 value at $\varepsilon = 0.03$ are double precision. The crossings therefore stay above the reconstructed GSA values 406/624/874 at every slack considered – at $\varepsilon = 0.03$ they are still $+5/+10/+17$ – so the shift of the crossing is not a property of the zero-slack convention alone. *Caveat.* The linear inequalities are imposed at every position including the detection block itself; condition (9) likewise reads the detection block's Gaussian-heuristic length off its model. Without the detection block's own constraint the prefix volume is unbounded (Section 3). The class is a statement about profiles; no basis with a planted vector is asserted to be block-admissible at the detection block.

6.3. **Simulators, the q-ary ceiling and the dual route.**

Simulators. The two standard BKZ simulators, Chen–Nguyen [4] and the probabilistic variant of Bai, Stehlé and Wen [3], as shipped in fpylll, were run on a Z-shaped LLL profile of each Kyber lattice (entries clamped to $[0, \log q]$, slope $2 \log 1.0219$, volume q^m), for every blocksize in a window about the crossing and every sample count on a stride, and the detection condition was read off the output in both forms: the last block's Gaussian-heuristic length from the prefix volume (the floor-side reading) and the entry ℓ_{d-b+1} (the 2016 estimate). The deterministic simulator stops on its own progress criterion (27–76 tours here); the probabilistic one has no such criterion and was given 50 tours, so its output is budget-limited. Table 2 lists the least passing blocksizes.

shape model (reading)	Kyber512	Kyber768	Kyber1024
GSA line, condition (9), reconstructed	406	624	874
the specification's script (GSA clipped at $\log q$, entry reading)	406	626	878
Chen–Nguyen, converged (both readings agree)	417	642	900
Bai–Stehlé–Wen, 50 tours (last-block reading / entry reading)	417 / 416	642 / 642	900 / 899
extremal profile of the class (certified, Section 6)	417	642	900
head-clipped extremal profile (q-aware, below)	417	642	900

Table 2. Least passing blocksize of the detection condition under six shape models. Scan domains: the GSA and extremal rows over all admissible m (the extremal row certified); the simulator rows over a window of 48 blocksizes from 8 below the printed value, with m on a stride of 4 (Chen–Nguyen) and 32 (Bai–Stehlé–Wen; 64 on $[896, 904]$ at Kyber1024); the clipped row over $[b_{\text{printed}} - 12,\ b_{\text{printed}} + 36]$ with m on a stride of 2. Double precision except the certified row.

The agreement of the Chen–Nguyen row with the certified row has a structural explanation, though it is not an identity.

Proposition 6.1 (fixed points of the Chen–Nguyen tour)**.** *A tour of the Chen–Nguyen simulator [4], in the form implemented in fpylll, visits the positions $k = 1, \dots, d-1$ in order with a flag that is clear at the start of the tour. At position k it computes $G_k := \log GH(B_k)$ from the current profile – the block of dimension $\beta_k = \min\{\beta, d-k+1\}$, with the constant $\log \hat{c}_{\beta_k}$ when $\beta_k > 45$ and a tabulated constant from averages of HKZ-reduced bases when $\beta_k \le 45$ – and sets $\ell_k := G_k$ if the flag is set or $G_k < \ell_k$, setting the flag in the latter case; the last entry absorbs the volume. Thus before the first change of a tour an entry is only ever lowered, and from the first change on every entry is set to its block's value, whether that raises or lowers it. (i) A profile is a fixed point – a tour changes nothing – if and only if $\ell_k \le \log GH(B_k)$ for every $k \le d-1$, with the simulator's constants. (ii) The all-tight profile of Section 2 with those constants in place of $\hat{c}_n$, $n \le 45$, is a fixed point, and it is the only fixed point that satisfies the $(\beta, 0)$-admissibility inequalities with those constants.*

Proof. (i) The flag is set at the first k with $G_k < \ell_k$, and a tour changes nothing exactly when that never happens; every G_k is then computed from the unchanged profile, which is the stated inequality, and conversely. (ii) Equality in every inequality is the tight system, whose solution at given volume is unique by the back-substitution behind Theorem 2.2; a fixed point satisfying admissibility has $\ell_k \ge \log GH(B_k)$ as well as $\ell_k \le \log GH(B_k)$ at every k, hence equality. $\square$

Which fixed point the iteration reaches is not part of the proposition; it was computed. From the Z-shaped start at $(b, m) = (417, 520)$ and $(642, 700)$ the simulator stops by its own criterion after 58 and 40 tours at a profile satisfying (i) to 10^{-14}: it has 3 and 2 head entries at $\log q$, none above it, with $\log GH(B_k) - \ell_k = 1.8 \cdot 10^{-2}$, $1.0 \cdot 10^{-2}$, $2.5 \cdot 10^{-3}$ and $7.4 \cdot 10^{-3}$, $1.7 \cdot 10^{-3}$ at those entries, and equality below 10^{-6} at every other position. The head-clipped tight profile of the next paragraph, with the constants $\hat{c}_n$ in the tail, satisfies the same inequalities with the same deficits at the same clipped positions and equality elsewhere, and the two profiles differ by at most $4.3 \cdot 10^{-3}$ and $2.9 \cdot 10^{-3}$ in the first $d - 45$ entries, the residual being the two tail conventions propagated through the volume. So the profile the simulator converges to is, up to its tail convention, the head-clipped tight profile: the two rows of Table 2 read the detection condition off nearly the same profile, which explains their agreement, while the coincidence of the three integer crossings is a numerical finding consistent with this, not a theorem. The reading of Table 1 that follows is the candid one: the floor-side crossings 417/642/900 are the crossings a practitioner obtains by replacing the GSA line by the standard simulator, and the shift $+11/+18/+26$ over the GSA chain is that discrepancy; what this paper adds is the statement that no profile obeying the block inequalities does better, with a certificate, and the sensitivity analysis under head and uniform slack. The probabilistic simulator, which models head concavity, moves the crossing by at most one blocksize, in agreement with Proposition 4.1. The q-ary ceiling. The standard basis of the primal lattice – q times the first m unit vectors, then the rows (a_i, e_i) and the embedding row – has Gram–Schmidt norms q and then 1, so its profile lies under $\log q$, and the standard reduction pipeline keeps it there.

Lemma 6.2 (the ceiling of the pipeline)**.** *Size reductions leave the Gram–Schmidt norms unchanged; a Lovász swap at $(k, k+1)$ replaces $(\|b_k^*\|, \|b_{k+1}^*\|)$ by two norms each at most $\|b_k^*\|$; and an insertion at position κ of a vector v of the block lattice whose projection orthogonally to $b_1, \dots, b_{\kappa-1}$ is shorter than b_κ^* – the only vectors BKZ inserts – realised as in fplll by placing v before b_κ and running LLL on the extended block until the dependency is removed, starts from Gram–Schmidt norms each at most the block's previous maximum and proceeds by swaps and size reductions. Hence every basis produced from the standard basis by LLL and by BKZ with such insertions has $\max_i \ell_i \leq \log q$.*

Proof. Size reduction changes no b_i^*. For the swap, with $\mu := \mu_{k+1,k}$, the new norms are $n_k := \|b_{k+1}^* + \mu b_k^*\|$, which is below $\|b_k^*\|$ because the Lovász condition fails at $(k, k+1)$, and $n_{k+1} = \|b_k^*\| \, \|b_{k+1}^*\|/n_k$, which is at most $\|b_k^*\|$ because $n_k \geq \|b_{k+1}^*\|$. For the insertion, the Gram–Schmidt norm of v in the extended list is the length of its projection orthogonally to $b_1, \dots, b_{\kappa-1}$, which is below $\|b_\kappa^*\|$ by the insertion hypothesis (membership in the block lattice alone would not bound it); each b_j of the block is now projected orthogonally to a larger span, which does not increase its norm (one becomes zero); LLL on the extended list then only swaps and size-reduces, and removing the zero vector changes no other norm. Vectors after the block are unaffected. □

This is a property of the pipeline, not of the lattice: $(M, 1)$, $(M+1, 1)$ is a basis of $\mathbb{Z}^2$ whose first Gram–Schmidt norm is about M. The class of Definition 2.1 has no ceiling, and its extremal profile can have a head above $\log q$, at which point it is the profile of no basis the pipeline produces. At the three certified crossings the extremal head is 8.1053, 8.1065 and 8.1068 against $\log q = 8.1104$, with no entry above the ceiling, and the sample counts 517, 698, 890 of the crossings are exactly the largest ceiling-respecting ones at their blocksizes (the extremal head is at most $\log q$ for $m \in [321, 517]$, $[515, 698]$, $[775, 890]$): the crossings sit on the ceiling's edge, inside it. That is structural, not accidental. For a profile of GSA type with slope $\log \delta$ the detection margin at fixed b, as a function of the sample count m with $d = m + kn + 1$, has derivative $(kn+1)\log q/d^2 - \log \delta$, which vanishes where $d^2 \log \delta = (kn+1)\log q$, while the head $\frac{m}{d} \log q + (d-1)\log \delta$ equals $\log q$ where $d(d-1)\log \delta = (kn+1)\log q$; the two continuous optima differ by about half a unit in d. So the margin-maximising sample count and the ceiling point coincide to within one sample count for the GSA line and, the extremal profile's slope not being constant, to within a few for it (at the three sets the two counts agree to within the scan's stride of 4). This is also why the maximum-margin witness of Section 6 is ill-conditioned: it sits where the class stops being realisable. Where the extremal profile would exceed the ceiling the axiom and the ceiling are incompatible, so a q-aware variant of the class was also computed: $\ell_i \leq \log q$ together with $\ell_i \geq \min\{\log GH(B_i), \log q\} - \varepsilon$, which exempts the blocks whose Gaussian-heuristic value exceeds q. The exemption is a modelling choice made for feasibility – the projected lattice at position i contains a vector of norm at most q, but the block lattice need not – and the variant restricts the model; it is not a property of every basis. Under it the profile of least prefix volume among those with a head at $\log q$ followed by an admissible suffix is, by Theorem 3.1 applied to the suffix, the head-clipped tight profile: k entries at $\log q$ followed by the all-tight profile of the remaining dimension and volume, with k least such that that suffix's head is at most $\log q$ (the tail of a tight profile rises, so the construction is head-clipped and its entries are checked separately; none exceeds the ceiling at the points reported). Its least passing blocksizes are 417 (at $m = 520$, $k = 3$), 642 ($m = 706$, $k = 8$) and 900 ($m = 1064$, $k = 174$): the ceiling does not move the crossings, and the head-clipped extremal profile is, up to the tail convention, the profile the Chen–Nguyen simulator converges to (the computation following Proposition 6.1).

The dual route. The specification's script prices a second route, the dual attack, by the length l of the first vector after BKZ-b on a lattice of dimension $d = kn + m$ and volume q^{kn}, read off the unclipped GSA line ($\log l = S/d + (d-1)\log\delta(b)$), through $\tau = ls/q$, $\log_2\varepsilon = -2\pi^2\tau^2/\ln 2$, $\log_2 R = \max\{0, -2\log_2\varepsilon - 0.2075\,b\}$ and the cost $0.292\,b + \log_2 R$ bits, minimised over (b, m). The head floor of Theorem 2.2 is a certified lower bound on $\log l$ over the class, and the cost is increasing in l, so substituting the floor gives a lower bound on the model's cost over the class. At Kyber512 the script's optimum $(b, m) = (403, 513)$ has $l = 3294 < q$ and costs 117.9 bits; under the floor the optimum moves to $(410, 531)$, $l = 3327$, 120.2 bits, an increase of 2.3 bits, and the extremal head there is $6\cdot 10^{-4}$ below $\log q$ while at the script's optimum it is 0.047 above it, so the class contains no profile the pipeline can realise there. At Kyber768 and Kyber1024 the script's own optima have $l = 5002$ and $5918 > q$ – the script flags this itself – so its model predicts a first vector longer than the q-vectors the lattice contains and the comparison lies outside the model's domain; for the record the floor gives 185.4 bits at $(634, 703)$ against the model's 181.3 at Kyber768, and 261.1 bits at $(892, 933)$ against 253.9 at Kyber1024, with the extremal head above $\log q$ in every one of these four optima. All figures of this paragraph are double precision (sense (iv)).

6.4. **Provenance of every blocksize triple.** Table 3 reconciles the blocksize triples that appear in this paper.

Table 3: Every blocksize triple in this paper, by attack model, target, dimension, scan domain and certification mode.

triple	what it is	target / condition	dimension	scan domain and mode
403/625/877	printed in the specification's Table 4	(9), GSA	1003/1424/1885	the specification's script
406/624/874	(9) reconstructed from the printed parameters	(9), $\delta(b)$	999/1427/1867 (smallest passing; intervals in Section 6)	least $b \in [300, 1100]$, double precision over all admissible m
413/645/908	least β whose zero-slack floor reaches $\delta_{\mathrm{GSA}}(b_{\mathrm{spec}})$	fixed target, exact δ_{GSA}	printed d	every $\beta \in [2, \text{crossing}]$ decided rigorously; least $\beta \geq 15, 15, 14$
409/640/902, 401/629/889	the same at slack $\varepsilon = 0.01, 0.03$	fixed target	printed d	first passing β in $[400, 415]$, $[622, 685]$, $[874, 937]$, every β up to it decided rigorously

continued on the next page

triple	what it is	target / condition	dimension	scan domain and mode
408/626/876	(9) with δ_0^{floor} in place of $\delta(b)$ (GSA shape kept)	(9), floor head	1022/1457/1891	crossings and all (b, m) with $b \in [b_{\text{GSA}} - 3, b^* - 1]$ decided rigorously; least in that range
417/642/900	prefix-volume form: $\log(\varsigma\sqrt{b}) \le \ell_{d-b+1}^{\text{tight}}$	detection block's GH	1030/1467/1915 (maximum-margin witnesses m = 517/698/890; passing $m \in [481, 555]$, $[651, 747]$, $[833, 949]$ in double precision)	crossings rigorous; each certified least in $[b_{\text{GSA}} - 3, b^*]$ over every admissible m
415/639/897, 411/634/891	prefix-volume form at uniform slack $\varepsilon = 0.01, 0.03$: $\log(\varsigma\sqrt{b}) \le \ell_{d-b+1}^{\text{tight}}(\varepsilon) + \varepsilon$	detection block's GH, slack ε	995/1441/1863, 996/1442/1869 (first passing m)	least b in $[b_{\text{GSA}} - 20, b^* + 5]$, double precision over all admissible m; the Kyber512 crossings certified least in $[403, 415]$ and $[403, 411]$, the Kyber768 crossing at $\varepsilon = 0.01$ certified least in $[621, 639]$
418/654/924	first passing β under primal and full-size dual constraints, target $\delta_{\text{GSA}}(b_{\text{spec}})$; least among $\beta \ge 15, 15, 14$ by the primal failures below	fixed target	printed d	exact two-family certificates from 410, 640, 905 to the crossing and primal witnesses at the crossings (Appendix B)

7. What block reduction actually does: an applicability audit

Theorem 2.2 is a theorem about the class; whether the bases produced by block reduction belong to it is a separate question, and at the sizes where it can be examined by full enumeration the answer is no at zero slack. *Protocol.* The lattices are random q-ary lattices of dimension d with $q = 2^{16} + 1$: the row lattice of $\left(\begin{smallmatrix} qI_{d/2} & 0 \\ A & I_{d/2} \end{smallmatrix}\right)$ with A uniform modulo q from a seeded generator, LLL-reduced. A *strict tour* visits the positions $k = 1, \ldots, d-1$ in order; at each it enumerates the projected block B_k without pruning inside the ball of squared radius $0.99\,\|b_k^*\|^2$ (the reduction library's convention, the factor being its LLL parameter δ), on Gram–Schmidt data held in double precision for $d < 150$ and in long double precision for $d \ge 150$; inserts the vector found, if any – so a vector is inserted exactly when the enumeration finds one shorter than $\sqrt{0.99}\,\|b_k^*\|$; and completes the block by an LLL reduction confined to the block's rows. The enumeration is unpruned, but its bounds are evaluated in floating point and no rigorous completeness certificate is attached to them: every block minimum below is the output of such an enumeration, a measurement and not a certified exact value, and the statement that the outputs leave the class is a statement about these measured minima. A tour in which no insertion occurs is *clean*; at a clean tour every block's first vector is within a factor $0.99^{-1/2}$ of its measured minimum, so $R(B) \le -\frac{1}{2}\log 0.99 \sum_i y_i$, which is below 0.014 at $(75, 30)$, and the measured R at the clean tours is at most 10^{-4}. Runs stop at the first clean tour or after 512 tours. Every $\lambda_1(B_i)$ quoted for a final basis was obtained by unpruned enumeration of its block inside a ball whose squared radius starts at $(1.05\ \mathrm{GH}(B_i))^2$ and is multiplied by 1.5 until a vector is found. The head ℓ_1, the volume S and the positional deficits ν_k of an archived integer basis need no block minimum and are exact functions of the basis: the Gram–Schmidt squared norms are the ratios of consecutive leading principal minors of the Gram matrix, exact rationals, so $\ell_1 - S/d - h_{d,\beta}(0)$, each ν_k and the weighted deficit are enclosed in ball arithmetic from exact rational inputs, with no enumeration and none of the non-rigorous floating-point Gram–Schmidt data used for enumeration; the statements below that a basis lies below the zero-slack floor, and the values of ν_1 and of the weighted deficits, are certified in sense (ii) of Section 5, while the clean-tour status, the non-tightness R and the block minima are measurements. Three seeds were run at each of $(d, \beta) = (75, 30)$, $(100, 40)$, $(150, 30)$, $(225, 30)$.

Findings. At $(75, 30)$ all three runs reach a clean tour (at tours 29, 131, 223), with heads 0.0054, 0.0788, 0.0751 *below* the zero-slack floor (certified: the three differences $\ell_1 - S/d - h_{75,30}(0)$ are enclosed in balls of radius below 10^{-40} whose upper endpoints are negative), i.e. weighted profile deficits of those sizes. At $(100, 40)$ no run is clean by 512 tours; the residual is $R \approx 0.02$, and the heads are 0.1327, 0.1385, 0.1328 below the floor (certified likewise; weighted sub-Gaussian masses 0.15 to 0.17). At $(150, 30)$ and $(225, 30)$ the heads are still above the floor after 512 tours ($+0.0034$ to $+0.0250$ and $+0.0809$ to $+0.1507$, certified likewise) with $R = 0.06$–0.16; a fixed tour count compares different stages of convergence across d, and no convergence statement is made at these sizes. So the clean outputs that could be produced lie outside the zero-slack class – a certified statement about their profiles – and the 512-tour outputs at $(100, 40)$ do so more.

Where the deficit sits decides what it means for the chain. Reading the profile deficits ν_k position by position: at $(100, 40)$ after 512 tours the first position has $\nu_1 = +0.128, +0.125, +0.132$ with weight $y_1 = 1.015$, so $y_1\nu_1 = +0.130, +0.127, +0.134$ against the weighted profile deficits $+0.133, +0.138, +0.133$ – the weighted deficit of positions 2 to $d-1$ is $+0.003, +0.011, -0.001$; at $(75, 30)$ at the clean tours, $\nu_1 = +0.043, +0.100, +0.105$ and the rest $-0.039, -0.023, -0.032$. In these samples the deficit is head-dominated: at $(100, 40)$ it is the first block's own ratio to within 0.01 (the first vector is 12% shorter than the Gaussian-heuristic length of its block), and at $(75, 30)$ the remaining positions partly cancel it. The head block is the one position whose denominator never changes; its numerator is re-formed every tour, so the head retains the best of T draws. An independent-draw heuristic for T draws from the Poisson law above predicts a head ratio of about $(\log(T/2) + \gamma)/\beta$, γ Euler's constant; the observed ν_1 is 0.85 of that value at $(100, 40)$, $T = 512$, and 0.40–0.63 of it at $(75, 30)$. This is the head-concavity phenomenon of Bai, Stehlé and Wen [3, §3–4], measured here with unpruned block enumeration. Two consequences. First, if the heuristic's scaling $\log T/\beta$ extrapolates – a heuristic, not a measurement at $\beta \approx 400$ – then for T polynomial in d the deficit is $O(\log d/\beta)$, under one blocksize at Kyber512 for 20 tours and 2–5 for 2^{20} tours at the local conversion rate. Second, and independently of any extrapolation: by the prefix identity of Section 4, a profile deficit ν_1 at the head alone, with every other inequality tight, lowers P_{d-b} by $w_1^{(d-b)}\nu_1$ and hence raises the quotient's Gaussian-heuristic upper bound by $w_1^{(d-b)}\nu_1/b$. At the Kyber512 crossing $(1030, 417)$, $w_1^{(613)} = 0.4058$, so a head-only deficit of 0.13 changes that bound by $0.4058 \cdot 0.13/417 = 1.3 \cdot 10^{-4}$ in the logarithm, against the 0.06 by which the extremal entry differs from the GSA line's – two hundredths of a blocksize at the local rate. That is a hypothetical head-only sensitivity. The measured effect on the audited bases is the weighted deficit $W := \sum_i w_i^{(d-\beta)}\nu_i$ at their own detection position $m = d - \beta$, which by the identity $P_m(\ell) = P_m(\ell^{\text{tight}}) - W$ is exactly the amount by which the basis's prefix volume falls short of the extremal profile's (double precision from the exact Gram–Schmidt norms; the identity closes to 10^{-13} on every basis). It is negative on all twelve: $W = -1.08, -0.90, -1.07$ at $(100, 40)$, $-1.96, -1.53, -1.35$ at $(75, 30)$, $-5.6, -5.4, -5.8$ at $(150, 30)$ and $-9.3, -9.9, -9.4$ at $(225, 30)$, against head terms $w_1\nu_1$ of only $+0.05$, $+0.02$ to $+0.04$, $+0.015$ and $+0.01$. The violated blocks (positive deficits, 19 to 36 per basis) all lie in the prefix – the head-concavity region, largest deficit 0.18 – but they are outweighed by the shrinking tail, whose blocks have first vectors longer than their Gaussian heuristic (tail contribution -1.1 to -2.0 on every basis), and at $d/\beta \geq 5$ by the body as well (-4.1 to -8.2 at $(150, 30)$ and $(225, 30)$, bases that had not reached a clean tour). So on these bases the prefix volume *exceeds* the extremal minimum and the last block's Gaussian-heuristic bound sits *below* the floor's, by $W/\beta = -0.023$ to -0.027 at $(100, 40)$ and -0.05 to -0.33 elsewhere: the measured deficits move the detection bound in the direction that keeps the floor a floor, and the head's violation is the smallest of the three contributions. These are bases of $d \leq 225$; nothing here extrapolates them to cryptographic sizes. Proposition 4.1 makes the sensitivity exact and certifiable: over the head-slack class the detection bound rises by exactly $b\varepsilon_1/(d(b-1))$, and with $\varepsilon_1 = 1$ – eight times the largest measured head deficit – every $b \in [403, 416]$ at Kyber512 and every $b \in [621, 641]$ at Kyber768 still fails the detection condition at every admissible sample count (sense (iii), every m decided), so 417 and 642 remain the least passing blocksizes in the scanned intervals $[403, 417]$ and $[621, 642]$ under every head slack $\varepsilon_1 \leq 1$. At Kyber1024 every $b \in [871, 898]$ fails likewise, and since the bound increases with ε_1 a rigorous failure at $\varepsilon_1 = 1$ is a failure at every smaller slack; but $b = 899$, whose zero-slack condition misses by only $1.9 \cdot 10^{-4}$ in the logarithm at $m = 890$ against a shift of $5.2 \cdot 10^{-4}$ per unit of head slack, fails at every admissible m under $\varepsilon_1 = 0.3$ and passes at 34 sample counts under $\varepsilon_1 = 1$ (both certified). So 900 remains the least passing blocksize in $[871, 900]$ under every head slack $\varepsilon_1 \leq 0.3$, and under $\varepsilon_1 = 1$ the least passing blocksize in that interval is 899: the one place where a head slack of the certified size moves a crossing, and by one blocksize. Under the independent-draw model, if the head block's minimum is redrawn T times from the law $\Pr[\lambda_1 > t\ \mathrm{GH}] = \exp(-t^\beta/2)$ and the head keeps the shortest, then $\Pr[\nu_1 > x] = 1 - \exp(-Te^{-\beta x}/2) \leq Te^{-\beta x}/2$, so $\nu_1 \leq (\log(T/2) + s)/\beta$ with probability at least $1 - e^{-s}$: at $(1003, 403)$ with $T = 2^{20}$ and $s = 5$ this is $\nu_1 \leq 0.045$, far below the certified slack. The conditional crossings of Section 6 are defined independently of tour behaviour; what the audit shows is that direct applicability of the class to reduction outputs fails at the head by the amount stated, and that a deficit spread over many positions of the body – which the measured bases do not show – would be needed to move the prefix-volume optimum materially.

8. Related work

work	statement	relation to this paper
Hanrot–Pujol–Stehlé [9, Lemma 4, Lemma 7, Theorem 1]	under their sandpile model assumption (each reduced block attains the worst-case Hermite constant with equality) a tour of their BKZ variant is an affine map whose fixed point solves a triangular system in the block averages; its convergence bounds the head after polynomially many tours	the same triangular tight system as ℓ^{tight} with worst-case constants in place of $L(\beta)$; direction opposite (achievability under an equality assumption versus a floor under inequalities); they note that the equalities can fail
Li–Nguyen [11]	a complete analysis of BKZ by the dynamical-systems method: proven upper bounds on the output quality after polynomially many tours	the upper direction again, in the worst case; no floor and no profile class
Chen–Nguyen [4]	a simulator whose tour lowers an entry to its block's Gaussian-heuristic value until the first change and sets every later entry to it, with tabulated constants for block dimensions ≤ 45	its fixed points are exactly the profiles satisfying the reverse block inequalities, the tight profile the only one inside the class (Proposition 6.1); its converged crossings coincide with the certified ones (Table 2); it uses tabulated constants where the class uses $\hat{c}_n$, $n \leq 45$
Bai–Stehlé–Wen [3, §3–4]	head concavity of BKZ profiles below the simulation and the GSA, explained by the probabilistic Gaussian heuristic and a best-of-draws effect; a refined simulator	Section 7 measures the phenomenon with unpruned block enumeration and locates the deficit at the head block
Gama–Nguyen [8]	the GSA as an empirical model of reduced bases; worst-case bounds for slide reduction	the class of Definition 2.1 replaces the line by inequalities; worst-case results are a different kind of statement
Gama–Howgrave-Graham–Koy–Nguyen [7]	Rankin's constant bounds Schnorr's worst-case quality constant of semi-block reduction from below (cited from the published abstract)	a limitation on worst-case guarantees, orthogonal to an average-case profile floor
Alkim–Ducas–Pöppelmann–Schwabe [1]; the round-3 specification [10]; the Lattice Estimator [2]	the core-SVP convention; condition (9); current estimates	the objects re-evaluated here; the estimator's current attack models differ from the round-3 chain and are not audited

Table 4. The nearest results and the delta.

We have not found – in the works of Table 4, in the proven analyses of BKZ [9, 11], or in the estimator literature [2] – a formulation of the block-Gaussian-heuristic inequalities as a linear programme over all profiles, an exact positive dual certificate for its optimum, the prefix-volume floor, or the identification of the prefix volume of the detection block as the quantity on which the primal criterion depends; the fixed point itself, the GSA slope $\hat{c}_\beta^{1/(\beta-1)}$ and the upturned tail are standard. The lower-bound results we know of for blockwise reduction are worst-case (Mordell's inequality, [7], [8]); we know of no floor side of the average-case cost models.

9. Escapes and limitations

excluded operation	why the floor does not apply	status
an output that is not block-admissible	the class is a hypothesis on the final basis	clean enumeration tours at $(75, 30)$ leave it at the head, the head's position below the floor certified from the archived bases; the weighted deficit of the other positions is small or negative in the samples; no proof for any algorithm class
non-uniform slack	the class has one ε	the head identity of Section 4 accounts exactly for any pattern of deficits; the prefix identity uses different weights; measured patterns are head-dominated
dimensions for free, sieving list output, progressive blocksizes	the class speaks of a final basis, not of how it was made	unmodelled; a list oracle would change what one call buys but not the class
rerandomisation, non-tour schedules	likewise	unmodelled; the identities apply to whatever basis results
dual, hybrid and other embeddings	only the round-3 primal chain is re-evaluated	not audited
the detection block's own inequality	the prefix floor needs the constraint at every block	without it the prefix volume is unbounded; (9) reads the same block off its own model
the cost of an SVP call	the paper prices nothing	the core-SVP exponent differences quoted are a conversion convention

Table 5. What the model forbids or omits, and what is known about each.

Nothing in this paper is a lower bound on the cost of attacking ML-KEM or any scheme. The theorems are about a class of profiles; the certified numbers are about the reconstructed round-3 chain under the class's extremal profile in place of the GSA line; the audit says where and by how much real block reduction leaves the class at small sizes. In summary: the specification's own reasoning, carried through with the floor side of the heuristic it already uses instead of a straight line, gives crossings at 417/642/900 rather than 406/624/874 (and 415/639/897 under a uniform slack of 0.01, the first two certified, the third in double precision); the measured departure of real reduction from the class is head-concentrated and, inserted at the head alone, does not change the conditional crossing (certified within the scanned intervals for every head slack up to 1 at Kyber512 and Kyber768 and up to 0.3 at Kyber1024, where a head slack of 1 moves the crossing from 900 to 899); and whether the class describes any reduction algorithm's outputs at cryptographic sizes remains open.

Appendix A. The identities

Pairing $e_1 = \sum_i y_i a_i + z\mathbf{1}$ with $\ell_j = j$: the left side is 1; $\langle a_i, \ell\rangle = (1-1/\beta_i)i - \frac{1}{\beta_i}\sum_{j=i+1}^{i+\beta_i-1} j = -(\beta_i-1)/2$; and $z\sum_j j = (d+1)/2$. Hence $1 = -\frac{1}{2}\sum_i(\beta_i-1)y_i + (d+1)/2$, i.e. $\sum_i(\beta_i-1)y_i = d-1$. For κ: $h_{d,\beta}(0) = \sum_i y_i L(\beta_i)$ and, by the identity just proved, $g_{d,\beta} = (d-1)f(\beta) = \sum_i(\beta_i - 1)y_i f(\beta)$; subtracting, $\kappa_{d,\beta} = \sum_i y_i(\beta_i-1)[f(\beta) - f(\beta_i)]$, in which the full-size terms vanish. The ε-dependence of δ_0^{floor} follows from $h(\varepsilon) = h(0) - \varepsilon\sum_i y_i$. The per-basis and prefix identities of Section 4 are the pairings $\langle e_1, \ell\rangle = \sum_i y_i \langle a_i, \ell\rangle + S/d$ and $\langle \mathbf{1}_{\le m}, \ell\rangle = \sum_i w_i^{(m)} \langle a_i, \ell\rangle + mS/d$ with $\langle a_i, \ell\rangle = L(\beta_i) - \nu_i$.

Appendix B. The two-sided floor

For a lattice M let $M^\vee := \{x \in \operatorname{span} M : \langle x, M\rangle \subseteq \mathbb{Z}\}$ be its dual. If the block-Gaussian-heuristic lower bound is also imposed on the dual of every consecutive projected block – as the outputs of a self-dual reduction would have to satisfy for a floor to apply to them – a mirrored family results. For the block $B^{(e)}$ ending at e, of size $n = \min(\beta, e)$, the vector $b_e^*/\|b_e^*\|^2$ lies in $(B^{(e)})^\vee$ with norm $1/\|b_e^*\|$, and $\operatorname{vol}(B^{(e)})^\vee = (\operatorname{vol} B^{(e)})^{-1}$; hence $\lambda_1((B^{(e)})^\vee) \ge e^{-\varepsilon}\hat{c}_n(\operatorname{vol}(B^{(e)})^\vee)^{1/n}$ implies the linear constraint

$$(D_e) \qquad \langle \tilde{d}_e, \ell\rangle := \operatorname{avg}_{[e-n+1,\, e]} \ell - \ell_e \;\ge\; L(n) - \varepsilon, \qquad 2 \le e \le d,$$

with $\tilde{d}_e$ the displayed coefficient vector (coordinate sum 0). The *full-size* family imposes (D_e) for $e \ge \beta$ only; the *all-block* family adds the shrinking head blocks $e < \beta$. The GSA line satisfies every full-size dual constraint with equality and violates the shrinking head blocks by $L(e) - (e-1)L(\beta)/(\beta-1)$, up to 0.53 at $(1003, 403)$; the primal all-tight profile violates (D_d) by $v_\beta = L(\beta) - \sum_{n=2}^{\beta} L(n)/(n-1) = 1.08, 1.31, 1.38, 1.36$ at $\beta = 10, 20, 40, 50$ (its tail upturn). The two-sided minimum of ℓ_1 under (A_i) and a chosen dual family is bounded below by any exact two-family certificate (y, w, z) with $e_1 = \sum_i y_i a_i + \sum_e w_e \tilde{d}_e + z\mathbf{1}$, $y, w \ge 0$: pairing with ℓ gives $\ell_1 \ge \sum_i y_i(L(\beta_i) - \varepsilon) + \sum_e w_e(L(n_e) - \varepsilon) + zS$. The certificate is obtained by solving for the multipliers on a set of $d-1$ rows chosen from a double-precision linear programme – the rows of positive dual value, rank-filtered by a pivoted QR decomposition and completed by independent tight rows, or, when these do not suffice, by rows of least slack – then verifying $y, w \ge 0$ exactly in rational arithmetic and enclosing the value in ball arithmetic; whichever rows are chosen, the exact identity and nonnegativity make the result a valid lower bound. A dual certificate alone certifies that a target is *not* reached, never that it is. A passing β in a scan is therefore attained by a *primal witness*: a rational profile ℓ^{w} (the double-precision LP solution rounded to rationals and refined on its tight rows, then shifted along the steepening direction $s_j := (d+1)/2 - j$, which raises the slack of every constraint of size n by exactly $t(n-1)/2$ and the head by $t(d-1)/2$, by a rational $t \ge 0$ that is a directed upper bound on the shift needed to make every constraint's slack, evaluated as a ball, have a nonnegative lower endpoint, enlarged by a fixed margin) whose root-Hermite ball lies below the target ball. The shifts of the two-sided lower bound above the primal floor at the three printed tuples are $+0.0345$, $+0.0392$, $+0.0457$ in ℓ_1 for the full-size family and $+0.174$, $+0.698$, $+1.201$ for the all-block family (certificates archived). With the full-size family at $\varepsilon = 0$, two-family certificates were computed for β from 410, 640 and 905 upwards – every smaller $\beta \ge 15, 15, 14$ already fails the primal floor of Table 1, and the two-sided class is smaller – and exclude every β below 418, 654, 924, each of which is attained by a witness; at $\varepsilon = 0.01$ the scans from 409, 640, 902 give 414, 649, 917, and at 0.03 the scans from 401, 629, 889 give 406, 638, 903, every intermediate value excluded by a certificate. So the zero-slack values are the least passing values among all $\beta \ge 15, 15, 14$, while the positive-slack values are the first passing values in their scanned intervals (the primal exclusions of Table 1 at positive slack cover only the intervals stated there). The full-size shift is not monotone in β at fixed d: at $d = 1003$ and $\varepsilon = 0.03$ it is 0.03349 at $\beta = 401$, 0.03350 at 402–404 and 0.03349 again at 405. The two-sided floor is a diagnostic for a class that no tested output satisfies: on BKZ-2.0 outputs at $(100, 40)$ and $(125, 50)$ after at most eight tours (the implementation's auto-abort may stop earlier) about half of the full-size dual blocks have a shortest dual vector below the dual Gaussian-heuristic length – measured by unpruned enumeration on floating-point Gram–Schmidt data of the reversed dual basis: for a basis matrix $\mathbf{B}$ of a lattice containing $q\mathbb{Z}^d$ the projected blocks of the reversed rows of $q\mathbf{B}^{-\mathsf{T}}$ are q times the dual blocks – and plain BKZ does not dual-reduce; the all-block class is not the GSA-flat-head model of BKZ outputs at all.

Appendix C. The idealised recurrence

Consider the profile recurrence in which a step at position i sets $\ell_i \leftarrow L(\beta_i) + \mathrm{avg}_{[i,i+\beta_i-1]}\,\ell$ (the block's Gaussian-heuristic value) whenever this is smaller, and distributes the removed mass m uniformly over the other $\beta_i - 1$ positions of the block. A step at i is inert iff $\langle a_i, \ell\rangle \le L(\beta_i)$, the reverse of (A_i) at $\varepsilon = 0$; hence a linearly $(\beta, 0)$-admissible profile is a fixed point of every step iff every (A_i) holds with equality, i.e. iff it is $\ell^{\mathrm{tight}}(d, \beta, 0; S)$, and profiles outside the class may be fixed points without being tight. Each step lowers the potential $\Phi := \sum_i (d-i)\ell_i$ by exactly $m\beta_i/2$, since it moves the mass m from position i to the average position $i + \beta_i/2$ of the block. Simulators in the style of Chen and Nguyen [4] have as fixed points exactly the profiles with $\ell_i \le \log GH(B_i)$ at every block, the tight profile being the only one inside the class (Proposition 6.1); at the converged profiles of Section 6.3 the inequality is strict only at head entries pinned at $\log q$. The head deficits of Section 7 have the same sign as such a pinned entry's but arise with no ceiling in force; the deterministic simulator does not model their mechanism, the probabilistic one [3] does. Real completions – the LLL of the block after an insertion – need not preserve the inequalities, and the potential Φ is not monotone under them: only F_i is forced smaller by an insertion at i, and the intermediate flag members' volumes can grow.

Appendix D. Reproducibility

Every certificate reported here is regenerated deterministically from its stated parameters by the code in the public repository [5], in the times given in Section 5; the archive stores every directed decision (pass, fail or undecided) with the parameters it concerns and the scope of each scan, the crossings with their sample counts and dimensions, and the independent checker of Section 5, which re-derives each certificate from the stated parameters alone and re-decides the comparison. The exact multipliers and ball enclosures themselves are not stored, since either program reproduces them. The bases of Section 7 are archived in their final state with their seeds and tour counts, together with the statistics recorded at the intermediate checkpoints, and every block minimum quoted was obtained by unpruned enumeration on floating-point Gram–Schmidt data as described there; the certified head positions and deficits of Section 7 are recomputed from the archived integer bases by exact rational Gram–Schmidt and ball arithmetic, and their enclosures are archived with exact rational endpoints. No number in the paper depends on a random choice that is not recorded.

References

[1] E. Alkim, L. Ducas, T. Pöppelmann, P. Schwabe, Post-quantum key exchange – a new hope, *USENIX Security 2016*, 327–343.

[2] M. R. Albrecht, R. Player, S. Scott, On the concrete hardness of Learning with Errors, *J. Math. Cryptol.* 9 (2015), 169–203; the Lattice Estimator, `github.com/malb/lattice-estimator`.

[3] S. Bai, D. Stehlé, W. Wen, Measuring, simulating and exploiting the head concavity phenomenon in BKZ, *ASIACRYPT 2018*, LNCS 11272, 369–404; IACR ePrint 2018/856.

[4] Y. Chen, P. Q. Nguyen, BKZ 2.0: better lattice security estimates, *ASIACRYPT 2011*, LNCS 7073, 1–20.

[5] S. Ward, Maestro, *factorlab and latticelab*: code, tests and archived results behind the reports, version 2026.09, `https://github.com/iGentAI/factorlab`.

[6] National Institute of Standards and Technology, Module-Lattice-Based Key-Encapsulation Mechanism Standard, FIPS 203, August 2024.

[7] N. Gama, N. Howgrave-Graham, H. Koy, P. Q. Nguyen, Rankin's constant and blockwise lattice reduction, *CRYPTO 2006*, LNCS 4117, 112–130.

[8] N. Gama, P. Q. Nguyen, Predicting lattice reduction, *EUROCRYPT 2008*, LNCS 4965, 31–51.

[9] G. Hanrot, X. Pujol, D. Stehlé, Analyzing blockwise lattice algorithms using dynamical systems, *CRYPTO 2011*, LNCS 6841, 447–464; IACR ePrint 2011/198.

[10] R. Avanzi, J. Bos, L. Ducas, E. Kiltz, T. Lepoint, V. Lyubashevsky, J. M. Schanck, P. Schwabe, G. Seiler, D. Stehlé, CRYSTALS-Kyber: algorithm specifications and supporting documentation (version 3.01), 31 January 2021.

[11] J. Li, P. Q. Nguyen, A complete analysis of the BKZ lattice reduction algorithm, IACR Cryptology ePrint Archive, Report 2020/1237 (2020); *J. Cryptology*, to appear.

[12] R. Avanzi, J. Bos, L. Ducas, E. Kiltz, T. Lepoint, V. Lyubashevsky, J. M. Schanck, P. Schwabe, G. Seiler, D. Stehlé, CRYSTALS-Kyber security estimates, `github.com/pq-crystals/security-estimates`, commit `75c26949` (16 March 2021).

iGent AI
Email address: `sward@igent.ai`

iGent AI