Provable Integer Factorization via Multilevel Descent: Approaching Heuristic Bounds

The KeyBreak Cryptanalysis Agent

(Harness design and prompt by Qi Wang[1], Zhian Qian[1], Changhao Yuan[1], Juanru Li[2], Xiaogang Zhou[3], Shi Bai[1], and Yu Yu[1]. Base model: GPT-6 Astra)

[1]Shanghai Jiao Tong University

[2]East China Normal University

[3]China Telecom Quantum Information Technology Group Co., Ltd

November 26, 2026

Abstract

The usual complexity analysis of the number field sieve for integer factorization relies on several heuristics, in particular concerning the smoothness of polynomial values. We give a variant of the number field sieve based on a multilevel descent. Assuming the generalized Riemann hypothesis for Dirichlet L-functions, its expected running time is $L_N[1/3, 2.97]$ bit operations. We also give an unconditional bound of $L_N[1/3, 10^{13}]$. The main difficulty in the analysis of the descent is that the prime ideals to be eliminated are determined by the relations found at earlier levels. We estimate both the number of relations available for a given prime ideal and the number of times it occurs in relations at the preceding level, under the same distribution on the defining polynomial. The conditional analysis uses smaller search regions and a bounded hyperelliptic smoothness test. In both cases, the algorithm factors any integer $N \geq 2$, returns its prime factors with multiplicities, and terminates on every random tape. The bounds include the cost of all auxiliary computations.

1 Introduction

Integer factorization is a fundamental problem in computational number theory. It is of particular interest in cryptography, since the ability to factor an RSA modulus suffices to recover the corresponding private key [58]. Many of the algorithms developed for this problem are based on the construction of a congruence of squares, an approach that underlies the continued fraction method [51], the random-squares algorithm of Dixon [23], the quadratic sieve [56], and the number field sieve (NFS) [40, 15]. We refer to [54, 61] for introductions to these methods.

The general number field sieve (GNFS) is the standard method for factoring large integers without a special algebraic form. The most notable achievements of NFS include the factorization of RSA-768 (768 bits) in 2009 [34], RSA-240 (795 bits) in 2019 and RSA-250 (829 bits) in 2020 [11]. The latter two factorizations were carried out using CADO-NFS [16], an open-source implementation of NFS. In September 2026, Lu and Weis reported the factorizations of RSA-260 (862 bits) [46] and RSA-896 (896 bits) [65], completed on 3 and 19 September, respectively. Both reports describe GPU implementations derived from CADO-NFS. For a survey of classical computational records and their implications for cryptography, see Boudot et al. [12]. These computations demonstrate the practical effectiveness of the NFS; they do not, however, establish an asymptotic provable running-time bound.

Throughout this paper, log denotes the natural logarithm, and we use the usual notation

$$L_N[\alpha, c] = \exp\bigl((c + o(1))(\log N)^{\alpha}(\log\log N)^{1-\alpha}\bigr), \qquad 0 < \alpha < 1.$$

For an upper bound expressed in this notation, the asymptotic assertion is that the limsup of the logarithm of the cost, divided by $(\log N)^{\alpha}(\log\log N)^{1-\alpha}$, is at most c.

Lenstra and Pomerance [44] proved an unconditional expected running time of $L_N[1/2, 1]$ for integer factorization. The classical analysis of GNFS gives the heuristic bound

$$L_N[1/3, (64/9)^{1/3}].$$

See [15, Section 11] and the refined analysis of Le Gluher, Spaenlehauer and Thomé [37]. The latter work studies the lower-order terms in this estimate, retaining the usual smoothness heuristics; it also explains why the leading constant alone is not a reliable estimate of the cost for cryptographic input sizes. An asymptotic improvement in the leading constant of an upper bound should therefore not be interpreted as a practical speedup.

For an odd composite integer N, suppose that we have found two units X, Y modulo N such that $X^2 \equiv Y^2 \pmod{N}$. If $X \not\equiv \pm Y \pmod{N}$, then $\gcd(X - Y, N)$ is a nontrivial factor of N. For an RSA modulus $N = pq$, this gives both prime factors; for general inputs, recursive splitting and primality testing suffice. The NFS constructs such square congruences using relations between elements whose factorizations on the rational and algebraic sides are known. Two distinct difficulties arise in proving a running-time bound: one must show that sufficiently many suitable relations can be found, and that the resulting congruence is nontrivial with sufficiently large probability. Lee and Venkatesan [38, Theorems 2.1 and 2.3] proved that a randomized number field sieve constructs a square congruence in expected time $L_N[1/3, (64/9)^{1/3}]$. Their bound for obtaining a nontrivial congruence when $N = pq$ with $p \equiv q \equiv 3 \pmod{4}$ uses Conjecture 7.1 of their paper.

Our aim is to obtain an NFS-type algorithm whose expected running time can be bounded without the usual smoothness heuristics. We do not analyze the standard GNFS itself; instead, we modify the way in which the relations are constructed and combined. Our approach starts with squares of randomly chosen units and uses a succession of relations to express these squares in terms of rational primes and prime ideals in a factor base. In the course of this descent, a relation for a large prime ideal may introduce several smaller prime ideals, which must in turn be eliminated. It does not suffice to show that almost all prime ideals admit many relations: the ideals that are introduced by the descent need not be uniformly distributed. The purpose of our analysis is to control this dependence without changing the polynomial within a relation tree. We combine this argument with a square-class certificate and a conditional probability argument for the final gcd.

1.1 Main results

Put $\Lambda = \Lambda(N) = (\log N)^{1/3}(\log\log N)^{2/3}$. We count bit operations throughout; in particular, we include the cost of testing smoothness, factoring auxiliary integers, and performing the algebraic computations. We use elementary integer and polynomial arithmetic for our upper bounds; for background on multiple-precision arithmetic, see Brent and Zimmermann [14]. The input N is fixed, and the expectation is taken only over the random choices of the algorithm.

Theorem 1.1 (Unconditional factorization)**.** *There is a fixed classical randomized algorithm that, for every integer $N \geq 2$, terminates on every random tape and returns the distinct prime*

factors of N with their multiplicities. Its expected running time, including the factorization of all auxiliary integers, satisfies

$$\mathbb{E}T(N) \le \exp\bigl((10^{13} + o(1))\Lambda(N)\bigr).$$

The error term is uniform over inputs of the same bit length.

By the generalized Riemann hypothesis (GRH) we mean the Riemann hypothesis for every primitive Dirichlet L-function, including the Riemann zeta function at modulus one.

Theorem 1.2 (Factorization under GRH)**.** *The algorithm in Theorem 1.1 can be chosen so that, under GRH, its expected running time also satisfies*

$$\mathbb{E}T(N) \le \exp\bigl((2.97 + o(1))\Lambda(N)\bigr).$$

The proofs use two procedures, denoted by U and G. In the first, the numerical tests are performed by recursive calls to the factoring algorithm, with strictly smaller inputs. In the second, we use GRH to analyze relation searches in smaller regions and employ a smoothness test with a prescribed running-time bound. Lemma 7.1 shows how to interleave the two procedures so that the same algorithm satisfies both theorems. Correctness of the output and termination on every random tape are unconditional; GRH is needed for the smaller time bound.

1.2 Outline of the proof

The analysis of the descent requires two estimates for a common random polynomial. First, for a prescribed large prime ideal, we estimate the number of pairs (a, b) for which the rational value and the algebraic cofactor have the required smoothness properties. Second, we estimate the weighted number of relations at one level that contain a prescribed prime ideal at the next level. The corresponding statements are Theorems 3.1 and 3.2.

For the first estimate, we decompose the indicator of smoothness of an algebraic cofactor into a conditional smoothness density and a remainder. The conditional density depends on the size and the valuations at small primes of the cofactor; its definition does not depend on the choice of a second point in a moment calculation. Moments of increasing order bound the variation of this density. In the second moment of the remainder, the rational and algebraic character sums are associated with the determinant of the same pair of points. Summing jointly over their conductors is what allows us to obtain the required error bound.

For the second estimate, we weight the prime ideals at the next level by the conditional expectation of the number of relations in which they occur. We compare these weights with deterministic functions of the prime and bound the normalized second moment of the number of occurrences. The comparison can be iterated through the entire descent. This gives a bound on the fraction of relations that must be discarded because at least one of the smaller prime ideals cannot be eliminated. In particular, the argument remains valid when the number of levels tends to infinity with N.

Each successful descent gives a row of a matrix over $\mathbb{F}_2$. A linear dependence between the rows, together with the quadratic character conditions, allows us to compute a square root of the corresponding product. The other square root is obtained from the original randomly chosen units. Since the latter have not been used in choosing the relations or the dependence, their conditional distribution can be analyzed even after all these choices have been made. This gives a constant lower bound for the probability of a nontrivial factor (Section 5).

1.3 Related work

There is a substantial literature on the analysis and implementation of the NFS. Coppersmith [20] proposed modifications to its asymptotic complexity. For polynomial selection, we refer to Bai, Brent and Thomé [4], Bai, Bouvier, Kruppa and Zimmermann [5], and Barbulescu and Lachand [9]. See also Bai and Zimmermann [8] on size optimization of sextic polynomials and Prest and Zimmermann [55] on the selection of two nonlinear polynomials. The filtering step used in the RSA-240 and RSA-250 computations is described by Bouillaguet and Zimmermann [13]. Other developments include the algorithm of Thomé [62] for computing vector generating polynomials in the block Wiedemann algorithm, and his study of algebraic square-root algorithms [63]. The refined asymptotic analysis of Le Gluher, Spaenlehauer and Thomé [37] concerns the lower-order terms in the heuristic complexity of NFS. In contrast, we seek an NFS variant for which the relation search and the subsequent factor extraction admit a rigorous expected-time analysis. Table 1 summarizes the bounds and their hypotheses.

Table 1: Expected running-time bounds for integer factorization and for the construction of square congruences. Only the Lee–Venkatesan entry concerns a square congruence rather than factoring.

Result	Expected time	Assumptions
Classical GNFS [15, 37]	$L_N[1/3, (64/9)^{1/3}]$	Heuristic
Lenstra–Pomerance [44]	$L_N[1/2, 1]$	Unconditional
Lee–Venkatesan [38]	$L_N[1/3, (64/9)^{1/3}]$	Unconditional
Cogentic [19]	$L_N[1/3, 105]$	Unconditional (manuscript)
Theorem 1.1	$L_N[1/3, 10^{13}]$	Unconditional
Theorem 1.2	$L_N[1/3, 2.97]$	GRH

In a concurrent manuscript, Cogentic [19, Theorem 1.1] states an unconditional expected-time bound of $L_N[1/3, 105]$ and a deterministic space bound of $L_N[1/3, 40]$ for integer factorization. Its algorithm uses two random polynomials and a weight supported on smooth integers. The analysis conditions on the polynomial values previously encountered along a path, and bounds the probability of failure by an argument from the last level to the first [19, Lemmas 4.4 and 4.6]. In contrast, we count ordinary smooth relations and the occurrences of the prime ideals that they introduce, under a fixed polynomial distribution. Both algorithms choose the public square root without using the original private roots. The unconditional time constant in [19] is smaller than ours; our bound with constant 2.97 assumes GRH. We do not compare space bounds, nor do we claim a bound on the best constant obtainable by adapting their algorithm under GRH.

Descents through progressively smaller prime ideals are familiar from the computation of individual discrete logarithms by NFS; see, for example, Joux and Lercier [32]. The descent in the present work builds on the constructions in the manuscript collection [33]. In particular, [33, `multidescent`, Sections 2–3; `pc3`, Section 2] introduces weights for requests between levels, and [33, `mfU_theorem`, Sections 7–9] gives the private-root construction, square certification and recursion on smaller inputs. We use these constructions and prove the relation-count and incidence estimates needed to analyze them under the same polynomial distribution. Appendices B and C identify the fixed-moment and character-sum estimates in that collection that are extended here.

The smoothness estimates used in this paper belong to the theory of integers without large prime factors, developed notably by de Bruijn [22], Canfield, Erdős and Pomerance [17], and Hildebrand and Tenenbaum [31, 30, 29]; see Granville [26] for a survey of the applications to computational number theory. The distribution of smooth numbers in arithmetic progressions is

studied, among others, by Soundararajan [59], Harper [28], Drappeau [24, 25], and Pascadi [52]. For our unconditional proof, the character-sum estimates of Drappeau [24] and Harper [27] are of particular importance. For the conditional proof we also use the bounds of Lagarias and Soundararajan [36] and the short-interval estimates of Younis [67]. The preceding references provide context; the precise analytic inputs and their ranges of validity are specified in the proofs.

The elliptic curve method of Lenstra [45] and its genus-two analogue of Lenstra, Pila and Pomerance [42, 43] are particularly suited to finding small prime factors. For the development and implementation of ECM, see Zimmermann and Dodson [69]. We use the distribution of genus-two Jacobian orders from [43] to analyze a bounded smoothness test under ordinary RH. The leading constant 2 was predicted in the earlier work on the hyperelliptic curve method [42, 43]. An analysis under RH giving this constant is described in the manuscript collection [33, `genus2rh`, Theorem R and Corollary R2]; see also the manuscript of Zhou [68]. The proof below uses the published order-fibre bound and Younis's short-interval estimate. Our analysis bounds the probability of rejecting a smooth input and the number of calls to this test during a descent. The use of a bounded test here is distinct from batch smoothness testing via product trees [10].

1.4 Organization of the paper

Section 2 describes the algorithm and the random polynomial experiment. Section 3 states the two main arithmetic estimates, and Section 4 uses them to analyze a descent of increasing depth. Sections 5 and 6 treat square certification, factor extraction, and the unconditional running-time bound. The changes required for the smaller constant under GRH are presented in Section 7. The appendices contain the parameter choices, the proofs of the counting estimates, and the explicit algorithms and bit-complexity bounds. In particular, Appendices B and C establish the unconditional estimates, while Appendices G and H give the conditional analysis.

2 Description of the algorithm

Let G and $\mathcal{F}$ be the rational linear form and the algebraic form, respectively. Let B be the smoothness bound for the factor base. Suppose that a relation involves a prime ideal above a rational prime $\ell > B$, with odd valuation. To eliminate this prime ideal, we seek a relation on the corresponding special-prime lattice, with valuation one at the prescribed prime ideal and with smaller prime factors elsewhere. More precisely, at level i of the descent we seek a primitive pair (a, b) satisfying

$$|G(a,b)| \in \mathcal{S}_B, \qquad |\mathcal{F}(a,b)| = \ell \prod_{\nu=1}^{t} q_\nu^{e_\nu}, \qquad q_\nu \le Y_{i+1} < \ell,$$

where $\mathcal{S}_y$ denotes the set of positive y-smooth integers, i.e., those whose prime factors are at most y. We impose the congruence corresponding to the given root of $\mathcal{F}$ modulo ℓ, so that the two relations together have even valuation at the prescribed prime ideal. Every remaining nonterminal prime ideal with odd valuation must in turn be eliminated. We call such an occurrence a *request*; it is specified by a prime q_ν and a root modulo q_ν, in the same number field. The original relation can be used only if all these requests are satisfied. The congruence is the same as in special-prime lattice sieving [53], and the recursive use of relations is related to the individual-logarithm descent [32]. Here it is used to construct parity relations for integer factorization.

An estimate for a uniformly distributed prime ideal does not by itself suffice. Indeed, if a set of prime ideals has small probability ϵ under the uniform distribution, the distribution induced by the preceding relations could be supported entirely on this set, with density $1/\epsilon$ there. We must therefore estimate both the number of relations for a prescribed ideal and the number of times that ideal occurs in relations at the preceding level. The latter selection can affect the distribution of the pair defining G as well as that of the prime ideal. We formulate the comparison on the joint probability space of this pair, the polynomial and the prime ideal.

2.1 Construction of a congruence of squares

For each row, choose an independent private root $h \bmod N$ and release only the residue $R = h^2 \bmod N$ to the relation-finding procedure. Reduction of a two-dimensional lattice gives a short pair (u, v) with $u \equiv Rv \pmod N$. If both coordinates have the prescribed factorization, their prime divisors form the first level of the descent. We retain the row if all its descents succeed. The quadratic characters used to certify squares in the number field are chosen before relation collection. Linear algebra over $\mathbb{F}_2$ and square-root extraction give a square root Y of the product of the selected residues. At this point we multiply the corresponding private roots to obtain X, and compute $\gcd(X - Y, N)$. Distinct rows retain independent private roots, even if their public data coincide.

Sections 4–6 analyze relation collection, factor extraction and the expected running time, including unsuccessful searches and recursive factorizations. A trial division procedure ensures termination even on exceptional random tapes. Any nontrivial divisor encountered during the computation is verified and returned immediately. The complete algorithm is given in Appendix F.9.

2.2 Choice of parameters and polynomials

Write $n = \mathrm{b}(N)$, $m_b = \mathrm{b}(n)$, where $\mathrm{b}(x) = \lfloor \log_2 x \rfloor + 1$, and $D = \log N / \log\log N$. The degree d is the least odd integer at least three with $d^3 m_b \geq n$. Each prescribed factorization of a private coordinate has $k = \kappa d$ prime factors, and $\log B \sim b_0 \Lambda$, for fixed positive constants κ, b_0. Appendix F.1 specifies their values, the cutoff recursion and the integer tests.

After preprocessing, a remaining input is odd, is not a prime power, and has no prime divisor at most B. Set $M = \lfloor N^{1/d} \rfloor$. Choose independently a uniform prime $Q \in (3M/4, M]$ and a uniform integer $r \in (M/2, 3M/4]$. If Q divides N, return it. Otherwise put

$$g = (r, Q), \qquad G(X, Y) = QX - rY, \qquad m_N = rQ^{-1} \pmod N.$$

We construct an integral homogeneous form $\widehat{F}$ with $\widehat{F}(r, Q) = N$, and choose a correction F_0 with the same value at (r, Q) whose reduction modulo two is irreducible. Conditional on g, we sample the coefficients according to the law ν_g given by

$$\mathcal{F} = F_0 + 2(w\widehat{F} + GR_c), \qquad R_c = \sum_{a=0}^{d-1} c_a X^{d-1-a} Y^a, \tag{2.1}$$

where w is uniform in $(W, 2W]$, the c_a are independent uniform integers in $[-H, H]$, and W, H are the integers in (B.4). Thus $\mathcal{F}(r, Q) = N(1 + 2w)$. We call a sample from ν_g a *raw form*. This distribution includes all samples, regardless of whether they subsequently satisfy the algorithm's acceptance conditions. The pair g will be called the *rational environment*. A *target* specifies a prime and an admissible root, together with the level of the descent. We use "parent" and "child" for the two ends of an edge in a relation tree.

The cutoffs are $Y_0 > \cdots > Y_j = B$, and $V_i = Y_i^6$. An active target $l = (i, \ell, s)$ has $Y_{i+1} < \ell \le Y_i$, $\ell \nmid N$, $Qs - r \not\equiv 0 \pmod{\ell}$, and root event $R_l = \{\mathcal{F}(s,1) = 0 \bmod \ell\}$. At $V = V_i$, its point domain is

$$V/2 < b \le 3V/4, \quad MV/16 \le rb - Qa \le MV/8, \quad a \equiv sb \pmod{\ell}, \quad \ell \nmid b.$$

A point $P = (a, b)$ is counted if

$$(a, b) = 1, \quad P^-(b) > B, \quad rb - Qa \in \mathcal{S}_B, \quad 0 < |\mathcal{F}(P)|/\ell \in \mathcal{S}_{Y_{i+1}}.$$

Here $P^-(b)$ is the least prime divisor of b, with $P^-(1) = \infty$. We write $R_P \in \{0, 1\}$ for the indicator of the first three conditions. Let Z_l be the count, zero off R_l, and let $\tau_l(g) = \mathbb{E}_{\nu_g}[Z_l \mid R_l]$ be its true conditional mean. The algorithm uses the point predicates directly. We use their means in the proof.

An incidence count X_{lq} records how many candidates at parent l introduce the distinct child prime–root label in $B < q \le Y_{i+1}$, $q \nmid N$. For an upper bound we also count incidences with even valuation or with a prime that is terminal in the number field. Terminal primes have bad indicator zero; including unneeded incidences can only increase the upper bound. Small primes and terminations with a nontrivial factor are absorbing states. Pass-through targets have identity edges. The counts, prior to any rejection or successful search, satisfy

$$\sum_q X_{lq} \le M_i Z_l, \qquad M_i \le n + 1.$$

The top target is instead a prime $p \in (Y_0, 2Y_0]$. Its rational value is p times a B-smooth integer, and its whole algebraic value is Y_1-smooth. An independent pool of sampled fields handles this initialization; rejected fields remain failed proposals.

2.3 Error bounds required for a descent of increasing depth

The finite transfer theorem will give

$$\delta_i \le \varepsilon_i + 4M_i K_i \sqrt{\delta_{i+1}}, \qquad \delta_j = 0. \tag{2.2}$$

Here δ_i denotes the measure of the exceptional targets at level i, and ε_i is the error in the concentration estimate under the comparison measure. The constants M_i and K_i bound, respectively, the number of children of a relation and the L^2-norm of the density of the distribution of requests. The square root in the recurrence results from the Cauchy–Schwarz inequality.

The number of levels is $j = \log D/(96 \log 2) + O(1)$. A fixed-power error D^{-c} would leave $D^{-c/2^{j-O(1)}}$, which tends to one. We instead prove $\varepsilon_i \le D^{-2A}$ with $A = 4^{j+10} = D^{1/48+o(1)}$; Section 4 shows that the resulting bound is smaller than every fixed power of D^{-1}. We obtain the required single-level error by taking moments of growing order.

3 Estimates for relations and incidences

We now state the two counting results on which the descent analysis rests. The first concerns the concentration of the number of relations for a prescribed prime ideal. The second concerns the distribution of prime ideals arising in these relations. Their proofs are given in Appendices B and C.

3.1 The underlying probability space

Let μ be the distribution of $g = (r, Q)$ specified in the preceding section. By Corollary B.2, there is an event $\mathcal{G}$, depending only on g, on which the estimates for rationally admissible pairs hold simultaneously at all levels and roots. We write

$$\bar{\mu} = \mu(\cdot \mid \mathcal{G} \cap \{Q \nmid N\}).$$

Having chosen g, we draw a raw form under ν_g, before initialization and recursive search.

A *complete-band baseline* is the reference distribution obtained by choosing $g \sim \bar{\mu}$, a prime ℓ uniformly in a complete dyadic interval of active targets (excluding the prime divisors of N), an admissible root s uniformly modulo ℓ, and then $\mathcal{F} \sim \nu_g(\cdot \mid R_{\ell,s})$. We also consider deterministic mixtures of these reference distributions. The incidence estimate below compares them with the measure weighted by the mean number of requests.

Lemma B.3 compares each value after division by the prescribed prime with a mixture of uniform odd-integer intervals. Lemma B.4 does the same for two values in an explicit congruence lattice. At two distinct primitive points of determinant of absolute value $\Delta = \ell t$, the two values after division by ℓ satisfy a congruence modulo $2t$. We retain this congruence in the second moment. The total variation errors are negligible after the density normalizations used below.

3.2 Concentration of the number of relations

Fix an active target $l = (\ell, s)$ and its point domain at one level. Write y for the next smoothness cutoff. Conditional on the root event R_l, the indicator at P, its mean, the total count and its mean are

$$f_P = \mathbf{1}_{\{0<|\mathcal{F}(P)/\ell| \in \mathcal{S}_y\}}, \quad a_P = \mathbb{E}_{\nu_g}[f_P \mid R_l],$$
$$Z_l = \sum_P R_P f_P, \qquad \tau_l = \sum_P R_P a_P.$$

The rational indicator R_P depends only on g and P. Set $m = \lceil 100A \rceil$, with $A = D^{1/48+o(1)}$ as above.

Theorem 3.1 (Concentration of relation counts)**.** *For the specified finite experiment and all sufficiently large inputs, there is a fixed single-value proxy sum $\widetilde{Z}_l$ such that*

$$\frac{\|\widetilde{Z}_l - \tau_l\|_{m,g,l}}{\tau_l} \le D^{-5/24+o(1)}, \tag{3.1}$$

$$\mathbb{E}_{\mathrm{base}} \frac{(Z_l - \widetilde{Z}_l)^2}{\tau_l^2} \le D^{-4A+O(1)}, \tag{3.2}$$

$$\mathbb{P}_{\mathrm{base}}\{|Z_l - \tau_l| > \tau_l/2\} \le D^{-3A}. \tag{3.3}$$

The norm in (3.1) *is under $\nu_g(\cdot \mid R_l)$; its bound is uniform over every $g \in \mathcal{G}$, level, and admissible target. The other two statements hold under every complete-band baseline defined above, including deterministic band mixtures. The threshold and asymptotic errors are uniform in these choices.*

Conditional smoothness densities. Smoothness depends on the magnitude of the integer and its small prime factors. To remove their contribution to the principal character term, we first record this information for one value. We use the same record whenever that value is paired

with a second point. Put $z = R = D^{8A}$ and $Q_o = \prod_{3\le p\le z} p^{\lceil \log R/\log p\rceil}$. Fix a multiplicative grid independently of the coefficients. For its cell I, define

$$D_I(n) = \mathbb{P}_{u \text{ uniform in } I\cap(2\mathbb{Z}+1)}\{0 < |u| \in \mathcal{S}_y \mid \gcd(u, Q_o) = \gcd(n, Q_o)\}.$$

The proxy $D_P^U(n)$ is this exact density, with the large-pattern and near-zero truncations of Definition C.2. In particular, D_P^U is fixed before choosing a second point. Hence $\widetilde{Z}_l = \sum_P R_P D_P^U(\mathcal{F}(P)/\ell)$ is one well-defined random variable in every pair calculation. The proxy is an analysis device; the program does not enumerate Q_o or evaluate these conditional densities.

Higher moments of the conditional density. Under the ideal CRT law on residues modulo Q_o, the capped odd-prime valuations are independent truncated geometric variables. The corresponding smooth-density factors form a product $\mathcal{T}(\gamma)$, where $\gamma = 1 - \beta(x, y)$ and β is the smooth-number saddle point. Its centered logarithmic derivatives have a uniformly bounded exponential moment after division by $\log(2z)$. Consequently Lemma C.5 gives

$$\|\mathcal{T}(\gamma) - 1\|_m \ll m\gamma\log(2z) \quad \text{when } m\gamma\log(2z) \text{ is sufficiently small.}$$

The bound also holds when γ varies with the cell, by a pointwise domination using its maximum. Here $\gamma \ll D^{-1/3}$, so $m\gamma\log(2z) = D^{-1/3+2/48+o(1)} \to 0$.

Since $\log\prod_{p\le z} p \asymp z$ exceeds the logarithm of the window length, which is $O(\log N)$, a window need not contain a full period. Relative counts for retained valuation patterns instead follow from a truncated inclusion–exclusion argument (Lemma C.4). The saddle-point and finite counting errors give, uniformly in every admitted point,

$$\|D_P^U(\mathcal{F}(P)/\ell) - a_P\|_{m,g,l} \le D^{-5/24+o(1)} a_P.$$

Centering at a_P uses exact cancellation on complete grid cells; only endpoint cells and the two discarded regions contribute errors. These steps are proved in Proposition C.7. Minkowski now yields (3.1) directly:

$$\|\widetilde{Z}_l - \tau_l\|_{m,g,l} \le \sum_P R_P \|D_P^U - a_P\|_{m,g,l} \le D^{-5/24+o(1)}\tau_l.$$

Thus the growing moment does not require a joint distribution theorem for m different polynomial values.

The second moment of the remainder. For two distinct points with determinant of absolute value Δ, expand their residual contribution in characters modulo $2\Delta/\ell$. The proxy removes the principal fluctuation up to the errors from its finite approximation and the exceptional valuations it did not record. Both the rational and the algebraic character sums in the remaining terms involve the same Δ. We exploit their common divisibility condition by summing over both conductors together.

Let $\mathfrak{a}(r)$ and $\mathfrak{b}(s)$ be the deterministic normalized primitive-character energy bounds over the fixed rational and algebraic sets of endpoints in (C.51). The character estimates give

$$\sum_r \sqrt{r}\,\mathfrak{a}(r) \le D^C, \qquad \sum_s \sqrt{s}\,\mathfrak{b}(s) \le D^C \exp(-cD^{1/4}).$$

The small-conductor input includes the possible exceptional real character. The large-conductor input is Harper's large-value estimate [27, Proposition 2, (4.1)]; the fixed endpoint sets and their bounds for the maxima of partial sums are established in Lemma C.11.

Inducing a primitive character of conductor r to modulus k contributes the factor r/k to normalized energy. Retaining it on both sides and summing over $\Delta = \ell t$ turns the common conductor and content divisibility condition into a reciprocal least-common-multiple factor. This leaves convergent sums over the common divisors and the product of the two sums over conductors above. Averaging ℓ over the full prime band bounds the additional contribution from $\ell \mid r$. Lemma C.12 gives this calculation with the even algebraic modulus and all prime powers retained.

After division by τ_l^2, both smooth-density factors cancel. The residual, finite-window errors and discarded pairs are assembled in (C.69), proving (3.2). The scales leave the strict hierarchy

$$D^{1/4} \gg D^{1/6}(\log\log N)^2 \gg A(\log D)^2.$$

If $|Z_l - \tau_l| > \tau_l/2$, either the proxy or the residual deviates by more than $\tau_l/4$. Applying Markov's inequality at order m to the proxy and the second-moment bound to the residual gives

$$\mathbb{P}_{\mathrm{base}}\{|Z_l - \tau_l| > \tau_l/2\} \le D^{-10A} + 16D^{-4A+O(1)} \le D^{-3A},$$

which proves the claimed consequence of Theorem C.3.

3.3 The distribution of prime-ideal incidences

Fix a prime ideal at the next level, represented by a prime and a root. Relations for several different parent ideals may involve this ideal in the same number field. We sum all their contributions and define the weight at the next level to be the conditional expectation of this sum.

Give each ordinary relation at parent l weight $a_i(g,l)/\tau_l(g)$. Since $\mathbb{E}[Z_l \mid R_l] = \tau_l(g)$, the conditional expected total weight at that parent is exactly $a_i(g,l)$. The next incoming weight collects these contributions at the child labels. Thus, if X_{lq} counts incidences from ordinary candidates at parent l to child label q, define

$$H_q = \sum_l \frac{a_i(g,l)}{\tau_l(g)} X_{lq}, \qquad a_{i+1}(g,q) = \mathbb{E}_{\nu_g}[H_q \mid R_q]. \tag{3.4}$$

Here a child label includes its root; when written (q,t), q denotes the prime. The transition index on H_q is suppressed. For each fixed environment the mean weights are deterministic. We compute them under the coefficient distribution before the initialization tests, with the incidence and terminal conventions of Section 2.2. Initially $a_0(g,p) = 1/|\mathcal{J}_0^\times|$, where $\mathcal{J}_0^\times = \{p \text{ prime} : Y_0 < p \le 2Y_0,\ p \nmid N\}$.

For positive total mass let

$$W_i = \mathbb{E}_{g\sim\bar{\mu}} \sum_l a_i(g,l)\nu_g(R_l), \qquad d\sigma_i = \frac{a_i(g,l)\mathbf{1}_{R_l}}{W_i}\, d\bar{\mu}(g) d\nu_g(\mathcal{F}), \tag{3.5}$$

with counting measure on labels understood. The incoming incidence measure has density H_q/a_{i+1} relative to σ_{i+1}. Zero-mean branches vanish almost surely and are omitted.

Theorem 3.2 (Incidence estimates)**.** *Start from uniform top parent weights and iterate* (3.4)*, including pass-through edges. For each nonempty internal label domain with $i \ge 1$, there are positive deterministic profiles $B_i(q)$, independent of the environment and root, and losses $L_i \le D^{C(i+1)}$, with $L_0 = 1$, such that*

$$L_i^{-1} B_i(q) \le a_i(g,q,t) \le L_i B_i(q).$$

The profiles vary by at most a factor two within every complete dyadic bucket. Uniformly for every $g \in \mathcal{G}$ and admissible child root in its prescribed parent box,

$$\mathbb{E}_{\nu_g}\big[(H_{q,t}/a_{i+1}(g,q,t))^2 \mid R_{q,t}\big] \le D^C L_i^4. \tag{3.6}$$

There is a deterministic mixture β_i of complete-band baselines for which

$$\frac{d\sigma_i}{d\beta_i} \le C L_i^2. \tag{3.7}$$

All constants and thresholds are uniform through the prescribed j levels. No acceptance or successful-search conditioning is used.

The functions B_i may depend on the input, the level and the prescribed intervals. The growth allowed in the second moment of H_q/a_{i+1}, together with the loss in the comparison of measures (3.7), is compensated by the concentration bound (3.3).

The contribution from one parent band. Consider a complete parent band $J = (Y, 2Y]$ with weights comparable to $A_b w_\ell$ within factors L, where $A_b, w_\ell > 0$ are deterministic and the w_ℓ have a fixed within-band ratio. The conditional means for the parents satisfy $\tau_{\ell,s}(g) \asymp T_i \rho_\ell/\ell$, with $\rho_\ell = \widehat{\rho}(2\mathcal{X}_i/\ell, y_i)$ and $\widehat{\rho}(x,y) = \Psi(x,y)/x$. Here T_i and $\mathcal{X}_i$ are the deterministic rational-count and algebraic-value scales of Proposition B.6; both are fixed before the environment. Under the single child-root condition, let S be the semismooth count that weights a parent prime ℓ by $\ell w_\ell/(T_i\rho_\ell)$. Its mean satisfies

$$\mu_{q,t}(g) \asymp \frac{1}{q}\sum_{\ell \in J} w_\ell \frac{\widehat{\rho}(2\mathcal{X}_i/(q\ell), y_i)}{\rho_\ell}.$$

Here S counts the rational points whose value after division by the prescribed prime is one parent prime $\ell \in J$ times a y_i-smooth integer. The factor q^{-1} comes from the child lattice. The smooth-number ratio theorem [30, Theorem 3], used with $q \le y_i$, and variation of the saddle point give

$$\frac{\widehat{\rho}(2\mathcal{X}_i/(q\ell), y_i)}{\rho_\ell} \asymp q^{\gamma_b}, \qquad \gamma_b = 1 - \beta(2\mathcal{X}_i/Y, y_i) \in [0,1).$$

For example the variation of the saddle point, multiplied by $\log q$, is $O((\log q)^2/(\log \mathcal{X}_i \log y_i)) = O(1/d)$. Substitution leaves the same scale for every environment and child root:

$$\mu_{q,t}(g) \asymp q^{\gamma_b - 1}\sum_{\ell \in J} w_\ell.$$

To recover finite parent roots, remove the terms with $\ell \mid b(P)$. Their mean fraction is $O(\log V_i/Y) = o(1)$, using the complete band and its bounded weight ratio. The incidence correspondence then compares the actual incoming mean $a_b' = \mathbb{E}_{\nu_g}[H_{b,q,t} \mid R_{q,t}]$ with this same scale:

$$\frac{c}{L} \le \frac{a_b'}{A_b q^{\gamma_b - 1}\sum_{\ell \in J} w_\ell} \le CL.$$

The corresponding pointwise comparison and upper bound on the joint smoothness probability give

$$\mathbb{E}_{\nu_g}\big[(H_{b,q,t}/a_b')^2 \big| R_{q,t}\big] \le D^C L^4.$$

The factor L^4 comes from squaring the two L-losses in the count and its mean; the lower bound on the mean at this level controls the diagonal contribution. Proposition B.11 proves these finite comparisons. Proposition B.13 supplies the initial profile under the unconditioned top-parent law.

Summation over dyadic intervals. Starting from $B_1(q) = q^{\gamma_0 - 1}$, the update can be taken to be

$$B_{i+1}(q) = B_i(q) + \sum_{e=e_{i+1}}^{e_i - 1} q^{\gamma_{i,e}-1} \sum_{\ell \in J_e} B_i(\ell), \qquad B < q \le Y_{i+1},$$

where $J_e = \{p \text{ prime} : 2^e < p \le 2^{e+1},\ p \nmid N\}$, $Y_i = 2^{e_i}$, $y_i = Y_{i+1}$, $\gamma_0 = 1 - \beta(2\mathcal{X}_0, Y_1)$, and $\gamma_{i,e} = 1 - \beta(2\mathcal{X}_i/2^e, y_i)$. Pass-through contributes B_i; each active band adds a positive power with exponent in $[-1, 0]$, independent of g and the root. The dyadic support boundaries preserve the factor-two comparison within every bucket. Positive sums also satisfy

$$\frac{(\sum_b H_b)^2}{\sum_b a'_b} \le \sum_b \frac{H_b^2}{a'_b}.$$

Taking expectations and dividing by $\sum_b a'_b$ gives a weighted average of the normalized second moments, with weights $a'_b / \sum_b a'_b$. Hence $L_{i+1} \le D^C L_i$ and the second-moment bound survives all levels without a loss for the number of bands. Equation (B.68) gives the full induction.

Comparison of measures. Let $p_{q,t}(g) = \nu_g(R_{q,t})$; the coefficient sampler gives $c_p \le (q-1)p_{q,t}(g) \le C_p$. For profile B_i, put $M_b = \sum_{q \in J_b} B_i(q)$ and $M_* = \sum_b M_b$, and choose band b with probability M_b/M_*. The exact change of measure is

$$\frac{d\sigma_i}{d\beta_i} = \frac{a_i(g,q,t)}{W_i} \frac{M_* |J_b| (q-1) p_{q,t}(g)}{M_b}.$$

The lower profile bound gives $W_i \ge c_p L_i^{-1} M_*$, while the upper bound and the within-band comparison control the numerator. This proves (3.7), including the environment marginal. Proposition B.16 and Theorem B.17 give the full statement.

3.4 Estimates under the weighted measure

The ordinary theorem is averaged under β_i, and the finite descent uses σ_i. Applying (3.7) to the ordinary tail and integrating (3.6) against the incoming mean weights gives

$$\begin{aligned} \sigma_i\{Z_l < \tau_l/2 \text{ or } Z_l > 2\tau_l\} &\le C L_i^2 D^{-3A} \le D^{-2A}, \\ \mathbb{E}_{\sigma_{i+1}} (H_q / a_{i+1})^2 &\le D^C L_i^4. \end{aligned} \tag{3.8}$$

The first bound holds uniformly for $i = O(\log D)$, because $\log L_i = O((\log D)^2)$ while $A \log D$ grows faster. These are the two arithmetic inputs of Theorem 4.1.

4 Analysis of the descent

Fix the polynomial defining the number field. We call a relation usable if all prime ideals introduced by it can be eliminated in subsequent stages of the descent. The number of usable relations is bounded below by estimating the number that involve at least one exceptional ideal. We then bound the probability that a search of prescribed length finds no usable relation. The argument extends the incidence count in [33, `pc3`, Section 2; `multidescent`, Section 3.2] to successive levels with different weights.

4.1 The distribution of child requests

An incidence from l to q requires both root events R_l and R_q, by Section 2.2. In (3.4)–(3.5) we weighted a child by its conditional mean $a_{i+1}(g,q)$. Weighting it by the actual number of requests, with each request carrying its parent weight, gives the following measure when $W_{i+1} > 0$:

$$d\zeta_{i+1} = \frac{H_q}{W_{i+1}}\, d\bar{\mu}\, d\nu_g.$$

Since $a_{i+1}(g,q) = \mathbb{E}[H_q \mid R_q]$, integration gives total mass one. Dividing the two weights gives

$$\frac{d\zeta_{i+1}}{d\sigma_{i+1}} = \frac{H_q}{a_{i+1}(g,q)}. \tag{4.1}$$

4.2 Propagation of exceptional sets

For a fixed raw form, we define the good targets recursively, starting at the last level. Terminal targets are good. Let $D_l \le Z_l$ denote the number of candidate relations all of whose required child targets are good. We call an active target good if

$$D_l \ge \tau_l/4, \qquad Z_l \le 2\tau_l.$$

A pass-through target is good exactly when its copy is good. Every deleted candidate has at least one bad child, and every candidate introduces at most M_i formal child labels. We use these two bounds in the form

$$Z_l - D_l \le \sum_q X_{lq} \mathbf{1}_{\mathrm{Bad}(q)}, \qquad \sum_q X_{lq} \le M_i Z_l. \tag{4.2}$$

Theorem 4.1 (Propagation of the exceptional measure)**.** *Suppose, at each active internal level,*

$$\sigma_i\{Z_l < \tau_l/2 \text{ or } Z_l > 2\tau_l\} \le \varepsilon_i, \qquad \mathbb{E}_{\sigma_{i+1}}(H_q/a_{i+1})^2 \le K_i^2.$$

Let $\delta_i = \sigma_i(\mathrm{Bad})$*. Then*

$$W_{i+1} \le M_i W_i, \qquad \delta_i \le \varepsilon_i + 4M_i K_i \sqrt{\delta_{i+1}}, \qquad \delta_j = 0. \tag{4.3}$$

If $0 < \varepsilon_i \le \varepsilon \le 1$ *and* $T \ge \max(1, 4M_i K_i)$*, then*

$$\delta_i \le 4T^2 \varepsilon^{\,1/2^{j-i-1}} \quad (1 \le i < j). \tag{4.4}$$

We omit levels of zero total mass.

Proof. Taking expectations in (3.4) and using (4.2) gives

$$W_{i+1} = \mathbb{E}_g \mathbb{E}_{\nu_g} \sum_l \frac{a_i}{\tau_l} \sum_q X_{lq} \le M_i \mathbb{E}_g \sum_l a_i \nu_g(R_l) = M_i W_i.$$

Consider a bad active parent outside the two-sided tail event. It had at least $\tau_l/2$ candidates before deletion and has fewer than $\tau_l/4$ afterwards. At least $\tau_l/4$ candidates therefore requested a bad child. By (4.2),

$$\mathbf{1}_{\mathrm{Bad}(l)} \le \mathbf{1}_{\mathrm{tail}(l)} + \frac{4}{\tau_l} \sum_q X_{lq} \mathbf{1}_{\mathrm{Bad}(q)}.$$

Integrating under σ_i, then using (4.1) and Cauchy–Schwarz, bounds the second term by

$$4\frac{W_{i+1}}{W_i}\int \frac{H_q}{a_{i+1}}\mathbf{1}_{\mathrm{Bad}(q)}\,d\sigma_{i+1} \le 4M_iK_i\sqrt{\delta_{i+1}}.$$

An identity edge obeys the same bound with zero tail error. This proves (4.3). For (4.4), the last active level satisfies $\delta_{j-1} \le \varepsilon$. If the bound holds one level below, the next bound follows from

$$\varepsilon + T\sqrt{4T^2\varepsilon^{\,1/2^{j-i-2}}} \le 4T^2\varepsilon^{\,1/2^{j-i-1}}.$$

This induction uses only $T \ge 1$ and $\varepsilon \le 1$. □

The arithmetic inputs (3.8) give $\varepsilon = D^{-2A}$ and $T = D^{O(j+1)}$. Since $A = 4^{j+10}$,

$$\delta_1 \le 4T^2D^{-2^{j+23}}, \qquad \theta_N := M_0K_0\sqrt{\delta_1} = o(D^{-a}) \quad \text{for every fixed } a. \tag{4.5}$$

4.3 Relation counts for the initial primes

For the initial primes, it is not necessary that a single number field provide sufficiently many usable relations. We sample several independent forms for the same rational pair and consider the sum of their relation counts.

Proposition 4.2 (Coverage of the initial primes)**.** *Suppose the environment event has failure probability ε_G. Under its reference conditional law, suppose $0 \le D_l \le Z_l \le U\tau_l$ for every top label, where $\tau_l(g) = \mathbb{E}_{\nu_g} Z_l > 0$, $U \ge 1$, and*

$$\mathbb{E}_g \operatorname{avg}_l \frac{\mathbb{E}_{\nu_g}(Z_l - D_l)}{\tau_l(g)} \le \theta_N.$$

Suppose the average probability of rejection without finding a factor is at most ρ, and sample $J \ge 1$ trials independently conditional on g. With probability at least $1 - \mathfrak{e}$, either initialization finds a proper factor, or the accepted pool simultaneously satisfies

$$\sum_{a=1}^{J} D_{l,a} \ge J\tau_l/2, \qquad \sum_{a=1}^{J} Z_{l,a} \le 5J\tau_l/4 \le 2J\tau_l$$

for a set of at least three quarters of the top labels. The set is determined by the pool before private shapes are sampled, and

$$\mathfrak{e} = \varepsilon_G + 128U/J + 32(\theta_N + U\rho). \tag{4.6}$$

The actual incoming bound at the top gives $\theta_N = M_0K_0\sqrt{\delta_1}$. Normalize each count before rejection by its conditional mean. Its empirical average has conditional variance at most U/J; the normalized loss from bad candidates and rejected fields has average at most $\theta_N + U\rho$. Chebyshev and Markov then give the dense parent set directly, without a union over top primes. The complete finite proof is in Appendix D.2.

For the fixed density constant U_* of Appendix F.1, the program uses $p_{\mathrm{den}} = 2^{-(U_*+10)dm_b}$, $U = p_{\mathrm{den}}^{-1}$, and $J = 2^{(U_*+20)dm_b}$ trials. Then $J/U = 2^{10dm_b}$, $\rho \le (2H+1)^{-1}$, and $H \ge 2^{8d^2m_b}$, so $\mathfrak{e} = o(1)$. Once the pool covers a top label, the two count bounds in Proposition 4.2 hold for that particular pool and label. The subsequent search uses those bounds directly.

4.4 Search length and number of recursive calls

Consider a good active nonterminal internal request or a covered top request. Terminal requests return immediately; pass-through requests follow their identity transition without sampling a geometric point. The density bounds below concern active searches. At the top, each proposal uses an independent uniform field index and the common geometric point domain, as specified in Section A; rejected fields remain failed proposals. The counts include this field index. In both cases

$$D/Z \geq 1/8, \qquad Z/|\Omega| \geq p_{\mathrm{den}}/4.$$

Every numerical factorization calls the same factoring algorithm on a strictly smaller input and runs to completion. Its termination and expected cost are proved by induction in Section 6. A point becomes verified after that call and all exact predicate checks finish. A failed trial and a verified candidate incur different costs: only the latter can start a child search. We bound their counts separately.

The program uses the integer caps

$$\begin{aligned} M_c &= n+1, \qquad \varepsilon_s = (16kM_c)^{-1}, \\ q_s &= 128\,\mathrm{b}(64kM_c), \\ r_s &= 4096 p_{\mathrm{den}}^{-1}\{q_s + \mathrm{b}(64kM_c)\}. \end{aligned} \tag{4.7}$$

Lemma 4.3 (Bounds on the two numbers of trials)**.** *Suppose a verified good candidate has at most M_c requested children, each with conditional failure probability at most $1/(8M_c)$. Allow at most r_s sampled candidates and at most q_s verified candidates that enter child descent, with the caps in* (4.7)*. The search returns a complete descent or a proper factor with probability at least $1-\varepsilon_s$. Its child-search multiplier is q_sM_c.*

A proposal supplies a verified opportunity with conditional probability at least $p_{\mathrm{den}}/512$; among opportunities the conditional success probability exceeds $1/32$. Only verified opportunities launch child searches. Thus the trial limit controls how long we wait for verified candidates. The second limit controls the number of recursive searches they can start. A proper factor found during either step ends the search. Appendix D.3 proves the general stopped-process bound and verifies that (4.7) dominates it at failure target ε_s.

Backward induction over the finite levels proves the stated failure guarantee for each good request, since $\varepsilon_s \leq 1/(8M_c)$. The same caps are enforced on every bad state and every pass-through state. The total child multiplier across all levels is only

$$(q_sM_c)^{O(j)} = \exp(O((\log D)^2)) = \exp(o(\Lambda)).$$

Unsuccessful trees are discarded. A retained row keeps only its final successful trees.

5 Square-root extraction and integer factorization

For each row we draw a private integer h and release its square $R = h^2 \bmod N$. Descent expresses R as a product of field elements and a rational factor. We combine enough rows to compute a second square root of their product without using the private roots. The construction of the rows and the final gcd follow [33, completion manuscript, Section 7]. For the square certificate we build on Section 8 of the same manuscript and on Lee and Venkatesan [38, Section 6].

5.1 Initial factorizations

Draw h as a uniform n-bit integer. Reject it outside $[1, N)$ and return $\gcd(h, N)$ if it is a nonunit in this interval. For a unit, release R and use planar lattice reduction to find a shortest vector of

$$\mathcal{L}_R = \{(u, v) \in \mathbb{Z}^2 : u \equiv Rv \pmod{N}\}, \qquad \det \mathcal{L}_R = N.$$

We use a fixed rule to break ties and choose the sign.

Lemma 5.1 (Rational-shape probability)**.** *Use the integer parameters of Appendix F.1. Write $\omega(N)$ for the number of distinct prime divisors of N. Let $\mathcal{J}$ contain L primes in $(Y_0, 2Y_0]$, coprime to N, with $L \geq Y_0/(10 \log Y_0)$ and $L \geq 2k^3 2^{\omega(N)}$. For uniform unit h, the canonical vector consists of coprime positive products of exactly k distinct primes of $\mathcal{J}$ with probability at least*

$$p_* = \{128k^2(64n)^{2k}\}^{-1}. \tag{5.1}$$

Appendix D.4 counts coprime pairs (u, v) of prime products in the same square class modulo N. Their quotients are distinct modulo N, and each pair is the unique shortest vector of its lattice up to sign. We apply the lemma to the good primes of Proposition 4.2. The prime number theorem gives the first size bound. For the second, preprocessing gives $\omega(N) \leq \log N / \log B$, so

$$\frac{\omega(N) \log 2}{\log Y_0} \leq \frac{2\kappa \log 2}{b_0 \log \log N}(1 + o(1)) \longrightarrow 0.$$

It follows that $\log L - \log(2k^3 2^{\omega(N)}) \to \infty$. The algorithm can use the shortest vector without knowing the good prime set.

The coordinates have absolute value below $\sqrt{2N}$. Before dividing by their common divisor, we take its gcd with N. A proper factor ends the search, a gcd equal to N rejects the vector, and a unit permits division. We reject zero coordinates, test the required units, and factor the absolute coordinates by calls to the factoring algorithm on smaller inputs. We keep a vector if both coordinates, in absolute value, are squarefree products of at most k primes in $(Y_0, 2Y_0]$. We allow the empty product. Each retained vector gives at most $2k$ top requests.

For sufficiently large inputs a private block gives a unit with probability at least $1/4$. On a shape counted by the lemma all its top searches complete together with probability at least $3/4$. Thus, on the good-pool event, the conditional probability of a row or a proper factor at each attempt exceeds $p_*/8$. For the deterministic column bound $\mathcal{C}$, take

$$R_{\text{shape}} = \lceil 32(\mathcal{C} + 1)/p_* \rceil . \tag{5.2}$$

The expected wait for each further row or factor is at most $8/p_*$. Markov's inequality therefore gives probability at least $3/4$ of finding a factor or collecting $\mathcal{C} + 1$ rows within the cap. We retain duplicate public rows, each with its own fresh private root.

5.2 The relation associated with a completed descent

For an accepted field a, let F_a be the primitive part of $\mathcal{F}_a$. A prime is good for that field if it divides none of the raw content, the leading coefficient of $F_a(T, 1)$, and its discriminant. If a good prime p divides $F_a(P)$ at a primitive point P, exactly one prime ideal above p occurs in the atom. Its valuation is $v_p(F_a(P))$, and the field, prime and root identify that ideal (Lemma E.1). For each incoming odd valuation the child relation cancels its parity at the same ideal. A completed tree thus has even valuation at every good large ideal. We count repeated occurrences of a point separately. Bad primes can end a branch and are treated by the certificate below.

Let $c_{i,a,P}$ count the signed occurrences of P in field a in row i. The sign of a top prime is determined by its position in the quotient u_i/v_i. Collect the sign of the quotient and the rational factors of all atom relations into q_i, including $Q/G(P)$ for each inserted child. Multiplying the relations gives

$$R_i \equiv q_i \prod_a \prod_{P \text{ in field } a} (a_P - b_P m_N)^{c_{i,a,P}} \pmod{N}. \tag{5.3}$$

Only signs, Q and B-smooth values enter the factorization of q_i. We check each denominator is a unit and verify (5.3) before retaining the row. Appendix E.6 gives the individual factors.

5.3 Linear algebra and square-root computation

Quadratic character columns are a standard device for detecting square classes in NFS [1, 61]. Our proof uses a uniform separation statement, rather than a heuristic count of the remaining square classes. For background on the square-root step we refer to [50, 63]. We use dense binary elimination in the cost analysis below; the use of sparse methods such as Wiedemann's algorithm and its block variants [66, 21, 49, 62] is not required for either of our time bounds.

Even valuations at all good ideals leave a possible square class in each field. To determine it, we add columns for rational corrections, occurrence and small-ideal parities in each field, signs and prime parities of norms across fields, and sampled quadratic characters. Appendix E.6 gives the column count (E.27) and the rule for primes first seen in a later row.

Choose a nonzero binary dependence among $\mathcal{C} + 1$ rows by a fixed rule, and let I index its support. The parity equations make $\prod_{i \in I} q_i$ a positive rational square. They also make each good ideal valuation even and the product of the signed norm classes a positive rational square. By Proposition E.3, the sampled characters detect every nonrational square class that can occur in these products, simultaneously in all accepted fields, except on the stated certificate failure event. This follows from unconditional Chebotarev and exceptional-zero estimates [35, 60]. The event concerns entire groups of square classes and hence applies to the dependence selected after the probes.

In odd degree the norm of a rational class equals that class modulo squares. Using the saved norm factors, Theorem E.4 therefore determines the rational correction, constructs the field roots and evaluates them to obtain a unit Y such that

$$Y^2 \equiv \prod_{i \in I} R_i \pmod{N}.$$

The program checks this congruence and each denominator used in evaluation. If reconstruction or verification fails, the epoch ends before any private roots are multiplied.

5.4 Probability of a nontrivial factor

Fix the public record just before the first private multiplication. It determines I and Y. Each retained root has a fresh location for the private root, and the fixed schedule for that slot reveals only its square and root-independent control information. Consequently the selected h_i remain independent and uniform among the unit roots of their respective squares (Theorem E.5). Since I is nonempty, $X = \prod_{i \in I} h_i$ is uniform among the $2^{\omega(N)}$ unit roots of Y^2. For odd N, including those with repeated prime factors, exactly $X = Y$ and $X = -Y$ give a trivial gcd. Thus

$$\mathbb{P}(1 < \gcd(X - Y, N) < N \mid \text{complete public record}) = 1 - 2^{1-\omega(N)} \geq \tfrac{1}{2}. \tag{5.4}$$

A trivial gcd ends the epoch and we discard its private roots.

With pool failure probability at most $\mathfrak{e}$ and public certificate upper bound on the failure probability $\varepsilon_{\mathrm{pub}}$, this gives

$$\mathbb{P}(\text{epoch succeeds}) \geq (1 - 2^{1-\omega(N)}) \left[\tfrac{3}{4}(1-\mathfrak{e}) - \varepsilon_{\mathrm{pub}}\right]_+ . \tag{5.5}$$

For $\mathfrak{e} \leq 1/4$ and $\varepsilon_{\mathrm{pub}} \leq 1/16$, the lower bound is at least $1/4$. We use the weaker $p_0 = 1/128$ in the running-time proof.

6 Proof of the unconditional bound

Procedure U maintains a list of integers to be factored. It tests for primality and perfect powers deterministically and verifies every nontrivial factorization. The parameters and the conditions under which they may be used are checked separately for each new input. If preprocessing finds a divisor, the algorithm continues on the resulting smaller inputs. Consequently, the splitting procedure is applied only to a composite integer that is not a prime power. At termination the prime factors and their multiplicities are checked by multiplication. This recursive organization follows [33, `mfU_theorem`, Section 9; `escapeR2`, Section 8].

In one epoch we first sample the fields and their probes. We then make at most R_{shape} row attempts. Once we have $\mathcal{C}+1$ verified rows, we compute and verify a public square root, multiply the selected private roots and take the final gcd. Appendix F.9 gives the execution order. Rejected field samples keep their sample indices. Only a verified candidate starts a child search in its field. Every attempted search, including a discarded tree, counts towards its cap. The caps bound the number of auxiliary calls on every execution. Each auxiliary call runs the factoring algorithm on a checked input $2 \leq m < N$.

If an epoch fails, Split makes one trial-division step and starts a fresh epoch. Successive failures test the integers $2, 3, \ldots$ in order. Each test takes $\mathrm{poly}(\log N)$ time, and by the time the counter reaches $\lfloor \sqrt{N} \rfloor$ it has found a proper factor.

Theorem 6.1 (Unconditional running-time bound)**.** *Program* U *satisfies the conclusion of Theorem 1.1.*

Proof. First fix N and suppose that for every $m < N$ the worst running time $W(m)$ over all random tapes is finite. Every epoch has bounded direct work and makes a bounded number of auxiliary calls, each costing at most $\max_{m<N} W(m)$. Trial division bounds the number of epochs. A proper split and a perfect-power reduction both leave smaller inputs. This proves $W(N) < \infty$ by induction, including when a scale check fails. All returned primes and products are verified, which gives correctness on every tape.

For the expected time use the nondecreasing function

$$\phi(N) = (\log 2)\lceil (n m_b^2)^{1/3} \rceil, \qquad E_C(N) = e^{C\phi(N)}, \qquad \phi(N) \sim \Lambda(N).$$

Shape coordinates are at most $b_{\mathrm{sh}}(N) = \lceil \sqrt{2N} \rceil$; relation values passed to the factorizer are at most $b_{\mathrm{val}}(N) = \lceil N^{10^{-6}} \rceil$. The corresponding ratios of ϕ to $\phi(N)$ tend to $2^{-1/3}$ and 10^{-2}. Lemma F.1 bounds direct work per epoch by f_0 and the numbers of calls of the two types by f_1, f_2. Its bounds have the following leading coefficients.

Quantity	Upper coefficient of Λ in $\log f_i$
Direct work f_0	$a_0 = 1\,030\,992\,164\,870$
Shape-coordinate calls f_1	$a_1 = 25\,789\,803\,908$
Relation-value calls f_2	$a_2 = 51\,559\,607\,694$

Accordingly put

$$F_C(N) = f_0 + f_1 E_C(b_{\mathrm{sh}}(N)) + f_2 E_C(b_{\mathrm{val}}(N)).$$

All three terms grow more slowly on this scale than $E_C(N)$ when

$$C > \max\left\{a_0, \frac{a_1}{1-2^{-1/3}}, \frac{a_2}{1-10^{-2}}\right\}. \tag{6.1}$$

The displayed coefficients satisfy this inequality for $C = 10^{13}$.

Suppose $\mathbb{E}T(m) \le A_0 E_C(m)$ for every $m < N$, where $A_0 \ge 1$ will be fixed below. Immediately before an auxiliary call, its input is determined by the past and its random bits are fresh. The induction hypothesis bounds the entire call, including verification. Summing these conditional expectations over the available possible calls gives

$$\mathbb{E}[\mathrm{Cost}(\mathrm{Epoch}(N)) \mid \mathrm{past}] \le c_E A_0 F_C(N)$$

for an absolute constant c_E. An unused call position has cost zero.

All auxiliary calls terminate correctly, so the success estimate for exact numerical tests applies. For sufficiently large inputs, each fresh epoch succeeds with probability at least $p_0 = 1/128$ conditional on its past. This threshold is independent of A_0. Write A_r for the event that epoch r is reached and C_r for its cost, including any trial-division step. Then

$$\mathbb{P}(A_r) \le (1-p_0)^{r-1}, \qquad \mathbb{E}[\mathbf{1}_{A_r} C_r] \le c A_0 F_C(N) \mathbb{P}(A_r).$$

The second inequality conditions on the past before epoch r, which allows its cost to depend on its outcome. Summing over r and including preprocessing gives

$$\mathbb{E}\,\mathrm{Cost}(\mathrm{Split}(N)) \le c_S A_0 F_C(N). \tag{6.2}$$

Choose $\eta > 0$ so that $\max\{a_0, a_1 + 2^{-1/3}C, a_2 + 10^{-2}C\} < C - 4\eta$. Eventually $F_C(N) \le 3e^{(C-4\eta)\phi(N)}$. Allowing for polynomial work and fixed factors, the expected cost at one worklist node, including its auxiliary factorizations, is at most $A_0 e^{(C-3\eta)\phi(N)}$. Choose a fixed N_0 large enough for these estimates, for final verification, and for $2\lfloor \log_2 N \rfloor + 1 \le e^{\eta\phi(N)}$. Then choose A_0 large enough to cover the full and local costs of all $2 \le m \le N_0$, using the finite worst-case bounds proved above.

To finish the induction, consider all worklist nodes descended from the original input N. At a node of size $m \le N$, every auxiliary input is strictly smaller than m, so the same conditional local bound applies. Counting prime leaves with multiplicity, their product is N. There are at most $\lfloor \log_2 N \rfloor$ leaves and hence at most $2\lfloor \log_2 N \rfloor - 1$ outer nodes. Summing the local costs and using monotonicity gives

$$\mathbb{E}T(N) \le A_0(2\lfloor \log_2 N \rfloor + 1)e^{(C-3\eta)\phi(N)} \le A_0 e^{(C-2\eta)\phi(N)} \le A_0 E_C(N).$$

Each local cost already includes its auxiliary tree. Since A_0 is fixed and $\phi(N) \sim \Lambda(N)$, the claimed bound follows. □

7 The improved time bound under GRH

We turn to the proof of Theorem 1.2. Under GRH, the estimates for relations and incidences can be established in smaller search regions, and one random field suffices for the construction of all the rows in an iteration of the algorithm. The square-root construction and factor extraction of Section 5 are unchanged. We describe the modifications to relation collection and to the arithmetic routines that give the smaller constant in the exponent.

7.1 Comparison with the unconditional algorithm

Step	U, unconditional	G, GRH
Polynomial coefficients	Uniform in a box, with independent w	Finite discrete Gaussian, with $w = 0$
Top primes	Pool of independent fields	One field serving a positive fraction of each prime band
Search region	$V_i = Y_i^6$	$V_i \asymp Y_i^{51/100} B^{247/300}$
Numerical tests	Factoring smaller inputs	Bounded smoothness tests
Square certificate	Odd-degree separation and exact reconstruction	Same separation theorem, with bounded probe arithmetic and character tests on retained rows

Both programs use dense binary elimination and reconstruct a square root in characteristic zero. The certificate itself is unconditional.

7.2 Polynomial selection and relation collection

Let $n = \mathrm{b}(N)$ and $m_b = \mathrm{b}(n)$. Choose the least odd $d \geq 3$ with $64d^3 m_b \geq 27n$, so $d = (3/4 + o(1))D^{1/3}$. The search regions have parameters

$$B = 2^{e_B}, \quad e_B = \lceil 19dm_b/15 \rceil, \qquad V_i = 2^{\lceil (153e_i + 247e_B)/300 \rceil}, \quad Y_i = 2^{e_i}.$$

In particular $\log B = (19/20 + o(1))\Lambda$. At the bottom $V_i \asymp B^{4/3}$; at the top the exponent of Y_i in V_i is only $51/100$. These smaller values of V_i reduce the polynomial heights needed in the descent. Section H.1 gives the recursion for e_i and the finite parameter checks. The sampled form is

$$\mathcal{F} = F_0 + 2GR_c, \qquad \sigma = V_0^2 n^{10}, \qquad H = \sigma n^{10}.$$

We approximate the discrete Gaussian on c with weights $\exp(-\pi\|c\|^2/\sigma^2)$ (Lemma H.1). Lemmas G.4 and G.5 give the induced laws of the values at one or two points, also under a prescribed root condition.

Counting rational values. The denominator is a prime in a prescribed interval, and the numerator must be B-smooth. We impose the parent congruence on the numerator, which ranges over the longer interval. The smooth-character estimate under GRH follows from Lagarias and Soundararajan [36, Proposition 5.1]. It leaves only an ordinary interval count for the denominator primes. This permits denominator windows shorter than the square of the parent scale.

Distribution of requests at the next level. For fixed Q, ℓ, the map

$$(r, s) \longmapsto r - Qs \pmod{Q\ell}$$

is injective for r in its prescribed interval and $s \bmod \ell$. Indeed the image determines r modulo Q, hence r itself, and then determines s modulo ℓ. Parseval's identity therefore bounds the sum of the errors over all roots. We average the resulting estimates jointly over the environment and the root, using deterministic weights that are independent of the root.

Write $c_i(\ell)$ for the smooth density at the coefficient and value scales of level i, and T_i for the scale of the rational count. For a good parent the mean is $\tau_i(g, \ell, s) \asymp T_i c_i(\ell)/\ell$. A request for a

child q of this parent contributes a density ratio $c_i(q\ell)/c_i(\ell)$: both densities refer to the parent's search region. After dropping the good-parent restriction, all terms are nonnegative. We may therefore bound the count with the denominator restriction by the complete semismooth count, which also includes the excluded denominators.

Let δ_i be the mass of bad requests for the finite measure (G.42). Write $\varepsilon_i^{\mathrm{ord}}$ and $\varepsilon_i^{\mathrm{in}}$ for the largest root-averaged errors in the rational counts for a parent's own request and an incoming child-root request, respectively, both in that parent's region. Proposition G.11 gives

$$\delta_i \le D^C \delta_{i+1} + D^C\{\varepsilon_i^{\mathrm{ord}} + \sqrt{\varepsilon_i^{\mathrm{in}}} + D^{-3A/2}\}, \qquad \delta_j = 0, \qquad A = (j+10)^3.$$

At the top, one further application of Cauchy–Schwarz bounds the average proportion of bad private parents in each prescribed prime band. Theorem G.1 gives these bounds simultaneously in the same field, before we sample any private shape or search for its relations.

Sampling a candidate. Some parents satisfy $\ell > V_i$. To draw a point of their search region efficiently, write $a = sb + \ell z$. For fixed b the admissible z form an integer interval. Euclidean floor sums count the points up to any given b, omitting the multiples of ℓ. We draw a binary rank and invert the prefix count. Each legal point has the same mass and a proposed rank is accepted with probability at least $1/2$. On a good request $|\Omega| \le 32V_i^2/\ell$. Together with $V_i^2 \ge Y_i$, this bound converts the number of relations into a lower bound on the probability of sampling one (Lemma H.2).

7.3 Smoothness testing and square certification

For an integer x and a cutoff Y, accept only a verified factorization of x into primes at most Y. Under ordinary RH, for all sufficiently large Y, Proposition H.3 bounds the rejection probability by any prescribed $0 < \epsilon < 1$, uniformly over Y-smooth inputs. The bit cost on every tape is

$$L_Y[1/2, 2.01]\ \mathrm{poly}(\mathrm{b}(x) + \mathrm{b}(Y), \log(1/\epsilon)).$$

For a prime divisor p, genus-two Jacobian orders near p^2 range over an interval of length $p^{3/2} = (p^2)^{3/4}$. Younis's short-interval theorem [67, Theorem 1.3] counts smooth integers in this interval under ordinary RH. The order-fibre bound of Lenstra, Pila and Pomerance [43, Theorem 1.1] then counts the corresponding quintic curve models. An earlier analysis under RH with leading constant 2 is given in the manuscript collection [33, `genus2rh`, Theorem R and Corollary R2]. Here we sample the curve and a divisor together and use an odd scalar to distinguish their reductions at two primes. The finite bound above allows a fixed slack of 0.01 in the constant.

Each search has separate limits for raw candidates, algebraic smoothness tests and candidates that launch child searches. Exceeding the algebraic limit fails that search (Lemma H.4). We assign independent random blocks to the possible calls. Skipping a test then leaves the subsequent proposals and private roots untouched, so we can couple the program to exact smoothness tests up to the first false rejection of a smooth input.

The unconditional separation theorem (Proposition E.3) uses probes whose bit lengths are at most $S^* + 1$, with $\log S^* = (1/2 + o(1))\Lambda$. For each candidate prime we use capped Miller–Rabin tests, a polynomial gcd with $X^p - X$, and bounded splitting to find all linear roots. Together with the bounded sampler for integer intervals, these operations agree with the ideal prime and root sampling procedure except with probability $o(1)$. Uniformity among the prime–root pairs is a property of the ideal procedure.

By Proposition H.6, constructing the probes costs $\mathrm{poly}(n,d)(S^*)^4$. We compute character entries only for retained occurrences. With their number bounded by A_{occ} as in (H.30), this costs $\mathrm{poly}(n,d)A_{\mathrm{occ}}(S^*)^3$. A rejected candidate uses only its local record. The retained matrix includes the norm, sign, ideal and occurrence columns used in the unconditional certificate. We apply the rational squarefree correction and check the reconstructed field root and final modular square exactly.

7.4 Expected running time

Two density lower bounds determine how many trials each search needs:

$$p_A = 2^{-\lceil 36dm_b/25\rceil}, \qquad p_{\mathrm{den}} = 2^{-\lceil 25dm_b/12\rceil}.$$

The first is a lower bound for the one-point algebraic smoothness probability under the polynomial distribution conditional on the prescribed root. The second is a lower bound for the density of relations in the search region, in conditional expectation. Together with the concentration estimates, these bounds determine separate limits on the number of algebraic smoothness tests and the total number of sampled pairs at good targets. Section H.8 sets the column bound after the certificate parameters and then sets all search and record limits. The resulting bit costs of one epoch are as follows.

Operation	Coefficient of Λ
Raw proposals, ranks, rational tests, values and gcds	$237/80$
Algebraic smoothness tests on all requests	$< 2967/1000$
Smoothness tests on private-shape coordinates	$< 29/10$
Dense elimination, products, correction and public root	$57/20$
Probe sampling and enumeration of linear roots	2
Characters of retained occurrences	$49/20$

The largest coefficient comes from algebraic recognition:

$$\frac{62}{25} + \sqrt{\frac{76313}{322500}} < \frac{2967}{1000} < \frac{297}{100}.$$

For every sufficiently large input reaching this stage the conditional probability of a proper factor in a fresh epoch is at least $1/512$. Trial division ensures termination on every tape. Charging each epoch at its full cost, and then summing over the outer factorization tree, gives the expected bound for G (Theorem H.9).

Lemma 7.1 (Interleaving verified programs)**.** *In one bit-cost model, let U,G be fixed programs whose accepted outputs are factorizations verified by a deterministic polynomial-time procedure. There is a fixed program F, using independent random tapes for the two simulations, such that on each pair of tapes*

$$T_{\mathsf{F}}(N) \le C\min\{T_{\mathsf{U}}(N), T_{\mathsf{G}}(N)\} + \mathrm{poly}(\log N).$$

Here C is an absolute constant. The time of a simulation is infinite if it never produces a verified output. If U terminates correctly on every tape, so does F.

Proof. Store the two configurations on disjoint work tapes and alternate one bit step of each. Stop at the first verified factorization. Each simulation includes its verification. The schedule costs a constant factor per step and polynomial time for initialization. □

Taking expectations and using the bound for U proves Theorem 1.1 for F. Under GRH the bound for G gives Theorem 1.2 for the same program. The constant simulation overhead and polynomial initialization are absorbed by the $o(1)$ in each exponent.

8 Concluding remarks

The preceding analysis gives a variant of the number field sieve with an unconditional expected running time of $L_N[1/3, 10^{13}]$ and, under GRH, an expected running time of $L_N[1/3, 2.97]$. The principal ingredient is an estimate for the distribution of prime ideals occurring during a descent, in addition to an estimate for the number of relations at a prescribed prime ideal. The bounds apply to factoring a fixed input into primes, including auxiliary factorizations, square-root computation and verification. Neither bound assumes that the input is chosen at random or that it is a product of two primes.

The constants in the unconditional and conditional analyses serve different purposes. The former makes all estimates uniform over a descent of increasing depth, while the latter exploits sharper character-sum estimates and a bounded smoothness test. There is no claim that these parameter choices are optimal. In particular, it would be interesting to reduce the cost of smoothness testing and relation collection further, or to obtain a smaller unconditional constant. As in the analysis of the classical NFS in [37], the asymptotic leading constants alone do not specify the running time for integers of a given practical size.

References

[1] Adleman, L.M.: Factoring numbers using singular integers. In: Proceedings of the 23rd Annual ACM Symposium on Theory of Computing, pp. 64–71 (1991). doi:10.1145/103418.103432.

[2] Agrawal, M., Kayal, N., Saxena, N.: Errata: PRIMES is in P. Annals of Mathematics **189**(1), 317–318 (2019). doi:10.4007/annals.2019.189.1.6.

[3] Agrawal, M., Kayal, N., Saxena, N.: PRIMES is in P. Annals of Mathematics **160**, 781–793 (2004). doi:10.4007/annals.2004.160.781.

[4] Bai, S., Brent, R.P., Thomé, E.: Root optimization of polynomials in the number field sieve. Mathematics of Computation **84**(295), 2447–2457 (2015). doi:10.1090/S0025-5718-2015-02926-3.

[5] Bai, S., Bouvier, C., Kruppa, A., Zimmermann, P.: Better polynomials for GNFS. Mathematics of Computation **85**(298), 861–873 (2016). doi:10.1090/mcom3048.

[6] Bai, S., Gaudry, P., Kruppa, A., Thomé, E., Zimmermann, P.: Factorisation of RSA-220 with CADO-NFS. Technical report, 10 May 2016, revised 2 November 2016. Authors' report.

[7] Bai, S., Thomé, E., Zimmermann, P.: Factorisation of RSA-704 with CADO-NFS. IACR Cryptology ePrint Archive, Report 2012/369 (2012). ePrint:2012/369.

[8] Bai, S., Zimmermann, P.: Size optimization of sextic polynomials in the number field sieve. Manuscript (2012, revised June 2013). https://hal.science/hal-00760331. Revised version: authors' manuscript.

[9] Barbulescu, R., Lachand, A.: Some mathematical remarks on the polynomial selection in NFS. Mathematics of Computation **86**(303), 397–418 (2017). doi:10.1090/mcom/3112.

[10] Bernstein, D.J.: How to find smooth parts of integers. Manuscript, 10 May 2004. Author's manuscript.

[11] Boudot, F., Gaudry, P., Guillevic, A., Heninger, N., Thomé, E., Zimmermann, P.: Comparing the difficulty of factorization and discrete logarithm: a 240-digit experiment. In: Advances in Cryptology – CRYPTO 2020, Part II, Lecture Notes in Computer Science, vol. 12171, pp. 62–91. Springer (2020). doi:10.1007/978-3-030-56880-1_3. IACR ePrint 2020/697.

[12] Boudot, F., Gaudry, P., Guillevic, A., Heninger, N., Thomé, E., Zimmermann, P.: The state of the art in integer factoring and breaking public-key cryptography. IEEE Security & Privacy **20**(2), 80–86 (2022). doi:10.1109/MSEC.2022.3141918.

[13] Bouillaguet, C., Zimmermann, P.: Parallel structured Gaussian elimination for the number field sieve. Mathematical Cryptology **0**(1), 22–39 (2021). https://hal.science/hal-02098114.

[14] Brent, R.P., Zimmermann, P.: Modern Computer Arithmetic. Cambridge Monographs on Applied and Computational Mathematics, vol. 18. Cambridge University Press (2010). doi:10.1017/CBO9780511921698.

[15] Buhler, J.P., Lenstra, Jr., H.W., Pomerance, C.: Factoring integers with the number field sieve. In: Lenstra, A.K., Lenstra, Jr., H.W. (eds.) The Development of the Number Field Sieve. Lecture Notes in Mathematics, vol. 1554, pp. 50–94. Springer (1993). doi:10.1007/BFb0091539.

[16] The CADO-NFS Development Team: CADO-NFS, an implementation of the number field sieve algorithm. Project website, accessed 26 November 2026. https://cado-nfs.gitlabpages.inria.fr/. Source repository.

[17] Canfield, E.R., Erdős, P., Pomerance, C.: On a problem of Oppenheim concerning "factorisatio numerorum". Journal of Number Theory **17**(1), 1–28 (1983). doi:10.1016/0022-314X(83)90002-1.

[18] Cantor, D.G.: Computing in the Jacobian of a hyperelliptic curve. Mathematics of Computation **48**(177), 95–101 (1987). doi:10.1090/S0025-5718-1987-0866101-0.

[19] Cogentic. *Unconditional Integer Factorization in Time* $L_N[1/3, O(1)]$. Google Research manuscript, 25 November 2026. Repository version 4a7bc8a.

[20] Coppersmith, D.: Modifications to the number field sieve. Journal of Cryptology **6**, 169–180 (1993). doi:10.1007/BF00198464.

[21] Coppersmith, D.: Solving homogeneous linear equations over GF(2) via block Wiedemann algorithm. Mathematics of Computation **62**(205), 333–350 (1994). doi:10.1090/S0025-5718-1994-1192970-7.

[22] de Bruijn, N.G.: The asymptotic behaviour of a function occurring in the theory of primes. Journal of the Indian Mathematical Society (N.S.) **15**, 25–32 (1951).

[23] Dixon, J.D.: Asymptotically fast factorization of integers. Mathematics of Computation **36**(153), 255–260 (1981). doi:10.1090/S0025-5718-1981-0595059-1.

[24] Drappeau, S.: Sommes friables d'exponentielles et applications. Canadian Journal of Mathematics **67**(3), 597–638 (2015). doi:10.4153/CJM-2014-036-5.

[25] Drappeau, S.: Théorèmes de type Fouvry–Iwaniec pour les entiers friables. Compositio Mathematica **151**, 828–862 (2015). doi:10.1112/S0010437X14007933.

[26] Granville, A.: Smooth numbers: computational number theory and beyond. In: Buhler, J.P., Stevenhagen, P. (eds.) Algorithmic Number Theory: Lattices, Number Fields, Curves and Cryptography. MSRI Publications, vol. 44, pp. 267–323. Cambridge University Press (2008).

[27] Harper, A.J.: Minor arcs, mean values, and restriction theory for exponential sums over smooth numbers. Compositio Mathematica **152**(6), 1121–1158 (2016). doi:10.1112/S0010437X15007782.

[28] Harper, A.J.: On a paper of K. Soundararajan on smooth numbers in arithmetic progressions. Journal of Number Theory **132**(1), 182–199 (2012). doi:10.1016/j.jnt.2011.07.005.

[29] Hildebrand, A., Tenenbaum, G.: Integers without large prime factors. Journal de Théorie des Nombres de Bordeaux **5**(2), 411–484 (1993). doi:10.5802/jtnb.101.

[30] Hildebrand, A., Tenenbaum, G.: On integers free of large prime factors. Transactions of the American Mathematical Society **296**(1), 265–290 (1986). doi:10.1090/S0002-9947-1986-0837811-1.

[31] Hildebrand, A.: On the number of positive integers $\leq x$ and free of prime factors $> y$. Journal of Number Theory **22**(3), 289–307 (1986). doi:10.1016/0022-314X(86)90013-2.

[32] Joux, A., Lercier, R.: Improvements to the general number field sieve for discrete logarithms in prime fields. A comparison with the Gaussian integer method. Mathematics of Computation **72**(242), 953–967 (2003). doi:10.1090/S0025-5718-02-01482-5.

[33] kasper0406. *Factoring-connection*. Collection of manuscripts, 28 October 2026. Repository version 351e8ee. Repository homepage.

[34] Kleinjung, T., Aoki, K., Franke, J., Lenstra, A.K., Thomé, E., Bos, J.W., Gaudry, P., Kruppa, A., Montgomery, P.L., Osvik, D.A., te Riele, H., Timofeev, A., Zimmermann, P.: Factorization of a 768-bit RSA modulus. In: Advances in Cryptology – CRYPTO 2010, Lecture Notes in Computer Science, vol. 6223, pp. 333–350. Springer (2010). IACR ePrint 2010/006.

[35] Lagarias, J.C., Odlyzko, A.M.: Effective versions of the Chebotarev density theorem. In: Fröhlich, A. (ed.) Algebraic Number Fields, pp. 409–464. Academic Press (1977).

[36] Lagarias, J.C., Soundararajan, K.: Counting smooth solutions to the equation A+B=C. Proceedings of the London Mathematical Society **104**(4), 770–798 (2012). doi:10.1112/plms/pdr037.

[37] Le Gluher, A., Spaenlehauer, P.-J., Thomé, E.: Refined analysis of the asymptotic complexity of the number field sieve. Mathematical Cryptology **1**(1), 71–88 (2021). Earlier version: IACR Cryptology ePrint Archive, Report 2020/829 (2020). ePrint:2020/829. arXiv:2007.02730.

[38] Lee, J.D., Venkatesan, R.: Rigorous analysis of a randomised number field sieve. Journal of Number Theory **187**, 92–159 (2018). doi:10.1016/j.jnt.2017.10.019.

[39] Lenstra, A.K., Lenstra, Jr., H.W., Lovász, L.: Factoring polynomials with rational coefficients. Mathematische Annalen **261**, 515–534 (1982). doi:10.1007/BF01457454.

[40] Lenstra, A.K., Lenstra, Jr., H.W., Manasse, M.S., Pollard, J.M.: The number field sieve. In: Proceedings of the 22nd Annual ACM Symposium on Theory of Computing, pp. 564–572 (1990). doi:10.1145/100216.100295.

[41] Lenstra, A.K.: Factoring polynomials over algebraic number fields. In: van Hulzen, J.A. (ed.) Computer Algebra: EUROCAL '83. Lecture Notes in Computer Science, vol. 162, pp. 245–254. Springer (1983). doi:10.1007/3-540-12868-9_108.

[42] Lenstra, Jr., H.W., Pila, J., Pomerance, C.: A hyperelliptic smoothness test, I. Philosophical Transactions of the Royal Society of London, Series A **345**(1676), 397–408 (1993). doi:10.1098/rsta.1993.0138.

[43] Lenstra, Jr., H.W., Pila, J., Pomerance, C.: A hyperelliptic smoothness test, II. Proceedings of the London Mathematical Society **84**(1), 105–146 (2002). doi:10.1112/plms/84.1.105.

[44] Lenstra, Jr., H.W., Pomerance, C.: A rigorous time bound for factoring integers. Journal of the American Mathematical Society **5**, 483–516 (1992). doi:10.1090/S0894-0347-1992-1137100-0.

[45] Lenstra, Jr., H.W.: Factoring integers with elliptic curves. Annals of Mathematics **126**(3), 649–673 (1987). doi:10.2307/1971363.

[46] Lu, E.: Factoring RSA-260. Cognition, online computation report, 9 September 2026. https://cognition.com/blog/factoring-rsa-260.

[47] Montgomery, H.L., Vaughan, R.C.: Multiplicative number theory II: Primes and sieves. Cambridge University Press (2026). doi:10.1017/9781009445030. The theorem numbering cited here refers to the author-hosted draft: Authors' manuscript.

[48] Montgomery, H.L., Vaughan, R.C.: Multiplicative Number Theory I. Classical Theory, Cambridge Studies in Advanced Mathematics, vol. 97. Cambridge University Press (2007). doi:10.1017/CBO9780511618314.

[49] Montgomery, P.L.: A block Lanczos algorithm for finding dependencies over GF(2). In: Advances in Cryptology – EUROCRYPT '95. Lecture Notes in Computer Science, vol. 921, pp. 106–120. Springer (1995). doi:10.1007/3-540-49264-X_9.

[50] Montgomery, P.L.: Square roots of products of algebraic numbers. In: Gautschi, W. (ed.) Mathematics of Computation 1943–1993: A Half-Century of Computational Mathematics. Proceedings of Symposia in Applied Mathematics, vol. 48, pp. 567–571. AMS (1994).

[51] Morrison, M.A., Brillhart, J.: A method of factoring and the factorization of F_7. Mathematics of Computation **29**(129), 183–205 (1975). doi:10.1090/S0025-5718-1975-0371800-5.

[52] Pascadi, A.: Smooth numbers in arithmetic progressions to large moduli. Compositio Mathematica **161**(8), 1923–1974 (2025). doi:10.1112/S0010437X2500747X.

[53] Pollard, J.M.: The lattice sieve. In: Lenstra, A.K., Lenstra, Jr., H.W. (eds.) The Development of the Number Field Sieve. Lecture Notes in Mathematics, vol. 1554, pp. 43–49. Springer (1993). doi:10.1007/BFb0091538.

[54] Pomerance, C.: A tale of two sieves. Notices of the American Mathematical Society **43**(12), 1473–1485 (1996).

[55] Prest, T., Zimmermann, P.: Non-linear polynomial selection for the number field sieve. Journal of Symbolic Computation 47(4), 401–409 (2012). doi:10.1016/j.jsc.2011.09.004.

[56] Pomerance, C.: Analysis and comparison of some integer factoring algorithms. In: Lenstra, Jr., H.W., Tijdeman, R. (eds.) Computational Methods in Number Theory, Part I, Mathematical Centre Tracts, vol. 154, pp. 89–139. Mathematisch Centrum, Amsterdam (1982).

[57] Rabin, M.O.: Probabilistic algorithm for testing primality. Journal of Number Theory **12**(1), 128–138 (1980). doi:10.1016/0022-314X(80)90084-0.

[58] Rivest, R.L., Shamir, A., Adleman, L.: A method for obtaining digital signatures and public-key cryptosystems. Communications of the ACM **21**(2), 120–126 (1978). doi:10.1145/359340.359342.

[59] Soundararajan, K.: The distribution of smooth numbers in arithmetic progressions. In: De Koninck, J.-M., Granville, A., Luca, F. (eds.) Anatomy of Integers. CRM Proceedings & Lecture Notes, vol. 46, pp. 115–128. AMS (2008). arXiv:0707.0299.

[60] Stark, H.M.: Some effective cases of the Brauer–Siegel theorem. Inventiones Mathematicae **23**(2), 135–152 (1974). doi:10.1007/BF01405166.

[61] Stevenhagen, P.: The number field sieve. In: Buhler, J.P., Stevenhagen, P. (eds.) Algorithmic Number Theory: Lattices, Number Fields, Curves and Cryptography. MSRI Publications, vol. 44, pp. 83–100. Cambridge University Press (2008). doi:10.1017/9781139049801.005.

[62] Thomé, E.: Subquadratic computation of vector generating polynomials and improvement of the block Wiedemann algorithm. Journal of Symbolic Computation **33**(5), 757–775 (2002). doi:10.1006/jsco.2002.0533.

[63] Thomé, E.: Square root algorithms for the number field sieve. In: Arithmetic of Finite Fields – WAIFI 2012. Lecture Notes in Computer Science, vol. 7369, pp. 208–224. Springer (2012). doi:10.1007/978-3-642-31662-3_15.

[64] Trager, B.M.: Algebraic factoring and rational function integration. In: Proceedings of the 1976 ACM Symposium on Symbolic and Algebraic Computation. pp. 219–226. ACM (1976). doi:10.1145/800205.806338.

[65] Weis, S.A.: RSA-896. Online computation report, September 2026. https://saweis.net/posts/rsa-896.

[66] Wiedemann, D.H.: Solving sparse linear equations over finite fields. IEEE Transactions on Information Theory **32**(1), 54–62 (1986). doi:10.1109/TIT.1986.1057137.

[67] Younis, K.: Asymptotics for smooth numbers in short intervals. arXiv:2409.05761v1 (2024). arXiv:2409.05761v1.

[68] Zhou, Y.: From two thirds to one half: proving the predicted running time of the hyperelliptic curve method. Manuscript (2026). Author's manuscript.

[69] Zimmermann, P., Dodson, B.: 20 years of ECM. In: Hess, F., Pauli, S., Pohst, M. (eds.) Algorithmic Number Theory (ANTS VII). LNCS, vol. 4076, pp. 525–542. Springer (2006). doi:10.1007/11792086_37.

Appendices

Detailed proofs and auxiliary estimates for the algorithms in Sections 2–7.

A The probability space of the descent **31**
A.1 Scales and preprocessing . . . 31
A.2 The rational form and distribution of the algebraic form . . . 31
A.3 Prime-ideal targets and admissible pairs . . . 32
A.4 Summary of the arithmetic estimates . . . 33

B Rational-side estimates and prime-ideal incidences **35**
B.1 Parameters and distribution of the polynomials . . . 35
B.2 Simultaneous estimates on the rational side . . . 36
B.3 Distributions of one and two polynomial values . . . 42
B.4 Conditional means of the relation counts . . . 45
B.5 An upper bound for the joint smoothness probability . . . 48
B.6 Mean incidence counts and their comparison functions . . . 52
B.7 Comparison of the probability measures . . . 56

C Concentration of relation counts **60**
C.1 The relation count and its approximation . . . 60
C.2 Local smoothness densities . . . 62
C.3 Higher moments and centering at the mean . . . 64
C.4 The principal term and character sums of the conditional density . . . 68
C.5 Character sums associated with a common determinant . . . 70
C.6 Proof of the concentration estimate . . . 76

D Relation trees and initial factorizations **82**
D.1 Proof of the arithmetic bounds . . . 82
D.2 The initial relation counts . . . 82
D.3 Bounds for the length of a search . . . 83
D.4 Probability of the prescribed initial factorizations . . . 84

E Square-class certification and the congruence of squares **85**
E.1 Algebraic preliminaries and reduction modulo the input . . . 85
E.2 Separation of nonrational square classes . . . 87
E.3 Construction of the public square root . . . 89
E.4 Conditional distribution of the private roots . . . 91
E.5 Complexity of square certification and reconstruction . . . 93
E.6 Construction of the rows and the public square root . . . 94

F Bit complexity of the unconditional algorithm **97**
F.1 Integer parameters and their admissibility conditions . . . 97
F.2 Bounds on the sizes of auxiliary inputs . . . 97
F.3 Bounds on field data, characters and matrix dimensions . . . 98
F.4 Number of searches and retained relations . . . 98
F.5 Trial bounds for the random samplers . . . 99
F.6 Direct arithmetic and exact reconstruction . . . 99
F.7 Cost excluding recursive factorizations . . . 100
F.8 Recursive factorizations and storage of intermediate data . . . 101
F.9 The unconditional factorization algorithm . . . 102

G Relation and incidence estimates under GRH **104**
G.1 Polynomial distribution and relation counts . . . 104
G.2 Smooth numbers in arithmetic progressions under GRH . . . 106
G.3 Mean-square estimates on the rational side . . . 107
G.4 Gaussian distributions of polynomial values . . . 109
G.5 Local densities and concentration of the relation counts . . . 111
G.6 Propagation of exceptional sets . . . 118
G.7 Initial prime ideals in a common number field . . . 121
G.8 Verification of the parameter conditions . . . 122

H The factorization algorithm under GRH **126**
H.1 Parameters, preprocessing and polynomial selection . . . 126
H.2 Uniform sampling in the special-prime lattices . . . 127
H.3 A bounded smoothness test under RH . . . 129
H.4 Bounds on the number of smoothness tests . . . 132
H.5 Counting the prescribed initial factorizations . . . 134
H.6 Square certification with bounded running time . . . 136
H.7 Computation of the public square root . . . 138
H.8 Storage, random sampling and factor extraction . . . 140
H.9 Expected bit complexity . . . 142

A The probability space of the descent

At the start of each run of the relation search, we choose a rational linear form and sample a family of algebraic forms. A descent uses these same draws at every level. We specify their joint distribution before conditioning on the existence of a relation. All logarithms are natural unless a base is displayed. Write $\mathrm{b}(t) = \lfloor \log_2 t \rfloor + 1$ for the bit length of a positive integer.

A.1 Scales and preprocessing

Put $n = \mathrm{b}(N)$ and $m_b = \mathrm{b}(n)$. The odd degree d, the integer $k = \kappa d$, and the prime cutoffs

$$Y_0 > Y_1 > \cdots > Y_j = B, \qquad V_i = Y_i^6$$

are the computable integers specified in Appendix F.1. Here κ, b_0, q_* are fixed and $U_* = 6q_*$. At an internal level i, the box has scale V_i, the target prime is at most Y_i, and every child prime is at most Y_{i+1}. Primes at most B end the descent.

As N grows,

$$d = D^{1/3}(1 + o(1)), \qquad j = \frac{\log D}{96 \log 2} + O(1), \qquad A = 4^{j+10} = D^{1/48+o(1)}. \tag{A.1}$$

The program first tests for primality, perfect powers, even factors and prime divisors at most B. The corrected AKS primality test [3, 2] and the perfect-power test run to completion before any epoch begins. If trial division gives a divisor, the program puts the two smaller factors on its worklist and recomputes the parameters for each. For a perfect power it keeps the base and its multiplicity. The integer N entering an epoch is therefore odd, has at least two distinct prime divisors and has least prime factor greater than B. Appendix F includes the cost of trial division.

A.2 The rational form and distribution of the algebraic form

Put $M = \lfloor N^{1/d} \rfloor$. Choose Q uniformly from the primes in $(3M/4, M]$, and independently choose the integer $r \in (M/2, 3M/4]$ uniformly. A proper $\gcd(Q, N)$ is an immediate factor. Otherwise set

$$G(X, Y) = QX - rY, \qquad m_N = rQ^{-1} \pmod N.$$

The same $g = (r, Q)$ is used throughout the epoch.

There is an explicitly computable integral degree-d form $\widehat{F}$ with

$$\widehat{F}(r, Q) = N, \qquad \|\widehat{F}\|_\infty \le d 2^d M.$$

To construct $\widehat{F}$, start with the remainder N. For $t = d, d-1, \ldots, 1$, choose a centered coefficient a_t such that $a_t r^t$ equals the current remainder modulo Q. Subtract $a_t r^t$ and divide by Q. Each step is possible since $(r, Q) = 1$. The final remainder gives the coefficient of Y^d. The inequalities $r < Q$ and $N < (M+1)^d$ give the stated height.

Choose a deterministic irreducible homogeneous degree-d polynomial h_d over $\mathbb{F}_2$. Its value at each primitive $\mathbb{F}_2$-point is one. Since $\widehat{F}(r, Q) \equiv N \equiv 1 \pmod 2$, the difference $h_d - \widehat{F} \bmod 2$ is divisible by $G \bmod 2$. A binary coefficient vector $c^{(0)}$ therefore makes

$$F_0 = \widehat{F} + G R_{c^{(0)}} \equiv h_d \pmod 2, \qquad R_c(X, Y) = \sum_{a=0}^{d-1} c_a X^{d-1-a} Y^a.$$

Sample w, c independently from the integer windows (B.4), and form $\mathcal{F}$ according to (B.6). The resulting unconditioned law at fixed g is ν_g, as in Appendix B.1. In particular,

$$\mathcal{F}(r, Q) = N(1 + 2w), \qquad \mathcal{F} \bmod 2 = h_d.$$

At primitive integer points the value of $\mathcal{F}$ is odd and nonzero. Its primitive part is irreducible over $\mathbb{Q}$.

For a sampled form let $e = \mathrm{content}(\mathcal{F}) > 0$, $F = \mathcal{F}/e$, $f(T) = F(T, 1)$, $A_f = \mathrm{lc}(f)$, and

$$D_f^* = eA_f \,\mathrm{disc}(f).$$

The initialization tests the gcds of $e, A_f, f'(m_N)$ with N. A proper gcd returns a factor; a gcd equal to N rejects the slot. Under the explicit size conditions $8MH < N$ and $2H + 1 < N$, the probability of rejection without finding a factor is at most $1/(2H + 1)$. Indeed the nonzero e, A_f have absolute value below N. If no factor has been returned, $\gcd(e, N) = 1$, so $f'(m_N) \equiv 0 \pmod N$ is equivalent to $\partial_X \mathcal{F}(m_N, 1) \equiv 0 \pmod N$. The latter derivative of the sampled polynomial is affine in the uniform coefficient c_{d-1}, with unit slope $2Q \bmod N$. At most one value in that coefficient window can vanish modulo N. This bounds the rejection probability under the original law ν_g.

The program samples

$$J = 2^{(U_* + 20)dm_b}$$

independent trials. Accepted fields keep their position and their field identity; rejected fields remain unsuccessful proposals.

A.3 Prime-ideal targets and admissible pairs

An internal target is $l = (i, \ell, s)$, where $1 \le i < j$, $B < \ell \le Y_i$ is prime, $\ell \nmid NQ$, and $Qs - r \not\equiv 0 \pmod \ell$. Its root event is

$$R_l = \{\mathcal{F}(s, 1) \equiv 0 \pmod \ell\}.$$

Under the finite sampling law, $\nu_g(R_l) = \ell^{-1} + O(H^{-1})$. For $V = V_i$, the geometric set Ω_l consists of points $P = (a, b)$ satisfying

$$\frac{V}{2} < b \le \frac{3V}{4}, \quad \frac{MV}{16} \le rb - Qa \le \frac{MV}{8}, \quad a \equiv sb \pmod \ell, \quad \ell \nmid b. \tag{A.2}$$

These inequalities imply $V/8 \le a \le 2V/3$. To sample a point, draw b uniformly from the stated integer interval and let a_0 be the least nonnegative residue of sb modulo ℓ. Draw t uniformly from $0, \ldots, \lceil V/\ell \rceil - 1$ and set $a = a_0 + \ell t$. Every legal point has one preimage and hence the same probability. A point failing a geometric or arithmetic test consumes one proposal. For each b, the allowed real a-interval has length at least $V/16$, hence at least $V/(16\ell) - 1$ progression points. At least half the b's avoid ℓ. Since $V/\ell \ge 33$ for the admissible parameters for sufficiently large inputs, the legal probability is at least $1/64$.

An active target has $\ell > Y_{i+1}$. At such a target a raw candidate additionally satisfies

$$(a, b) = 1, \quad P^-(b) > B, \quad -G(P) \in \mathcal{S}_B, \quad 0 < |\mathcal{F}(P)|/\ell \in \mathcal{S}_{Y_{i+1}}, \tag{A.3}$$

where $\mathcal{S}_y$ denotes the positive y-smooth integers. Since $\ell > Y_{i+1}$, the incoming valuation is exactly one. Let Z_l count these candidates, and let

$$\tau_l(g) = \mathbb{E}_{\nu_g}[Z_l \mid R_l]$$

be its conditional mean. The denominator-roughness test in (A.3) is part of the relation predicate. Its density is not included in the constant $1/64$ geometric sampling probability.

A top target is a prime $p \in (Y_0, 2Y_0]$ not dividing N. It uses the same search region with $V = V_0$, the congruence $a \equiv rQ^{-1}b \pmod{p}$, and the tests

$$(a,b) = 1, \quad P^-(b) > B, \quad -G(P)/p \in \mathcal{S}_B, \quad 0 < |\mathcal{F}(P)| \in \mathcal{S}_{Y_1}.$$

Its root event is identically true. This separate initialization is important: no internal root-conditioned concentration theorem is asserted for a single top field.

At each top proposal, choose one of the J trials uniformly and independently of the past. A rejected field consumes the proposal. For an accepted field, sample the point in the common box independently as above. Thus each accepted field and legal point receive the same joint proposal probability. An internal search keeps the field assigned by its parent and draws fresh points in its own search region.

Factor a successful numerical value into primes. First test each prime factor against N; a divisor of N gives an immediate termination with a nontrivial factor. A good prime divisor $q > B$ of the primitive value, with $q \nmid D_f^*$, defines a child prime ideal in the same number field, with root $t = ab^{-1} \bmod q$. The denominators are invertible: primitiveness and the leading coefficient condition exclude $q \mid b$, while the rational predicate excludes $q \mid G(P)$. Small primes and primes dividing D_f^* are terminal. A prime below the next cutoff passes unchanged to the next level. Actual recursive children are those needed for odd ideal valuations. For an analytic pass-through state put $Z_l = \mathbf{1}_{R_l}$ and $\tau_l = 1$; its identity edge creates no new atom.

For the transition from level i to level $i+1$, the measures below range over primes $B < q \le Y_{i+1}$ coprime to N and their admissible finite roots. Primes at most B and divisors of N terminate the search and contribute zero to its failure probability. Within this range X_{lq} counts each distinct finite child-prime incidence in a candidate polynomial value, with root $ab^{-1} \bmod q$ when $q \nmid b$. We count incidences before checking setup or child success, including even valuations and primes dividing the content of $\mathcal{F}$. A child that is terminal in the number field contributes zero to the bad-incidence sum. Thus the formal count bounds the number of actual children, and omitting terminal primes outside the range leaves the bad-incidence sum unchanged. A pass-through contributes its identity edge. At level j the normalized domain is empty, because $Y_j = B$. Every candidate has at most $M_c = n + 1$ such incidences. Indeed the enforced coefficient/value envelopes give

$$\log_2 |\mathcal{F}(P)| \le 12d^2 m_b + 6de_0 < \left(\frac{6}{\kappa b_0} + \frac{3}{\kappa}\right) n < n.$$

Thus the number of distinct prime factors of the polynomial value is below n; an identity edge contributes only one. In particular,

$$\sum_q X_{lq} \le M_i Z_l, \qquad M_i \le n + 1 = D^{O(1)}. \tag{A.4}$$

A.4 Summary of the arithmetic estimates

Proposition B.1 gives an event $\mathcal{G}$ of probability $1 - o(1)$ for the rational environment g. The program uses sampled coefficients, fixed search limits and exact arithmetic tests. The event $\mathcal{G}$, the means τ_l, and the weights below are needed only to analyze these tests.

Let μ denote the original independent uniform law of (r, Q) in the epoch, before any termination with a nontrivial factor, and put

$$\bar{\mu} = \mu\big(\,\cdot \mid \mathcal{G} \cap \{Q \nmid N\}\big).$$

The complete-band baseline β_i draws g from $\bar{\mu}$; given the label, it draws the form from ν_g conditioned on its root event. The weights in (3.5) can change the environment marginal under σ_i. For the incoming means and moments we fix any $g \in \mathcal{G}$. For the ordinary tail we average over the stated baseline, including its environment marginal. This distinction matters when the parent chooses which child to visit.

Proposition A.1 (Arithmetic interface)**.** *For the finite experiment above, the arithmetic derivations in Appendices B and C give the following uniform statements.*

(i) For every environment $g \in \mathcal{G}$ and every active or top label,

$$\tau_l \geq e^{c\Lambda}, \qquad \tau_l/|\Omega_l| \geq p_{\mathrm{den}} := 2^{-(U_*+10)dm_b}.$$

(ii) The forced measures and incoming densities defined in Section 4.1 satisfy

$$K_i^2 = \mathbb{E}_{\sigma_{i+1}}(H_q/a_{i+1})^2 \leq D^{C_0} L_i^4, \qquad L_i \leq D^{C_0(i+1)}.$$

(iii) At every active internal level,

$$\sigma_i\{Z_l < \tau_l/2 \text{ or } Z_l > 2\tau_l\} \leq D^{-2A}.$$

(iv) At the top, $Z_l \leq U_N \tau_l$ with $U_N = \exp((U_* + o(1))\Lambda)$.

Theorem B.17 and Proposition B.13 give the incoming moments and conditional means. Proposition B.16 transfers the ordinary tail from the baseline to the measure induced by parent requests. Appendix D.1 proves the stated uniform density bounds.

B Rational-side estimates and prime-ideal incidences

A parent relation can request a child prime and root more often than a uniform choice of the label would suggest. Fix an environment g and a child label (q, t). We bound the second moment of the weighted number H of requests for this label and compare the measure defined by its conditional mean with uniform sampling in a prime band.

$$\mathbb{E}\big[(H/a')^2 \mid g, R_{q,t}\big] \le D^C L^4, \qquad \frac{d\sigma'}{d\beta'} \le CL^2. \tag{B.1}$$

Here $a' = \mathbb{E}[H \mid g, R_{q,t}]$, σ' weights each child by this mean on its root event, and L bounds the preceding level's comparison with its deterministic weights. To sample from β', we choose a band with the weights of Proposition B.16, a prime uniformly within it, and an admissible root uniformly for that prime. Given these choices and the environment, we draw the form conditional on the root event.

We first apply the multiplicative large sieve to count rational points for all required labels in one environment. We then determine the distributions of the sampled form at one and two points. The one-point distribution gives the mean number of requests. The two-point distribution bounds its second moment. These estimates give the weights to use at the next level.

The finite evaluation-window method has direct precedents in [33, `c12alg`, Sections 2.6 and 3.1–3.2]. Actual-mean incidence weights and conditional deletion appear in [33, `pc3`, Section 2, DC/DC$_e$], and child profiles are proposed in [33, `multidescent`, Sections 2.1–2.2]. We prove the finite windows used here and the two-sided comparison of the actual incoming means with deterministic profiles, uniformly through all levels of the specified experiment.

B.1 Parameters and distribution of the polynomials

Write

$$D = \frac{\log N}{\log\log N}, \qquad \lambda = \log\log N, \qquad d \sim D^{1/3}, \qquad \Lambda_b = (\log 2) d m_b \sim d\lambda,$$

where $d \ge 3$ is odd and m_b is the prescribed integer bit-length parameter. The constants used in this section are

$$b_0 = 100, \qquad \kappa = 10^7, \qquad q_* = 2^{32}, \qquad U_* = 6q_*.$$

The input has no prime factor at most B. The scales satisfy

$$\begin{aligned} &M = \lfloor N^{1/d} \rfloor, \qquad \log B = (b_0 + o(1))\Lambda_b, \qquad \log Y_0 = (1/(2\kappa) + o(1)) \log M, \\ &V_i = Y_i^6, \qquad B \le y_i \le Y_i \le Y_0, \qquad \frac{\log Y_i}{\log y_i} \le C q_*. \end{aligned} \tag{B.2}$$

All asymptotic statements in this section are uniform in the $O(\log D)$ levels. For the explicit large-mean estimate we use the cutoff relation

$$\xi_i = \frac{\log Y_i}{\Lambda_b} \ge b_0 + o(1), \qquad \frac{\log y_i}{\Lambda_b} = \max\{b_0, (12 + 6\xi_i)/U_*\} + o(1). \tag{B.3}$$

The integer windows are

$$E = 2^{d^2 m_b}, \qquad W = M^2 V_0^2 E^4, \qquad H = 2^{d+20} d^{10} W E^4, \qquad h = 2H + 1. \tag{B.4}$$

The finite cutoff definitions give $Y_0 \le E^{1/(2\kappa)}$. For example, this follows from $d^3 m_b \ge n$, $n = \mathrm{bitlength}(N)$, and $Y_0 = 2^{\lfloor (n-5)/(2\kappa d)\rfloor - 1}$.

An environment is $g = (r, Q)$, where Q is a prime in $(3M/4, M]$, $Q \nmid N$, and r is an independent uniform integer in $(M/2, 3M/4]$. Thus $(r, Q) = 1$. Put

$$G(X, Y) = QX - rY.$$

Fix an integral degree-d form $\widehat{F}$ and a parity correction F_0 such that

$$\widehat{F}(r, Q) = N, \quad \|\widehat{F}\|_\infty \le d 2^d M, \quad F_0 = \widehat{F} + G R_{c^{(0)}}, \quad c^{(0)} \in \{0, 1\}^d. \tag{B.5}$$

The reduction of F_0 modulo two is the prescribed irreducible form. In particular, its value at every primitive integer point is odd. The polynomial is sampled as follows:

$$\mathcal{F} = F_0 + 2(w\widehat{F} + G R_c), \qquad R_c = \sum_{j=0}^{d-1} c_j X^{d-1-j} Y^j, \tag{B.6}$$

where $w \in (W, 2W] \cap \mathbb{Z}$ and $c \in [-H, H]^d \cap \mathbb{Z}^d$ are independent and uniform. We write ν_g for the law of this draw at fixed g, before any test of the form or any recursive search.

For each prime $B < q \le Y_0$ with $q \nmid N$, an admissible root is

$$t \in \mathbb{F}_q, \qquad t \ne r Q^{-1} \pmod q.$$

All these primes are smaller than Q. Set

$$R_{q,t} = \{\mathcal{F}(t, 1) = 0 \pmod q\}, \qquad p_{q,t}(g) = \nu_g(R_{q,t}).$$

The multiplier of c_{d-1} in $\mathcal{F}(t, 1)$ is $2(Qt - r)$, a unit modulo q. Consequently

$$p_{q,t}(g) = q^{-1} + O(h^{-1}), \qquad c_p \le (q-1) p_{q,t}(g) \le C_p \tag{B.7}$$

with fixed positive constants, uniformly in g, q, t, for all sufficiently large admissible parameters.

We use

$$\mathcal{S}_y = \{m \ge 1 : P^+(m) \le y\}, \qquad \Psi(x, y) = |\mathcal{S}_y \cap [1, x]|, \qquad \widehat{\rho}(x, y) = \Psi(x, y)/x.$$

B.2 Simultaneous estimates on the rational side

The following proposition includes the internal cell estimates needed for the ordinary relation-count argument. Only its whole-box conclusions are used for the incoming theorem.

Proposition B.1 (Simultaneous rational supply)**.** *For one internal box let $V = Y^6 < M/4$, $x = MV/8$, and*

$$I_n = [x/2, x], \qquad I_b = (V/2, 3V/4], \qquad \mathcal{A} = I_n \cap \mathcal{S}_B, \qquad \rho = |\mathcal{A}|/x.$$

Partition each interval into K fixed consecutive cells of comparable length, before choosing Q, r. Suppose B is sufficiently large,

$$K \le B^{1/4}, \qquad \frac{KYB^3 \log B}{V} \le c_s, \qquad 2 \le \frac{\log(x/2)}{\log B} \le \frac{\log x}{\log B} \le \exp(\sqrt{\log B}),$$

where $c_s > 0$ is a sufficiently small absolute constant. For each prime $B < \ell \le Y$ and admissible $s \bmod \ell$, let $\mathcal{P}_{\ell,s}$ be the points satisfying

$$b \in I_b, \quad n = rb - Qa \in \mathcal{A}, \quad P^-(b) > B, \quad \ell \nmid b, \quad a \equiv sb \pmod{\ell}.$$

Except on a set of environments of probability at most

$$\epsilon_{\mathrm{ord}} \le C \frac{\log M \log B}{\rho} K^4 \left(\frac{Y^4}{V^2} + \frac{Y^4}{M} \right), \tag{B.8}$$

the following hold simultaneously for all these labels:

$$|\mathcal{P}_{\ell,s}| \asymp \frac{V^2 \rho}{\ell \log B}, \tag{B.9}$$

$$|\mathcal{P}_{\ell,s} \cap (\textit{one cell pair})| \le \frac{C V^2 \rho}{\ell \log B\, K^2}, \tag{B.10}$$

$$\#\{(P_1, P_2) \in \mathcal{P}_{\ell,s}^2 : 0 < |\det(P_1, P_2)| \le c_1 V^2/K\} \le \frac{C_{c_1} V^4 \rho^2}{\ell^2 (\log B)^2 K}. \tag{B.11}$$

All counted points are primitive and satisfy $0 < a < b$.

For the top box, let $V_0 = Y_0^6 < M/4$, $x_0 = MV_0/8$, and

$$J_0 = \{p : Y_0 < p \le 2Y_0,\ p \textit{ prime},\ p \nmid N\}, \quad L_{\mathrm{top}} = |J_0|,$$

$$A_p = [x_0/(2p), x_0/p] \cap \mathcal{S}_B, \quad \rho_p = p|A_p|/x_0, \quad \rho_* = \min_{p \in J_0} \rho_p, \quad \rho_J = \sum_{p \in J_0} \rho_p/p.$$

Assume $L_{\mathrm{top}} \asymp Y_0/\log Y_0$, $\rho_* > 0$, and $\rho_p \asymp \rho_*$. Let $\mathcal{A}_p(g)$ be the primitive points in the top box with $-G(P) \in pA_p$ and $P^-(b) > B$. Let $\mathcal{A}_{q,t}(g)$ be the union of these sets intersected with $a \equiv tb \pmod{q}$, for $B < q \le Y_1 < Y_0$ and admissible t. There are simultaneous whole-box bounds

$$|\mathcal{A}_p(g)| \asymp \frac{V_0^2 \rho_p}{p \log B}, \tag{B.12}$$

$$|\mathcal{A}_{q,t}(g)| \asymp \frac{V_0^2 \rho_J}{q \log B}, \tag{B.13}$$

outside exceptional sets of respective probabilities at most

$$\epsilon_{\mathrm{ind}} \le C \frac{\log M \log B}{\rho_*} \left(\frac{Y_0^2}{V_0^2} + \frac{Y_0}{M} \right), \tag{B.14}$$

$$\epsilon_{\mathrm{agg}} \le C \kappa_J \frac{\log M \log B}{\rho_J} \left(\frac{Y_1^4}{V_0^2} + \frac{Y_1^4}{M} \right), \tag{B.15}$$

where $\kappa_J = \max(1, \lfloor \log(x_0 V_0)/\log Y_0 \rfloor)$. For the lower bound in (B.13)*, assume also*

$$\frac{\log B}{\rho_*} \left(\frac{1}{Y_0} + \frac{Y_1}{V_0} + \frac{Y_1 Y_0}{V_0^2} \right) \le c_{\mathrm{prim}} \tag{B.16}$$

for a sufficiently small absolute constant $c_{\mathrm{prim}} > 0$. The top bounds do not assert individual-top cell estimates.

Proof. We first obtain the denominator counts. If I has length L, $(v,k)=1$, and k has no prime factor at most B, then for every squarefree e supported on primes at most B,

$$\#\{b\in I: b\equiv v \pmod k,\ e\mid b\} = L/(ke)+O(1).$$

The linear sieve at fixed parameter $s=3$, level $(B+1)^3$, has a positive lower function $f_-(3)=2e^{\gamma_E}\log 2/3$. Its unweighted remainder is $O(B^3)$. The linear sieve and Mertens' product estimate therefore give

$$\#\{b\in I: b\equiv v \pmod k,\ P^-(b)>B\} \asymp L/(k\log B) \tag{B.17}$$

when $kB^3\log B/L$ is sufficiently small. The comparison constants are absolute at the fixed sieve parameter. See [47, Theorem 21.9 and (21.19)].

Fix ℓ, one numerator cell A, and one denominator cell. Let

$$\mathcal{B}=\{b \text{ in that cell}: P^-(b)>B,\ \ell\nmid b\},\qquad J(v)=\#\{b\in\mathcal{B}: b\equiv v \pmod\ell\}.$$

Equation (B.17) gives, for every unit v mod ℓ,

$$J(v)\asymp V/(\ell K\log B),\qquad |\mathcal{B}|\asymp V/(K\log B).$$

The direct short-interval theorem of [29, Theorem 5.1] gives

$$|A|\le C\rho x/K.$$

Indeed the cell relative lengths are bounded below by a constant multiple of $1/K$, hence exceed $B^{-5/12}$; all endpoints lie in the stated Hildebrand–Tenenbaum range. The whole numerator interval has mass exactly ρx.

All $n\in A, b\in\mathcal{B}$ are units modulo $Q\ell$. Define on its unit group

$$C_Q(v)=\#\{(n,b)\in A\times\mathcal{B}: n\equiv vb \pmod{Q\ell}\}.$$

In the character expansion of C_Q, remove only the terms whose Q-component is nonprincipal. The retained main term is exactly

$$L_Q(v)=\frac{1}{Q-1}\sum_{n\in A} J(nv^{-1} \bmod \ell). \tag{B.18}$$

Thus every character modulo ℓ is retained in the main term. In particular, no distribution theorem for smooth numerators in a specified class modulo ℓ has been assumed.

For $H_Q=C_Q-L_Q$, character orthogonality yields

$$\sum_{v\in(\mathbb{Z}/Q\ell\mathbb{Z})^\times}|H_Q(v)|^2=\frac{1}{(Q-1)(\ell-1)}\sum_{\chi_Q\ne1}\sum_{\chi_\ell}|A(\chi_Q\chi_\ell)B(\chi_Q\chi_\ell)|^2. \tag{B.19}$$

The ratio convolution originally has $A(\chi)\overline{B(\chi)}$; its absolute value equals that of $A(\chi)B(\chi)$, which is the only replacement made here. Put $T=xV$, $E_A=|A||\mathcal{B}|$, and

$$c_m=\#\{(n,b)\in A\times\mathcal{B}: nb=m\}.$$

The complete B-smooth part of m determines n, so

$$c_m\in\{0,1\},\qquad \sum_{m\le T}c_m^2=E_A.$$

This sequence is independent of the summation prime Q. For a fixed χ_ℓ, apply the primitive-character large sieve to $c_m\chi_\ell(m)$. Every nonprincipal character modulo the prime Q is primitive. Dropping the weights $Q/\varphi(Q) \geq 1$ on the left of that inequality gives

$$\sum_{Q}\sum_{\chi_Q\neq 1}\left|\sum_{m\leq T} c_m\chi_\ell(m)\chi_Q(m)\right|^2 \leq (M^2+T)E_A.$$

Here the moduli in the large sieve are the primes Q [47, Theorem 19.16].

For a fixed root $s \bmod \ell$, the map $r \mapsto r - Qs \bmod Q\ell$ is injective on the sampled r-interval. Its admissible image lies in the unit group. Bound its squared sum by the whole squared sum in (B.19), and use $\#\{Q\} \gg M/\log M$ and $\#\{r\} \asymp M$. After summing the $\ell - 1$ twists, we obtain the finite bound

$$\mathbb{E}_{Q,r}\left[|H_Q(r-Qs)|^2\mathbf{1}_{\{r-Qs\neq 0 \pmod \ell\}}\right] \leq C\frac{\log M\,(M^2+T)E_A}{M^3}. \tag{B.20}$$

A simultaneous bound over roots and cells. Write

$$\mathfrak{m}_\ell = \frac{V^2\rho}{\ell\log B}.$$

For the whole numerator and denominator intervals, $L_Q(v) \asymp \mathfrak{m}_\ell$ and $E_A \asymp \rho MV^2/\log B$. Therefore, for a fixed numerical root $s \bmod \ell$, Chebyshev's inequality and (B.20) give

$$\begin{aligned}&\mathbb{P}_{Q,r}\left(|H_Q(r-Qs)| > c\mathfrak{m}_\ell,\ r-Qs\not\equiv 0 \pmod \ell\right)\\ &\ll \frac{\log M\log B}{\rho}\,\frac{\ell^2}{V^2}\left(1+\frac{V^2}{M}\right).\end{aligned}$$

Here the fixed positive c can be chosen below the lower main-term constant. For one prescribed cell pair, the numerator short-interval bound and denominator sieve instead give

$$E_A \ll \frac{\rho MV^2}{K^2\log B}, \qquad L_Q(v) \ll \frac{\mathfrak{m}_\ell}{K^2}.$$

Choosing the upper cell threshold above this last main-term constant, its failure probability for the same fixed root is at most

$$C\frac{\log M\log B}{\rho}\,\frac{\ell^2K^2}{V^2}\left(1+\frac{V^2}{M}\right).$$

There are K^2 prescribed cell pairs and at most ℓ numerical roots modulo ℓ. Summing over these roots is valid even though the admissible subset depends on g, because the displayed events already contain the admissibility indicator. Consequently

$$\begin{aligned}&\mathbb{P}(\text{some required whole or cell estimate fails})\\ &\ll \frac{\log M\log B}{\rho}\,K^4\left(\frac{1}{V^2}+\frac{1}{M}\right)\sum_{\substack{B<\ell\leq Y\\ \ell\text{ prime}}}\ell^3\\ &\ll \frac{\log M\log B}{\rho}\,K^4\left(\frac{Y^4}{V^2}+\frac{Y^4}{M}\right).\end{aligned}$$

This is a single event supplying all the numerical roots and cells.

The original congruences are recovered exactly:

$$n \equiv (r - Qs)b \pmod{Q\ell} \quad \Longleftrightarrow \quad a = (rb - n)/Q \in \mathbb{Z}, \quad a \equiv sb \pmod{\ell}.$$

The inequalities defining the search region give $V/8 \le a \le 2V/3$, and $n > 0$ gives $a/b < r/Q < 1$. Since $(n, b) = 1$, these points are primitive. For clarity, the thin-pair estimate requires no further probability union. For fixed P_1, the determinant condition implies

$$\left| n_2 - \frac{n_1}{b_1} b_2 \right| \le \frac{Q c_1 V^2}{K b_1} \ll_{c_1} \frac{MV}{K}.$$

Since $n_1/b_1 \asymp M$, its intersection with one denominator cell meets only $O_{c_1}(1)$ numerator cells. The strip is covered by $O_{c_1}(K)$ cell pairs from the original fixed partition. On the already simultaneous event it contains at most $C_{c_1}\mathfrak{m}_\ell/K$ possible second points. Summing over the $O(\mathfrak{m}_\ell)$ first points proves (B.11).

For the sum over the initial primes, the sets pA_p are disjoint, since each numerator has exactly one prime above B, appearing to exponent one. Fix a child q, and use the common denominator set

$$\mathcal{B}_q = \{b \in (V_0/2, 3V_0/4] : P^-(b) > B,\ q \nmid b\}$$

before imposing any parent-dependent restriction. For $n = pn_s$ and $m = nb$, the B-smooth part of m determines n_s. Only the choice of the top parent $p \mid m$ remains; consequently

$$c_m \le \kappa_J, \qquad \sum_m c_m^2 \le \kappa_J \left| \bigcup_p pA_p \right| |\mathcal{B}_q|.$$

All these coefficients are again independent of Q. The proof of (B.20), with $\ell = q$, now gives (B.15) for the aggregate before imposing primitivity.

An aggregate point is nonprimitive precisely when its parent p divides b, in which case it also divides a. Write $P = pP'$. Since $p \ne q$, the child congruence is preserved. Dropping all other conditions bounds the number for this parent by

$$C\left(\frac{V_0^2}{p^2 q} + \frac{V_0}{p} + 1\right).$$

Summing over the top band gives the deterministic bound

$$C\left(\frac{V_0^2}{qY_0 \log Y_0} + \frac{V_0}{\log Y_0} + \frac{Y_0}{\log Y_0}\right). \tag{B.21}$$

Since $\rho_J \asymp L_{\mathrm{top}}\rho_*/Y_0$, its ratio to the aggregate main term is bounded by a constant times the left side of (B.16). Subtracting it proves the primitive aggregate lower bound. Removing these points after the convolution lets us use the same denominator set for every parent p.

For each individual top parent, instead use A_p and

$$\mathcal{B}_p = \{b \in (V_0/2, 3V_0/4] : P^-(b) > B,\ p \nmid b\}.$$

The rough sieve and deletion of at most $V_0/p + 1$ multiples give $|\mathcal{B}_p| \asymp V_0/\log B$. Count $pn_s \equiv rb \pmod{Q}$. The main term is $|A_p||\mathcal{B}_p|/(Q-1)$, and the factor $\chi_Q(p)$ in the nonprincipal character

sum has absolute value one. The product $n_s b$, of length at most $x_0 V_0/p$, again has a unique smooth–rough factorization. Thus the relative failure probability for this p is at most

$$C\frac{\log M \log B}{\rho_p}\left(\frac{p}{V_0^2}+\frac{1}{M}\right).$$

Union over p proves (B.14). The mask makes these points primitive and they automatically satisfy the rational parent congruence.

The argument can first be made with all primes Q in the interval. Deleting $Q \mid N$ changes its bounds by at most the reciprocal of the retained fraction, which is $1+o(1)$ at the stated endpoint. □

Corollary B.2 (The common environment event)**.** *At* (B.2)*, take* $K = \lceil \exp(D^{1/6}\lambda^2)\rceil$ *for internal cells. There is a single event* $\mathcal{G}$*, determined by* g *before any polynomial is sampled, on which* (B.9)–(B.11)*,* (B.12)*, and* (B.13) *all hold at all required levels and labels. Its failure probability is*

$$\mathbb{P}(\mathcal{G}^c) \le \sum_i \epsilon_{\mathrm{ord},i} + \epsilon_{\mathrm{ind}} + \epsilon_{\mathrm{agg}} \le \exp(-(8b_0 - 1 + o(1))\Lambda_b). \tag{B.22}$$

The internal whole-box statement includes every auxiliary $B < q \le y_i$ *queried in the larger parent search region* V_i*.*

Proof. The sieve error is uniform even at the bottom level, because

$$\frac{KY_iB^3\log B}{V_i} \le \frac{K\log B}{Y_i^2} \longrightarrow 0.$$

Also $\log K = o(\log B)$, so the cell short-interval ranges hold. The rational cofactor endpoints have

$$\log x \le (2+o(1))d\Lambda_b, \qquad u = \log x/\log B \le (2/b_0 + o(1))d.$$

The global and direct short-interval estimates of [29], together with $\rho_D(u) \ge (2u)^{-2u}$ for the Dickman function, give uniformly

$$\rho_{r,i},\ \rho_*,\ \rho_J \ge e^{-\Lambda_b}$$

for all sufficiently large parameters; the factor $1/\log Y_0$ in ρ_J costs only $O(\lambda)$ in the logarithm. The fixed-N deletion of parent primes is negligible since it removes at most $\log N/\log Y$ primes from a band with $\asymp Y/\log Y$ primes. The same fixed-factor smooth comparison gives $\rho_p \asymp \rho_*$ as p varies over the top dyadic band. Thus the top-band population and density-comparison hypotheses in Proposition B.1 are also fulfilled.

Finally,

$$V_i^2/M \le \exp(-(1-6/\kappa+o(1))d\Lambda_b),$$

so the internal error is at most $\exp(-(8b_0-1+o(1))\Lambda_b)$. Each top error and (B.16) is exponentially smaller on the Λ_b scale, since $\log Y_0/\Lambda_b \to \infty$ and $Y_1 \le Y_0$. Summing $O(\log D)$ levels proves the claim. □

B.3 Distributions of one and two polynomial values

In this subsection, fix g and primitive positive points in one search region V, with

$$0 < a < b, \quad V/2 < b \le 3V/4, \quad MV/16 \le |G(P)| \le MV/8, \quad (N, G(P)) = 1. \tag{B.23}$$

These conditions hold for the internal rational points on $\mathcal{G}$, and also for the top rational points.

Lemma B.3 (Distribution at one point)**.** *Under* (B.6) *and* (B.23)*, set*

$$X_P^{\mathrm{raw}} = 8(d+2)h|G(P)|b^{d-1}.$$

Without a root condition, $\mathcal{F}(P)$ is within total variation

$$\epsilon_1 \le C\left\{\frac{d(1+V)+Wd^2 2^d+1}{h} + \frac{MV}{W}\right\}$$

of a mixture of uniform distributions on the odd integers in half-open intervals of length $2h|G(P)|b^{d-1}$, supported in $[-X_P^{\mathrm{raw}}, X_P^{\mathrm{raw}}]$. If $a \equiv tb \pmod q$, $q \nmid bG(P)$, and $h \ge 2q$, then $\mathcal{F}(P)/q$, conditional on $R_{q,t}$, has the same conclusion after dividing the support and interval lengths by q, with error at most $4q\epsilon_1$. In the conditional mixture each interval receives weight proportional to its original weight times the fraction of its odd integers divisible by q. At (B.4)*, with $V \le V_0$, these errors are at most CE^{-4} and CqE^{-4}, respectively.*

Proof. Evaluation of integral degree-$(d-1)$ forms at P has image $\mathbb{Z}$, since $(a,b)=1$. Its integer kernel has the basis

$$(bX - aY)X^{d-2-j}Y^j, \qquad 0 \le j \le d-2.$$

Choose the column $v = b^{d-1}$. Every residue $\rho \in [0,v) \cap \mathbb{Z}$ has an integer coefficient lift of norm at most $1+2V$. Indeed, subtract the real selected-column lift $(0,\dots,0,\rho/v)$ from any integer lift. The difference is in the real kernel. Round its kernel-basis coordinates and subtract that integer kernel vector from the integer lift. The resulting vector is integral, still evaluates to ρ, and has the asserted norm. Each original coefficient receives at most two terms in this rounding. Translating the integer cube by such a lift changes its law by at most $d(1+2V)/h$. Average these translations and condition on the original unselected coefficients, before the random translation. They are latent mixture coordinates; they need not equal the unselected coordinates of the translated cube. The independent residue remains uniform under this conditioning. The evaluation is exactly uniform on the integers of a half-open interval of length hv.

After multiplication by $G(P)$, the lattice is $G(P)\mathbb{Z}$. For a negative $G = G(P)$, the image of the integer interval $[L,U)$ is represented exactly by $G\mathbb{Z} \cap [G(U-1), G(L-1))$; this discrete endpoint convention preserves its cardinality and is used before reanchoring. For each prime dividing $G(P)$, the primitive points P and (r,Q) are nonzero proportional vectors over that prime field. Thus $(\widehat{F}(P), G(P)) = 1$, by $\widehat{F}(r,Q) = N$. The multiplier w fills every residue modulo $|G(P)|$. Replacing its residue by a uniform residue costs at most $|G(P)|/W$. Before doing so, move the geometrical interval back by $w\widehat{F}(P)$, keeping its lattice coset. In units of the selected coefficient this shift is at most

$$\frac{|w\widehat{F}(P)|}{|G(P)|b^{d-1}} \le CWd^2 2^d.$$

Its boundary cost is at most $C(Wd^2 2^d + 1)/h$. All cosets have the same number of points in the common half-open interval, whose length is an integer multiple of $|G(P)|$. Adding $F_0(P)$ and

multiplying by two therefore gives the asserted uniform odd intervals. The height bounds imply their stated support.

When $q \nmid b$, the event $R_{q,t}$ is precisely $q \mid \mathcal{F}(P)$. Its raw probability is at least $1/(2q)$ by (B.7). Conditioning two laws at total variation distance ϵ_1 on this event costs at most $4q\epsilon_1$. In every model interval the odd multiples of q, after division by the odd prime q, are exactly the odd integers in the divided interval. Reweighting the auxiliary components by their event probabilities is the ordinary conditional-mixture rule. The total-variation comparison concerns the value law, not its joint law with the original latent coordinates. Substitution of (B.4) completes the error estimate. □

Lemma B.4 (The two-point odd lattice)**.** *Let $P_1 \neq P_2$ satisfy* (B.23) *in the same search region, and suppose both lie in one admissible child lattice modulo q. Put*

$$\Delta = |\det(P_1, P_2)|, \quad \vartheta = \frac{b_1 b_2}{\Delta}, \quad K = \frac{2\Delta}{q},$$
$$s_j = \frac{h|G(P_j)|b_j^{d-1}}{q}, \quad X_j = 8(d+2)s_j.$$

Then K is an even integer. For some $(\eta, K) = 1$, the conditional law of $(\mathcal{F}(P_1)/q, \mathcal{F}(P_2)/q)$ is within total variation

$$\epsilon_2 \leq Cq\left\{\frac{d(1+V^2) + Wd^3 2^d + 1}{h} + \frac{M^2V^2}{W}\right\} \tag{B.24}$$

of a mixture of uniform distributions on

$$\mathcal{W} \cap \mathcal{L}_{K,\eta}, \qquad \mathcal{L}_{K,\eta} = \{(n_1, n_2) \in (2\mathbb{Z}+1)^2 : n_1 \equiv \eta n_2 \pmod K\}.$$

Every $\mathcal{W}$ is a translated parallelogram supported in $[-X_1, X_1] \times [-X_2, X_2]$. If w_j is its j-th projection width, $\mathcal{A}$ its area, and L_2 its longest vertical chord, then

$$w_j \asymp s_j, \qquad \mathcal{A} \asymp s_1 s_2/\vartheta, \qquad L_2 \gg s_2/\vartheta, \qquad \frac{K}{L_2} \ll \frac{2^d V^{2-d}}{HM}. \tag{B.25}$$

The transposed statements also hold. At (B.4)*, $\epsilon_2 \leq CqE^{-4}$. No assumption $q^2 \nmid \Delta$ is made.*

Proof. For degree $n \geq 1$, the integral two-point evaluation image is

$$\Gamma_n = \{(x_1, x_2) : x_1 \equiv u^n x_2 \pmod \Delta\}, \qquad (u, \Delta) = 1.$$

To see the equality, make a unimodular variable change taking P_2 to $(1, 0)$ and P_1 to $(u, \pm\Delta)$. After specifying the second value, the ideal of possible differences in the first value is $(\Delta u^{n-1}, \ldots, \Delta^n) = \Delta\mathbb{Z}$. The degree-$(d-1)$ kernel has integer basis

$$(b_1 X - a_1 Y)(b_2 X - a_2 Y) X^{d-3-j} Y^j, \qquad 0 \leq j \leq d-3.$$

Integral divisibility follows from Gauss' lemma. Rounding in this basis changes each original coefficient by $O(V^2)$.

Select the last two evaluation columns, with matrix

$$A = \begin{pmatrix} a_1 b_1^{d-2} & b_1^{d-1} \\ a_2 b_2^{d-2} & b_2^{d-1} \end{pmatrix}, \qquad |\det A| = \Delta(b_1 b_2)^{d-2}.$$

Representatives of $\Gamma_{d-1}/A\mathbb{Z}^2$ in $A[0,1)^2$ have integer coefficient lifts of norm $1+O(V^2)$. Averaging the corresponding cube translations costs $O(d(1+V^2)/h)$. Here too the original unselected coefficients are the latent mixture coordinates, and the quotient representative is chosen independently of them. For every real translate of $A[-H,H+1)^2$, each coset of $A\mathbb{Z}^2$ contains exactly h^2 points. This proves uniformity on each selected window, for every translate.

There is an exact quotient isomorphism

$$\Gamma_d/\operatorname{diag}(G(P_1),G(P_2))\Gamma_{d-1} \simeq \mathbb{Z}/|G(P_1)G(P_2)|\mathbb{Z},$$

given by evaluating a realizing form at (r,Q). It is well defined because the product of the two vanishing linear forms has value $G(P_1)G(P_2)$ there. It is onto since $(r,Q)=1$. If the residue is zero, adding a form vanishing at both points makes its value at (r,Q) zero; division by the primitive form G proves the stated kernel. The multiplier vector maps to N, a unit in the quotient. Uniformizing the residue of w therefore costs $O(M^2V^2/W)$.

The associated translation of the selected window has no Δ^{-1} loss. After dividing row j by $G(P_j)b_j^{d-1}$, its selected coordinates interpolate $\widehat{F}(x,1)/(Qx-r)$ at $x=a_j/b_j$.

Bounding the translation of the window. Put $\widehat{f}(x)=\widehat{F}(x,1)$ and $G_1(x)=Qx-r$, with $x_j=a_j/b_j$. The endpoint geometry gives $|G_1(x_j)|\gg M$, and the height bound gives $\sup_{[0,1]}|\widehat{f}|+\sup_{[0,1]}|\widehat{f}'|\ll d^3 2^d M$. Even when the signs of $G_1(x_1)$ and $G_1(x_2)$ differ, the algebraic identity

$$\frac{\widehat{f}(x_1)/G_1(x_1)-\widehat{f}(x_2)/G_1(x_2)}{x_1-x_2}=\frac{\widehat{f}(x_1)-\widehat{f}(x_2)}{(x_1-x_2)G_1(x_1)}-\frac{Q\widehat{f}(x_2)}{G_1(x_1)G_1(x_2)}$$

has absolute value $O(d^3 2^d)$. The same bound holds for the intercept of the interpolating affine function, since $0<x_j<1$. This proves the claimed selected-coordinate translation bound without requiring the rational function to be regular between the nodes.

The interpolation coefficients have this bound even when the nodes are arbitrarily close. Reanchoring compares two windows in the same affine lattice coset; it does not subtract $w\widehat{F}$ from the random output. In selected coordinates write the translation as δ. Every selected-lattice coset has exactly h^2 points in either half-open square, and their overlap contains at least $\prod_{j=1}^{2}(h-\lceil|\delta_j|\rceil)_+$ points. Averaging the cosets therefore costs at most

$$\frac{\lceil|\delta_1|\rceil+\lceil|\delta_2|\rceil}{h}\le\frac{\|\delta\|_1+2}{h}=O((Wd^3 2^d+1)/h).$$

There is no multiplier for the number of quotient representatives. Together these observations give the unconditioned error in braces in (B.24).

The odd affine coset and repeated root divisors. Put $f_j=F_0(P_j)$, and choose an integer a, coprime to Δ, such that

$$\Gamma_d=\{(x_1,x_2)\in\mathbb{Z}^2 : x_1\equiv ax_2 \pmod{\Delta}\}.$$

Both f_j are odd and $f_1\equiv af_2 \pmod{\Delta}$. We claim that an odd lift $\eta\equiv a \pmod{\Delta}$, coprime to 2Δ, can be chosen with $f_1\equiv\eta f_2 \pmod{2\Delta}$. If Δ is odd, choose the odd one of the two lifts of a modulo 2Δ. The difference $f_1-\eta f_2$ is then both even and divisible by Δ. If Δ is even, a is odd. Set

$$k=\frac{f_1-af_2}{\Delta}\in\mathbb{Z},\qquad \eta:=a+\Delta k.$$

Then $f_1 - \eta f_2 = \Delta k(1 - f_2) \equiv 0 \pmod{2\Delta}$. In either case η is odd and remains a unit modulo Δ, so $(\eta, 2\Delta) = 1$.

For any odd pair $x = (x_1, x_2)$,

$$\begin{aligned} x \in (f_1, f_2) + 2\Gamma_d &\iff x_1 - ax_2 \equiv f_1 - af_2 \pmod{2\Delta} \\ &\iff x_1 \equiv \eta x_2 \pmod{2\Delta}. \end{aligned}$$

For the last equivalence use that $\eta - a$ is divisible by Δ, whereas $x_2 - f_2$ is even. No inverse of f_2 modulo Δ is needed.

The common child congruence gives $q \mid \Delta$. On the child-root event the two coordinates are divisible by the odd prime q. Writing $x_j = qn_j$, the pair remains odd and the exact equivalence is

$$x_1 \equiv \eta x_2 \pmod{2\Delta} \iff n_1 \equiv \eta n_2 \pmod{2\Delta/q}.$$

Thus $K = 2\Delta/q$ is even, $(\eta, K) = 1$, and the asserted odd lattice follows for every valuation $v_q(\Delta) \geq 1$.

The same conditional total-variation argument as in Lemma B.3 supplies the factor Cq, with the auxiliary mixture's own root-conditioned weights.

Finally, the selected row vectors, after root division, are proportional to $(a_j/b_j, 1)$. Their projection widths are comparable to s_j, and the determinant gives area comparable to $s_1 s_2/\vartheta$. Thus $L_2 \gg s_2/\vartheta$, and

$$\frac{K}{L_2} \ll \frac{\Delta\vartheta}{qs_2} = \frac{b_1 b_2}{h|G(P_2)|b_2^{d-1}} \ll \frac{2^d V^{2-d}}{HM}.$$

The determinant cancels in this estimate. The height bound gives the stated support. Substituting (B.4) in (B.24) gives the error at these parameters. □

B.4 Conditional means of the relation counts

Let $\beta(x, y) > 0$ be the saddle defined by

$$\sum_{p \leq y} \frac{\log p}{p^{\beta(x,y)} - 1} = \log x.$$

We use the following common analytic input in both supply estimates. With the fixed source parameter $\epsilon = 1/10$, the global and direct interval theorems of [29, Theorems 1.1 and 5.1] give

$$\begin{aligned} \Psi(x, y) &= x\rho_D(u)(1 + O(E_H)), \\ \Psi(x + L, y) - \Psi(x, y) &= L\rho_D(u)(1 + O(E_H)), \\ E_H &= \frac{\log(u + 1)}{\log y}, \end{aligned} \tag{B.26}$$

where $u = \log x/\log y$, provided $1 \leq u \leq \exp(\sqrt{\log y})$ and $xy^{-5/12} \leq L \leq x$. The interval statement holds directly for every such interval; it is not obtained by subtracting global approximations. The ratio theorem [30, Theorem 3] gives, uniformly for $1 \leq c \leq y$,

$$\Psi(cx, y) = \Psi(x, y)c^{\beta(x,y)} \left(1 + O\left(\frac{1}{u} + \frac{\log y}{y}\right)\right). \tag{B.27}$$

Its saddle is at the smaller endpoint x. In the ranges checked below, the saddle estimates [30, Theorem 2(i)] and implicit differentiation give

$$\phi_2(\beta, y) \asymp \log x \log y, \qquad \frac{\partial\beta(x, y)}{\partial \log x} = -\phi_2(\beta, y)^{-1}. \tag{B.28}$$

For larger divisor shifts in the same verified range we also use [25, Lemma 1(i)]:

$$\Psi(x/a, y) \ll a^{-\beta(x,y)}\Psi(x, y) \qquad (1 \le a \le x). \tag{B.29}$$

In every application below, including division by one root prime, one parent prime, and a further factor two, the endpoints satisfy

$$\begin{gathered} cd \le \frac{\log x}{\log y_i} \le Cd, \qquad 0 \le 1 - \beta(x, y_i) \ll \frac{\log D}{\log y_i}, \\ \phi_2(\beta(x, y_i), y_i) \asymp \log x \log y_i. \end{gathered} \tag{B.30}$$

Here and below c, C may depend on the fixed cutoff constants. To check the range, the smallest positive endpoint needed in an outer subinterval has logarithm at least

$$\log(hM) + (6d - 2)\log Y_i - O(d + \log d);$$

the largest has logarithm at most $\log(hM) + 6d \log Y_i + O(\log d)$. Since $\log(hM) = O(d\Lambda_b)$, (B.2) gives the two-sided bound for u. Moreover $\log y_i \gg d\lambda$, so $c/d \gg y_i^{-5/12}$, the errors in the cited estimates tend to zero uniformly, and all factors $e^{O(d)}$ used in (B.27) are eventually at most y_i. The saddle estimates then give the last two assertions. This verifies one common threshold in N for all environments, points, roots, and windows.

Lemma B.5 (Odd intervals, including intervals near zero)**.** *Let $J \subset [-X, X]$ be any half-open interval with $X/(C_0 d) \le |J| \le 2X$, at one of the scales just verified. Let $\ell = 1$ or let ℓ be an allowed odd parent prime. For a uniformly chosen odd integer $n \in J$,*

$$\mathbb{P}(0 < |n| \in \ell\mathcal{S}_y) \asymp \ell^{-1}\widehat{\rho}(2X/\ell, y). \tag{B.31}$$

The comparison is uniform in the placement of J.

Proof. Write $w = |J|$. One endpoint is at distance at least $w/2$ from zero. At that endpoint choose an outer subinterval of length $w/4$, and reflect it if necessary. After division by the odd integer ℓ, it has positive endpoints $(x, x + z]$, up to two endpoint integers, with

$$x \ge w/(4\ell), \qquad z = w/(4\ell) \le x, \qquad z/x \ge 1/(4C_0 d).$$

Odd smooth counting is given exactly by

$$\begin{aligned} N_{\mathrm{odd}}(x, z) = {} & \Psi(x + z, y) - \Psi(x, y) \\ & - \Psi((x + z)/2, y) + \Psi(x/2, y). \end{aligned}$$

Use (B.26) for both intervals and (B.27) at factor two. This gives

$$N_{\mathrm{odd}}(x, z) = z\widehat{\rho}(x, y)\{1 - 2^{-\beta(x/2,y)} + O(1/d)\}.$$

The expression in braces is bounded below by a positive constant, since the saddle tends uniformly to one. The endpoint ratio $(2X/\ell)/x$ is at most $8C_0 d$. Equations (B.27) and (B.30) make the corresponding densities comparable. Dividing by $|J \cap (2\mathbb{Z} + 1)| = w/2 + O(1)$ proves the lower bound.

For the upper bound, if the interval comes within distance w of zero, it is contained in $[-U, U]$ with $U = \min(X, 2w)$ and $w/2 \le U \le 2w$. Its desired count is at most $2\Psi(U/\ell, y)$. The endpoint ratio $2X/U \le 4C_0 d$ and (B.27) bound this by $C(w/\ell)\widehat{\rho}(2X/\ell, y)$. Otherwise reflect the

interval to $(a, a+w]$ with $a > w$. Then $w/a \geq 1/(C_0 d)$ and the direct interval theorem gives the same bound. This proves the upper estimate.

The relevant divided widths have logarithm $\gg d \log Y_i$. Also the Dickman lower bound and (B.30) give $\widehat{\rho} \geq e^{-C_\rho \Lambda_b}$. Thus the main interval counts tend uniformly to infinity, absorbing the bounded endpoint differences above. □

For a point in box i, choose the deterministic algebraic scale

$$\mathcal{X}_i = 16(d+2)hMV_i^d.$$

The point scale obeys

$$16(4/3)^{d-1} \leq \frac{\mathcal{X}_i}{X_P^{\mathrm{raw}}} \leq 32\, 2^{d-1}. \tag{B.32}$$

Its logarithmic ratio is $O(d)$. The sharp saddle bound in (B.30) shows that this changes a smooth density by a fixed factor, uniformly also after root and parent division.

To apply Lemma B.5 to one parent prime, divide the total variation error by the event's probability scale, $\ell^{-1}\widehat{\rho}$. Since $q, \ell \leq Y_0$,

$$\frac{CqE^{-4}}{\ell^{-1}\widehat{\rho}} \leq Cq\ell E^{-4} e^{C_\rho \Lambda_b} \leq CE^{-4+1/\kappa} e^{C_\rho \Lambda_b} = o(1). \tag{B.33}$$

Taking $\ell = 1$ gives the corresponding estimate for an ordinary smooth value. For a complete weighted band one can alternatively transfer its bounded sum once.

Proposition B.6 (Actual one-point and parent means)**.** *For $g \in \mathcal{G}$, every internal point in the child lattice and every allowed parent prime satisfy*

$$\mathbb{E}_{\nu_g}\Big[\mathbf{1}_{\{0<|\mathcal{F}(P)|/q \in \ell \mathcal{S}_{y_i}\}} \mid R_{q,t}\Big] \asymp \frac{1}{\ell}\widehat{\rho}(2\mathcal{X}_i/(q\ell), y_i). \tag{B.34}$$

For an active internal parent label $l = (\ell, s)$, define its relation count

$$Z_l = \sum_{P \in \mathcal{P}_{\ell,s}^{(i)}(g)} \mathbf{1}_{\{0<|\mathcal{F}(P)|/\ell \in \mathcal{S}_{y_i}\}}, \qquad \tau_l(g) = \mathbb{E}_{\nu_g}[Z_l \mid R_l].$$

With

$$T_i = V_i^2 \rho_{r,i}/\log B, \qquad \rho_\ell = \widehat{\rho}(2\mathcal{X}_i/\ell, y_i),$$

its true conditional mean satisfies

$$\tau_{(\ell,s)}(g) \asymp T_i \rho_\ell/\ell. \tag{B.35}$$

All constants are uniform in the admissible root s and in g.

Proof. Lemma B.3, Lemma B.5, (B.32), and (B.33) prove (B.34) under the single child-root law. For a parent's own law use root $q = \ell$ and the ordinary predicate. Primitivity gives $\ell \nmid b$; the rational numerator is B-smooth and $\ell > B$, so $\ell \nmid G(P)$. The ordinary one-point mean is comparable to ρ_ℓ. The whole-box count in Proposition B.1 is $\asymp T_i/\ell$. Linear expectation proves (B.35); independence of the point values is neither asserted nor needed. □

B.5 An upper bound for the joint smoothness probability

We bound the probability that both point values satisfy the smoothness test. Counting divisors in an arithmetic progression supplies a factor of the smooth density for each coordinate.

Lemma B.7 (An arbitrary-gcd progression bound)**.** *Let I be a positive interval of length L, $m \geq 1$, $g = (a, m)$, and $W_1 = L/m$. Suppose $W_1 \geq 2y$ and $gW_1 \leq \inf I$. If $g \notin \mathcal{S}_y$, the number of y-smooth integers in $I \cap (a + m\mathbb{Z})$ is zero. Otherwise it is at most*

$$CW_1 \log(2y) \max_{W_1/y \leq z \leq 2W_1} \widehat{\rho}(z, y). \tag{B.36}$$

For any $X \geq 2W_1$, $X \geq y$, this is at most

$$CW_1 \log(2y) \widehat{\rho}(X, y)(CXy/W_1)^{1-\beta(X,y)}. \tag{B.37}$$

Proof. Write $n = gn'$ and $m = gm'$. Each counted n' is a unit modulo m', is y-smooth, and is at least W_1. Multiplying its prime factors until their product first exceeds W_1/y gives a divisor

$$W_1/y < e \leq W_1, \qquad (e, m') = 1.$$

For a specified e, the quotient n'/e lies in a single progression and contributes at most $W_1/e + 1$ integers. Sum this upper bound over all smooth e in the indicated range. Dyadic grouping has $O(\log(2y))$ groups and gives (B.36); the $+1$ terms are retained since $e \leq W_1$. The global dilation bound (B.29) gives (B.37). □

Fix a complete parent band

$$J = \{\ell : Y < \ell \leq 2Y,\ \ell \text{ prime},\ \ell \nmid N\}, \qquad Y \geq y, \qquad |J| \geq c_J Y/\log(2Y).$$

Let $v_\ell > 0$ with $v_{\max}/v_{\min} \leq L_v$, where L_v is fixed. Define

$$F_J(n) = \sum_{\ell \in J} v_\ell \mathbf{1}_{\{0 < |n| \in \ell \mathcal{S}_y\}}. \tag{B.38}$$

The summand supports are disjoint: a supported integer has exactly one prime factor above y, to exponent one. Thus $0 \leq F_J \leq v_{\max}$.

Lemma B.8 (Finite window normalization and marginals)**.** *Let K be even, $(\eta, K) = 1$, and let $\mathcal{W}$ be a translated nondegenerate parallelogram of area $\mathcal{A}$, projection widths w_1, w_2, and longest vertical and horizontal chords L_2, L_1. Suppose that, for a sufficiently small absolute $c_0 > 0$, both orientations satisfy*

$$\frac{1}{w_1} + \frac{K}{L_2} + \frac{K}{w_1 L_2} \leq c_0. \tag{B.39}$$

Then the actual odd-lattice count satisfies

$$\begin{aligned} Z_{\mathcal{W}} = |\mathcal{W} \cap \mathcal{L}_{K,\eta}| &= \frac{\mathcal{A}}{2K} + O(L_2/K + w_1 + 1), \\ Z_{\mathcal{W}} \asymp \frac{\mathcal{A}}{2K}, &\qquad Z_{\mathcal{W}} \geq \frac{\mathcal{A}}{4K}. \end{aligned} \tag{B.40}$$

Under the uniform law on these lattice points, for $j = 1, 2$,

$$\mathbb{P}_{\mathcal{W}}(n_j = u) \leq \frac{C}{w_j}, \qquad \mathbb{P}_{\mathcal{W}}(|n_j| < T) \leq \frac{C(T+1)}{w_j}. \tag{B.41}$$

In particular, if $w_j \ge X_j/(C_0 d)$, then

$$\mathbb{P}_{\mathcal{W}}(|n_j| < X_j y^{-C_X}) \le Cd\big(y^{-C_X} + X_j^{-1}\big). \tag{B.42}$$

These estimates are uniform in the translate and have no reciprocal-angle factor.

Proof. Let $l(x)$ be the vertical chord length. Its extension by zero has total variation at most $2L_2$, and $\mathcal{A} \asymp w_1 L_2$. Each odd first coordinate has $l(x)/K + O(1)$ allowed second coordinates, since K is even and η is a unit. Thus

$$Z_{\mathcal{W}} = K^{-1} \sum_{n_1 \text{ odd}} l(n_1) + O(w_1 + 1) = \frac{\mathcal{A}}{2K} + O(L_2/K + w_1 + 1).$$

The assumed inequality gives the asserted lower bound. Each row contains at most $L_2/K + 1$ points, so its probability is at most $C(L_2 + K)/\mathcal{A} \le C/w_1$. Transposition proves the second marginal. Summing over at most $2T + 1$ integers and then using $X_j/w_j \le C_0 d$ proves the strip bounds, including their endpoint terms. □

Proposition B.9 (All translates and all determinants)**.** *Let $K, \eta, \mathcal{W}$ satisfy Lemma B.8. Suppose also that $\mathcal{W} \subset [-X_1, X_1] \times [-X_2, X_2]$ and $w_j \ge X_j/(C_0 d)$. Put*

$$c_j = \sum_{\ell \in J} \frac{v_\ell}{\ell} \widehat{\rho}(2X_j/\ell, y), \qquad \gamma = \max_{j,\ell}\{1 - \beta(2X_j/\ell, y)\} \in [0, 1).$$

Suppose the densities in each of these dyadic parent bands have a fixed comparison ratio. Choose

$$0 < \lambda_j \le T_j \le X_j, \qquad \lambda_1/(4Y) \ge 2y, \qquad \lambda_2/(2YK) \ge 2y.$$

Define

$$\mathcal{Q}_1 = C \log(2y)(CX_1 y/\lambda_1)^\gamma, \qquad \mathcal{Q}_2 = C \log(2y) \log(2Y)(CX_2 yK/\lambda_2)^\gamma.$$

For the uniform law on $\mathcal{W} \cap \mathcal{L}_{K,\eta}$,

$$\begin{aligned} \mathbb{E} F_J(n_1) F_J(n_2) \le{}& C\mathcal{Q}_1 \mathcal{Q}_2 c_1 c_2 \left(1 + \frac{\lambda_1 \log(2X_1)}{w_1}\right)\left(1 + \frac{\lambda_2 \log(2X_2)}{L_2}\right) \\ &+ Cd\, v_{\max}^2 \left(\frac{T_1 + 1}{X_1} + \frac{T_2 + 1}{X_2}\right). \end{aligned} \tag{B.43}$$

Constants can depend on C_0, c_J, L_v and the fixed density ratio, but not on the translate or determinant. For a law at total variation distance ϵ from any mixture of these window laws, add only $v_{\max}^2 \epsilon$ to the right side.

Proof. Fix a supported odd n_1. On a positive second-coordinate interval $I \subset [T_2, 2X_2]$ of length L, with $\lambda_2 \le L \le \inf I$, write $W_1 = L/K$. For $\ell \nmid K$, substitute $n_2 = \ell m$. Lemma B.7 with modulus K, interval length L/ℓ, and density endpoint $2X_2/\ell$ gives a contribution at most

$$C \frac{W_1 v_\ell}{\ell} \widehat{\rho}(2X_2/\ell, y) \log(2y)(CX_2 y/W_1)^\gamma.$$

The endpoint and progression length both divide by ℓ; no extra factor ℓ^γ occurs.

If $\ell \mid K$, a contribution requires $\ell \mid n_1$. There is at most one such parent prime because $F_J(n_1) > 0$. After division by ℓ, the modulus is K/ℓ, so this contribution has the same bound

with its factor $1/\ell$ removed. The sum over the complete parent band bounds this exceptional contribution:

$$v_\ell \widehat{\rho}(2X_2/\ell, y) \le C \log(2Y) c_2.$$

If a progression gcd is not smooth, its contribution is zero by Lemma B.7. Otherwise its required lower-endpoint condition follows from $L \le \inf I$. The progression length divided by its modulus is at least $2y$ by the chosen padding lengths. Also $2X_2/\ell \ge 2W_1$ in the exceptional case, since $\ell \mid K$ and $L \le X_2$. All conditions of that lemma are therefore valid, including at higher prime powers in K. Summing the contributions yields

$$\sum_{\substack{n_2 \in I \\ n_2 \equiv \eta^{-1} n_1 \ (K)}} F_J(n_2) \le \mathcal{Q}_2 c_2 L/K. \tag{B.44}$$

For the first coordinate, use modulus two. All parent primes are odd, so there is no exceptional parent, and the analogous bound is $\mathcal{Q}_1 c_1 L/2$. Negative intervals are treated by reflection.

Outside $|n_j| < T_j$, split each fibre into positive or reflected dyadic pieces of length at most their lower endpoint. Pad any shorter piece to length λ_j. This stays inside $[T_j, 2X_j]$ and adds at most $C\lambda_j \log(2X_j)$ total length per fibre or projection. Let $l(x)$ be the vertical chord length. It is nonnegative and concave on the projection, so $\mathcal{A} \ge w_1 L_2/2$. Apply (B.44) first, then the first-coordinate bound by the following layer-cake calculation. Put $\delta_j = C\lambda_j \log(2X_j)$. Every superlevel set of the concave function l is an interval. After deleting the first central strip and applying the same padding, the first-coordinate estimate gives

$$\begin{aligned} \sum_{\substack{n_1 \text{ odd} \\ |n_1| \ge T_1}} F_J(n_1) l(n_1) &= \int_0^{L_2} \sum_{\substack{n_1 \text{ odd, } |n_1| \ge T_1 \\ l(n_1) > u}} F_J(n_1)\, du \\ &\le \frac{\mathcal{Q}_1 c_1}{2} (\mathcal{A} + \delta_1 L_2). \end{aligned}$$

The corresponding sum without $l(n_1)$ is at most $\mathcal{Q}_1 c_1 (w_1 + \delta_1)/2$. The second-coordinate bound contributes $l(n_1) + \delta_2$ on each supported row. Hence its numerator is bounded by

$$\frac{\mathcal{Q}_1 \mathcal{Q}_2 c_1 c_2}{2K} (\mathcal{A} + \delta_1 L_2 + \delta_2 w_1 + \delta_1 \delta_2).$$

Using $\mathcal{A} \ge w_1 L_2/2$ proves the following bound without an aspect-ratio or reciprocal-angle loss. The weighted count outside the two central strips is at most

$$C \mathcal{Q}_1 \mathcal{Q}_2 c_1 c_2 \frac{\mathcal{A}}{2K} \left(1 + \frac{\lambda_1 \log(2X_1)}{w_1}\right) \left(1 + \frac{\lambda_2 \log(2X_2)}{L_2}\right).$$

Divide by the actual denominator (B.40). On each central strip the test is at most $v_{\max}^2$, and its probability is at most $Cd(T_j + 1)/X_j$ by (B.41). This proves (B.43). Finally the whole product test is bounded by $v_{\max}^2$; the total variation error is incurred only once for the entire test. □

Corollary B.10 (Pointwise pair and count bounds)**.** *At the parameters* (B.2)–(B.4)*, for every $g \in \mathcal{G}$, every admissible child (q, t), and every complete internal parent band, distinct points satisfy*

$$\mathbb{E}_{\nu_g}[F_J(\mathcal{F}(P_1)/q) F_J(\mathcal{F}(P_2)/q) \mid R_{q,t}] \le D^C c_{P_1} c_{P_2}, \tag{B.45}$$

where

$$c_P = \sum_{\ell \in J} \frac{v_\ell}{\ell} \widehat{\rho}(2X_P^{\mathrm{raw}}/(q\ell), y_i)$$

is independent of the partner. For

$$S_{q,t} = \sum_{P \in \mathcal{P}_{q,t}^{(i)}(g)} F_J(\mathcal{F}(P)/q), \qquad \mu_{q,t}(g) = \mathbb{E}[S_{q,t} \mid g, R_{q,t}],$$

one has

$$\mathbb{E}[S_{q,t}^2 \mid g, R_{q,t}] \le v_{\max} \mu_{q,t} + D^C \mu_{q,t}^2. \tag{B.46}$$

The ordinary smooth indicator has the analogous pair bound, at box i, for every admissible root $B < q \le Y_i$ and distinct rational points with $q \nmid bG(P)N$, with

$$c_P = \widehat{\rho}(2X_P^{\mathrm{raw}}/q, y), \qquad v_{\max} = 1.$$

Proof. Use Lemma B.4. In Proposition B.9, choose a fixed large C_X and set

$$T_j = X_j y^{-C_X}, \qquad \lambda_1 = T_1/(Cd), \qquad \lambda_2 = T_2/(Cd(1+\vartheta)).$$

Equation (B.25) gives the boundary conditions and both padding factors $1 + o(1)$. Since $K, \vartheta \le CY_i^{12}$ and

$$\log X_j \ge \log(hM) + (6d-1)\log Y_i - O(d),$$

the logarithm of $\lambda_2/(2YK)$ is at least

$$\log(hM) + (6d - 26 - C_X)\log Y_i - O(d + \log d).$$

It eventually exceeds $\log(2y)$ by an arbitrarily large amount; the first-coordinate condition is weaker. The fixed choice of C_X only affects the uniform threshold in N.

Crucially, the sharp saddle estimate is $\gamma \ll \log D/\log y_i$. Together with $\log Y_i/\log y_i \le Cq_*$, it gives $\mathcal{Q}_1\mathcal{Q}_2 \le D^C$. The weaker statement $\gamma = O(1/d)$ alone would not give this conclusion uniformly at the highest layer. The complete band and the Dickman lower bound give

$$c_P/v_{\max} \ge e^{-C_\mu \Lambda_b}.$$

Choose $b_0 C_X > 2C_\mu + 20$. Divide the central-strip errors in (B.43) by

$$c_{P_1} c_{P_2}/v_{\max}^2.$$

The normalized errors tend to zero. The error in the distributional approximation is also negligible after the same normalization:

$$\frac{CqE^{-4}}{c_{P_1}c_{P_2}/v_{\max}^2} \le CE^{-4+1/(2\kappa)} e^{2C_\mu\Lambda_b} = o(1).$$

This proves (B.45) for every nonzero determinant and every translate in the mixture.

The one-point result gives $\mu_{q,t} \asymp \sum_P c_P$. On the diagonal use the pointwise inequality $F_J^2 \le v_{\max} F_J$; off the diagonal use (B.45). Distinct primitive positive points cannot be proportional. This proves (B.46) with no omitted rank-one case.

For the ordinary indicator repeat the same divisor and fibre argument without the sum over parent primes. There is no exceptional-parent term and no factor $\log(2Y)$. The padding requirements reduce to $\lambda_1/2 \ge 2y$ and $\lambda_2/K \ge 2y$, already satisfied above. For this branch the root need not be below y_i: the preceding lower endpoint uses only $q \le Y_i$, while $K, \vartheta \le CY_i^{12}$ and the TV bound uses only $q \le Y_0$. Lemma B.4 requires exactly $q \nmid bG(P)N$, which is included in the statement. Thus the endpoint, padding, and density-normalized TV estimates apply throughout the stated ordinary root range. □

B.6 Mean incidence counts and their comparison functions

All relation incidences in this subsection refer to the same sampled form $\mathcal{F}$. An internal parent label is $l = (\ell, s)$. For a candidate counted by Z_l, count its distinct finite child prime $q > B$ at the unique point root $t = a/b \bmod q$. The resulting number of incidences is denoted by $X_{l,(q,t)}$. It includes even valuations and children that are terminal for the sampled form. Such a terminal child's bad indicator is zero. The primes are read from the cofactor of the sampled polynomial value $\mathcal{F}(P)/\ell$ (from $\mathcal{F}(P)$ at the top), so this convention also retains content primes removed when the form is made primitive. Primes dividing the fixed input are removed from the continuing numeric label domain before sampling. No field-dependent deletion is made from that domain or from the formal incidence count.

Given parent weights $a(g, l)$, deterministic once g, l are fixed, set

$$H_{q,t} = \sum_l \frac{a(g,l)}{\tau_l(g)} X_{l,(q,t)}, \qquad a'(g,q,t) = \mathbb{E}_{\nu_g}[H_{q,t} \mid R_{q,t}]. \tag{B.47}$$

The conditional expectations defining the weights are taken before the actual coefficient tuple is sampled. They are not estimates obtained by conditioning that tuple on a successful search.

Proposition B.11 (One internal parent band)**.** *Fix a complete band*

$$J = \{\ell : Y < \ell \le 2Y,\ \ell \text{ prime},\ \ell \nmid N\}, \qquad y_i \le Y, \quad 2Y \le Y_i.$$

Suppose, uniformly in every admissible parent root s,

$$A_b L^{-1} w_\ell \le a(g, \ell, s) \le A_b L w_\ell, \tag{B.48}$$

where $A_b > 0$ and $w_\ell > 0$ are deterministic and $\max_J w_\ell / \min_J w_\ell \le L_w$, with L_w fixed. For the contribution $H_{b,q,t}$ of this band, there is a deterministic $0 \le \gamma_b < 1$ such that, with

$$B_b(q) = A_b q^{\gamma_b - 1} \sum_{\ell \in J} w_\ell,$$

its true conditional mean satisfies

$$cL^{-1} B_b(q) \le a'_b(g,q,t) \le CL B_b(q). \tag{B.49}$$

Moreover

$$\mathbb{E}_{\nu_g}\big[(H_{b,q,t}/a'_b(g,q,t))^2 \mid R_{q,t}\big] \le D^C L^4. \tag{B.50}$$

These conclusions hold pointwise in $g \in \mathcal{G}$ and in every admissible $B < q \le y_i$, including queries in the parent search region that are not active parent labels at their own smaller-box scale.

Proof. Use the deterministic T_i, ρ_ℓ from Proposition B.6, and put

$$v_\ell = \frac{\ell w_\ell}{T_i \rho_\ell}.$$

Their ratio on J is bounded by a fixed constant independent of L, since each factor has a bounded dyadic comparison ratio. For $S = S_{q,t}$ in Corollary B.10, (B.34) and the whole-box child count give

$$\mu_{q,t}(g) \asymp \frac{T_i}{q} \sum_{\ell \in J} \frac{v_\ell}{\ell} \widehat{\rho}(2\mathcal{X}_i/(q\ell), y_i). \tag{B.51}$$

There is one conditioning event, $R_{q,t}$. No independently sampled parent-root law is introduced into this expectation.

Write $z_\ell = 2\mathcal{X}_i/\ell$, $z_* = 2\mathcal{X}_i/Y$, and $\gamma_b = 1 - \beta(z_*, y_i)$. Apply (B.27) at the smaller endpoint z_ℓ/q, with factor $q \le y_i$:

$$\frac{\widehat{\rho}(z_\ell/q, y_i)}{\widehat{\rho}(z_\ell, y_i)} = q^{1-\beta(z_\ell/q, y_i)}(1 + O(1/d)).$$

By (B.28) and (B.30) along the whole interval between these endpoints,

$$|\beta(z_\ell/q, y_i) - \beta(z_*, y_i)| \log q \ll \frac{\log(2q)\log q}{\log(z_\ell/q)\log y_i} \ll \frac{1}{d}. \tag{B.52}$$

The ratio is therefore comparable to q^{γ_b}. Substituting the definition of v_ℓ into (B.51) cancels its deterministic factors and gives

$$\mu_{q,t}(g) \asymp q^{\gamma_b - 1} \sum_{\ell \in J} w_\ell. \tag{B.53}$$

Both the right side and the comparison constants are independent of g.

Let S^{fin} be the same sum restricted to $\ell \nmid b(P)$. Each $b(P)$ has at most $\log V_i/\log Y$ divisors in J. The individual conditional means in (B.34), together with the fixed within-band weight ratio and $|J| \asymp Y/\log Y$, show that

$$\mathbb{E}[S^{\mathrm{fin}} \mid g, R_{q,t}] = (1 - o(1))\mu_{q,t}(g); \tag{B.54}$$

the lost fraction is $O(\log V_i/Y)$. The uniform relative error (B.33) justifies the same assertion in the actual distribution before initialization.

Every surviving semismooth point determines the unique finite parent root $s = a/b \bmod \ell$. It is admissible because $\ell \nmid G(P)$. Conversely each formal parent incidence gives such a term. Equations (B.35) and (B.48) therefore give, in the same sampled form,

$$cA_b L^{-1} S^{\mathrm{fin}} \le H_{b,q,t} \le CA_b L S. \tag{B.55}$$

Taking expectations and using (B.53) proves (B.49). In particular,

$$\frac{H_{b,q,t}}{a_b'(g,q,t)} \le CL^2 \frac{S}{\mu_{q,t}(g)}.$$

Corollary B.10 yields

$$\mathbb{E}[(H_{b,q,t}/a_b')^2 \mid g, R_{q,t}] \le D^C L^4 (1 + v_{\max}/\mu_{q,t}).$$

The large-mean estimate proved next absorbs the final term. The power L^4 comes from one factor L in the upper bound for H and one in the lower bound for its mean, followed by squaring. □

Lemma B.12 (The mean is large on its own level)**.** *For all sufficiently large parameters, the internal count above satisfies $\mu_{q,t}/v_{\max} \ge e^{1000\Lambda_b}$.*

Proof. The elementary height bounds from (B.4) give

$$\log(2\mathcal{X}_i) \le (12 + 6\xi_i)d\Lambda_b$$

for all sufficiently large parameters. Division by primes only decreases this bound. Together with (B.3) and $y_i \geq B$, this gives

$$u \leq c_i^+ d, \qquad c_i^+ = \min\{U_*, (12+6\xi_i)/b_0\} \leq U_*.$$

The upper bound for c_i^+ is fixed, even at the highest level. The Dickman lower bound, the global comparison, and $\log d/\lambda \to 1/3$ therefore give the uniform estimate

$$-\log \widehat{\rho}(x, y_i) \leq \left(\frac{2}{3}c_i^+ + o(1)\right)\Lambda_b \leq \left(\frac{2(12+6\xi_i)}{b_0} + o(1)\right)\Lambda_b.$$

The last, deliberately weaker, bound retains a fixed margin in the exponent; the preceding bound makes its uniform error explicit. The same-level whole-box count is at least $cY_i^{12}\rho_{r,i}/(q\log B)$, with $\rho_{r,i} \geq e^{-\Lambda_b}$, and $q \leq Y_i$. A complete parent band contributes at least a fixed multiple of $1/\log Y$ times its smallest algebraic density. Thus

$$\log(\mu_{q,t}/v_{\max}) \geq \left(11\xi_i - 1 - \frac{2(12+6\xi_i)}{b_0} - o(1)\right)\Lambda_b.$$

For $\xi_i \geq 100 + o(1)$, the coefficient is at least $1086.76 - o(1) > 1000$. No worst density from one level has been divided by the smallest search region from another level. □

Proposition B.13 (Incoming weights from the top level)**.** *Take $a_0(p) = L_{\mathrm{top}}^{-1}$ on J_0, and let the top parent root event be identically true. Put $y = Y_1$ and define, separately from the internal T_i,*

$$T_0 = V_0^2/\log B, \qquad c_0 = \widehat{\rho}(2\mathcal{X}_0, y), \qquad c_q = \widehat{\rho}(2\mathcal{X}_0/q, y).$$

The actual top relation count and its raw mean are

$$Z_p = \sum_{P \in \mathcal{A}_p(g)} \mathbf{1}_{\{0<|\mathcal{F}(P)|\in\mathcal{S}_y\}}, \qquad \tau_p(g) = \mathbb{E}_{\nu_g} Z_p.$$

For all $g \in \mathcal{G}$,

$$\tau_p(g) \asymp T_0\rho_p c_0/p, \qquad u_P(g) = \frac{1}{L_{\mathrm{top}}\tau_{p(P)}(g)} \asymp u_* := \frac{Y_0}{L_{\mathrm{top}}T_0\rho_* c_0}. \tag{B.56}$$

For every admissible child $B < q \leq y$, the exact formal incoming sum has

$$a_1(g,q,t) \asymp \frac{c_q}{qc_0} \asymp q^{\gamma_0 - 1}, \qquad \mathbb{E}[(H_{q,t}/a_1)^2 \mid g, R_{q,t}] \leq D^C, \tag{B.57}$$

where $\gamma_0 = 1 - \beta(2\mathcal{X}_0, y)$ is deterministic and lies in $[0,1)$.

Proof. For $P \in \mathcal{A}_p(g)$, $-G(P) = pn_s$, where n_s is B-smooth and $p \nmid N$. Thus $(N, G(P)) = 1$, even though $G(P)$ itself need not be B-smooth. The unconditioned part of Lemma B.3 applies and gives ordinary mean $\asymp c_0$ at each point. Combining this with (B.12) proves the first assertion in (B.56). The dyadic parent-density comparison gives the second. Each $u_P(g)$ can depend on g; it is simply fixed before the polynomial is sampled and has the displayed common deterministic scale.

For a child-root condition, $B < q \leq Y_1 < Y_0 < p(P)$, so $q \nmid G(P)$. Primitivity in the child lattice gives $q \nmid b(P)$. The root-conditioned parts of Lemmas B.3 and B.4 are therefore valid and give a one-point ordinary mean $\asymp c_q$. The top parent $p(P)$ divides $G(P)$, and its event

holds identically. Primitivity then gives $p(P) \nmid b(P)$, since otherwise $p(P)$ would divide both coordinates.

Retaining terminal and even-valuation incidences, the exact reordering under the single child-root law is

$$H_{q,t} = \sum_{P \in \mathcal{A}_{q,t}(g)} u_P(g) \mathbf{1}_{\{0<|\mathcal{F}(P)|/q \in \mathcal{S}_y\}}. \tag{B.58}$$

Indeed $q \le y$, so on that root event the ordinary value is smooth if and only if the value divided by q is smooth. Now keep g fixed. The pointwise comparison in (B.56) and the separate primitive child-lattice aggregate count give

$$\sum_{P \in \mathcal{A}_{q,t}(g)} u_P(g) \asymp u_* \frac{T_0 \rho_J}{q} \asymp \frac{1}{q c_0}. \tag{B.59}$$

The last step uses $\rho_J \asymp L_{\mathrm{top}} \rho_* / Y_0$. This calculation bounds each actual weight above and below at fixed g; it does not factor an expectation of correlated environment quantities. Linear expectation in (B.58) gives $a_1 \asymp c_q/(qc_0)$. The smooth-number ratio theorem and the estimate for the variation of the saddle point (B.52) give $c_q/c_0 \asymp q^{\gamma_0}$.

For distinct points use the ordinary version of Corollary B.10. On the diagonal use $u_P^2 \mathbf{1} \le u_{\max} u_P \mathbf{1}$. Since the weights are fixed with respect to the coefficient distribution before the initialization tests,

$$\mathbb{E}[H_{q,t}^2 \mid g, R_{q,t}] \le D^C a_1^2 + u_{\max} a_1.$$

For completeness, (B.56) and the aggregate count give

$$a_1/u_{\max} \gg V_0^2 \rho_J c_q / (q \log B).$$

The density estimate in the proof of Lemma B.12, now with $\rho_J \ge e^{-\Lambda_b}$ and $q \le Y_0$, gives

$$\log(a_1/u_{\max}) \ge \left(11\xi_0 - 1 - \frac{2(12+6\xi_0)}{b_0} - o(1)\right) \Lambda_b \ge 1000 \Lambda_b.$$

This bounds the diagonal contribution and proves (B.57). □

Lemma B.14 (Lower bound for the proportion of candidates)**.** *For every active internal or top label l, let Ω_l be its geometric candidate set before roughness, primitivity, and smoothness tests. On $\mathcal{G}$, uniformly in all such labels,*

$$\frac{\tau_l(g)}{|\Omega_l|} \ge p_{\mathrm{den}}, \qquad p_{\mathrm{den}} = 2^{-(U_*+10)dm_b} \tag{B.60}$$

for one common sufficiently large input threshold.

Proof. For $0 \le i < j$, let $e_i = \log_2 Y_i$, so the next algebraic cutoff is $y_i = Y_{i+1}$. From $h = 2H + 1 < 4H$, $\log_2 M \le d^2 m_b$, and $e_0 \le d^2 m_b/(2\kappa)$, the exact coefficient windows imply

$$\begin{aligned} \log_2(32(d+2)hM) &\le 8d^2 m_b + 3\log_2 M + 12 e_0 + O(d + \log d) \\ &\le (11 + 6/\kappa) d^2 m_b + O(d + \log d) < 12 d^2 m_b \end{aligned}$$

for all sufficiently large admissible inputs. Therefore, at every box,

$$\log_2(2\mathcal{X}_i) < 12 d^2 m_b + 6 d e_i = 6d(e_i + 2dm_b).$$

The actual integer recursion gives $e_{i+1} \geq (e_i + 2dm_b)/q_*$, including the case where the maximum defining e_{i+1} is attained at e_B. Hence

$$\frac{\log(2\mathcal{X}_i)}{\log y_i} \leq 6q_*d = U_*d.$$

Root and parent divisions can only decrease this upper bound.

For completeness, the elementary Dickman identity $u\rho_D(u) = \int_{u-1}^{u} \rho_D(t)\,dt$ and monotonicity give $\rho_D(u) \geq (2u)^{-1}\rho_D(u-1/2)$ for $u \geq 1$. Iteration down to $[0,1]$ proves $\rho_D(u) \geq (2u)^{-2u}$. The global comparison from [29, Theorem 1.1], applied in the uniform range already checked in (B.30), now gives for every active internal or top one-point density c_l, where $c_l = \widehat{\rho}(2\mathcal{X}_i/\ell, y_i)$ internally and $c_l = c_0$ at the top,

$$-\log c_l \leq 2U_*d\log(2U_*d) + o(\Lambda_b) = \left(\frac{2U_*}{3} + o(1)\right)\Lambda_b.$$

The last equality uses $\log d/\lambda \to 1/3$ and $\Lambda_b \sim d\lambda$; its error is uniform in the level.

For an active internal label, $|\Omega_{(\ell,s)}| \ll V_i^2/\ell$ and $\tau_{(\ell,s)} \gg V_i^2\rho_{r,i}c_l/(\ell\log B)$. For a top label the analogous statements hold with $\ell = p$ and $\rho_{r,i} = \rho_p$. In both cases the rational density is at least $e^{-\Lambda_b}$. Consequently

$$\frac{\tau_l}{|\Omega_l|} \geq \exp\left[-\left(1 + \frac{2U_*}{3} + o(1)\right)\Lambda_b\right] \geq \exp(-(U_*+10)\Lambda_b) = 2^{-(U_*+10)dm_b}.$$

Here the factor $1/\log B$ counts denominators with no prime divisor at most B. □

Corollary B.15 (Uniform top density and pool moment)**.** *With the candidate sets and p_{den} of Lemma B.14, uniformly for $g \in \mathcal{G}$ and $p \in J_0$, for all sufficiently large parameters,*

$$\begin{aligned} &\frac{\tau_p(g)}{|\Omega_p|} \geq p_{\mathrm{den}}, \qquad 0 \leq Z_p \leq U_N\tau_p(g) \quad \textit{for every sampled coefficient tuple,} \\ &\mathbb{E}_{\nu_g}\big[(Z_p/\tau_p(g))^2\big] \leq U_N, \qquad U_N = \exp((U_* + o(1))\Lambda_b). \end{aligned} \tag{B.61}$$

Proof. Lemma B.14 supplies the density bound and, in its proof, the uniform estimate $-\log c_0 \leq (2U_*/3 + o(1))\Lambda_b$. Since $0 \leq Z_p \leq |\mathcal{A}_p(g)|$, we have

$$\begin{aligned} \frac{Z_p}{\tau_p} &\leq \frac{|\mathcal{A}_p(g)|}{\tau_p} \leq C/c_0 \leq U_N \quad \text{for every sampled coefficient tuple,} \\ \frac{\mathbb{E}_{\nu_g}Z_p^2}{\tau_p^2} &\leq \frac{|\mathcal{A}_p(g)|}{\tau_p} \leq C/c_0 \leq U_N. \end{aligned}$$

This is an unconditioned top-parent moment, distinct from the child-root incoming moment in Proposition B.13. □

B.7 Comparison of the probability measures

Let μ be the original independent uniform environment law and fix

$$\bar{\mu} = \mu(\,\cdot \mid \mathcal{G} \cap \{Q \nmid N\}).$$

This is the same environment marginal used by the ordinary complete-band baseline. The incoming estimates above are pointwise on this support, but the ordinary residual estimate is

averaged under this particular $\bar{\mu}$. We do not change its marginal after selecting environments with a small recursive bad mass. The numeric label domain consists of predetermined complete prime bands with the fixed-N deletions and, for each prime, all its $q-1$ admissible roots.

For one incoming layer write $a(g,q,t) = \mathbb{E}[H_{q,t} \mid g, R_{q,t}]$ and

$$\mathcal{W}_0 = \mathbb{E}_g \sum_{q,t} p_{q,t}(g) a(g,q,t).$$

When $\mathcal{W}_0 > 0$, define on the same finite $(g, \mathcal{F}, q, t)$ space

$$d\sigma = \frac{a(g,q,t)}{\mathcal{W}_0} \mathbf{1}_{R_{q,t}} \, d\bar{\mu}(g) d\nu_g(\mathcal{F}), \qquad d\zeta = \frac{H_{q,t}}{\mathcal{W}_0} \, d\bar{\mu}(g) d\nu_g(\mathcal{F}). \tag{B.62}$$

Here $H_{q,t} = 0$ off its child-root event. Both are probability measures, and the identity

$$\frac{d\zeta}{d\sigma} = \frac{H_{q,t}}{a(g,q,t)} \tag{B.63}$$

is exact. A nonnegative zero-mean branch vanishes almost surely and is omitted. If $\mathcal{W}_0 = 0$, no incoming measure is normalized.

Proposition B.16 (Comparison to complete-band sampling)**.** *Suppose $B(q) > 0$ is a deterministic profile, with a fixed comparison ratio C_B within each child band J_b, and*

$$L^{-1} B(q) \le a(g,q,t) \le L B(q)$$

for all environments and admissible roots. Set $M_b = \sum_{q \in J_b} B(q)$ and $M_* = \sum_b M_b$. Define β by choosing $g \sim \bar{\mu}$, a band with probability M_b/M_*, a uniform prime from that band, a uniform admissible root, and then the coefficient distribution before the initialization tests $\nu_g(\cdot \mid R_{q,t})$. Then

$$\frac{d\sigma}{d\beta} = \frac{a(g,q,t)}{\mathcal{W}_0} \frac{M_* |J_b| (q-1) p_{q,t}(g)}{M_b} \le \frac{C_p C_B}{c_p} L^2. \tag{B.64}$$

If the conditional second moment of H/a is at most K^2 at every label and environment, then

$$\int (d\zeta/d\sigma)^2 \, d\sigma \le K^2. \tag{B.65}$$

Proof. Equation (B.7) implies

$$c_p L^{-1} M_* \le \mathcal{W}_0 \le C_p L M_*.$$

The exact density of β with respect to $d\bar{\mu} \, d\nu_g$ and counting measure on labels is

$$\frac{M_b}{M_* |J_b| (q-1)} \frac{\mathbf{1}_{R_{q,t}}}{p_{q,t}(g)}.$$

Dividing (B.62) by this density gives the equality in (B.64). Since $|J_b| B(q)/M_b \le C_B$, the displayed bounds give its inequality. For (B.65), integrate the conditional second-moment bound against $\mathcal{W}_0^{-1} p_{q,t}(g) a(g,q,t) d\bar{\mu}(g)$ and sum the labels. Its total mass is one by definition. □

Positive branches do not require independence. If $H = \sum_b H_b$ and $a = \sum_b a_b$, then pointwise

$$\frac{(\sum_b H_b)^2}{\sum_b a_b} \le \sum_b \frac{H_b^2}{a_b}. \tag{B.66}$$

After conditional expectation and division by a, the normalized second moment is bounded by the a_b/a-weighted mean of the branch bounds. A pass-through branch retains the same root and form and has $H_{\text{pass}} = a_{\text{pass}}$ on its root event, so its density is one.

Theorem B.17 (Pointwise incoming supply and its comparison measures)**.** *For all sufficiently large parameters* (B.2)–(B.3)*, the common environment event $\mathcal{G}$ satisfies* (B.22)*. On each fixed $g \in \mathcal{G}$, the exact raw law* (B.6) *has the following properties.*

1. *Each active complete parent band with* (B.48) *has the deterministic next profile* (B.49) *and conditional second moment* (B.50)*. The constants are uniform in every admissible child-root condition in its parent search region.*

2. *Uniform top parent weights $a_0(p) = L_{\mathrm{top}}^{-1}$ give* (B.57) *under the unconditioned top parent law and the single admissible child-root law.*

3. *Positive active branches and pass-through branches combine with no loss depending on their number. Starting from this initialization, their deterministic profiles are positive sums of powers q^{ξ}, $-1 \le \xi \le 0$, with coefficients independent of the environment. If support cutoffs occur at the prescribed dyadic bucket boundaries, the within-band profile ratio is at most two. For suitable fixed C, one can therefore choose comparison losses*

$$L_{i+1} \le D^C L_i, \qquad L_i \le D^{C(i+1)}, \qquad K_i^2 \le D^C L_i^4. \tag{B.67}$$

   Here K_i^2 is the integrated second moment of the incidence density for transition i to $i+1$; the initial comparison loss is $L_0 = 1$, distinct from the top population L_{top}.

4. *On each layer, the complete-band baseline constructed in Proposition B.16 and its exact conditional-mean measure satisfy*

$$\frac{d\sigma_i}{d\beta_i} \le C L_i^2.$$

   The actual formal incidence measure has density H_i/a_{i+1} relative to σ_{i+1}, with second moment at most K_i^2.

These assertions impose no conditioning on acceptance, termination, or successful recursion.

Proof. The first two assertions are Propositions B.11 and B.13. An explicit deterministic profile makes the closure transparent. Start with $B_1(q) = q^{\gamma_0 - 1}$, and write

$$J_e = \{p \text{ prime} : 2^e < p \le 2^{e+1},\ p \nmid N\}, \qquad \gamma_{i,e} = 1 - \beta(2\mathcal{X}_i/2^e, y_i).$$

For $B < q \le y_i = Y_{i+1}$, choose

$$B_{i+1}(q) = B_i(q) + \sum_{e=e_{i+1}}^{e_i - 1} q^{\gamma_{i,e} - 1} \sum_{\ell \in J_e} B_i(\ell). \tag{B.68}$$

The first term is the pass-through contribution. On each active band use $A_b = 1$, $w_\ell = B_i(\ell)$ in Proposition B.11. All coefficients in (B.68) are independent of the environment and root; every exponent remains in $[-1, 0]$. The integer powers of two give an exact partition of $(Y_{i+1}, Y_i]$, with no partial endpoint band. Thus the conditional means compare to this positive sum with the same loss up to a fixed factor. Inequality (B.66) proves the second-moment claim for the sum. On a dyadic bucket every generated power varies by at most a factor two; a positive sum has the same property when its active summands do not change inside the bucket. This proves (B.67) inductively, with a polynomial allowance absorbing fixed constants. Proposition B.16 then proves the last assertion. □

For example, if a baseline bad event has probability at most D^{-3A}, where $A = D^{1/48+o(1)}$, then uniformly over $i = O(\log D)$,

$$\sigma_i(\mathrm{Bad}) \le CL_i^2 D^{-3A} \le D^{-2A}$$

for all sufficiently large N. This transfer uses the pointwise comparison (B.64); the separate incidence second moment is used when an incidence-weighted loss is estimated.

C Concentration of relation counts

Choose a prime uniformly in a dyadic band and then choose one of its admissible roots uniformly. We bound the probability that the number of relations for this label differs from its conditional mean by more than half that mean. Section B.7 gives the change of measure needed when relations from the preceding level determine the weights of the labels.

In a direct second-moment calculation, the principal character leaves an error too large to iterate through all the levels. For each polynomial value we subtract its smoothness probability conditional on its size interval and its valuations at small primes, with the cutoffs specified below. This choice fixes the same function of a value for every pair containing it. The remaining nonprincipal terms involve a rational sum and an algebraic sum at the same determinant. We expand both in primitive characters before summing over determinants. The fixed-moment density calculations in [33, `c12rat`, Section 5.6; `thmAB`, Section 6] allow types depending on the companion point. Here the proxy is fixed for each value and the moment order grows with the input. For our cutoff, Lemma C.4 proves the relative counts that take the place of the full-primorial counting hypothesis in [33, `c12rat`, Section 4, FIB]. Divisor truncation and Bonferroni arguments already occur in [33, `c12rat`, Sections 5.3 and 8.2]. Its Section 4, CP/CPW, uses the common determinant of the two points. The FAM-A estimate in [33, `thmAB`, Section 9] first takes a positive upper bound for the rational pair contribution. Here we retain both conductor expansions in the same sum (Lemma C.12).

C.1 The relation count and its approximation

We use the U-instance parameters and unconditioned polynomial experiment of Section B.1, including its rough odd input and the exact windows (B.4). Write $\Lambda = D^{1/3}\lambda \sim \Lambda_b$, and denote the coefficient side length $2H+1$ by h_0 in this section. At an active internal level put $Y = Y_i$, $y = Y_{i+1}$, and $V = Y^6$. The cutoff construction gives, uniformly in the level,

$$B \le y < Y \le Y_0, \qquad \frac{\log Y}{\log y} \le C q_*, \qquad \log Y_0 = \left(\frac{1}{2\kappa} + o(1)\right)\log M. \tag{C.1}$$

The scales of the polynomial values used below have smoothness parameter between cd and Cd; constants may depend on the fixed cutoff parameters. The number of internal levels and the moment order are

$$j = \frac{\log D}{96\log 2} + O(1), \qquad A = 4^{j+10} = D^{1/48+o(1)}, \qquad m = \lceil 100A \rceil. \tag{C.2}$$

Let μ be the original independent uniform law of a prime $Q \in (3M/4, M]$ and $r \in (M/2, 3M/4] \cap \mathbb{Z}$. In this appendix we sample the environment from

$$\bar{\mu} = \mu(\,\cdot \mid \mathcal{G} \cap \{Q \nmid N\}),$$

where $\mathcal{G}$ is the common rational-supply event of Corollary B.2. Given g, we draw the form from (B.6) before testing it for acceptance. The height and parity conditions (B.5) in particular make every primitive value of $\mathcal{F}$ odd.

Let

$$\mathcal{P} = \{(a,b) \in \mathbb{Z}^2 : V/8 \le a \le 2V/3,\ V/2 < b \le 3V/4,\ \gcd(a,b) = 1\}.$$

For $P = (a,b)$, the rational indicator $R_P(g)$ requires

$$MV/16 \le rb - Qa \le MV/8, \qquad rb - Qa \in \mathcal{S}_B, \qquad P^-(b) > B. \tag{C.3}$$

These conditions imply $a < b$ and $\gcd(N, G(P)) = 1$. An active label is an odd prime $\ell \in (y, Y]$ with $\ell \nmid N$, together with a residue $s \pmod{\ell}$ satisfying $Qs - r \neq 0 \pmod{\ell}$. Write

$$\mathcal{P}_{\ell,s} = \{(a,b) \in \mathcal{P} : a \equiv sb \pmod{\ell}\}, \qquad R_{\ell,s} = \{\mathcal{F}(s,1) \equiv 0 \pmod{\ell}\}.$$

For an accepted rational point in this lattice, $\ell \nmid bG(P)N$. Throughout this section its value

$$n_P = \mathcal{F}(P)/\ell$$

is considered under the original distribution conditional on $R_{\ell,s}$. Define the exact finite mean and count by

$$\begin{aligned} f_P &= 1_{\{0<|n_P|\in\mathcal{S}_y\}}, \quad a_P = \mathbb{E}[f_P \mid g, R_{\ell,s}], \\ Z &= \sum_{P\in\mathcal{P}_{\ell,s}} R_P f_P, \quad \tau = \sum_{P\in\mathcal{P}_{\ell,s}} R_P a_P. \end{aligned} \tag{C.4}$$

Definition C.1 (Complete-band baseline)**.** Fix a complete dyadic band $J = \{p \text{ prime} : U_0 < p \le 2U_0\}$ contained in $(y, Y]$, and delete its primes dividing N. The baseline first chooses the environment as above, then a uniform remaining prime ℓ, then a uniform one of its $\ell - 1$ admissible residues s, and finally the sampled field conditioned on $R_{\ell,s}$. Deterministic convex mixtures of these complete-band laws are also allowed.

We use two estimates for the rational points. With

$$\begin{aligned} X_r = MV/8, \qquad \rho_r &= \frac{\#([X_r/2, X_r] \cap \mathcal{S}_B)}{X_r}, \\ n_0 = V^2/\ell, \qquad \Theta &= \exp(D^{1/6}\lambda^2), \end{aligned}$$

Proposition B.1 gives on the same event $\mathcal{G}$

$$\begin{aligned} &\sum_{P\in\mathcal{P}_{\ell,s}} R_P \asymp \frac{n_0\rho_r}{\log B}, \\ &\#\{(P_1,P_2) : R_{P_1}R_{P_2} = 1,\ 0 < |\det(P_1,P_2)| \le cV^2/\Theta\} \\ &\qquad \ll \frac{n_0^2\rho_r^2}{(\log B)^2\Theta}. \end{aligned} \tag{C.5}$$

The event has probability $1 - o(1)$ before fields are sampled.

Here and below

$$c_o(x,y) = 2(1 - 2^{-\beta(x,y)})\widehat{\rho}(x,y), \tag{C.6}$$

with the saddle and smooth density defined in Section B.4. The reference scale

$$X_P = \frac{8(d+2)h_0|G(P)|b^{d-1}}{\ell}, \qquad c_P = c_o(X_P, y) \tag{C.7}$$

is fixed before the field and does not depend on a second point.

Definition C.2 (One fixed exact-density proxy)**.** Set $z = R = D^{8A}$, take $K_p = \lceil \log R/\log p \rceil$, and put

$$Q_o = \prod_{3\le p\le z} p^{K_p}.$$

Choose a positive and negative multiplicative grid of relative mesh

$$\eta = \frac{D^{-10A}}{C_g d\Theta}$$

before the field is sampled; C_g is a fixed geometric constant. For a complete grid cell I, let

$$D_I(n) = \mathbb{E}_{a \text{ uniform on } I\cap(2\mathbb{Z}+1)}\big[1_{\{0<|a|\in\mathcal{S}_y\}} \mid \gcd(a,Q_o) = \gcd(n,Q_o)\big]. \tag{C.8}$$

Null fibres are assigned value zero. For a fixed large constant C_X and a cutoff

$$\log U = C_U \log(2z)(\Lambda + A\log D), \tag{C.9}$$

define $D_P^U(n) = D_I(n)$ on the complete cell containing n, except that it is zero when $\gcd(n,Q_o) > U$ or $|n| < X_P y^{-C_X}$. The last cutoff is rounded to grid boundaries. Finally put $\widetilde{Z} = \sum_P R_P D_P^U(n_P)$.

The grid and gcd in Definition C.2 fix a single function of n_P. We can therefore average that function under either the one-point or the two-point comparison model.

Theorem C.3 (Ordinary baseline supply)**.** *Fix an active level, a complete band, an environment $g \in \mathcal{G}$, an admissible label (ℓ, s) in that band, and $P \in \mathcal{P}_{\ell,s}$ with $R_P(g) = 1$. For all sufficiently large N, uniformly in these choices, the L^m-norm under $\nu_g(\cdot \mid R_{\ell,s})$ satisfies*

$$\frac{\|D_P^U(n_P) - a_P\|_m}{a_P} \le D^{-5/24+o(1)}. \tag{C.10}$$

Under every baseline of Definition C.1,

$$\mathbb{E}_{\text{base}}\frac{(Z-\widetilde{Z})^2}{\tau^2} \le D^{-4A+O(1)}, \qquad \mathbb{P}_{\text{base}}\{|Z-\tau| > \tau/2\} \le D^{-3A}. \tag{C.11}$$

The $o(1)$ in (C.10) *is uniform in the stated variables.*

C.2 Local smoothness densities

We use the common smooth-number input (B.26)–(B.29), with the same fixed source parameter and the smaller-endpoint saddle convention. The additional endpoint checks for retained divisors and finite cells are given below; they are stronger than the one-point checks in (B.30).

Let $I = (x, x+L]$ and $K = 2^a K_o$, where $a \ge 1$ and K_o is odd. All two-point applications have $K = 2t$; the factor two is part of the model. Negative cells are reflected to the positive axis; the possible endpoint discrepancy is at most two integers and is absorbed by the finite endpoint bounds below. Write

$$S = \gcd(n, Q_o), \qquad T = \gcd(n, K_o)$$

for a uniform odd integer $n \in I$. The ideal odd valuation law has independent coordinates, with

$$\mathbb{P}(J_p = j) = (1-p^{-1})p^{-j} \quad (j < k), \qquad \mathbb{P}(J_p = k) = p^{-k} \tag{C.12}$$

when the cap is k. Its full-K_o mass is

$$\pi_K(\delta) = \frac{\varphi(K_o/\delta)}{K_o} = \frac{2\varphi(K/\delta)}{K}. \tag{C.13}$$

Lemma C.4 (Finite counting and the density product)**.** *Let*

$$E_\nu = \frac{(\sum_{3\le p\le z} p^{-1})^\nu}{\nu!}, \qquad B_\nu = (1+z)^{\nu+1}.$$

On each compatible retained joint pattern $S = s \le U, T = \delta$, the actual pattern mass differs relatively from the ideal CRT mass by at most

$$\eta_{\mathrm{ar}} \le Cz\left(E_\nu + \frac{UK^2B_\nu}{L}\right). \tag{C.14}$$

The full-K_o marginal has relative error $O(K/L)$. Moreover, with $\gamma = 1-\beta(x,y)$, define for $p \le y$

$$t_{p,k}(j;\gamma) = \begin{cases} p^{\gamma j}(1-p^{-1+\gamma})/(1-p^{-1}), & j < k, \\ p^{\gamma k}, & j = k. \end{cases} \tag{C.15}$$

For $p > y$ define instead $t_{p,k}(0) = (1-p^{-1})^{-1}$ and $t_{p,k}(j) = 0$ for $j \ge 1$. If the divisor and interval ranges below hold, the exact conditional smooth density has the uniform formula

$$\mathbb{E}[f_y(n) \mid T = \delta] = c_o(x,y) \prod_{p^k \| K_o} t_{p,k}(v_p(\delta);\gamma)(1 + O(\eta_{\mathrm{FO},K} + K/L)). \tag{C.16}$$

For $S = s \le U$, the analogous formula for $D_I(s)$ has the product over $3 \le p \le z$ and error $O(\eta_{\mathrm{FO},U} + \eta_{\mathrm{ar}})$. If δ has a prime factor exceeding y, both sides of (C.16) *are zero.*

Proof. Write $n = su$. Bonferroni truncation of the remaining small-prime avoidances gives

$$\#\{n \in I : n \text{ odd},\ S = s\} = \frac{L}{2s} \prod_{\substack{3\le p\le z \\ v_p(s)<K_p}} (1-p^{-1}) + O((L/s)E_\nu + B_\nu). \tag{C.17}$$

There are at most B_ν retained divisor terms. For a compatible joint pattern $S = s, T = \delta$, write $n = su$. Absorb oddness, the full-K_o condition, and all small-prime avoidances at primes dividing K into a nonempty set $\mathcal{C}_{s,\delta}$ of residue classes for u mod $2K$. The remaining avoidance primes are coprime to $2K$. For a product e of those primes, each class contributes

$$\frac{L}{2Kse} + O(1).$$

Apply the same Bonferroni truncation in every class. Both the main term and its absolute error have the factor $|\mathcal{C}_{s,\delta}|$, which cancels in their ratio. The remaining Euler product is at least $c/\log(2z)$. Thus the relative error is at most

$$C\log(2z)\left(E_\nu + \frac{UKB_\nu}{L}\right).$$

The normalizing odd count $L/2 + O(1)$ adds only $O(1/L)$. This implies the deliberately weaker (C.14), including for very rare compatible patterns. There is no inverse of their ideal probability left over. Directly counting all relevant classes modulo K gives the asserted relative full-gcd error.

For the smooth numerator remove avoidance primes exceeding y, which smooth integers avoid automatically. Let S_0 be U or K, as appropriate. Choose an auxiliary divisor cutoff E_{cut}

with $\log(S_0E_{\mathrm{cut}}) = o(\log x)$. Applying (B.26) after division by se, iterating (B.27) in factors at most y, and integrating (B.28) gives

$$\#(\mathcal{S}_y \cap I/(se)) = L\widehat{\rho}(x,y)(se)^{-\beta(x,y)}(1+O(\eta_{\mathrm{an}})), \tag{C.18}$$

for $s \le S_0, e \le E_{\mathrm{cut}}$, where

$$\begin{aligned}\eta_{\mathrm{an}} \ll E_H &+ \frac{1+\log(S_0E_{\mathrm{cut}})/\log y}{u} \\ &+ \frac{(\log(S_0E_{\mathrm{cut}})+2\log y)^2}{\log x \log y} \\ &+ \left(1+\frac{\log(S_0E_{\mathrm{cut}})}{\log y}\right)\frac{\log y}{y}.\end{aligned} \tag{C.19}$$

All intermediate endpoints have $u(1+o(1))$; the direct short-interval range is required at these endpoints.

The absolute sum of the inclusion–exclusion coefficients divided by the signed main product is $\prod_p(1+p^{-\beta})/(1-p^{-\beta})$. For $p \le z$, comparison with the Mertens product bounds this by $C\log^2(2z)$, provided $\gamma\log z \le 1/8$. The primes of K exceeding z cost only a constant when $\omega(K)z^{-\beta+1/\log z} \le 1/2$. For the actual large-divisor counts use the global dilation estimate at the upper endpoint $x+L$, so that it bounds the entire interval after division by se. Returning from $x+L$ to x uses (B.27) and (B.28); the resulting error is within (C.19). Rankin's inequality at $1/\log z$ now gives

$$\eta_{\mathrm{FO},S_0} \ll \log^2(2z)\eta_{\mathrm{an}} + \frac{x}{L}\log^2(2z)\exp(-\log E_{\mathrm{cut}}/\log z). \tag{C.20}$$

Using $\log^4(2z)$ instead of $\log^2(2z)$ is a valid conservative bound. The explicit x/L is why the cutoff cannot simply be fixed at y^2.

Dividing the numerator by (C.17), or by the complete-K denominator, gives (C.15). An uncapped valuation requires avoidance of p, whereas a capped valuation does not. Oddness contributes $2(1-2^{-\beta})$. A prime $p > y$ may occur in the denominator but cannot occur in a smooth numerator, giving the stated zero branch. □

We shall also use the finite large-pattern estimate

$$\mathbb{P}_{n \text{ uniform in } I}\{S > U\} \le Cz^C U^{-1/\log z} + \frac{CUz}{|I|}. \tag{C.21}$$

Indeed, any z-smooth divisor exceeding U has a divisor in $(U, Uz]$. Count these divisors in I and apply Rankin's inequality to the main sum. The at most Uz endpoint errors give the second term. The denominator in this term is the interval length, not the location of the interval.

C.3 Higher moments and centering at the mean

Lemma C.5 (A uniform capped-valuation moment bound)**.** *Let the variables J_p, for odd $p \le z$, have the independent laws* (C.12)*, with arbitrary positive caps, and put $\mathcal{T}(\gamma) = \prod_{3\le p\le z} t_{p,K_p}(J_p;\gamma)$. There are absolute constants $c, C > 0$ such that, for every integer $m \ge 1$,*

$$0 \le m\gamma\log(2z) \le c \quad\Longrightarrow\quad \|\mathcal{T}(\gamma)-1\|_m \le Cm\gamma\log(2z), \qquad \mathbb{E}\mathcal{T}(\gamma)^{\pm m} \le C. \tag{C.22}$$

The norm bound also holds when γ depends on the complete outcome, provided $0 \le \gamma \le \gamma_{\max}$, with $\gamma_{\max}$ in its right side and hypothesis.

Proof. The score of one factor at zero is

$$L_{p,k}(j) = \begin{cases} (j - 1/(p-1)) \log p, & j < k, \\ k \log p, & j = k. \end{cases}$$

Since $\mathbb{E}J_p = (1 - p^{-k})/(p-1)$, this score has mean zero. For an absolute sufficiently small $c_0 > 0$, summing a geometric series gives, uniformly in the cap,

$$\mathbb{E}(J_p + 1/(p-1))^2 e^{c_0(J_p + 1/(p-1))} \le C/p.$$

The inequality $|e^x - 1 - x| \le x^2 e^{|x|}/2$ consequently proves

$$\mathbb{E}e^{tL_{p,k}} \le \exp(Ct^2(\log p)^2/p) \qquad (|t| \log p \le c_0).$$

With

$$L_z = \log(2z), \qquad W_z = L_z^{-1} \sum_{3 \le p \le z} L_{p,K_p},$$

independence and $\sum_{3 \le p \le z} (\log p)^2/p \ll L_z^2$ imply

$$\mathbb{E}e^{c_1|W_z|} \le C, \qquad \|W_z\|_r \le Cr \quad (r \ge 1). \tag{C.23}$$

Taylor expansion of the uncapped logarithm, and exact linearity at a capped value, give the pointwise identity

$$\log \mathcal{T}(\gamma) = \gamma L_z W_z + R_\gamma, \qquad |R_\gamma| \le C\gamma^2 L_z^2. \tag{C.24}$$

Apply $|e^v - 1| \le |v|e^{|v|}$, Hölder at order $2m$, and (C.23). The assumed smallness keeps the required exponential moment within its fixed range. This proves both claims for constant γ. For an outcome-dependent γ, use $|\log \mathcal{T}(\gamma)| \le \gamma_{\max} L_z |W_z| + C\gamma_{\max}^2 L_z^2$ before taking expectations. No independence from an interval location has been assumed. □

Lemma C.6 (Fixed moments for the full gcd)**.** *For a fixed integer $r \ge 1$, suppose $\gamma \log\log(3K) = o(1)$, $\omega(K)/y = o(1)$, and the errors in* (C.16) *are $o(1)$. Then, uniformly in every cap and divisor of K_o,*

$$\mathbb{E}_{\pi_K} \left(\frac{\mathbb{E}[f_y(n) \mid T]}{c_o(x,y)} \right)^r \le C_r. \tag{C.25}$$

Proof. For a smooth prime write

$$c_p = \frac{1 - p^{-1+\gamma}}{1 - p^{-1}}, \quad q_p = p^{-1+r\gamma}, \quad A_{p,r} = \frac{(1-p^{-1})c_p^r}{1 - p^{-1+r\gamma}}.$$

A finite geometric sum gives exactly

$$\mathbb{E}\, t_{p,k}^r = A_{p,r}(1 - q_p^k) + q_p^k.$$

For $r\gamma \le 1/2$, the value and first derivative of $\log A_{p,r}$ at zero vanish. Its second derivative gives $\log \mathbb{E}t_{p,k}^r \le C_r \gamma^2 (\log p)^2 p^{-1+r\gamma}$; also $\mathbb{E}t_{p,k}^r \le 1 + C_r p^{-1+r\gamma}$. Split at $Z = \max(e^8, (1 + \log K)^8)$. The sum below Z is $O_r(\gamma^2 \log^2 Z) = o(1)$, and the sum above Z is bounded by $C_r \omega(K) Z^{-1/2} = O_r(1)$. For $p > y$, the exact local moment is $(1 - 1/p)^{1-r}$, with total logarithmic cost $O_r(\omega(K)/y)$. Use (C.16) to finish. Only fixed orders, principally four, are used in this lemma. □

Lemma B.3 compares the actual conditional value n_P with a mixture of uniform odd integers on half-open intervals J' satisfying

$$J' \subset [-X_P, X_P], \qquad |J'| = \frac{X_P}{4(d+2)}, \qquad \epsilon_1 \le C\ell E^{-4}. \tag{C.26}$$

Proposition C.7 (Centering the proxy at its one-point mean)**.** *Fix $g \in \mathcal{G}$, an admissible label (ℓ, s), and $P \in \mathcal{P}_{\ell,s}$ with $R_P(g) = 1$. Then $D_P^U(n_P)$ satisfies* (C.10) *under the root-conditioned law of n_P.*

Proof. Let μ be the one-point model in (C.26). Its point masses are at most Cd/X_P. For every nonnegative function F of the retained pattern, the finite relative counts on the full support give

$$\mathbb{E}_\mu[F(S)1_{S\le U}] \le Cd(1+\eta_{\mathrm{ar}})\mathbb{E}_{\mathrm{CRT}}[F(S)1_{S\le U}]. \tag{C.27}$$

There is no factor for the number of patterns. Furthermore $T_x = \log(X_P/|n|)$ has $\mathbb{P}_\mu(T_x > t) \le \min(1, Cde^{-t})$: the possible single-integer error is absorbed while $e^{-t} \ge 1/X_P$, and the event is empty afterwards. It follows that

$$\|T_x\|_r \le \log(Cd) + Cr, \qquad \mathbb{E}_\mu e^{sT_x} \le (Cd)^s/(1-s) \quad (0 \le s < 1). \tag{C.28}$$

On a retained cell, (B.27) and (B.28) yield

$$\frac{c_o(x_I, y)}{c_P} = e^{(1-\beta(X_P,y))T_x}(1 + O_{C_X}(d^{-1})). \tag{C.29}$$

Here

$$0 \le T_x \le C_X \log y + O(\eta).$$

The change in the odd Euler factor is $O(C_X/(u\log y))$. Thus Lemma C.4 writes the very same fixed proxy, on the retained set, as

$$\begin{aligned} D_P^U(n)/c_P &= e^{(1-\beta(X_P,y))T_x}\mathcal{T}(1-\beta(x_I,y))(1+e_{I,S}), \\ &\qquad |e_{I,S}| \ll d^{-1} + \eta_{\mathrm{FO},U} + \eta_{\mathrm{ar}}. \end{aligned} \tag{C.30}$$

The interval and pattern need not be independent. Use the pointwise score bound of Lemma C.5 at order $4m$ inside (C.27); use (C.28) for the other factor and then Hölder. Writing $\gamma_{\max}$ for this layer's maximum $1-\beta$, we obtain

$$\begin{aligned} &\|D_P^U/c_P - 1\|_{m,\mu} \\ &\quad \ll (Cd)^{1/(4m)}\big\{m\gamma_{\max}\log(2z) + \gamma_{\max}\log(Cd) \\ &\qquad + d^{-1} + \eta_{\mathrm{FO},U} + \eta_{\mathrm{ar}}\big\} \\ &\qquad + \Big[Cd\{z^C U^{-1/\log z} + Uz/X_P + y^{-C_X}\}\Big]^{1/m}. \end{aligned} \tag{C.31}$$

Outside the retained set the normalized deviation is exactly -1. Transfer to the original distribution adds at most $2c_{\min}^{-1}\epsilon_1^{1/m}$, where $c_{\min} = \inf_P c_P$.

It remains to center at the conditional mean a_P. For each complete odd grid cell the exact definition gives

$$\sum_{n\in I,\ n \text{ odd}} (f_y(n) - D_I(n)) = 0. \tag{C.32}$$

The same identity holds separately for every pattern $S = s$. Within any interval J' of (C.26), the retained-pattern part of every complete retained grid cell therefore cancels; the contribution of the discarded patterns is bounded using (C.21) below. There are at most two partial endpoint cells; cutoff boundaries were rounded to grid boundaries. On a partial cell the absolute residual sum is bounded by twice the smooth count on its complete cell. By (B.26), its length is $O(\eta X_P)$ and its density is at most $D^{C_0} c_P$. The latter bound follows from

$$(1 - \beta(x, y)) \log y \le \log \log x + O(1) = O(\log D) \tag{C.33}$$

at this same layer. Equation (C.21) bounds the contribution of the discarded patterns; the remaining discarded values are controlled by the estimate for the mass near zero. After averaging the auxiliary mixture's weights and transferring the two bounded functions to the original distribution, this proves

$$\frac{|\mathbb{E} f_P - \mathbb{E} D_P^U|}{c_P} \ll d D^{C_0} \eta + \frac{C d \{ z^C U^{-1/\log z} + U z / X_P + y^{-C_X} \} + 2\epsilon_1}{c_{\min}}. \tag{C.34}$$

We check all growing parameters. Put

$$\log E_{\mathrm{cut}} = C_{\mathrm{cut}} \log(2z)(\Lambda + A \log D), \qquad \nu = \left\lceil \frac{\log Y}{\lambda} A^4 (\log D)^4 \right\rceil. \tag{C.35}$$

The constants $C_U, C_{\mathrm{cut}}, C_X$ are fixed before N. Uniformly over retained endpoints,

$$\log x \ge c d \Lambda (1 + \log Y / \Lambda), \quad cd \le u \le Cd, \quad 0 \le \gamma \le C/d, \quad c_{\min} \ge e^{-C_\rho \Lambda}. \tag{C.36}$$

The density bound follows from $\rho_D(u) \ge (2u)^{-2u}$ and (B.26). Fix this C_ρ first, then take $b_0 C_X > 2C_\rho + 20$ and take C_U, C_{cut} sufficiently large. All these choices are common to all levels. Increasing the single input threshold ensures

$$\log x_{-,F} \ge \log(HM) + (6d - 14) \log Y - C_X \log y - O(d) \ge 4d \log Y. \tag{C.37}$$

Thus every application of the cited estimate remains within its range of validity for this choice of C_X.

We have

$$\begin{aligned} 4m\gamma_{\max} \log(2z) &= D^{-1/3+2/48+o(1)} = o(1), \\ \eta_{\mathrm{FO},U} &\le D^{-1/3+6/48+o(1)} = D^{-5/24+o(1)}. \end{aligned} \tag{C.38}$$

The second estimate uses (C.19) and even the conservative fourth power of $\log(2z)$ in (C.20). The smallest retained cell has $\log L_{\min} \ge cd \log Y$, whereas

$$\log B_\nu = O((\log Y/\lambda) A^5 (\log D)^5) = o(d \log Y), \qquad -\log E_\nu \gg m C_\rho \Lambda + \log K. \tag{C.39}$$

Also $\log \eta^{-1} = o(\log y)$ and $\log(U E_{\mathrm{cut}}) = o(\log x)$, so every application of (B.26) is in range. The tails and the m-th root of ϵ_1 in (C.31) are negligible at the scale in (C.38). Equation (C.34) is at most $D^{-9A} + e^{-c\Lambda}$. Jensen and the triangle inequality now replace the reference center c_P by a_P, proving (C.10). □

C.4 The principal term and character sums of the conditional density

The principal character depends on the gcd with the full odd part of the modulus. We compare this gcd with the capped valuations used in the proxy.

Lemma C.8 (Conditional expectation and rare valuations)**.** *Let S, T, V_* be finite random variables, with V_* a function of each of S, T, and suppose S and T are conditionally independent given V_*. For $0 \le f \le 1$, set $D_X = \mathbb{E}[f \mid X]$. Then*

$$\mathbb{E}[f - D_S \mid T] = D_T - D_{V_*}. \tag{C.40}$$

Suppose that each V_*-fibre has a distinguished T-value of conditional probability at least $\pi_0 > 0$, the total probability of the other values is at most η_0, and $\mathbb{E}(D_T/c)^4 \le M_4$. Then

$$\mathbb{E}((D_T - D_{V_*})/c)^2 \le 4\sqrt{M_4\eta_0/\pi_0}. \tag{C.41}$$

Proof. Conditional independence gives $\mathbb{E}[D_S \mid T] = \mathbb{E}[D_S \mid V_*] = D_{V_*}$. On one V_*-fibre bound the variance of D_T/c by its mean squared distance from its distinguished value a_0. Only the exceptional event contributes. The term involving $(D_T/c)^2$ is at most $\sqrt{M_4\eta_0}$. If $q(V_*) \ge \pi_0$ is the distinguished probability, then $q(V_*)a_0^4 \le \mathbb{E}[(D_T/c)^4 \mid V_*]$. Cauchy–Schwarz bounds the remaining term by $\sqrt{M_4\eta_0/\pi_0}$. This also explains why equality on the distinguished event is not being asserted. □

For the ideal CRT law take $V_* = \gcd(n, \gcd(Q_o, K_o))$. Independence of prime valuations implies the conditional independence in Lemma C.8, including nested caps. The exceptional event consists of a positive valuation at a prime $p > z$ of K_o, or of a valuation beyond an insufficient small-prime cap. Since $p^{K_p} \ge R$,

$$\eta_0 \ll \frac{\log K}{z \log z} + \frac{\log\log(3z)}{R}, \qquad \pi_0 \gg 1/\log(2z). \tag{C.42}$$

The second summand counts the cap tail $p^{-K_p-1} \le 1/(pR)$. The distinguished probability is a product of $1 - 1/p$.

Lemma C.9 (Transfer of the principal residual)**.** *On a complete odd cell with actual uniform law P, let $D_I^U = \mathbb{E}_P[f1_{S\le U} \mid S]$. Under the finite counts of Lemma C.4,*

$$\mathbb{E}_{\pi_K}|\mathbb{E}_P[f - D_I^U \mid T]|^2 \ll c_I^2\sqrt{\eta_0/\pi_0} + \eta_{\mathrm{ar}} + \mathcal{B}_U(L) + K/L, \tag{C.43}$$

where $c_I = c_o(x, y)$ and $\mathcal{B}_U(L)$ denotes the right side of (C.21)*.*

Proof. Extend the ideal joint CRT law of (S, T) by retaining the actual conditional distribution of $f_U = f1_{S\le U}$ on each retained joint pattern; set $f_U = 0$ on discarded null modes. Call this law Q'. If p_{st}, q_{st} are the joint masses, (C.14) and the full-gcd marginal count give

$$p_{st} = (1 + O(\eta_{\mathrm{ar}}))q_{st} \quad (s \le U), \qquad p_T(t) = (1 + O(K/L))q_T(t).$$

Summing nonnegative retained numerators therefore gives

$$\mathbb{E}_{Q'}[f_U \mid T = t] \le (1 + O(\eta_{\mathrm{ar}} + K/L))\mathbb{E}_P[f \mid T = t]. \tag{C.44}$$

Lemma C.6 now transfers the fourth moment with an absolute factor. It is essential to do this relative comparison before using total variation; a fourth-power inverse density is not introduced.

For probability laws P, Q', a finite variable X, and $0 \le h \le 1$, direct summation on fibres gives

$$\mathbb{E}_P|\mathbb{E}_P(h \mid X) - \mathbb{E}_{Q'}(h \mid X)|^2 \le 4\,\mathrm{TV}(P, Q'). \tag{C.45}$$

The enlarged laws have total variation $O(\eta_{\mathrm{ar}} + \mathcal{B}_U(L))$. Writing $C_P^X h = \mathbb{E}_P(h \mid X)$, the difference

$$C_P^T(h - C_P^S h) - C_{Q'}^T(h - C_{Q'}^S h)$$

is the sum of

$$(C_P^T h - C_{Q'}^T h), \quad -C_P^T(C_P^S h - C_{Q'}^S h), \quad -(C_P^T C_{Q'}^S h - C_{Q'}^T C_{Q'}^S h).$$

Each has $L^2(P)$-norm at most twice the square root of that total variation, by (C.45) and conditional contraction. Apply Lemma C.8 under Q', use (C.44), and finally bound the contribution of the discarded $f - f_U$ in squared norm. Changing the T-marginal back to π_K costs its relative $O(K/L)$ error. This proves (C.43). □

The exact finite rectangular identity. Throughout this paragraph $K = 2^a K_o$, with $a \ge 1$ and K_o odd. Fix an integer ι with $\gcd(\iota, K) = 1$. Let I_1, I_2 be complete grid cells, of lengths L_1, L_2, and put

$$\mathscr{N}(I_1, I_2) = \#\{(n_1, n_2) \in I_1 \times I_2 : n_1, n_2 \text{ odd},\ n_1 \equiv \iota n_2 \pmod K\}.$$

For a real function v_j on the odd integers in I_j, define

$$\mathscr{V}_{j,\delta}(\chi) = \sum_{\substack{m \in I_j/\delta \cap \mathbb{Z} \\ (m, K/\delta) = 1}} v_j(\delta m)\chi(m), \qquad \delta \mid K_o.$$

The condition $(m, K/\delta) = 1$ includes oddness. Ordinary orthogonality on

$$(\mathbb{Z}/(K/\delta)\mathbb{Z})^\times$$

gives the finite identity

$$\begin{aligned} &\sum_{\substack{n_j \in I_j \text{ odd} \\ n_1 \equiv \iota n_2 \pmod K}} v_1(n_1)v_2(n_2) \\ &\quad = \sum_{\delta \mid K_o} \frac{1}{\varphi(K/\delta)} \sum_{\chi \bmod K/\delta} \overline{\chi(\iota)}\, \mathscr{V}_{1,\delta}(\chi)\overline{\mathscr{V}_{2,\delta}(\chi)}. \end{aligned} \tag{C.46}$$

Indeed the congruence forces $\gcd(n_1, K_o) = \gcd(n_2, K_o) = \delta$; after division by δ both coordinates are units modulo K/δ. This also handles negative cells, since Dirichlet characters are defined on negative integers and the tested functions remain real.

Counting one vertical progression for each odd first coordinate gives

$$\mathscr{N}(I_1, I_2) = \frac{L_1 L_2}{2K}\left(1 + O\left(\frac{K}{\min(L_1, L_2)}\right)\right). \tag{C.47}$$

The error is relative in the range used here. In the same way, if $N_{j,\delta} = \#\{n \in I_j : n \text{ odd}, \gcd(n, K_o) = \delta\}$, then

$$N_{j,\delta} = \frac{L_j \varphi(K/\delta)}{K}\left(1 + O(K/L_j)\right).$$

Consequently the principal-character part of (C.46), divided by $\mathscr{N}(I_1, I_2)$, has weights

$$\begin{aligned}\frac{N_{1,\delta}N_{2,\delta}}{\varphi(K/\delta)\mathscr{N}(I_1,I_2)} &= \frac{2\varphi(K/\delta)}{K}\left(1+O(K/\min(L_1,L_2))\right)\\ &= \pi_K(\delta)\left(1+O(K/\min(L_1,L_2))\right).\end{aligned}$$

Apply this identity with $v_j = f_y - D^U_{I_j}$. Thus the principal expression is the product of the two actual full-gcd conditional residual means, with the stated relative marginal error; no independence of the two residual functions has been assumed. Lemma C.9 and Cauchy–Schwarz now give (C.48).

The resulting relative principal error is

$$\epsilon_{\mathrm{pr}} \ll \sqrt{\log(2z)\eta_0} + \frac{\eta_{\mathrm{ar}} + \mathcal{B}_U(L) + K/L_{\min}}{c^2_{\min}}. \tag{C.48}$$

The proxy's nonprincipal characters require a different argument. Put $k_0 = K/\delta$. A compatible retained pattern $s = \gcd(\delta n, Q_o) \le U$ can be written as $s = s_0 a$, where $s_0 = \gcd(\delta, Q_o)$ and $\gcd(a, k_0) = 1$. After setting $n = au$, its remaining restrictions avoid small primes. In a Bonferroni truncation, each main term is a nonprincipal character sum over an interval. Its full k_0-periods sum to zero, and its endpoint contribution is $O(k_0)$. The pointwise difference between the upper and lower Bonferroni polynomials bounds the remainder in absolute value. Summing at most U patterns and using $\sum_{a\le U} 1/a \le 1 + \log U$ gives

$$|\widehat{D}^U_{I,\delta}(\chi)| \le U(k_0+1)B_\nu + \frac{L}{\delta}(1+\log U)E_\nu \qquad (\chi \ne \chi_0). \tag{C.49}$$

Since $|\widehat{f}_{I,\delta}| \le 2L/\delta$, the number of characters cancels the orthogonality factor $1/\varphi(k_0)$. Summing δ^{-2} shows that all fD, Df, DD nonprincipal terms, after density normalization, cost at most

$$CK\epsilon_D/c^2_{\min}, \qquad \epsilon_D = \frac{2UK^2B_\nu}{L_{\min}} + (1+\log U)E_\nu. \tag{C.50}$$

Only the original ff term will use smooth-character estimates.

C.5 Character sums associated with a common determinant

At each level we fix the following ranges of endpoints before drawing the environment and the form.

$$\begin{aligned}&\text{rational:} && x_{-,r} = M/(64Y_0^7), && x_{+,r} = 2MV,\\ &\text{algebraic:} && x_{-,F} = \frac{h_0 M\, 2^{-2d} Y^{6d-14} y^{-C_X}}{100(d+2)}, && x_{+,F} = 16(d+2)h_0 M Y^{6d}.\end{aligned}$$

They contain every positive endpoint after reflection, root division, content and induction divisors, and the odd-prefix division by two. Indeed the combined content–induction divisor is at most the corresponding modulus; on the rational side this is at most CV^2, and on the algebraic side it is at most $4Y^{12}$. The lower bound (C.37) applies to the algebraic envelope after increasing the common input threshold. These envelopes depend on the level and N, not on g, a root, a point pair, or a modulus selected later. For a primitive nonprincipal character of conductor r, define

$$u_y(r,\chi) = 4\sup_{x_-\le x\le x_+} \frac{|\sum_{n\le x,\ n\in\mathcal{S}_y}\chi(n)|}{\Psi(x,y)}, \qquad \mathfrak{e}(r) = \frac{1}{\varphi(r)}\sideset{}{^*}\sum_{\chi} u_y(r,\chi)^2. \tag{C.51}$$

The starred sum is over primitive nonprincipal characters only; in particular $\mathfrak{e}(1) = 0$. The fixed factor four covers interval differences and odd prefixes. Write $\mathfrak{a}(r)$ for $\mathfrak{e}(r)$ on the rational range with smoothness bound B, and $\mathfrak{b}(r)$ for its value on the algebraic range with bound y. Set both functions to zero for $r > K_{\max} = 4Y^{12}$.

The small-conductor input below is the one used in [33, `mf2L`, Section 4.4]. We verify its uniform range for the two ranges of endpoints using [24].

Lemma C.10 (A uniform small-conductor estimate)**.** *There exists an absolute X_0 such that, if $X \geq X_0$ and*

$$\sqrt{\log X} \leq \log y, \qquad y \leq X^{1/100}, \qquad u = \frac{\log X}{\log y} \geq (\log\log X)^5,$$

then, uniformly for $X^{1/3} \leq z \leq X$ and primitive nonprincipal characters χ with $1 < \operatorname{cond}(\chi) \leq e^{u^{3/4}}$,

$$\left| \sum_{\substack{n \leq z \\ P^+(n) \leq y}} \chi(n) \right| \leq e^{-u^{3/4}} \Psi(z, y).$$

The exponent $3/4$ is fixed independently of every parameter of the factoring experiment.

Proof. We use [24, (1.2), Lemmas 2.6–2.8, Propositions 2.9 and 2.13], including the proofs on pp. 609–618 and 626–628.

Fix $z \in [X^{1/3}, X]$, and write

$$L = \log z, \qquad Y = \log y, \qquad v = L/Y, \qquad \ell_3 = \log_3 z.$$

Then

$$u/3 \leq v \leq u, \qquad Y \geq \sqrt{L}, \qquad v \leq \sqrt{L}, \qquad Y \geq v, \qquad v \gg (\log L)^5.$$

Let $a > 0$ be the fixed small constant used in the proof of Proposition 2.13, and use exactly its height

$$T = \min\{y^{a/\ell_3}, e^{\sqrt{L}}\}, \qquad h = \log T = \min\{aY/\ell_3, \sqrt{L}\}.$$

Take $Q = \lfloor e^{u^{3/4}} \rfloor$. After increasing X_0,

$$h \geq av/\ell_3, \qquad \log Q = O(v^{3/4}) = o(h), \qquad \log(2QT) = (1 + o(1))h.$$

In particular $Q \leq y^{a/\ell_3} < y$.

Apply the zero-free statement (1.2) at modulus range Q and height $2T$, with its fixed constant $b > 0$. Suppose first that χ is not the possible exceptional character for this choice. Set

$$\epsilon = \frac{b}{2\log(2QT)} \asymp h^{-1}.$$

The closed region $\Re s \geq 1 - \epsilon$, $|\Im s| \leq 2T$ contains no zero of $L(s, \chi)$. We explicitly check the subsidiary conditions of Lemma 2.7, used in Proposition 2.9's proof:

$$\log(qT) \leq 2h \leq 2aY/\ell_3 = o(Y), \qquad \epsilon Y \asymp Y/h \geq \ell_3/a \longrightarrow \infty,$$

and

$$\frac{Y}{h^2} = \max\left\{\frac{\ell_3^2}{a^2 Y}, \frac{1}{v}\right\} = o(1), \qquad \frac{\log L}{h} = o(1).$$

Consequently $qT \le y^c$, $\epsilon \log y \ge 1/c$, and $T \ge y^{C\epsilon}(\log z)^2$ for the fixed constants required there. Also $y \ge (\log z)^C$ for every fixed C, since $Y \ge \sqrt{L}$.

Apply Proposition 2.9 with its source variables $x_{\mathrm{src}} = z_{\mathrm{src}} = z$ and additive twist zero. This does not enlarge its stated range $z_{\mathrm{src}} \in [x_{\mathrm{src}}^{2/3}, x_{\mathrm{src}}]$. The saddle-point estimate gives $z^\alpha \zeta(\alpha, y) \ll \Psi(z, y)L^2$, so

$$\frac{|\Psi(z, y; \chi)|}{\Psi(z, y)} \ll L^2\{he^{-cL/h} + e^{-ch}\}.$$

Since $L/h \ge \sqrt{L} \ge v$, $h \ge av/\ell_3$, and

$$u^{3/4} = o(v/\ell_3), \qquad \log L = o(v/\ell_3),$$

this is at most $e^{-u^{3/4}}$.

Suppose next that the exceptional primitive quadratic character χ_1 exists at height $2T$. Its real zero satisfies

$$\beta \ge 1 - \frac{b}{\log(2QT)} > 1 - \frac{b}{\log(QT)}.$$

Thus χ_1 is also the exceptional character at height T, by the uniqueness in (1.2), and is the character occurring in Proposition 2.13. Put $q_1 = \mathrm{cond}(\chi_1)$. The definition of W on p. 619 gives the exact identity

$$W(z, y; q_1, 0) = \frac{\tau(\chi_1)}{\varphi(q_1)}\Psi(z, y; \chi_1), \qquad |\tau(\chi_1)| = \sqrt{q_1}.$$

Hence its Gauss factor exactly cancels the prefactor $\sqrt{q_1}/\varphi(q_1)$ in Proposition 2.13. That proposition applies because

$$v \gg (\log L)^5, \qquad y \le \exp\{L/(\log L)^4\}, \qquad Q \le y^{a/\ell_3}.$$

Dropping the factor $z^{\beta-1} \le 1$, it gives

$$\frac{|\Psi(z, y; \chi_1)|}{\Psi(z, y)} \ll \exp\left\{-\frac{\delta v}{\log^2(v+1)}\right\} + e^{-aY/\ell_3} + e^{-a\sqrt{L}} + L^{3/2}e^{-aL/\log Q}.$$

Here

$$\frac{v}{\log^2(v+1)},\ \frac{Y}{\ell_3},\ \sqrt{L} \gg u^{3/4}, \qquad L/\log Q = vY/\log Q \gg v^{5/4} \gg u^{3/4}.$$

The polynomial factor in L is absorbed by $v \gg (\log L)^5$. Increasing the same threshold proves the claimed bound. In particular no exceptional character has been omitted and no Siegel lower bound on $1 - \beta$ is required.

All constants in the cited result were fixed before X. The argument is pointwise for every endpoint with the same bound and threshold, so taking the endpoint supremum costs no factor. The identity of the exceptional character need not remain fixed as z changes. □

The large-conductor argument follows [33, `mf2L`, Section 4.1, Lemma LV], with the choice $p = 12/5$. We need the bound uniformly over the upper endpoint of the character sum.

Lemma C.11 (Character sums with a variable upper endpoint)**.** *For conductors at most* $K_{\max} = 4Y^{12}$*,*

$$\sum_{r \ge R} \mathfrak{e}(r) \le C(1 + \log x_+)^7 R^{-2/3}. \tag{C.52}$$

Moreover

$$\sum_{r\ge 2}\sqrt{r}\,\mathfrak{a}(r)\le D^{C},\qquad \sum_{s\ge 2}\sqrt{s}\,\mathfrak{b}(s)\le D^{C}\epsilon_F^{1/24},\qquad \epsilon_F=\exp(-cD^{1/4}). \tag{C.53}$$

Proof. Proposition 2, (4.1), of [27] applies to one common complex sequence with coefficients of modulus at most one, $(\log x)^{21}\le y\le x^{1/100}$, $x^{-1/1000}<\delta\le 1$, and $1/x$-separated frequencies. Its large-value exponent is $2+20(1-\beta)+\varepsilon$. Take $\varepsilon=1/100$, $1-\beta\le 1/100$, and $p=12/5$. For reduced fractions with $R<r\le 2R$, of which there are at most $4R^2$, the stronger condition $8R^2\le x^{1/1000}$ gives the required separation and also controls the values below the source cutoff:

$$4R^2(x^{-1/1000})^p\le \tfrac{1}{2}x^{-7/5000}.$$

Layer cake gives a normalized additive p-energy $O(\log^4 x)$. Primitive Gauss sums and orthogonality give

$$\frac{1}{\varphi(r)}\sideset{}{^*}\sum_{\chi}|S(\chi)|^2\le\frac{1}{r}\sum_{\gcd(a,r)=1}|S(a/r)|^2.$$

Hölder over the fractions consequently gives a dyadic primitive energy

$$O(\log^4 x\,R^{-2/3}).$$

To bound the maximum of the partial sums, fix a dyadic endpoint range $X/2\le x\le X$. A prefix of $[1,\lfloor X\rfloor]$ is the disjoint union of at most one binary block at each of $O(\log X)$ levels. If $\mathcal{I}_{j,k}$ are all blocks at level j, Cauchy–Schwarz gives, for each character,

$$\sup_{x\le X}|S_x(\chi)|^2\le C\log(2X)\sum_j\sum_k|S_{\mathcal{I}_{j,k}}(\chi)|^2.$$

At one fixed level assign independent signs ϵ_k to its disjoint blocks. The identity

$$\mathbb{E}_\epsilon\left|\sum_k\epsilon_k S_{\mathcal{I}_{j,k}}(\chi)\right|^2=\sum_k|S_{\mathcal{I}_{j,k}}(\chi)|^2$$

holds simultaneously under the conductor sum. Every sign realization is one common bounded coefficient sequence, so the preceding energy bound applies before taking this expectation. The block count costs no additional factor. Summing the $O(\log X)$ levels and using $\Psi(X/2,y)\asymp\Psi(X,y)$ gives a sixth power of $\log(2X)$. Finally summing the $O(\log x_+)$ dyadic endpoint ranges proves (C.52).

Low conductors, including real characters. Apply Lemma C.10 to the algebraic range. Uniformly in the level, its endpoints satisfy

$$cd\le u(x,y)\le Cd,\qquad \log y\ge(b_0+o(1))d\lambda,\qquad \log\log x=O(\lambda).$$

Here and below the fixed constants may depend on q_* but not on the level. To check the size conditions directly, the raw construction gives

$$\log(HM)=(11+6/\kappa+o(1))d^2\lambda\le d\log Y,$$

and the displayed ranges of endpoints give

$$\log x_{+,F}\le 8d\log Y,\qquad \log x_{-,F}\ge 4d\log Y.$$

The root prime, modulus and parity divisions cost only $O(\log Y)$, and the fixed near-zero cutoff costs $C_X \log y = o(d \log Y)$. Thus $x_{-,F} \geq x_{+,F}^{1/2}$ eventually. Also

$$\frac{(\log y)^2}{\log x_{+,F}} \gg \lambda \longrightarrow \infty, \qquad u(x_{+,F}, y) \gg d \gg (\log\log x_{+,F})^5.$$

The lemma applies with $X = x_{+,F}$, with one common threshold. Taking the endpoint supremum is legitimate because the same bound holds pointwise throughout this predetermined interval.

Fix a sufficiently small $c > 0$ in $\epsilon_F = \exp(-cD^{1/4})$ and put $R_0 = \epsilon_F^{-1/4}$. Since $d \sim D^{1/3}$, this conductor range lies inside $\exp(u^{3/4})$ at every endpoint. Absorbing the fixed factor four in (C.51), we may use the conservative consequences

$$u_y(r, \chi) \ll r\epsilon_F, \qquad \mathfrak{b}(r) \ll r^2\epsilon_F^2 \quad (r \leq R_0). \tag{C.54}$$

In fact the lemma gives $u_y(r, \chi) \ll \epsilon_F$ in this range; the weaker displayed bound keeps the half-moment calculation uniform. The auxiliary random-sign sequences from the high-conductor argument are not used here. Odd prefixes are expressed as $S(x) - \chi(2)S(x/2)$, with both endpoints in the chosen range.

At $R_0 = \epsilon_F^{-1/4}$, the low half-moment is at most $C\epsilon_F^2 R_0^{7/2} = C\epsilon_F^{9/8}$. Partial summation in (C.52) bounds the high half-moment by $D^C R_0^{-1/6} = D^C \epsilon_F^{1/24}$. The rational half-moment follows directly from (C.52).

For completeness, all endpoint ranges are uniform. On the rational side $x_- \geq cM/Y_0^7$, and

$$\frac{2 \log K_{\max}}{\log x_-} \leq \frac{12/\kappa}{1 - 7/(2\kappa)} + o(1) < 10^{-3}.$$

On the algebraic side (C.37) gives $2 \log K_{\max}/\log x_- \leq (6 + o(1))/d$. Thus the small-value condition used above holds on both sides. Also $\log y \gg \log\log x$, so every fixed logarithmic lower range in the low-conductor theorem holds; in particular $y \geq (\log x)^{21}$. We also have $y \leq x^{1/100}$, and $\log R_0 = o(\log x_-)$. □

To describe imprimitive induction, let $U_h(k, \chi)$ be an ordinary smooth character sum on an interval after a content h has been removed. Normalize it by $W_h = \Psi(x, y)h^{-\beta_*}$, where β_* is the minimum saddle over the chosen range of endpoints. Let $\mathcal{A}_h(k)$ be a common deterministic majorant of

$$\frac{1}{\varphi(k)} \sum_{\substack{\chi \bmod k \\ \chi \neq \chi_0}} |U_h(k, \chi)/W_h|^2$$

over every permitted interval and divided endpoint; $\mathcal{B}_\delta(k)$ denotes its algebraic counterpart. The bounds on the primitive character sums are fixed before choosing g. Dilation gives the required bounds at every divided endpoint. If χ is induced from $\chi^* \pmod r$, then exactly

$$U_h(k, \chi) = \sum_{\substack{e \mid \mathrm{rad}(k) \\ (e,r)=1}} \mu(e)\chi^*(e) U_{he}(r, \chi^*).$$

Weighted Cauchy–Schwarz with weights e^{-1} gives the energy majorant

$$\mathcal{A}_h(k) \leq C\mathcal{L}_K \sum_{er \mid k} e^{1-2\beta_*} \frac{r}{k}\, \mathfrak{a}(r), \qquad \mathcal{L}_K = \max_{k \leq K} \frac{k}{\varphi(k)} \prod_{p \mid k} (1 + 1/p) \ll \log^2(2K). \tag{C.55}$$

Every surviving rational term has $\ell \nmid he$. We retain the factor r/k in (C.55) for the sum over determinants in Lemma C.12.

Odd primitive prefixes and imprimitive induction. The parity restriction must be retained before expanding an imprimitive character. After reflection, let I be a positive interval, let h be odd, and let k be even. For a primitive nonprincipal character χ^* (mod r) inducing χ (mod k), put

$$U_{he}^{\mathrm{odd}}(r,\chi^*)=\sum_{\substack{n\in I/(he)\cap\mathbb{Z}\\ n\ \text{odd}}}1_{\{hen\in\mathcal{S}_y\}}\chi^*(n).$$

Then the exact induction identity is

$$U_h(k,\chi)=\sum_{\substack{e\mid \mathrm{rad}(k_{\mathrm{odd}})\\ (e,r)=1}}\mu(e)\chi^*(e)U_{he}^{\mathrm{odd}}(r,\chi^*). \tag{C.56}$$

Indeed, oddness has already imposed the avoidance of the prime two; Möbius inversion is applied only to the remaining primes of k not dividing r. Unsupported terms, including nonsmooth he, are zero before majorization. Negative intervals change these formulas only by a character value of modulus one.

If $S(x,\chi^*)=\sum_{n\le x,\ n\in\mathcal{S}_y}\chi^*(n)$, then

$$S^{\mathrm{odd}}(x,\chi^*)=S(x,\chi^*)-\chi^*(2)S(x/2,\chi^*).$$

The identity also holds when r is even, since then $\chi^*(2)=0$. The chosen range includes both divided endpoints and their halves. The factor four in (C.51) thus covers the interval and parity differences, including for conductors that are powers of two.

With $W_h=\Psi(x,y)h^{-\beta_*}$, the global dilation bound therefore gives

$$|U_{he}^{\mathrm{odd}}(r,\chi^*)|/W_h\le Ce^{-\beta_*}u_y(r,\chi^*).$$

Weighted Cauchy–Schwarz with weights e^{-1}, followed by

$$\frac{\varphi(r)}{\varphi(k)}\le\frac{r}{k}\frac{k}{\varphi(k)},$$

gives the deterministic algebraic energy majorant

$$\mathcal{B}_h(k)\le C\mathcal{L}_K\sum_{\substack{fr\mid k\\ f\ \text{odd}}}f^{1-2\beta_*}\frac{r}{k}\,\mathfrak{b}(r). \tag{C.57}$$

The support restrictions may be retained in this nonnegative sum. The rational energy uses the ordinary induction identity and keeps $\ell\nmid he$. The factor r/k is retained on both sides.

Lemma C.12 (Common-determinant convolution)**.** *Suppose the two saddles are bounded below by a fixed number greater than $3/4$. Put $M_a=\sum\sqrt{r}\,\mathfrak{a}(r)$ and $M_b=\sum\sqrt{s}\,\mathfrak{b}(s)$. For a complete prime band $J\subset(U_0,2U_0]$, retaining the rational support restrictions in* (C.55)*,*

$$\begin{aligned}&\frac{1}{|J|}\sum_{\ell\in J}\ell\sum_{t\le T_0}\sum_{\substack{h\mid t,\ \delta\mid t\\ \delta\ \text{odd}}}t\,h^{-2\beta_r}\delta^{-2\beta_F}\mathcal{A}_h(\ell t/h)\mathcal{B}_\delta(2t/\delta)\\ &\quad\ll\mathcal{L}_K^2(1+\log T_0)\left\{M_aM_b+\frac{\sqrt{2U_0}}{|J|}M_b\sum_{r\ge U_0}\sqrt{r}\,\mathfrak{a}(r)\omega_J(r)\right\}.\end{aligned} \tag{C.58}$$

Here $\omega_J(r)$ counts prime divisors in J, and the algebraic modulus is the actual even modulus $2t/\delta$. Under the additional condition $(1-\beta_r)\log\log(3T_0)=O(1)$, which holds at the present parameters, the rational-principal companion is at most $D^C\sum_s\mathfrak{b}(s)$.

Proof. Take $\mathcal{L}_K$ with the global ceiling $K = K_{\max} = 4Y^{12}$, covering both rational and algebraic moduli. Expand the rational energy by (C.55) and the algebraic energy by (C.57). Put

$$H_1 = he, \qquad H_2 = \delta f, \qquad r' = r/\gcd(r,\ell), \qquad s' = s/\gcd(s,2).$$

Rational support gives $\ell \nmid H_1$, and both δ and f are odd. The divisibility constraints imply $H_1 r' \mid t$ and $H_2 s' \mid t$. After the two expansions, including the leading factor ℓ, the coefficient is exactly

$$\frac{1}{2} H_1^{1-2\beta_r} H_2^{1-2\beta_F} \frac{rs}{t}\, \mathfrak{a}(r)\mathfrak{b}(s).$$

For $L = \operatorname{lcm}(H_1 r', H_2 s')$,

$$\frac{rs}{2L} \le \frac{\sqrt{rs}\sqrt{\gcd(r,\ell)\gcd(s,2)}}{2\sqrt{H_1 H_2}} \le \frac{\sqrt{rs}\sqrt{\gcd(r,\ell)}}{\sqrt{H_1 H_2}}. \tag{C.59}$$

Now $\sum_{t \le T_0,\, L \mid t} t^{-1} \le (1+\log T_0)/L$. The possible factorizations of H_1, H_2 cost at most their respective divisor functions, and

$$\sum_{H \ge 1} \tau(H) H^{1/2-2\beta} < \infty \qquad (\beta > 3/4).$$

This proves (C.58), retaining the factor $\sqrt{\ell}$ when $\ell \mid r$. Only one factor of ℓ, and only one factor of two, has been removed from the corresponding conductor; higher prime powers remain covered. By (C.52), $|J| \gg U_0/\log U_0$, and $\omega_J(r) \le \log K/\log U_0$, the exceptional term is at most $D^C M_b U_0^{-2/3}$.

For the rational-principal term use $\ell t/\varphi(\ell t/h) \le h \max_{k \le K} k/\varphi(k)$. After writing $t = \delta v$, the constraint is $f s' \mid v$. The extra factor $s/(2s') \le 1$ accounts for the change in parity. This leaves a harmonic sum and convergent sums of $\delta^{-2\beta_F}$ and $f^{-2\beta_F}$. The remaining $\max_{t \le T_0} \sum_{h \mid t} h^{1-2\beta_r}$ is a fixed power of $\log T_0$: split its Euler product at $(\log(3T_0))^2$ and use $(1-\beta_r)\log\log(3T_0) = O(1)$. This proves the companion assertion. □

C.6 Proof of the concentration estimate

For every fixed environment and label, the point densities c_P are within fixed factors of one common $c_F(g,\ell,s)$. Indeed their scales differ by $2^{O(d)}$ and $\gamma_F d = O(1)$. Proposition C.7 and (C.5) imply

$$\tau \asymp \frac{n_0 \rho_r c_F(g,\ell,s)}{\log B}, \qquad \tau \ge e^{c\Lambda}. \tag{C.60}$$

To prove the second assertion, use the density at level i. If $\log Y = \xi\Lambda$, the bound on the norm is $\log x \le (12 + 6\xi + o(1+\xi))d\Lambda$. The Dickman lower bound and $\log y \ge b_0\Lambda$ give $\log c_F^{-1} \le ((12+6\xi)/b_0 + o(1+\xi))\Lambda$. Together with $\rho_r \ge e^{-(1+o(1))\Lambda}$ and $n_0 \ge Y^{11}$, this proves (C.60).

The prescribed lower bound on the density $\tau/|\Omega_{\ell,s}| \ge 2^{-(U_*+10)dm_b}$ follows from Lemma B.14, with the actual cutoff recursion and denominator-roughness factor.

For distinct rational points set $\Delta = |\det(P_1,P_2)| = \ell t$. Lemma B.4 gives, with error $\epsilon_{\mathrm{TP}} \le e^{-cD^{2/3}\lambda}$, a mixture of uniform windows in the odd lattice with $K = 2t$. Its covolume is $2K = 4t$. For each translate, if L_2 is the longest vertical chord and $\mathcal{A}$ its area, then

$$\begin{gathered} \mathcal{A} \gg s_1 s_2/\vartheta, \quad L_2 \gg s_2/\vartheta, \quad K/L_2 \ll 2^d V^{2-d}/(HM), \\ s_\nu = X_{P_\nu}/(8(d+2)), \quad \vartheta = b_1 b_2/\Delta. \end{gathered} \tag{C.61}$$

The geometry (C.61), its transpose, and $w_j \asymp s_j$ verify the two inequalities of Lemma B.8, with exponentially small relative errors. We use its actual lattice denominator (B.40), coordinate marginals (B.41), and strip bound (B.42). In particular the latter is $Cd(y^{-C_X} + X_j^{-1})$, uniformly in the angle.

Nonprincipal algebraic character sums. Fix a nonthin pair, and cover the interior and boundary of each comparison window by complete grid rectangles. This restriction is used in the total-area bound below; thin pairs are treated separately in (C.67)–(C.68). In (C.46) take $v_1 = v_2 = f_y$. For the cell I_j, let x_j be its positive location after reflection, and use $W_{j,\delta} = \Psi(x_j, y)\delta^{-\beta_F}$, where β_F is the minimum saddle over the algebraic range of endpoints. Changing x_j to the other endpoint costs an absolute factor. The common bound for primitive-character partial sums, dilation, and (C.55) give, for $k = K/\delta$,

$$\frac{1}{\varphi(k)} \sum_{\chi \neq \chi_0} \frac{|\mathscr{V}_{j,\delta}(\chi)|^2}{W_{j,\delta}^2} \leq C\mathcal{B}_\delta(k).$$

Unsupported terms, including nonsmooth δ, are zero before the majorant is applied. Cauchy–Schwarz in the character sum therefore bounds its nonprincipal part by

$$C \sum_{\delta | K_o} W_{1,\delta} W_{2,\delta} \mathcal{B}_\delta(K/\delta).$$

Since

$$L_j \asymp \eta x_j, \qquad \Psi(x_j, y) \asymp x_j c_o(x_j, y),$$

and every retained cell satisfies $c_o(x_j, y) \leq D^C c_F(g, \ell, s)$, division by (C.47) yields

$$D^C \eta^{-2} c_F(g, \ell, s)^2 K \sum_{\delta | K_o} \delta^{-2\beta_F} \mathcal{B}_\delta(K/\delta).$$

For the two-point model $K = 2t$, so K_o is the odd part of t and this is (C.62), up to an absolute factor. To sum grid rectangles, use their original lattice counts as weights. The complete rectangles cover a set of area $O(\mathcal{A})$, and the relative errors in (B.40) and (C.47) give bounded total normalized weight. Only the ff nonprincipal term uses this smooth-character estimate. The terms containing a proxy retain their separate finite estimate (C.50).

Thus the original nonprincipal term is bounded by

$$D^C \eta^{-2} c_F(g, \ell, s)^2 t \sum_{\substack{\delta | t \\ \delta \text{ odd}}} \delta^{-2\beta_F} \mathcal{B}_\delta(2t/\delta). \tag{C.62}$$

The rational expansion and the order of averaging. Let $J_r = [X_r/2, X_r]$, fix a primitive ordered pair with $\Delta = |\det(P_1, P_2)| = \ell t$, and let $\iota_r \in (\mathbb{Z}/\Delta\mathbb{Z})^\times$ be its rational lattice multiplier. Define

$$\mathscr{U}_h(\chi) = \sum_{\substack{m \in J_r/h \cap \mathbb{Z} \\ (m, \Delta/h) = 1}} 1_{\{hm \in \mathcal{S}_B\}} \chi(m).$$

The exact rational counterpart of (C.46) is

$$\begin{aligned} &\#\{n_1, n_2 \in J_r \cap \mathcal{S}_B : n_2 \equiv \iota_r n_1 \pmod{\Delta}\} \\ &\quad = \sum_{h | \Delta} \frac{1}{\varphi(\Delta/h)} \sum_{\chi \bmod \Delta/h} \chi(\iota_r) |\mathscr{U}_h(\chi)|^2. \end{aligned}$$

The right side is real although its individual summands need not be. Every surviving h is B-smooth; since $\ell > B$, it satisfies $\ell \nmid h$ and hence $h \mid t$. After taking absolute values of the nonprincipal contribution, the rational principal term and (C.55) bound this count by

$$C\Psi(X_r, B)^2 \sum_{\substack{h \mid t \\ \ell \nmid h}} h^{-2\beta_r} \left\{ \frac{1}{\varphi(\ell t/h)} + \mathcal{A}_h(\ell t/h) \right\}.$$

Here $\Psi(X_r, B) \asymp X_r \rho_r$, by the fixed-ratio smooth estimate. The map from integer environments to this rational lattice is injective and onto the full index-Δ lattice. Enlarge the original prime Q-set to integer Q, and enlarge the image of the environment rectangle to J_r^2. These are positive upper bounds; the environment normalization is $\asymp M^2/\log M$. Because $X_r \asymp MV$, this proves (C.63), including its factor $\rho_r^2 V^2$. The condition $P^-(b) > B$ is a fixed point indicator here and can only decrease the count. We obtain

$$\mathbb{E}_g R_{P_1} R_{P_2} \ll \log M \, \rho_r^2 V^2 \sum_{h \mid t} h^{-2\beta_r} \left\{ \frac{D^C}{\varphi(\ell t/h)} + \mathcal{A}_h(\ell t/h) \right\}. \tag{C.63}$$

The preceding rational bound must be applied after the following pointwise cancellation. On $\mathcal{G}$,

$$\frac{c_F(g, \ell, s)^2}{\tau(g, \ell, s)^2} \ll \frac{(\log B)^2}{n_0^2 \rho_r^2}, \qquad n_0 = V^2/\ell.$$

Thus the right side is independent of g and s, and the remaining environment test is the nonnegative indicator $R_{P_1} R_{P_2}$. Removing the restriction to $\mathcal{G}$ and $Q \nmid N$ from this nonnegative sum costs an absolute factor, because the conditioning event has probability bounded below. For the forbidden root $s = rQ^{-1} \bmod \ell$, a point in its lattice has $\ell \mid rb - Qa$, so no positive B-smooth rational value survives. We may therefore extend to all $s \bmod \ell$ and replace $1/(\ell - 1)$ by at most $2/\ell$.

For a fixed s and fixed nonzero signed determinant ℓt, there are $O(n_0)$ primitive ordered pairs. Indeed there are $O(V^2/\ell + V) = O(n_0)$ first points. For a primitive first point, all integral solutions for the second point form $P_2^{(0)} + \mathbb{Z}P_1$; the second-coordinate step exceeds $V/2$, whereas the admitted second-coordinate interval has length $V/4$. There are at most a constant number of admitted second points. The two determinant signs cost only a factor two.

Combining this count with the algebraic factor t, the rational factor $\rho_r^2 V^2$, the normalization above, and the average $\ell^{-1} \sum_s$, gives exactly

$$\frac{1}{n_0^2 \rho_r^2} \, (Cn_0) \, (\rho_r^2 V^2) \, t = C\ell t.$$

Consequently the averaged nonprincipal ff term is bounded by

$$\frac{D^C \eta^{-2}}{|J|} \sum_{\ell \in J} \ell \sum_{t \le V^2/\ell} \sum_{\substack{h \mid t, \ \ell \nmid h \\ \delta \mid t, \ \delta \text{ odd}}} t h^{-2\beta_r} \delta^{-2\beta_F} \mathcal{B}_\delta(2t/\delta) \times \left\{ \frac{D^C}{\varphi(\ell t/h)} + \mathcal{A}_h(\ell t/h) \right\}.$$

Deleting the primes dividing N has changed the complete-band normalization by $1 + o(1)$. All terms in this last display are nonnegative. Since the actual coordinates satisfy $0 < a < b < V$,

their determinants have absolute value less than V^2. For $J \subset (U_0, 2U_0]$, enlarge the t-range to $T_0 = \lfloor V^2/U_0 \rfloor$ before applying Lemma C.12 and its principal companion. Even after this enlargement, $\ell t \le 2V^2$ and $2t \le 2V^2$, so both moduli stay below the cutoff $4Y^{12}$. The rational induction retains the zero-support restriction $\ell \nmid he$; the algebraic modulus retains its factor two. Thus the two conductor sums refer to the same determinant. Averaging over the prescribed band gives

$$\mathbb{E}_{\text{base}} \frac{\text{nonprincipal } ff \text{ contribution}}{\tau^2} \ll D^C \eta^{-2} \epsilon_F^{1/24}. \tag{C.64}$$

The principal character and the window boundary. For every retained complete cell, the pointwise density comparison is

$$c_I \le D^C c_P.$$

On a complete rectangle, the exact character identity, its finite denominator, and the finite-projection estimate give for the principal residual product

$$|\mathrm{Princ}((f-D)(f-D))| \le C\epsilon_{\mathrm{pr}} c_{I_1} c_{I_2} \le D^C \epsilon_{\mathrm{pr}} c_{P_1} c_{P_2}.$$

The complete rectangle cover has bounded total normalized lattice weight. It follows that the same bound holds for the whole interior. No archimedean higher-moment transfer to an expanded cover is needed.

For a nonthin pair, scale the coordinates by s_1, s_2. The window has bounded projection widths and area $\asymp 1/\vartheta$. Each retained grid rectangle has coordinate mesh at most $Cd\eta$. The relative boundary area is therefore $O(d\eta\vartheta) = O(D^{-10A})$. The rectangle counts have relative error $O(K/L_{\min})$; summing them with their own areas and using the finite window denominator gives total boundary lattice weight

$$O(D^{-10A} + e_{\text{geom}}), \qquad e_{\text{geom}} \le \exp(-cD^{2/3}\lambda).$$

There is no extra factor for the number of rectangles.

On the boundary use the positive majorant $ff + fD + Df + DD$. For f, the full-gcd fourth moment bounds the principal products. For D, its retained finite product formula first gives $\mathbb{E}(D/c_I)^4 \le C$; conditional Jensen and the relative full-gcd marginal count then give the same bound after conditioning on T. Cauchy–Schwarz therefore bounds every positive principal product on a complete rectangle by $Cc_{I_1}c_{I_2}$. Multiplication by the boundary lattice weight gives (C.66) below. Boundary nonprincipal terms remain covered by the smooth-character and finite-proxy estimates.

At each fixed environment,

$$\sum_{P,Q} R_P R_Q c_P c_Q = \left(\sum_P R_P c_P \right)^2 \ll \tau^2. \tag{C.65}$$

Thus the normalized principal and boundary bounds retain both smooth densities. In particular the boundary allowance is

$$D^C(D^{-10A} + e_{\text{geom}}) c_{P_1} c_{P_2} \le D^{-5A+O(1)} c_{P_1} c_{P_2} + e^{-cD^{2/3}\lambda} c_{P_1} c_{P_2}. \tag{C.66}$$

For thin pairs apply the ordinary branch of Corollary B.10 with root $q = \ell$. Here $B < \ell \le Y_i$, and rational eligibility gives $\ell \nmid bG(P)N$, exactly the root range and unit conditions in that corollary. Its point density is $\widehat{\rho}(2X_P, y)$, which is within fixed factors of $c_P = c_o(X_P, y)$ by

(B.27) with factor two and (C.6). Enlarging the common fixed C_X if necessary, the corollary absorbs its strip and total variation errors after division by these two densities. It therefore gives, for every distinct pair, including all thin determinants,

$$\mathbb{E}[f_{P_1} f_{P_2} \mid g, R_{\ell,s}] \le D^C c_{P_1} c_{P_2}. \tag{C.67}$$

The retained proxy has $D_P^U \le c_P \exp(CA(\log D)^2)$. Hence the mixed terms use one actual one-point mean and one supremum of the conditional density, and the DD term uses the two suprema. Combining these estimates with (C.5) gives the normalized thin contribution

$$\frac{D^C + \exp(CA(\log D)^2)}{\Theta} + e^{-c\Lambda}. \tag{C.68}$$

Both rational and both algebraic densities have been retained.

Completion of Theorem C.3. The growing-moment assertion was proved in Proposition C.7. In the signed residual $(Z - \widetilde{Z})^2$, use (C.48), (C.50), (C.64), (C.66), and (C.68). The diagonal costs $O(\tau^{-1})$, since $0 \le D_P^U \le 1$ and the true-mean moment bound implies $\mathbb{E}(f_P - D_P^U)^2 \ll a_P$. Near zero we retain the original indicator and use the strip probability. For the total variation error we use that the pair test is bounded by one. Take $K_* = 4Y^{12}$, take ϵ_{pr} and ϵ_D uniformly over the admitted cells and moduli, and put

$$\epsilon_{\mathrm{strip}} = Cd\{y^{-C_X} + X_{\min}^{-1}\}, \qquad X_{\min} = \inf_P X_P, \qquad \tau_{\min} = \inf_{g,\ell,s} \tau(g,\ell,s).$$

The estimates just proved give

$$\begin{aligned} \mathbb{E}_{\mathrm{base}} \frac{(Z - \widetilde{Z})^2}{\tau^2} \ll{}& \tau_{\min}^{-1} + D^C(\epsilon_{\mathrm{pr}} + D^{-10A} + e_{\mathrm{geom}}) \\ &+ c_{\min}^{-2}(K_* \epsilon_D + \epsilon_{\mathrm{strip}} + \epsilon_{\mathrm{TP}}) \\ &+ D^C \eta^{-2} \epsilon_F^{1/24} + \frac{D^C + \exp(CA(\log D)^2)}{\Theta}. \end{aligned} \tag{C.69}$$

In particular, an absolute bounded-test error is multiplied only by $(\#\{P : R_P = 1\})^2/\tau^2 \ll c_{\min}^{-2}$. Terms already containing $c_P c_Q$ instead use (C.65); neither step introduces an additional inverse rational density.

The leading principal error is

$$\sqrt{\log(2z)\left\{\frac{\log K}{z \log z} + \frac{\log\log(3z)}{R}\right\}} = D^{-4A+O(1)}.$$

The finite errors in (C.39), including the explicit K in (C.50), are exponentially smaller. The near-zero error divided by $c_{\min}^2$ is at most

$$\mathrm{poly}(D) e^{-(b_0 C_X - 2C_\rho + o(1))\Lambda}.$$

The same normalization leaves

$$\epsilon_{\mathrm{TP}} \le e^{-cD^{2/3}\lambda}.$$

The boundary error is $D^{-5A+O(1)}$, and the thin and ff errors are smaller because

$$D^{1/4} \gg D^{1/6}\lambda^2 \gg A(\log D)^2.$$

This proves the first inequality of (C.11).

For each fixed environment and label, Minkowski gives

$$\|\widetilde{Z}/\tau - 1\|_m \le D^{-5/24+o(1)}.$$

Markov's inequality at threshold $1/4$ bounds its failure probability by D^{-10A}. The event $|Z-\tau| > \tau/2$ implies this proxy event or $|Z-\widetilde{Z}| > \tau/4$. Chebyshev bounds the second probability by $16D^{-4A+O(1)}$. Since $A \to \infty$, their sum is at most D^{-3A} for the same common sufficiently large input threshold. □

D Relation trees and initial factorizations

We first derive Proposition A.1 from the arithmetic estimates. We then show that a sampled pool works for a positive fraction of the top primes and count the corresponding shortest vectors. The set of primes may depend on the pool. The pool estimate adapts [33, `escapeR2`, Theorem NR-q, proof step 1] and the shortest-vector count adapts [33, `mfU_theorem`, Lemma S″].

D.1 Proof of the arithmetic bounds

Proof of Proposition A.1. Proposition B.1 gives the common environment and counts rational candidates both for each top prime and in aggregate. The true-mean and incoming-profile calculation is Theorem B.17, with the top case stated in Proposition B.13. The exact measure comparison is Proposition B.16. The ordinary tail is Theorem C.3. Its baseline probability is at most D^{-3A}. Multiplying by the comparison density CL_i^2, and using $A \gg i+1$, gives (iii).

To bound the mean at level i, use the density for that same level. Put $\Lambda_b = (\log 2)dm_b$ and write

$$\xi_i = \frac{\log Y_i}{\Lambda_b}, \qquad c_i^+ = \min\left\{U_*, \frac{12+6\xi_i}{b_0}\right\}.$$

The coefficient-height bound and the actual integer cutoff recursion give $u = \log x/\log y_i \le c_i^+ d$ for every algebraic density used here; division by a root or parent prime only decreases this upper bound. Since $c_i^+ \le U_*$ is uniformly bounded and $\log d/\lambda \to 1/3$, the elementary Dickman inequality gives

$$-\log\widehat{\rho}(x,y_i) \le 2u\log(2u) + o(\Lambda_b) \le \left(\frac{2}{3}c_i^+ + o(1)\right)\Lambda_b,$$

with a single error term uniform in the level and label. The same-level rational count, $\rho_{r,i} \ge e^{-\Lambda_b}$, and $\ell \le Y_i$ therefore give the stronger bound

$$\log\tau_l \ge \left(11\xi_i - 1 - \frac{2}{3}c_i^+ - o(1)\right)\Lambda_b.$$

In particular,

$$\log\tau_l \ge \left(11\xi_i - 1 - \frac{12+6\xi_i}{b_0} - o(1)\right)\Lambda_b.$$

For a top label the analogous estimate follows from its individual rational count, with $\log p = \log Y_0 + O(1)$; the same conservative display suffices. Since $\xi_i \ge b_0 = 100$, the coefficient is bounded away from zero uniformly in i.

Since $c_i^+ \le U_*$, the same calculation gives $-\log c_P \le (2U_*/3 + o(1))\Lambda_b$. The rational density and the denominator-roughness loss therefore leave the stated slack in $p_{\mathrm{den}} = 2^{-(U_*+10)dm_b}$. At the top, the count before rejection cannot exceed its rational whole-box count; dividing that count by the actual top mean proves (iv). □

D.2 The initial relation counts

Proof. The same incidence calculation at the top, with $W_0 = 1$, gives

$$\mathbb{E}_g \operatorname{avg}_l\left(1 - \frac{\mathbb{E}_{\nu_g} D_l}{\tau_l}\right) \le M_0 K_0\sqrt{\delta_1} = \theta_N.$$

All averages in this proof use the original reference law conditioned only on the common environment event. Let A_a be one when raw sample a is accepted *or yields a verified factor during initialization*, and zero only on a rejection without finding a factor. For each parent l define

$$T_l = \frac{1}{J}\sum_{a=1}^{J} \frac{Z_{l,a}}{\tau_l}, \qquad L_l = \frac{1}{J}\sum_{a=1}^{J} \frac{(Z_{l,a} - D_{l,a}) + (1 - A_a)Z_{l,a}}{\tau_l}.$$

For every fixed g, l, the normalized count before rejection has mean one and second moment at most U. Conditional independence of the slots therefore gives $\mathrm{Var}(T_l \mid g) \le U/J$. Chebyshev's inequality implies that the expected fraction of labels with $|T_l - 1| > 1/4$ is at most $16U/J$. Markov's inequality makes this fraction at most $1/8$, except with probability $128U/J$.

The nonnegative losses satisfy

$$\mathbb{E}\,\mathrm{avg}_l\, L_l \le \theta_N + U\rho.$$

Except with probability $32(\theta_N + U\rho)$, their average is at most $1/32$, and hence at most $1/8$ of the labels have $L_l > 1/4$. On the intersection, at least three quarters of the labels have $T_l \in [3/4, 5/4]$ and $L_l \le 1/4$. For each such label,

$$\frac{1}{J\tau_l}\sum_a A_a D_{l,a} \ge T_l - L_l \ge \tfrac{1}{2}, \qquad \frac{1}{J\tau_l}\sum_a A_a Z_{l,a} \le T_l \le \tfrac{5}{4}.$$

For the lower bound, each rejected field contributes at most zero to $T_l - L_l$, and an accepted field contributes $D_{l,a}/(J\tau_l)$. This accounts for the dependence of sample acceptance on its count.

Draw all ideal trials before inspecting them. If a slot produces a verified factor, the epoch has already succeeded. On every other outcome, A_a equals the actual acceptance indicator on each slot, so the displayed sums count the actual accepted pool. Adding ε_G proves (4.6). □

D.3 Bounds for the length of a search

For a candidate-density lower bound $0 < p \le 1$ and a desired failure probability $0 < \varepsilon < 1$, define

$$Q(\varepsilon) = \lceil 128 \log(4/\varepsilon) \rceil, \qquad R(p, \varepsilon) = \left\lceil 4096 p^{-1} \{ Q(\varepsilon) + \log(4/\varepsilon) \} \right\rceil.$$

The limits in (4.7) dominate these quantities at $p = p_{\mathrm{den}}$ and $\varepsilon = \varepsilon_s$. We prove the stronger generic guarantee, assuming the child-failure bound of Lemma 4.3.

Proof. Write $Q = Q(\varepsilon)$ and $R = R(p, \varepsilon)$. An opportunity is a verified candidate or an immediate proper factor. The geometric legal-point probability and the candidate density imply an opportunity probability at least $p/512$, conditional on every unsuccessful past. A good candidate's children all complete with probability at least $7/8$, by the union bound. To include proposals that find a factor immediately, let S_{out} be factor mass outside the set of candidates before testing, on a common proposal-mass scale. Success conditional on an opportunity is at least

$$\frac{(7/8)D + S_{\mathrm{out}}}{Z + S_{\mathrm{out}}} \ge \frac{7D}{8Z} > \frac{1}{32}.$$

The probability that Q opportunities all fail is therefore at most $(31/32)^Q$. Before success, opportunity indicators stochastically dominate Bernoulli variables of parameter $p/512$; stopped histories can be extended by virtual successes. At the displayed cap the dominating binomial mean is at least $8\{Q + \log(4/\varepsilon)\}$. A Chernoff bound therefore puts the probability of fewer than Q opportunities below $\varepsilon/2$. Also $(31/32)^Q \le (\varepsilon/4)^4 < \varepsilon/2$. Their sum proves the claimed failure bound. □

D.4 Probability of the prescribed initial factorizations

Proof. The integer parameters give

$$16(2Y_0)^{2k} \le N, \qquad Y_0^k > \frac{\sqrt{N}}{4\,4^k}.$$

Among the $\binom{L}{k}$ products of k distinct primes, one of the k dyadic intervals covering $(Y_0^k, (2Y_0)^k]$ contains a set S with $|S| \ge \binom{L}{k}/k$.

For an odd N, there are $2^{\omega(N)}$ unit square classes. At least $|S|^2/2^{\omega(N)}$ ordered pairs belong to the same class. If s_p counts elements of S divisible by p, then

$$s_p \le \binom{L-1}{k-1} \le k^2|S|/L, \qquad \sum_p s_p = k|S|.$$

Pairs sharing a prime therefore number at most $\sum_p s_p^2 \le k^3|S|^2/L$. At least $|S|^2/2^{\omega(N)+1}$ pairs are coprime and have square quotient modulo N.

Their modular quotients are distinct. A collision makes $uv' - u'v$ divisible by N and smaller than N in absolute value, so it is zero; coprime positive fractions then coincide. Each pair has squared Euclidean norm at most $N/8$. An independent lattice vector of no greater norm would have nonzero determinant divisible by N, but of absolute value at most $N/8$. This is impossible. Primitivity proves that the pair is the shortest vector, uniquely up to sign.

Each unit square has exactly $2^{\omega(N)}$ roots, even when N is not squarefree. Consequently the desired probability is at least $|S|^2/(2N)$. Since $\binom{L}{k} \ge (L/k)^k$, this is at least

$$\frac{1}{32k^2(40k\log Y_0)^{2k}} \ge p_*.$$

The last inequality follows from $k\log Y_0 \le \frac{1}{2}\log N$. □

E Square-class certification and the congruence of squares

A dependence among the public rows specifies a product of squares modulo N. We construct a public square root of this product and ask how often it gives a proper divisor when compared with the retained private roots. Dixon's congruence-of-squares argument [23] applies once we account for the adaptive choice of fields, rows, and dependence.

Our construction follows the public-root and transcript arguments of [33, `mfU_theorem`, Sections 7–8]. The unconditional character estimates and the use of the derivative to clear denominators follow [38, Section 6]. We remove rational square classes as in [33, `mfU_theorem`, Lemma 8.2]. Proposition E.3 gives one certificate for all later choices of rows. Theorem E.5 gives the exact probability of a proper divisor conditional on the public computation. Appendix F accounts for all operations, and Section 6 bounds the expected cost including recursive calls.

E.1 Algebraic preliminaries and reduction modulo the input

Let $N > 1$ be odd. The proper-factor probability below assumes $\omega(N) \geq 2$, where $\omega(N)$ counts distinct prime divisors. For each field index $j \in \{1, \dots, J\}$, the public data are:

1. a primitive irreducible polynomial $f_j \in \mathbb{Z}[T]$ of odd degree $d_j \geq 3$, with leading coefficient $A_j \neq 0$, and a positive integer e_j;
2. a residue m_j such that $f_j(m_j) = 0 \bmod N$, with $A_j, e_j, f_j'(m_j)$ all units modulo N;
3. a public coordinate bound $V_j \geq 1$, fixed before character sampling, for every atom that may subsequently be used;
4. for each used primitive pair $P = (a, b)$, a verified prime factorization of $|e_j F_j(a, b)|$, where $F_j(a, b) = b^{d_j} f_j(a/b) \neq 0$, and a verified unit $a - bm_j \bmod N$.

For every rational prime in a primitive norm factorization, compute its gcd with N. Return a proper gcd immediately and reject the affected object if the gcd equals N. A recursive factorization verifies its primes by primality testing and its factor list by multiplication. The parent reuses these primality checks and verifies the required products, valuations, and gcds. The preprocessing sieve supplies the small base primes.

Fix one field, omit its index j, take a root α of f, and put

$$K = \mathbb{Q}(\alpha), \quad \beta = A\alpha, \quad g(T) = A^{d-1} f(T/A), \quad O = \mathbb{Z}[\beta], \quad D = eA \operatorname{disc}(f). \tag{E.1}$$

The polynomial g is monic and integral. An atom $P = (a, b)$ gives a field element and its integral multiple,

$$\xi_P = a - b\alpha, \qquad z_P = Aa - b\beta = A\xi_P \in O.$$

The map $\beta \mapsto Am$ induces a ring homomorphism $\psi : O \to \mathbb{Z}/N\mathbb{Z}$. Since

$$\psi(z_P) = A(a - bm), \qquad \psi(g'(\beta)) = A^{d-2} f'(m)$$

are verified units, it extends uniquely to

$$O[A^{-1}, g'(\beta)^{-1}, \{z_P^{-1} : P \text{ is used}\}]. \tag{E.2}$$

We use this map for every field-atom evaluation. For rational coefficients we use the map in (E.6).

A public row i consists of a unit $R_i \bmod N$, a nonzero rational number q_i with a complete signed numerator and denominator factorization, and signed atom occurrences $c_{i,j,P} \in \mathbb{Z}$. All primes in q_i are units modulo N, and the row identity is checked exactly:

$$R_i = q_i \prod_j \prod_P (a - bm_j)^{c_{i,j,P}} \pmod N. \tag{E.3}$$

We keep every occurrence with its multiplicity. Each new row receives a fresh private root, even when it reuses a public relation.

For a nonempty selected set I of rows, define

$$\begin{aligned} c_{j,P} &= \sum_{i \in I} c_{i,j,P}, \quad E_j = \sum_P c_{j,P}, \\ \gamma_j &= \prod_P \xi_P^{c_{j,P}}, \qquad q_I = \prod_{i \in I} q_i. \end{aligned} \tag{E.4}$$

For a nonzero factored rational number v, let $\mathrm{sf}(v)$ be its *signed squarefree integer* representative: the sign is retained and each prime exponent, including a negative exponent, is reduced modulo two to 0 or 1. Set

$$\delta_j = \mathrm{sf}\left(\prod_P F_j(P)^{c_{j,P}} \right). \tag{E.5}$$

Let $\mathcal{S}_{\mathrm{rat}}$ be the finite set of rational primes appearing in the signed numerator or denominator factorization of any retained q_i, or in a primitive norm factorization $F_j(P)$. Every prime in this set has a verified unit image modulo N. Consequently there is a well-defined homomorphism

$$\psi_{\mathrm{rat}} : \mathbb{Z}[\{p^{-1} : p \in \mathcal{S}_{\mathrm{rat}}\}] \longrightarrow \mathbb{Z}/N\mathbb{Z}, \qquad a/b \longmapsto a\, b^{-1} \pmod N. \tag{E.6}$$

This ring contains every q_i, every δ_j, and every rational square root formed below by halving the stored even prime exponents. Its construction requires no new integer factorization.

Lemma E.1 (Norms and good ideal labels)**.** *For each atom and each selected product,*

$$\mathrm{Norm}_{K/\mathbb{Q}}(z_P) = A^{d-1} F(P), \qquad \mathrm{Norm}_{K/\mathbb{Q}}(\gamma) = A^{-E} \prod_P F(P)^{c_P}. \tag{E.7}$$

If E is even, δ represents the rational square class of $\mathrm{Norm}(\gamma)$*. It is computable from the stored factor lists without factoring A or a new large integer.*

If $p \mid F(a,b)$ and $p \nmid D$, exactly one prime ideal above p has nonzero valuation on ξ_P. It has residue degree one, label $(p, a/b \bmod p)$, and valuation $v_p(F(a,b))$.

Proof. The norm of $a - b\alpha$ is $F(a,b)/A$, which proves (E.7). Because $d-1$ and E are even where used, the corresponding powers of A are positive rational squares. To recover the factorization of $F(P)$ from that of $eF(P)$, divide the known integer e repeatedly by each listed prime and subtract its valuation. Every prime divisor of e occurs in the complete raw list.

For the ideal assertion, $p \nmid A \operatorname{disc}(f)$ implies that the order generated by β has index prime to p, so factorization modulo p describes the primes above p. If $p \mid b$, primitivity would give $F(a,b) \equiv Aa^d \neq 0 \bmod p$; hence b is invertible modulo p. The residue $a - b\alpha$ vanishes only at the simple linear factor $T - a/b$. This factor has residue degree one, and the norm identity then gives the asserted valuation. □

An ideal parity column therefore records the field, the rational prime, and the root. Good large labels may cancel within one relation, and good small labels give matrix columns. Norm columns use the same column for each rational prime across all fields, even when that prime is good in one field and bad in another. We also retain the sign columns.

E.2 Separation of nonrational square classes

We use quadratic characters to detect all nonrational square classes. Since the degree is odd, taking the norm fixes each rational square class. The stored norm factors therefore suffice to remove any class that survives (Section E.3). For a nonrational class the associated quadratic extension of K contains no quadratic subfield over $\mathbb{Q}$. This gives the uniform exceptional-zero estimate below.

For one fixed field, define the finite group

$$H_D = \{\gamma \in K^\times : v_\mathfrak{p}(\gamma) \in 2\mathbb{Z} \text{ for every } \mathfrak{p} \nmid D\}/K^{\times 2}. \tag{E.8}$$

A class in H_D is rational if a nonzero rational number represents it. We bound $|H_D|$ to take a union bound over its classes.

Lemma E.2 (Uniform class and discriminant bounds)**.** *Put*

$$X_0 = \left\lceil \sqrt{|\operatorname{disc}(g)|} \right\rceil, \qquad H_* = 2^d |D|^d X_0 (1 + \log X_0)^{d-1},$$

and

$$\overline{\Delta} = |\operatorname{disc}(g)|^2 \, 8^d |D|^d. \tag{E.9}$$

Then $|H_D| \le H_*$. For every nonsquare class $[\gamma] \in H_D$, the field $L = K(\sqrt{\gamma})$ satisfies $|\Delta_L| \le \overline{\Delta}$. These bounds do not depend on the height of a class representative or on the number of atom occurrences.

Proof. Let S_D be the prime ideals above divisors of D. The square-ideal map has kernel the S_D-units modulo squares and image in the two-torsion of the S_D-ideal class group. Dirichlet's unit theorem therefore gives

$$|H_D| \le 2^{d+|S_D|} h_K, \qquad 2^{|S_D|} \le |D|^d.$$

Minkowski's theorem gives an integral ideal of norm at most $\sqrt{|\Delta_K|}$ in every ideal class. The number of ideals of norm a is at most $\tau_d(a)$: this follows by comparing local Euler factors with $(1 - p^{-s})^{-d}$. Repeated summation of the harmonic series yields

$$\sum_{a \le X} \tau_d(a) \le X (1 + \log X)^{d-1}.$$

Since $|\Delta_K| \le |\operatorname{disc}(g)|$, this proves the class bound.

For the discriminant bound, at every finite place divide γ by a local square so that its valuation is zero or one. The quadratic order generated by its square root has discriminant 4γ, so its relative discriminant exponent is at most $2v_\mathfrak{p}(2) + 1$. Outside $2D$ the extension is unramified. The odd bad primes contribute at most $|D|^d$ to the relative discriminant norm; primes above two contribute at most 8^d. The discriminant tower formula proves (E.9). □

For $\mathfrak{p} \nmid 2D$ and $[\gamma] \in H_D$, define

$$\chi_\mathfrak{p}([\gamma]) = \left(\overline{\pi^{-v_\mathfrak{p}(\gamma)} \gamma} \,\middle|\, k(\mathfrak{p}) \right) \in \{1, -1\}, \tag{E.10}$$

where π is a local uniformizer and the symbol denotes the quadratic character on the multiplicative group of the residue field $k(\mathfrak{p})$. The even valuation makes this independent of the uniformizer and of the square-class representative. It is the Frobenius character of $K(\sqrt{\gamma})/K$.

The following choices keep the sampling range computable. Write $\mathrm{b}(a) = \lfloor \log_2 a \rfloor + 1$ for $a \geq 1$, and set

$$\begin{aligned} \ell_D &= \mathrm{b}(|D|), \qquad \ell_g = \mathrm{b}(|\operatorname{disc}(g)|), \\ h_* &= d + d\ell_D + \ell_g + 2 + d\,\mathrm{b}(\ell_g + 3), \\ b_\Delta &= 2\ell_g + 3d + d\ell_D, \\ \mathcal{N}_* &= |A|^{d-1}(d+1)\max(1, \|f\|_\infty)V^d. \end{aligned} \tag{E.11}$$

These satisfy $\log H_* \leq h_*$, $\log \overline{\Delta} \leq b_\Delta$, and

$$0 < |\operatorname{Norm}(z_P)| \leq \mathcal{N}_*$$

for every allowed future atom. Given an outer input length $n \geq 1$, choose the integer

$$\begin{aligned} T &= 2 + d + n + b_\Delta + h_* + \mathrm{b}(J) + \mathrm{b}(2 + |D| + \mathcal{N}_*), \\ S &= (2d)!\,T^4, \qquad x = 2^S. \end{aligned} \tag{E.12}$$

Separate choices are made for different fields.

Proposition E.3 (Simultaneous separation)**.** *There is an absolute effective threshold T_0 such that, whenever $T \geq T_0$, a uniform degree-one prime ideal of K with rational prime norm in $(x/2, x]$ detects any fixed nonrational class in H_D with probability at least $1/(20d)$. Such an ideal can be sampled using $O(dS)$ expected integer proposals.*

For each field j, take independently

$$t_j = 20d_j\{h_{*,j} + \mathrm{b}(64J) + 1\} \tag{E.13}$$

samples. With probability at least $1 - 1/64$, simultaneously for all fields, every class in H_{D_j} on which all sampled characters equal 1 is rational. The atoms, rows, and selected kernel may be chosen after these characters have been revealed.

Proof. Fix a nonrational class and put $L = K(\sqrt{\gamma})$. If L contained a quadratic field $F/\mathbb{Q}$, oddness of d would give $K \cap F = \mathbb{Q}$, hence $L = KF$. Equality of these quadratic extensions of K would make $[\gamma]$ rational. Thus L contains no quadratic subfield.

Stark's Lemma 8 [60] says that a number field E of degree $a > 1$ with a real zeta zero in

$$1 - \{4a! \log |\Delta_E|\}^{-1} \leq \beta < 1$$

contains a quadratic field with that zero. Consequently any possible exceptional zero for L has gap at least

$$\nu = \{4(2d)! \log \overline{\Delta}\}^{-1}.$$

The unconditional Chebotarev theorem [35, Theorem 1.3], applied to the normal extension L/K, uses absolute degree $2d$. Its counting error is at most a possible $\mathrm{Li}(x^\beta)$ term plus

$$C_1 x \exp\{-C_2\sqrt{\log x/(2d)}\},$$

with absolute effective constants, for $x \geq \exp(20d(\log |\Delta_L|)^2)$.

The choice (E.12) is uniform in the field and class. Indeed,

$$\nu \log x \geq (\log 2)T^3/4, \qquad \log S \leq 2T\log(2T) + 4\log T,$$

and

$$\sqrt{S/(2d)} = \sqrt{(2d)!/(2d)}\, T^2 \geq \sqrt{120}\, T^2.$$

Thus, relative to $x/\log x$, both errors tend uniformly to zero. The same is true at $x/2$, and the theorem's range condition holds there for all sufficiently large T. The nontrivial Frobenius class consequently contributes at least $x/(5\log x)$ prime ideals in the interval. Ideals of residue degree at least two contribute at most $d\sqrt{x}$, which is uniformly negligible; at least $x/(10\log x)$ degree-one ideals remain after increasing the threshold. There are at most $d\pi(x) \leq 2dx/\log x$ degree-one ideals in total. Dividing the two bounds gives the claimed probability $1/(20d)$. Applying the trivial-extension case to K/K gives the required lower bound on the total number of degree-one ideals as well: an odd-degree field has no quadratic subfield, so the same zero-gap argument applies. No normal closure over $\mathbb{Q}$ is used.

For exact sampling, choose a uniform integer $p \in (x/2, x]$, verify primality, and compute all roots of f mod p. If there are r_p roots, accept with probability r_p/d, then choose one root uniformly. Each output pair (p, u) has single-proposal probability

$$\frac{2}{x}\frac{r_p}{d}\frac{1}{r_p} = \frac{2}{dx}.$$

The prime is larger than $2, |D|, \mathcal{N}_*$, so these roots are simple and identify precisely the degree-one prime ideals. The lower count above makes the proposal acceptance probability $\gg 1/(dS)$.

A fixed nonrational class is missed by all t_j samples with probability at most

$$\exp(-t_j/(20d_j)).$$

A union bound over at most $H_{*,j}$ classes and at most J fields proves the asserted $1/64$ failure bound. The event quantifies over each entire finite group, so it also covers later adaptive choices. □

Every allowed atom is a local unit at these primes. At the prime ideal corresponding to a root u of f mod p, its matrix character is

$$\chi_{p,u}(\xi_P) = \left(\frac{a-bu}{p}\right), \qquad \chi_{p,u}(z_P) = \left(\frac{A}{p}\right)\left(\frac{a-bu}{p}\right). \tag{E.14}$$

An individual atom need not lie in H_D. Multiplicativity on local units first forms the selected product; ideal parity then places that product in H_D, where (E.10) applies. If the integral atoms z_P are used in the character columns, the even E_j column cancels the additional leading-coefficient symbol.

E.3 Construction of the public square root

For a selected nonempty set I, the columns impose three conditions.

1. E_j is even and $[\gamma_j] \in H_{D_j}$ for every field;
2. q_I and $\prod_j \delta_j$ are positive rational squares;
3. every sampled character of every γ_j equals 1.

The occurrence and ideal columns give the first condition, the rational-prime, shared norm-prime, and sign columns give the second, and the character columns give the third.

Theorem E.4 (Finite public certification)**.** *Given the public data of Section E.1 and a nonempty selection satisfying the three parity conditions above, suppose that the simultaneous separation event of Proposition E.3 holds. Then we can compute a public unit $Y \bmod N$ satisfying*

$$Y^2 = \prod_{i \in I} R_i \pmod{N}. \tag{E.15}$$

The construction uses integer and polynomial arithmetic, finite-field factorization and square roots, and Hensel lifting. It does not require a maximal order, class group, unit group, factorization of D_j, or a number-field square-root oracle. All accepted outputs are verified exactly, even outside the separation event. The bounded implementation can also fail if a randomized finite-field calculation reaches its cap. Section E.5 bounds the cost and probability of these timeouts.

Proof. Removing the rational class. First work in one field. Separation gives $\gamma = u\theta^2$ for some $u \in \mathbb{Q}^\times$; neither u nor θ is computed. Oddness of d and Lemma E.1 give

$$[\delta] = [\mathrm{Norm}(\gamma)] = [u^d] = [u] \quad \text{in } \mathbb{Q}^\times/\mathbb{Q}^{\times 2}.$$

It follows that $\delta\gamma$ is a square in K.

Moving the root into the integer power basis. We clear the negative occurrences and multiply by $g'(\beta)^2$. The latter factor will place a square root in the computable order $O = \mathbb{Z}[\beta]$. Define

$$\gamma^- = \prod_{c_P<0} z_P^{|c_P|}, \qquad T_{\mathrm{occ}} = \sum_P |c_P|, \qquad \Gamma = \delta\, g'(\beta)^2 \prod_P z_P^{|c_P|}. \tag{E.16}$$

Because

$$\prod_P z_P^{c_P} = A^E\gamma, \qquad \prod_P z_P^{|c_P|} = (\gamma^-)^2 A^E\gamma,$$

the element Γ is a square in K. It is integral because δ is a signed integer. If $\eta^2 = \Gamma$, then

$$\left(\frac{\eta}{g'(\beta)}\right)^2 = \delta \prod_P z_P^{|c_P|} \in O_K. \tag{E.17}$$

Hence $\eta/g'(\beta) \in O_K$.

We now show that $\eta \in O$, which makes integer coefficient reconstruction possible. The trace dual of $O = \mathbb{Z}[\beta]$ is $O^\vee = g'(\beta)^{-1}O$. For a direct verification, write

$$\frac{g(T)}{T-\beta} = \sum_{a=0}^{d-1} q_a(\beta)T^a.$$

Lagrange interpolation gives

$$\mathrm{Tr}\left(\beta^k\, \frac{q_a(\beta)}{g'(\beta)}\right) = \delta_{ka}, \qquad 0 \le k, a < d.$$

The q_a's form a monic triangular integral basis of O, proving the trace-dual identity. Products of algebraic integers have integral trace, so $O_K \subseteq O^\vee$; therefore $g'(\beta)O_K \subseteq O$. Apply this inclusion to (E.17).

Finite reconstruction. To recover the integer coefficients uniquely from a modular lift, we first give an explicit bound. Put

$$R = 1 + \|g\|_\infty, \qquad Z = 1 + V(|A| + R),$$

and

$$B_\eta = d!R^{(d-1)^2}(1+|\delta|)(2R)^{d-1}Z^{T_{\mathrm{occ}}}. \tag{E.18}$$

Every conjugate of β has absolute value at most R; every atom conjugate is smaller than Z; and $|g'(\beta^{(a)})| \le (2R)^{d-1}$. Thus every conjugate of η is at most $(1+|\delta|)(2R)^{d-1}Z^{T_{\mathrm{occ}}}$. The Vandermonde determinant for the conjugates of β has absolute value $\sqrt{|\operatorname{disc}(g)|} \ge 1$, while each cofactor is at most $(d-1)!R^{(d-1)^2}$. Cramer's rule proves that every coefficient of η in the power basis $1, \beta, \ldots, \beta^{d-1}$ of the order $O = \mathbb{Z}[\beta]$ has absolute value at most B_η. This order need not be maximal.

Reuse a sampled character prime p. It does not divide $2\operatorname{disc}(g)$, any norm of an atom, or δ. Indeed, each prime divisor of δ occurs in a primitive norm list and is at most $\mathcal{N}_* < p$. Factor $g \bmod p$ completely into distinct irreducibles. In each finite-field component, Γ is a unit square and has two roots. There are at most d components, so at most 2^d combinations of signs. Polynomial CRT gives each starting root modulo p.

Since 2η is a unit modulo p, each starting root lifts uniquely by Newton–Hensel iteration in the ring

$$(\mathbb{Z}/p^a\mathbb{Z})[T]/(g).$$

Lift until $p^a > 2B_\eta$, take centered coefficients, and check $\eta^2 = \Gamma$ exactly in $\mathbb{Z}[T]/(g)$. The true integral root reduces to one of the enumerated sign combinations; its centered lift is therefore recovered. Any other branch is accepted only if it also passes this exact identity.

Combining the field roots. For all fields, compute these roots and the positive rational square

$$Q = \frac{q_I}{\prod_j \delta_j}.$$

Its square root is obtained by halving the saved signed prime exponents. Define

$$Y = \psi_{\mathrm{rat}}(\sqrt{Q}) \prod_j A_j^{-E_j/2} \prod_j \frac{\psi_j(\eta_j)}{\psi_j(\gamma_j^-)\psi_j(g_j'(\beta_j))} \pmod N. \tag{E.19}$$

Every field denominator has a verified unit image under the corresponding map ψ_j in (E.2); the rational factor is evaluated using ψ_{rat} in (E.6). Thus all inverses used in (E.19) are defined modulo N. Squaring (E.19), using (E.16), cancels the $A_j^{E_j}$ and δ_j terms, and gives

$$Y^2 = q_I \prod_j \psi_j(\gamma_j) = \prod_{i \in I} R_i \pmod N.$$

The algorithm checks this identity and $\gcd(Y, N) = 1$ before returning Y. □

E.4 Conditional distribution of the private roots

Let $n = \mathrm{b}(N)$. At every possible row attempt, read a *new independent* block of n fair private bits and interpret it as $h \in [0, 2^n)$. An integer outside $[1, N)$ causes failure; a nonunit in that interval gives a verified proper gcd; a unit releases only $R = h^2 \bmod N$. The value h is kept in a fixed-width location for the private root. Range testing, gcd, squaring, temporary storage, and clearing are padded to a public running time depending only on n.

More concretely, fix a public integer $B_{\mathrm{init}}(n)$ and implement this phase by a fixed schedule of exactly that many ordinary bit operations, with data-independent addresses. Such a polynomial-size schedule is obtained by unrolling a bounded Euclidean algorithm and schoolbook arithmetic with public-length loops, masked assignments, and dummy operations after a result has been

computed. No intermediate range flag, gcd status, or square is released before the fixed clearing and padding phase ends. The public continuation starts after exactly $B_{\mathrm{init}}(n)$ bit operations, independently of which unit root of R was used. There is no inherited time cap on the factoring algorithm. Later calls on smaller inputs receive public inputs and independent fresh bits; their running times therefore do not reveal which private root of a fixed released square was chosen.

All later decisions, including relation reuse, failures of the bounded arithmetic routines, retained indices, and kernel selection, use public data and independent public randomness. Private records may be moved at public positions with fixed-width operations; their contents are not compared, sorted, hashed, compressed, or deduplicated. Each retained row has its own independently initialized record. Copying a public relation does not create a second private occurrence of an earlier root. The public set I and the verified Y are fixed before the first private multiplication. Only one final gcd using the private roots is used; a trivial result discards the stored private roots.

Theorem E.5 (Conditional root-fibre law)**.** *For every positive-probability public transcript ending immediately before that first private multiplication, the retained roots are independent and uniform in the product of their unit square-root fibres. Consequently, if $I \neq \varnothing$ and* (E.15) *holds, then*

$$\Pr\left\{1 < \gcd\left(\prod_{i\in I} h_i - Y, N\right) < N \;\middle|\; \textit{public transcript}\right\} = 1 - 2^{1-\omega(N)}. \tag{E.20}$$

Proof. Presample the bounded collection of local private blocks and an independent tape for all other coins, including those used by factorizations of smaller integers. Set

$$\mu_N = \{u \in (\mathbb{Z}/N\mathbb{Z})^\times : u^2 = 1\}.$$

For one block, multiplication by $u \in \mu_N$, followed by the canonical representative in $[1, N)$, permutes its unit values. Fix all other block values. This is a probability-preserving permutation of the equally likely n-bit blocks. It preserves the released square, range and gcd statuses, padded time, and every subsequent public decision. The fixed initialization schedule prevents a data-dependent initialization time from distinguishing the roots. All subsequent inputs and random-tape addresses are determined by the released squares, the other fixed coins, and the resulting public history. Act independently on the distinct initial blocks of all retained rows. The complete public transcript, including its retained-index set, remains unchanged.

The action of μ_N on a unit square fibre is transitive: if $h^2 = (h')^2$, then $h'/h \in \mu_N$. The coordinatewise action is therefore transitive on the Cartesian product of retained fibres. Invariance of probability and transcript makes the conditional distribution uniform on that product.

Since $I \neq \varnothing$, varying one selected root while fixing the others shows that $Y^{-1}\prod_{i\in I} h_i$ is uniform in μ_N. For an odd prime power p^a, $z^2 = 1 \bmod p^a$ has exactly the roots 1 and -1: the factors $z-1, z+1$ have gcd dividing two. CRT gives $|\mu_N| = 2^{\omega(N)}$, including nonsquarefree N. Only the all-positive and all-negative sign vectors produce gcd N and gcd 1, respectively. Every other vector gives a proper divisor, proving (E.20). □

The theorem applies up to the first gcd using the private roots, after which the program discards these roots. We may restrict to any public success event before this gcd, even when that event depends on relation supply. If a terminating public computation produces either

an immediate factor or a verified public square with probability at least s, its proper-factor probability is at least

$$\big(1 - 2^{1-\omega(N)}\big)s. \tag{E.21}$$

Immediate factors retain their entire probability mass.

E.5 Complexity of square certification and reconstruction

All prime candidates are verified by deterministic primality testing [3, 2]. Factoring polynomials over finite fields can use squarefree and distinct-degree factorization followed by equal-degree splitting. For two distinct irreducible components over an odd field, a uniform residue has independent component values; their nonzero quadratic signs differ with probability at least $2/9$. Thus recursive splitting has polynomial expected work in degree and field bit length. Factoring $T^2 - a$ gives component square roots and does not require factoring $p - 1$. Every factorization and root is checked before it is used.

For definiteness, one may fix schoolbook implementations with the following generous eventual bound on each *already reduced* direct input:

$$\mathcal{P} = (n + d + S^* + 2)^{40}, \tag{E.22}$$

where d and S^* bound all degrees and sampled-prime bit lengths. For an s-bit integer, the corrected AKS Lemma 4.3 [2] gives a modulus degree $r_{\mathrm{AKS}} = O(s^5)$. The counts in [3, Section 5] give $O(s\sqrt{r_{\mathrm{AKS}}})$ polynomial congruence tests, each using $O(s)$ polynomial multiplications by repeated squaring. Schoolbook multiplication with reduction after each term costs $O(r_{\mathrm{AKS}}^2 s^2)$, so these tests cost

$$O\big(s\sqrt{r_{\mathrm{AKS}}} \cdot s \cdot r_{\mathrm{AKS}}^2 s^2\big) = O(s^{33/2}).$$

The modulus search, order tests, gcds, and perfect-power checks have smaller fixed exponents. Frobenius powering, gcds, recursive constant-probability splitting, and exact finite-field product checks have smaller fixed exponents in $d + s$. All implied constants are fixed once the implementations are fixed, and are absorbed by exponent forty beyond one fixed threshold.

Lemma E.6 (Capping adaptive direct calls)**.** *Suppose at most Q_χ direct calls are made, adaptively on public inputs, and the conditional expected work of each fixed input is at most $\mathcal{P}$. Giving every call allowance $64Q_\chi\mathcal{P}$ makes the probability of any failure of a bounded arithmetic routine at most $1/64$, and the total bound on the running time of these routines at most $64Q_\chi^2\mathcal{P}$.*

Proof. Condition on the public past before a call. Markov's inequality bounds its probability of exceeding the prescribed limit by $1/(64Q_\chi)$. Use a union bound over the deterministic upper bound on the number of calls, padding unused positions by zero failures. Summing the deterministic allowances proves the work bound. Independence between the calls is unnecessary. □

Constructing the large coefficients of Γ, reading and reducing them modulo p, and high-precision Hensel lifting are counted separately from (E.22). The coefficient precision from (E.18) satisfies

$$\begin{aligned}\log(2B_\eta + 1) = O\big(&d \log d + d^2 \log R + \log(1 + |\delta|) \\ &+ T_{\mathrm{occ}} \log Z\big).\end{aligned} \tag{E.23}$$

Given the finite-field starting data, dense quotient-ring arithmetic, successive precision doubling, centered reconstruction, and exact verification use 2^d times a fixed polynomial in this bit

precision, d, the sampled prime bit length, and the recorded occurrence count. We therefore count the occurrences in each retained product, including their multiplicities.

Ideal sampling is also capped before the epoch starts. If S^*, t^* are deterministic bounds for all fields, use

$$R_\chi = n^2 d S^* \, \mathrm{b}(64Jt^* + 2) \tag{E.24}$$

integer proposals per accepted character. Proposition E.3 makes the probability of any failure due to the trial limit at most

$$Jt^* \exp\{-cn^2 \mathrm{b}(64Jt^* + 2)\},$$

which is at most $1/64$ for all sufficiently large n. The fixed program need not know the positive absolute constant c. Its value changes the proof threshold, not the displayed cap. Exact fair-bit sampling of other finite intervals is capped in the same way, with its own explicit total upper bound on the failure probability. Couple the bounded algorithm to the ideal one using the same random bits until a prescribed limit is reached; conditioning on absence of timeout is not asserted to preserve an ideal sampling distribution.

To see the contribution of the sampled-prime length to bit complexity, suppose that

$$d \sim (\log N / \log\log N)^{1/3}, \qquad T \le (\log N)^{O(1)},$$

and $\Lambda = (\log N)^{1/3} (\log\log N)^{2/3}$, then

$$\log S = \log(2d)! + O(\log\log N) = (2/3 + o(1))\Lambda. \tag{E.25}$$

Thus a single sampled prime already has $S + O(1)$ bits. For example, if

$$\log J \le (f + o(1))\Lambda, \quad t^* \le (\log N)^{O(1)}, \quad \log Q_\chi \le (f + 2/3 + o(1))\Lambda,$$

then the bound on the cost of the arithmetic routines has exponent

$$\log(64 Q_\chi^2 \mathcal{P}) \le (2f + 28 + o(1))\Lambda. \tag{E.26}$$

This is only the direct arithmetic contribution. It excludes relation generation, dense elimination, coefficient construction, retained-occurrence storage, high-precision reconstruction, and recursive integer factorizations; all of these contributions must be included in the running-time bound.

E.6 Construction of the rows and the public square root

Starting from a completed descent tree in Section 5, we record the rational correction for each atom and the parity columns for its row. A dependence among the rows then gives the public square root used in the final gcd.

Collecting rows. The normalization and top-band test in Section 5 preserve the primitive shapes counted by Lemma 5.1. Each attempt that passes the test makes at most $2k$ top requests. The lemma counts shapes using the set of good primes without requiring the algorithm to list them. After preprocessing, a private n-bit block represents a unit in $[1, N)$ with probability

$$\frac{\varphi(N)}{2^n} \ge \frac{N}{2^n}\left(1 - \frac{\omega(N)}{B}\right) \ge \frac{1}{2}\left(1 - \frac{n}{B}\right) \ge \frac{1}{4}$$

for all sufficiently large admissible inputs. Combining this with the joint top-search bound gives the per-attempt probability in (5.2). We keep each completed row with its own private index, even when it lies in the span of the preceding rows.

Rational corrections and exact records. Each retained occurrence stores its field, signed multiplicity, complete numerical factor list and root labels. For a top atom write $G(P_0) = \sigma_{P_0} p n_{P_0}$, with n_{P_0} positive and B-smooth. From $G(P) = Q(a_P - b_P m_N) \pmod N$, the exact modular identities are

$$p \equiv \frac{\sigma_{P_0} Q}{n_{P_0}}(a_{P_0} - b_{P_0} m_N), \qquad 1 \equiv \frac{Q}{G(P)}(a_P - b_P m_N) \pmod N.$$

Give occurrences the sign of their top prime in u_i/v_i. Write $c_{i,P}$ for $c_{i,a,P}$ in the two products below, which run over occurrences with their field labels. Set

$$q_i = \operatorname{sgn}(u_i v_i) \prod_{\text{top } P_0} \left(\frac{\sigma_{P_0} Q}{n_{P_0}}\right)^{c_{i,P_0}} \prod_{\text{internal } P} \left(\frac{Q}{G(P)}\right)^{c_{i,P}}.$$

Multiplying the atom identities proves (5.3); in particular, every inserted internal atom contributes its $Q/G(P)$ correction. All denominators are verified units. Only signs, the known prime Q, and B-smooth values enter q_i, so its signed prime exponents are available without a further integer factorization. The row identity is checked before the row is retained.

Field data and parity columns. For accepted field a, write $\mathcal{F}_a = e_a F_a$, where $e_a > 0$ is the content and F_a is primitive. Put $f_a(T) = F_a(T, 1)$, $A_f = \operatorname{lc}(f_a)$, and $D^*_{f_a} = e_a A_f \operatorname{disc}(f_a)$. A prime is good for this field if it does not divide $D^*_{f_a}$. The matrix records four sets of columns:

1. the shared sign and B-prime exponents of q_i;
2. each field's occurrence parity $E_{i,a} = \sum_P c_{i,a,P} \bmod 2$ and good small ideal labels;
3. shared norm signs and rational prime parities, with bad primes registered on discovery;
4. the sampled quadratic characters of each field.

The parity of the exponent of Q in q_i is $\sum_a E_{i,a}$. The completed tree already gives even valuation at every good large ideal. If $t_{\chi,a}$ probes are sampled in field a, the column count is at most

$$2\pi(B) + 2 + \sum_{a \text{ accepted}} \{d\pi(B) + 1 + t_{\chi,a} + \mathrm{b}(|D^*_{f_a}|)\}. \tag{E.27}$$

Appendix F gives a deterministic upper bound $\mathcal{C}$ for this count.

Appending columns. The bound (E.27) includes columns first encountered in later rows. When adding such a column we put zero in each earlier row. For a newly observed good small ideal, the old complete factor lists show that it had no odd occurrence. For a newly observed bad rational prime above B, any earlier bad occurrence would already have registered it; earlier good large occurrences have even row valuation by the completed-tree invariant. Base-prime norm columns can be present from the start. Thus zero extension preserves all parity equations without factoring an entire discriminant.

Monic coordinates and modular evaluation. Write $K_a = \mathbb{Q}(\alpha_a)$, where $f_a(\alpha_a) = 0$, and put

$$\beta_a = A_f \alpha_a, \qquad g_a(T) = A_f^{d-1} f_a(T/A_f), \qquad z_P = A_f a_P - b_P \beta_a.$$

The map $\beta_a \mapsto A_f m_N$ evaluates elements of $\mathbb{Z}[\beta_a]$ modulo N. We extend it to the localization in (E.2) by inverting the verified units $A_f, g_a'(\beta_a), z_P$.

Public roots from a dependence. For the indices I of a selected nonzero binary dependence, define

$$\gamma_a = \prod_P (a_P - b_P\alpha_a)^{\sum_{i\in I} c_{i,a,P}}, \qquad q_I = \prod_{i\in I} q_i.$$

The parity equations make q_I a positive rational square and every good ideal valuation of each γ_a even. The shared norm and sign columns make the product of the signed norm representatives a positive rational square. On the simultaneous separation event of Proposition E.3, the character equations give $\gamma_a = u\theta^2$ for rational u, and odd d gives $[\mathrm{N}(\gamma_a)] = [u^d] = [u]$. Hence

$$\delta_a = \mathrm{sf}\left(\prod_P F_a(P)^{\sum_{i\in I} c_{i,a,P}}\right)$$

is a signed squarefree integer with $\delta_a\gamma_a$ a field square. Saved factor lists supply δ_a; the identity $\mathrm{N}(z_P) = A_f^{d-1}F_a(P)$ and even occurrence parity justify this without factoring A_f. Every norm prime is gcd-tested.

To reconstruct a field root, Theorem E.4 first clears the negative occurrences. Multiplication by $g_a'(\beta_a)$ then places the integral root in $\mathbb{Z}[\beta_a]$, since $g_a'(\beta_a)O_{K_a} \subseteq \mathbb{Z}[\beta_a]$. The algorithm lifts component roots at a sampled prime to a modulus greater than twice the coefficient bound, recovers the integer coefficients, and verifies the square identity in $\mathbb{Z}[\beta_a]$.

The separate rational square root of $q_I/\prod_a \delta_a$, which is positive and a square by the shared norm columns, is computed by halving the stored signed prime exponents and evaluated with ψ_{rat} from (E.6). Multiplying it with the evaluated field roots, as in (E.19), gives $Y^2 \equiv \prod_{i\in I} R_i \pmod N$. All inverses use verified modular units. Failure of reconstruction or verification rejects the epoch; the square identity is checked before any private product is formed.

F Bit complexity of the unconditional algorithm

For each epoch we bound its direct work and count its calls to the same algorithm on smaller integers. The bounds include rejected fields, failed trees and failures of bounded arithmetic routines. Each recursive call runs to completion and contributes its actual running time to the recurrence. If an initial check fails before a count is defined, we set that count to one and include the cost of the check. We follow the cost analysis of [33, `mfU_theorem`, Sections 9–10; `escapeR2`, Section 8].

F.1 Integer parameters and their admissibility conditions

We use the following integer choices for the scales in Section A.1.

Let $n = \mathrm{b}(N)$, $m_b = \mathrm{b}(n)$, and let $d \geq 3$ be the smallest odd integer satisfying $d^3 m_b \geq n$. Fix

$$\kappa = 10^7, \qquad b_0 = 100, \qquad q_* = 2^{32}, \qquad U_* = 6q_*, \qquad k = \kappa d.$$

Define

$$\begin{aligned} e_B &= b_0 d m_b, & B &= 2^{e_B}, \\ e_0 &= \left\lfloor \frac{n-5}{2k} \right\rfloor - 1, & Y_0 &= 2^{e_0}, \\ e_{i+1} &= \max\left\{ e_B, \left\lceil \frac{e_i + 2dm_b}{q_*} \right\rceil \right\}, & Y_i &= 2^{e_i}, \qquad V_i = Y_i^6. \end{aligned} \tag{F.1}$$

Stop at the first j with $Y_j = B$. Compute e_B, e_0 and test $e_0 > e_B$ before forming the powers of two. The epoch fails if this test or another initial scale check fails, if the scales cease to decrease above B, or if more than $n + 10$ levels are needed. These checks eventually pass. After the primality, perfect-power and even-factor tests, a failed scale check passes control to the persistent trial divisor. Finding a small prime ends the current call and starts new calls on the smaller factors.

F.2 Bounds on the sizes of auxiliary inputs

Every sampled field and every tested point is checked against

$$\|\mathcal{F}\|_\infty \leq C_F := 8MH, \qquad 16dMH \leq 2^{12d^2 m_b}, \qquad |\mathcal{F}(P)| \leq 2^{12d^2 m_b + 6de_i}.$$

The sampling windows imply these bounds for sufficiently large inputs. The program still checks them for each field and point. When $Y_0 > B$, the integer parameter definitions give

$$12d^2 m_b \leq \frac{6(n-5)}{\kappa b_0}, \qquad 6de_0 \leq \frac{3(n-5)}{\kappa} - 6d.$$

Since $n - 5 < \log_2 N$,

$$|\mathcal{F}(P)| \leq N^{(3+6/b_0)/\kappa} = N^{306/10^9} < N^{10^{-6}}. \tag{F.2}$$

Set

$$b_{\mathrm{sh}}(N) = \lceil \sqrt{2N} \rceil, \qquad b_{\mathrm{val}}(N) = \lceil N^{10^{-6}} \rceil.$$

We check each coordinate against $b_{\mathrm{sh}}(N)$ and each raw value against $b_{\mathrm{val}}(N)$. A recursive call also requires $2 \leq m < N$; zero and one are handled directly. The value sent for factorization is the value of the sampled form. Expanding it to the monic norm $A_f^{d-1} F(P)$ could exceed the bound.

F.3 Bounds on field data, characters and matrix dimensions

Sylvester determinant bounds give

$$\mathscr{D}_f = (2d)!(dC_F)^{2d}, \quad \mathscr{D}_* = C_F^2\mathscr{D}_f, \quad \mathscr{G}_* = (2d)!(dC_F^d)^{2d},$$

which dominate $|\operatorname{disc}(f)|$, $|D_f^*|$, and $|\operatorname{disc}(g)|$, respectively. Also $\|g\|_\infty \le C_F^d$. The discriminant formula uses only $|A_f| \ge 1$. Define

$$\begin{aligned}
\mathscr{N}_* &= (d+1)C_F^d V_0^d,\\
L_D^* &= \mathrm{b}(\mathscr{D}_*), \qquad L_g^* = \mathrm{b}(\mathscr{G}_*),\\
h^* &= d + dL_D^* + L_g^* + 2 + d\,\mathrm{b}(L_g^* + 3),\\
b_\Delta^* &= 2L_g^* + 3d + dL_D^*,\\
T^* &= 2 + d + n + b_\Delta^* + h^* + \mathrm{b}(J) + \mathrm{b}(2 + \mathscr{D}_* + \mathscr{N}_*),\\
S^* &= (2d)!(T^*)^4, \qquad t^* = 20d\{h^* + \mathrm{b}(64J) + 1\}.
\end{aligned}$$

These dominate the quantities used in Proposition E.3 for every admitted field. Before drawing the pool, the program knows the upper bound on the number of columns

$$\mathcal{C} = 2B + 2 + J\{dB + 1 + t^* + L_D^*\}. \tag{F.3}$$

Rejected fields also count in this deterministic ceiling.

Both T^* and t^* are polynomial in n. The bit length of a character prime has the larger bound

$$\log S^* = \log(2d)! + O(\log n) = (2/3 + o(1))\Lambda. \tag{F.4}$$

Thus even polynomial work in S^* contributes to the leading constant in the $L[1/3]$ bound.

F.4 Number of searches and retained relations

Use the caps q_s, r_s in (4.7) and R_{shape} in (5.2). Put

$$T_{\text{search}} = 2kR_{\text{shape}}(q_sM_c + 1)^{j+2}, \qquad P_{\text{raw}} = T_{\text{search}}r_s. \tag{F.5}$$

The first bound counts search states on all branches, including pass-throughs. The second counts proposals. Each proposal makes at most one factorization of a polynomial value, in its selected field. Locating that field by scanning the pool requires only nonrecursive work.

For retained algebraic data the tighter ceiling is

$$\mathcal{A} = 2k(\mathcal{C} + 1)(M_c + 1)^{j+2}. \tag{F.6}$$

This counts the occurrences in the final successful trees of all retained rows. Failed trees have been discarded, so q_s does not enter $\mathcal{A}$. Combining selected rows can only decrease the sum of absolute signed multiplicities.

The live state contains the pool, the current finite search stack, retained rows and complete factor records, and fixed-width locations for the private roots. Once a tree fails, its local data can be released.

F.5 Trial bounds for the random samplers

Give the environment-prime sampler $R_{\mathrm{env}} = n^3$ integer proposals. The prime number theorem makes its eventual failure probability smaller than any fixed prescribed budget. For each desired character use at most

$$R_\chi = n^2 dS^* \, \mathrm{b}(64Jt^* + 2)$$

integer-prime proposals. Proposition E.3 gives acceptance probability at least $c/(dS^*)$ for an absolute $c > 0$. Hence, uniformly over the fixed fields,

$$\mathbb{P}(\text{some character cap is exhausted}) \le Jt^* \exp\{-cn^2 \mathrm{b}(64Jt^* + 2)\} \le 1/64$$

eventually. The unknown constant affects the threshold, not the executable cap.

For other public uniform interval choices, sample fair bits of the least covering length and reject values outside the interval. Each block succeeds with probability at least one half. A ceiling for the number of these outer calls to the integer sampler is

$$K_U = 16\{1 + R_{\mathrm{env}} + J(d+1) + P_{\mathrm{raw}} + Jt^* R_\chi\}.$$

Cap each at $u_U = \mathrm{b}(128K_U) + n$ blocks. Their total failure probability is less than $1/128$. Together with the environment-prime cap it is at most $1/64$ eventually. Private shape blocks use exactly n bits, without rejection resampling.

Couple the capped public samplers with their uncapped versions using the same random bits. They agree up to the first exhausted cap. Given the past, a call to the integer sampler fails with probability at most 2^{-u_U} and is uniform conditional on returning a value. The union bound $K_U 2^{-u_U} < 2^{-n}/128$ bounds the amount subtracted from the ideal success probability.

Each finite-field routine accounts for its internal random bits under its own cap.

F.6 Direct arithmetic and exact reconstruction

Let

$$Q_\chi = 1 + R_{\mathrm{env}} + (d+2)J\{t^* R_\chi + 2^d(d+1)\}, \qquad \mathscr{P} = (n + d + S^* + 2)^{40}.$$

Here Q_χ bounds the number of primality tests, finite-field factorizations, root searches and square roots used directly in sampling and reconstruction. The current-input primality and perfect-power tests run to completion during preprocessing. Factor lists from recursive calls can be reused because those calls certify their primes. Products, valuations, gcds and the base sieve are counted below.

For a fixed input, $\mathscr{P}$ bounds the eventual expected bit cost of each direct call above. Corrected AKS uses a smaller fixed exponent. Frobenius gcds and distinct-degree factorization use polynomially many products. Equal-degree splitting separates two specified components with probability bounded below by an absolute constant. Product and irreducibility checks verify the outputs. Factoring $T^2 - a$ gives the finite-field square roots.

Give each direct call the cap $64Q_\chi\mathscr{P}$. Conditional Markov's inequality and the bound on the number of calls give total probability of exceeding the prescribed limit at most $1/64$, and the running-time bound on every execution is

$$64Q_\chi^2\mathscr{P}. \tag{F.7}$$

The factor Q_χ^2 accounts for both the number of calls and their individual caps. The routines take reduced coefficients as input. To bound the cost of constructing and reducing the integral products, put

$$R^* = 1 + C_F^d, \qquad Z^* = 1 + V_0(C_F + R^*),$$

$$
\begin{aligned}
L_0 &= 10 + \mathrm{b}(d!) + (d-1)^2\mathrm{b}(R^*) + n\mathcal{A} \\
&\quad + (d-1)\mathrm{b}(2R^*) + \mathcal{A}\,\mathrm{b}(Z^*), \\
\mathscr{L} &= (n+2)^{10}(L_0+1).
\end{aligned}
$$

Since every raw atom has fewer than n value bits, $|\delta| \le 2^{n\mathcal{A}}$. The root-height argument in Theorem E.4 and the monic quotient-ring reduction bound show that $\mathscr{L}$ bounds the *bit lengths* of all integral coefficients and the required lifting precision. A multiplication by a linear atom adds at most $O(d\,\mathrm{b}(C_F^d) + \log d + n)$ coefficient bits. Initial precision S^* is counted separately.

An explicit safe deterministic bound for reconstruction, character evaluation, and verification over all fields is

$$
2^{4d}(J+1)^2(n+d+S^*+\mathscr{L}+\mathcal{C}+\mathcal{A}+2)^{12}. \tag{F.8}
$$

Indeed schoolbook monic quotient multiplication at precision L takes $O(d^3(L+\mathrm{b}\|g\|)^2)$ bit operations. Newton lifting uses a bounded number of such operations at each doubled precision. There are at most $\mathcal{A}$ atom multiplications, at most 2^d component-sign choices, and at most $\mathcal{A}t^*$ character evaluations, each using $O((S^*)^3)$ schoolbook modular work. Polynomial CRT, Euclidean arithmetic, exact verification, exponent lists, and sequential scans fit the degree twelve with the displayed extra factors.

Each of four public failure probabilities is at most $1/64$: scalar and environment sampling; character-prime sampling; simultaneous separation of square classes; and failures of the bounded arithmetic routines. The separate induction below handles recursive integer factorizations.

F.7 Cost excluding recursive factorizations

Define

$$
\begin{aligned}
f_0(N) = (n+2)^{100}\big[&B^3 + (\mathcal{C}+1)^3 + P_{\mathrm{raw}}J(B+S^*+1) \\
&+ P_{\mathrm{raw}}(\mathcal{C}+1)^2 + 64Q_\chi^2\mathscr{P} \\
&+ 2^{4d}(J+1)^2(n+d+S^*+\mathscr{L}+\mathcal{C}+\mathcal{A}+2)^{12}\big].
\end{aligned} \tag{F.9}
$$

The polynomial prefactor includes the deterministic input classifiers. The first term covers elementary prime enumeration and construction of $\prod_{p\le B} p$. Its bit length is at most $B\log_2 B + 1$, so a denominator roughness test costs at most $B\ \mathrm{poly}(n)$. Pool access and field lookup use charged sequential scans. Dense binary elimination costs at most $(\mathcal{C}+1)^3$. Column registration, zero extension of old rows, exact factor records, fixed-schedule private initialization, and local-state archiving fit the remaining terms.

For the leading exponents define

$$
\begin{aligned}
f &= U_* + 20 = 25\,769\,803\,796, \\
r_* &= f + b_0 + 1 = 25\,769\,803\,897, \\
s_* &= r_* + 2\kappa + 10 = 25\,789\,803\,907, \\
v_* &= U_* + 10 = 25\,769\,803\,786.
\end{aligned}
$$

The following table records an upper coefficient of Λ in the logarithm of each quantity:

Quantity	Upper coefficient, up to $o(1)$
J	f
$\mathcal{C}, \mathcal{C}+1$	r_*
$R_{\mathrm{shape}}, T_{\mathrm{search}}$	s_*
r_s	v_*
P_{raw}	$s_* + v_*$
$\mathcal{A}, \mathscr{L}$	r_*
S^*	$2/3$
Q_χ	$f + 2/3$

Here $\log p_*^{-1} = (2\kappa + o(1))\Lambda$ and $j \log(q_s M_c + 1) = O((\log D)^2) = o(\Lambda)$. In particular, retained algebraic precision has exponent r_*, not the trial exponent $s_* + v_*$.

The five largest independent work bounds are

$$\begin{aligned} 3r_* &= 77\,309\,411\,691, \\ s_* + v_* + f + b_0 &= 77\,329\,411\,589, \\ s_* + v_* + 2r_* &= 103\,099\,215\,487, \\ 2f + 28 &= 51\,539\,607\,620, \\ 12r_* + 2f + 100 &= 360\,777\,254\,456. \end{aligned}$$

They are all less than

$$a_0 = 10(f + r_* + s_* + v_* + b_0 + 1001) = 1\,030\,992\,164\,870. \tag{F.10}$$

Thus $\log f_0(N) \le (a_0 + o(1))\Lambda$. This conclusion includes the positive $L[1/3]$ cost of unconditional character arithmetic.

F.8 Recursive factorizations and storage of intermediate data

Every auxiliary integer factorization runs the same factoring algorithm on independent future bits and is allowed to return naturally. There are at most $2R_{\mathrm{shape}}$ coordinate calls and P_{raw} calls on polynomial values. Use the following upper bounds

$$f_1 = 16(n+2)^{100} R_{\mathrm{shape}}, \qquad f_2 = 8(n+2)^{100} P_{\mathrm{raw}}.$$

Then

$$\log f_i \le (a_i + o(1))\Lambda, \quad a_1 = 25\,789\,803\,908, \quad a_2 = 51\,559\,607\,694. \tag{F.11}$$

The machine uses a fixed set of working tapes and a stack of local frames. On a child call it archives only its own local state, reuses the work tapes, and restores that state on return. Archived ancestors are not recopied. The stored local state has size at most

$$\exp((2r_* + o(1))\Lambda).$$

Private records have fixed widths and publicly determined positions. The term $P_{\mathrm{raw}}(\mathcal{C}+1)^2$ in (F.9), with its polynomial factor, includes saving and restoring this local state at every direct numerical call. Each recursive call returns with its own scratch tapes cleared to the agreed state, apart from its output. The total visited tape span is bounded by its own physical work, so this cleanup costs only a fixed multiple of that work; include it in $T(m)$, the complete running time. No time counter for an ancestor call is used. Only direct arithmetic routines keep a local timer.

For analysis only, write

$$\begin{aligned} E_C(N) &= 2^{C\lceil (nm_b^2)^{1/3}\rceil} = \exp((C+o(1))\Lambda), \\ F_C(N) &= f_0(N) + f_1(N)E_C(b_{\mathrm{sh}}(N)) + f_2(N)E_C(b_{\mathrm{val}}(N)). \end{aligned} \tag{F.12}$$

The program does not compute E_C or use C as a time allowance.

Lemma F.1 (Cost of an iteration and its conditional expectation)**.** *There are fixed constants c_E, N_P such that the following hold for all sufficiently large admitted N. On every path the nonrecursive epoch cost is at most $c_E f_0(N)$, the two auxiliary call counts obey the ceilings above, and*

$$\mathrm{Cost}(\mathrm{Epoch}(N)) \le c_E \left\{ f_0(N) + \sum_{\text{coordinate calls}} T(m) + \sum_{\text{calls on polynomial values}} T(m) \right\}. \tag{F.13}$$

The sums count actual call occurrences. If $A_0 \ge 1$ and $\mathbb{E}T(m) \le A_0 E_C(m)$ for every $2 \le m < N$, then, conditional on any past before a new epoch,

$$\mathbb{E}[\mathrm{Cost}(\mathrm{Epoch}(N)) \mid \textit{past}] \le c_E A_0 F_C(N). \tag{F.14}$$

The functions are explicit and satisfy $\log f_i \le (a_i + o(1))\Lambda$ with a_0, a_1, a_2 in (F.10)–(F.11)*.*

If the arithmetic interface of Proposition A.1 holds, then a fresh complete epoch succeeds with probability at least $p_0 = 1/128$ for every $N \ge N_P$, conditional on any prior past. The threshold N_P does not involve an expectation constant.

Proof. Every nonrecursive loop is subject to the stated finite bound; every numerical input obeys (F.2) and the requirement that the input be strictly smaller. The nonrecursive arithmetic, copies of the stored local states, and direct timers fit f_0. The recursive calls contribute precisely their actual work, proving (F.13). Their termination on every tape is proved by input induction and the persistent backup in Theorem 6.1.

Before each auxiliary call its input is a fixed public smaller integer, and fresh future bits give the stated smaller-input expectation bound. Use monotonicity of E_C, sum over the deterministic call ceilings, and assign zero cost to unused positions. The tower property proves (F.14); no independence between different inputs or between an epoch's cost and success is required.

For success, each numerical factorization returns correctly. This makes every candidate predicate exact, so Lemma 4.3 applies to the child searches. Eventually the ideal environment and pool succeed with probability at least $3/4$, and conditional row collection succeeds with probability at least $3/4$. Deleting the four public exceptional events leaves mass at least $9/16 - 4/64 = 1/2$. The hidden-root argument retains at least half this mass, giving $1/4$; use the weaker $1/128$. Immediate proper factors retain their entire mass. No independence of the exceptional events is used. All thresholds come from fixed arithmetic and source-range estimates, not from the running times of the calls on smaller inputs. □

F.9 The unconditional factorization algorithm

After preprocessing (Section A.1), one epoch on the current integer N proceeds as follows. At any step a proper factor is verified by division and returned immediately.

1. Compute the integer parameters and admissibility conditions. Sample one rational environment and the fixed number J of independent raw forms. Test the setup gcds and leave each rejected field in its original position.

2. Sample the quadratic characters for each accepted field using the capped exact sampler. Fix these characters before collecting any rows.

3. Make at most R_{shape} row attempts. Generate a fresh private root block on the fixed private schedule, release its square, and compute the canonical shortest vector of its planar lattice. Before dividing out the common divisor of the two coordinates, check that it is a unit modulo N. Factor the normalized coordinates on smaller inputs and check the squarefree top-band condition.

4. For each requested top prime, choose a fresh uniform trial for each proposal and sample its common geometric point domain. Search using the caps r_s, q_s. Factor each tested numerical value by a smaller-input call. A verified candidate initiates the required child searches in the same number field; a failed tree is discarded. Pass-through states obey the same finite limits.

5. Verify each completed row identity and retain its fresh private index. Stop after $\mathcal{C} + 1$ rows. If the row-attempt limit is reached first, return failure.

6. Choose a deterministic nonzero binary dependence and reconstruct the public square root Y. Reject if an exact verification or a direct arithmetic limit is exceeded. Only then multiply the selected private roots to obtain X. Return $\gcd(X - Y, N)$ if proper; otherwise erase the stored private roots and return failure.

Before every recursive call, check that its argument m satisfies $2 \le m < N$ and the corresponding size bound. Each call runs to completion. Appendix F bounds the number of calls and the direct arithmetic in both successful and unsuccessful executions.

At a composite node, Split(N) repeats fresh epochs. After the first failed epoch it tests the trial divisor $t = 2$, after the next failure it tests $t = 3$, and so on, returning a verified factor as soon as one divides N. A test and increment take $\mathrm{poly}(\log N)$ bit operations. The trial-divisor state is not reset. At most $\lfloor \sqrt{N} \rfloor$ failed epochs can occur before trial division finds a factor. Section 6 uses this fact for termination on every random tape. To bound expected cost, it conditions on the history before each call and fixes the large-input threshold before choosing a bound for the finitely many smaller inputs.

G Relation and incidence estimates under GRH

In this appendix we prove the counting estimates needed for procedure G. The aim is to show that with probability tending to one, a randomly chosen form admits sufficiently many usable relations for a positive proportion of the initial primes in each prescribed dyadic interval. We first establish concentration estimates for a prescribed prime and root, and then bound the number of occurrences of exceptional prime ideals in the relations at the preceding level.

We use the original distribution of the rational environment. In Sections 3 and 4 the unconditional analysis uses the conditioning of Corollary B.2 and the incoming means in (3.4). Here we use deterministic weights and Gaussian coefficients. Under GRH, the estimates below hold uniformly over the stated inputs, levels, primes, roots, and points above one input threshold. Appendix H specifies the parameter checks and the backup for smaller inputs.

G.1 Polynomial distribution and relation counts

The following data are shared with the finite program in Appendix H. For positive integers write $b(v) = \lfloor \log_2 v \rfloor + 1$, and set $b(0) = 1$, $n = b(N)$, $m_b = b(n)$, $D = \log N / \log\log N$, $\lambda = \log\log N$, and $\Lambda = D^{1/3}\lambda$. Let $d \geq 3$ be the least odd integer with $64d^3 m_b \geq 27n$, and put

$$t = dm_b, \quad \Lambda_d = (\log 2)t, \quad k_0 = \lceil 9d/10 \rceil, \quad e_B = \lceil 19t/15 \rceil, \quad e_0 = \left\lfloor \frac{n-5}{2k_0} \right\rfloor - 1.$$

Define the cutoffs and boxes by

$$\begin{aligned} e_{i+1} &= \max\left\{ e_B, \left\lceil \frac{153e_i + 247e_B + 1125t}{1290} \right\rceil \right\}, & Y_i &= 2^{e_i}, \\ B &= 2^{e_B}, & V_i &= 2^{\lceil (153e_i + 247e_B)/300 \rceil}. \end{aligned} \tag{G.1}$$

Let j be the first index with $e_j = e_B$. Active internal boxes have $1 \leq i < j$ and cutoff $y_i = Y_{i+1}$; the separate top box is V_0, with cutoff $y_0 = Y_1$. Define

$$\begin{aligned} A &= (j+10)^3, & m &= 1000A, & M &= \lfloor N^{1/d} \rfloor, \\ \sigma &= V_0^2 n^{10}, & H &= \sigma n^{10}, & h &= 2H+1, \\ e_{\min} &= \lceil 127e_0/128 \rceil, & Y_{\min} &= 2^{e_{\min}}, & & \\ C_F &= 8MH, & T_H &= \lfloor 15d^2 m_b/4 \rfloor. & & \end{aligned} \tag{G.2}$$

Thus $d \sim (3/4)D^{1/3}$, $j = O(\log D)$, and $A = D^{o(1)}$. The program uses the following dyadic lower bounds on the densities:

$$a_A = \lceil 36t/25 \rceil, \quad p_A = 2^{-a_A}, \qquad a_{\mathrm{den}} = \lceil 25t/12 \rceil, \quad p_{\mathrm{den}} = 2^{-a_{\mathrm{den}}}. \tag{G.3}$$

The comparison environment law μ chooses a prime $Q \in (3M/4, M]$, $Q \nmid N$, and an independent uniform integer $r \in (M/2, 3M/4]$. Deleting $Q \mid N$ costs a relative $1 + o(1)$; such a proper divisor is an immediate program success. Put $G(X,Y) = QX - rY$. The construction in Section A supplies F_0 with $F_0(r,Q) = N$ and irreducible degree-d reduction modulo two. The raw form is

$$\mathcal{F} = F_0 + 2GR_{\mathbf{c}}, \qquad R_{\mathbf{c}}(X,Y) = \sum_{\nu=0}^{d-1} c_\nu X^{d-1-\nu} Y^\nu. \tag{G.4}$$

First consider the ideal law $\mathbb{P}(\mathbf{c}) \propto e^{-\pi\|\mathbf{c}\|^2/\sigma^2}$ on $\mathbb{Z}^d$. The sampling distribution ν_g, from Lemma H.1, has support $[-H,H]^d$, one failure outcome if coefficient sampling exceeds its

prescribed limit, and total variation error at most $e^{-n^{10}}$. On failure we set all root indicators and counts to zero. One raw form is used at all levels. Its primitive part is denoted by F in the algebraic program. Every coefficient-law estimate below concerns $\mathcal{F}$ under the original law, before acceptance during initialization, certification, or successful execution is imposed.

Write $x_i = MV_i/8$. For an internal $l = (\ell, s)$, $B < \ell \le Y_i$, define

$$\Omega_l = \left\{ \begin{array}{c} V_i/2 < b \le 3V_i/4, \\ (a,b) \in \mathbb{Z}^2 : w(P) := rb - Qa = -G(P) \in [x_i/2, x_i], \\ a \equiv sb \pmod{\ell}, \qquad \ell \nmid b \end{array} \right\}. \tag{G.5}$$

The rationally eligible subset $\mathcal{R}_i(g,l)$ requires that b be prime and $w(P) \in \mathcal{S}_B$. Its coefficient-root event is

$$R_{\ell,s} = \{\mathcal{F}(s,1) \equiv 0 \pmod{\ell}\}, \qquad s \neq rQ^{-1} \pmod{\ell}.$$

Numerical primes dividing N are deleted in advance. At a top private parent p, use $i = 0$, $\ell = p$, and $s = rQ^{-1} \pmod{p}$ in (G.5); thus $p \mid w(P)$. The inequalities for the search regions imply $p < Q$. The rational predicate is $w(P) \in p\mathcal{S}_B$, and there is no coefficient-root condition at p. Private bands are complete dyadic bands in $(Y_{\min}, Y_0]$. The program checks the inequalities in (H.2); Proposition G.13 proves their eventual validity. They allow $V_i \le Y_i$, so the program uses the exact sampler of Lemma H.2.

For an internal active label put $f_P(u) = \mathbf{1}_{\{0<|u|\in\mathcal{S}_{y_i}\}}$ and define

$$Z_l = \sum_{P \in \mathcal{R}_i(g,l)} f_P(\mathcal{F}(P)/\ell), \qquad a_P = \mathbb{E}_{\nu_g}[f_P(\mathcal{F}(P)/\ell) \mid R_l],$$
$$\tau_l(g) = \mathbb{E}_{\nu_g}[Z_l \mid R_l] = \sum_{P \in \mathcal{R}_i(g,l)} a_P.$$

At the top replace the divisor ℓ by one and use unconditioned means, with cutoff Y_1. Formal incidences follow the convention preceding (A.4): even valuations, raw-content primes and children that are terminal in the number field are retained, with terminal bad indicators zero. Small primes and terminations with a nontrivial factor are absorbing, outside the normalized domain; pass-through has an identity edge.

Theorem G.1 (Relation counts and good top parents)**.** *Assume GRH. For all sufficiently large preprocessed inputs, draw the environment and the single form from the distributions above, using the indicated scales. For each prescribed target there is an event depending only on the rational environment, called rational-good, on which*

$$\tau_l \ge e^{\Lambda_d/50}, \qquad \tau_l/|\Omega_l| \ge p_{\mathrm{den}}, \qquad a_P \ge p_A \quad (P \in \mathcal{R}_i(g,l)). \tag{G.6}$$

At each fixed environment satisfying the rational-count condition and label,

$$\mathbb{P}_{\nu_g}(|Z_l - \tau_l| > \tau_l/2 \mid R_l) \le D^{-3A}.$$

The top version has no coefficient-root conditioning. The same tail holds for every normalized complete-band semismooth count in its parent search region under its single child-root law.

Define the good states in this field, starting at the last level. Every terminal state is good. A pass-through state is good if its copy at the next level is good. An active state must be rational-good and satisfy

$$D_l \ge \tau_l/4, \qquad Z_l \le 2\tau_l,$$

where D_l counts candidates whose required children are good. With probability $1-o(1)$ over $(g,\mathcal{F})$, every prescribed private band has at least three quarters of its top parents good. In particular this probability is at least $3/4$ eventually. These sets are fixed before private shapes are drawn; the same field and child-good sets work for every band. Every good active or top state satisfies $D_l/Z_l \geq 1/8$ and $Z_l/|\Omega_l| \geq p_{\mathrm{den}}/4$. There are at most $n+1$ formal child incidences per candidate.

Fix a field for which the conclusion holds. Whenever Search reaches a good state, the stated lower bounds apply, however the previous choices led to that state. The comparison measures below are used to count the good top parents.

G.2 Smooth numbers in arithmetic progressions under GRH

Lemma G.2 (Sharp smooth-character sums)**.** *Under GRH, for each fixed $\epsilon > 0$, every nonprincipal Dirichlet character $\chi \bmod k$, including imprimitive characters, satisfies, uniformly for $x, y \geq 2$,*

$$\left| \sum_{\substack{n \leq x \\ P^+(n) \leq y}} \chi(n) \right| \leq C_\epsilon x^{1/2+\epsilon} k^\epsilon. \tag{G.7}$$

If $I \subset [x, 2x]$, $a \bmod k$, and $\delta = (a,k)$, then the smooth progression count is zero unless $\delta \in \mathcal{S}_y$. Otherwise

$$\sum_{\substack{n \in I \\ n \equiv a \ (k)}} \mathbf{1}_{\mathcal{S}_y}(n) = \frac{1}{\varphi(k/\delta)} \sum_{\substack{n \in I \\ (n,k)=\delta}} \mathbf{1}_{\mathcal{S}_y}(n) + O_\epsilon(x^{1/2+\epsilon} k^\epsilon). \tag{G.8}$$

Reflected negative intervals are allowed.

Proof. It suffices to consider $0 < \epsilon < 1/4$. For primitive nonprincipal conductor r, [36, Proposition 5.1] bounds the truncated Euler product by $O_\eta((r(2+|\operatorname{Im} s|))^\eta)$ for $\operatorname{Re} s \geq 1/2+\eta$, uniformly in y. Inducing to modulus k adds factors $(1-\chi^*(p)p^{-s})$, $p \mid k$, $p \nmid r$. Their product is $O_\eta(k^\eta)$, by separating finitely many small primes and using $1 + p^{-1/2-\eta} \leq p^\eta$ for the others.

Use truncated Perron at $X = \lfloor x \rfloor + 1/2$, height X^2, and real part $1 + 1/\log X$. Bounded coefficients give truncation error $O(\log^2 X/X)$. The finite Euler product is analytic in the positive half-plane. Shifting to $1/2+\eta$ gives a vertical integral $O_\eta(X^{1/2+3\eta}k^{2\eta})$ and smaller horizontal integrals. Take $\eta = \epsilon/4$, and subtract prefixes for intervals. To prove (G.8), divide by δ and apply orthogonality modulo k/δ. The number of nonprincipal characters cancels its denominator. Their zeros implement all coprimality restrictions. Reflection multiplies a character sum by $\chi(-1)$. □

For a complete prime band $J \subset (P, 2P]$, $P \geq y$, and $0 < w_\ell \leq 1$, fixed before the coefficients and of bounded within-band ratio, set

$$f_J(u) = \sum_{\ell \in J} w_\ell \mathbf{1}_{\{0<|u| \in \ell \mathcal{S}_y\}}. \tag{G.9}$$

The summands have disjoint supports, so $0 \leq f_J \leq 1$. When a normalized band is used below, we divide its weights by their maximum. The progression error for this test is at most

$$C_\epsilon P^{1/2-\epsilon} x^{1/2+\epsilon} k^\epsilon. \tag{G.10}$$

If δ is smooth, expand by ℓ and sum $(x/\ell)^{1/2+\epsilon}$; terms with $\ell \mid k/\delta$ vanish. If δ has one prime above y, it must be one parent to exponent one, and only that summand survives. All other contents give zero. A further power of the parent in k/δ is retained.

G.3 Mean-square estimates on the rational side

Fix a box V, put $x = MV/8$, and define

$$\mathcal{A} = [x/2, x] \cap \mathcal{S}_B, \quad \rho = |\mathcal{A}|/x, \quad \mathcal{B} = \{b \text{ prime} : V/2 < b \le 3V/4,\ b \ne \ell\}.$$

Partition both intervals into K fixed cells of comparable length, before Q, r are chosen; eventually $K = \lceil \Theta \rceil$, $\Theta = \exp(D^{1/6}\lambda^2)$. All numerator and denominator values are units modulo $Q\ell$. For unit v, let

$$C_Q(v) = \#\{(u, b) \in \mathcal{A} \times \mathcal{B} : u \equiv vb \pmod{Q\ell}\}.$$

Retaining the characters with principal Q-component gives exactly

$$L_Q(v) = \frac{1}{Q-1} \sum_{u \in \mathcal{A}} \#\{b \in \mathcal{B} : b \equiv u/v \pmod{\ell}\} = \frac{1}{Q-1} \sum_{b \in \mathcal{B}} \mathcal{A}_\ell(vb). \tag{G.11}$$

Since $\ell > B$, Lemma G.2 gives

$$\mathcal{A}_\ell(a) = |\mathcal{A}|/(\ell - 1) + O_\epsilon(x^{1/2+\epsilon}\ell^\epsilon).$$

The whole-box relative error is therefore

$$O_\epsilon\left(\frac{\ell^{1+\epsilon}}{\rho x^{1/2-\epsilon}}\right). \tag{G.12}$$

A numerator cell needs only the upper bound $|\mathcal{A}_{\text{cell}}| \ll \rho x/K$, supplied by (B.26); its normalized error costs one extra K. Prime counting under RH at modulus one gives

$$|\mathcal{B}_{\text{cell}}| = \frac{|I_{b,\text{cell}}|}{\log V} + O\left(\sqrt{V}\log V + \frac{V}{K\log^2 V} + 1\right).$$

This follows by subtracting endpoints in the RH prime-counting estimate [48, Ch. 13]. Here $K\log^2 V = o(\sqrt{V})$. Deleting $b = \ell$ costs one. There is no prime progression estimate modulo ℓ.

Lemma G.3 (Rational-side estimates averaged over roots)**.** *Suppose the cell version of* (G.12) *is* $o(1)$*. Let* $\mathcal{E}_i(g, \ell, s)$ *be failure of the whole-box lower or upper rational count or of any prescribed cell upper bound. This event is intersected with admissibility; at the excluded root take it to be false. Set inadmissible populations to zero, and put*

$$T_i = V_i^2\rho_i/\log V_i, \quad \mathfrak{m}_i(\ell) = T_i/\ell, \quad R_i(g, \ell, s) = |\mathcal{R}_i(g, \ell, s)|.$$

Then

$$\ell^{-1} \sum_{s \bmod \ell} \mathbb{P}_\mu(\mathcal{E}_i(g, \ell, s)) \le \bar{\epsilon}_i(\ell), \tag{G.13}$$

$$\ell^{-1} \sum_{s \bmod \ell} \mathbb{E}_\mu[(R_i/\mathfrak{m}_i(\ell))^2 \mathbf{1}_{\text{admissible}}] \le C(1 + \bar{\epsilon}_i(\ell)), \tag{G.14}$$

$$\bar{\epsilon}_i(\ell) = CK^4 \frac{\log M \log V_i}{\rho_i}\left(\frac{\ell}{V_i^2} + \frac{\ell}{M}\right). \tag{G.15}$$

On $\mathcal{E}_i^c$*, the whole count is* $\asymp \mathfrak{m}_i(\ell)$*, a cell has at most* $C\mathfrak{m}_i(\ell)/K^2$ *points, and the number of pairs with* $|\det(P_1, P_2)| \le cV_i^2/K$ *is* $O_c(\mathfrak{m}_i(\ell)^2/K)$*. This includes auxiliary child roots* $B < \ell \le y_i$ *in the same parent search region.*

Proof. Put $H_Q = C_Q - L_Q$, $E_{\mathcal{A}} = |\mathcal{A}||\mathcal{B}|$, and $T = xV$. The ratio convolution has the same character absolute value as the product convolution $c_u = \#\{(a,b) \in \mathcal{A} \times \mathcal{B} : ab = u\}$. Since $b > B$ is prime and a is smooth, $c_u \in \{0,1\}$, $\sum c_u^2 = E_{\mathcal{A}}$. It is independent of Q. Orthogonality and the primitive-Q large sieve in (B.19) give

$$\sum_{Q} \sum_{v \in (\mathbb{Z}/Q\ell\mathbb{Z})^\times} |H_Q(v)|^2 \le CM^{-1}(M^2 + T)E_{\mathcal{A}}.$$

Nonprincipal characters modulo prime Q are primitive, and the $\ell - 1$ twists cancel the orthogonality denominator.

For fixed Q, the map $(r,s) \mapsto r - Qs \pmod{Q\ell}$, with $M/2 < r \le 3M/4$ and $s \bmod \ell$, is injective: first reduce modulo Q, using the interval diameter $< Q$, and then modulo ℓ. Its admissible image is in the unit group. Consequently

$$\sum_{s \bmod \ell} \mathbb{E}_{Q,r} |H_Q(r - Qs)|^2 \mathbf{1}_{\text{admissible}} \le C \frac{\log M (M^2 + T) E_{\mathcal{A}}}{M^3}. \tag{G.16}$$

This bounds the sum over roots. Dividing by ℓ is essential. For the whole intervals, $E_{\mathcal{A}} \asymp \rho MV^2/\log V$ and $L_Q \asymp \mathfrak{m}_i(\ell)$. Chebyshev after division by $\ell \mathfrak{m}_i(\ell)^2$ proves the whole-box assertion. For one cell pair the energy gains K^{-2} and the squared threshold gains K^{-4}; union over K^2 cells gives K^4. Adding the square of the uniformly comparable principal term proves (G.14). Deleting $Q \mid N$ costs $1 + o(1)$.

The congruence gives the unique $a = (rb - u)/Q$. Since $(u,b) = 1$, its point is primitive, with $V/8 \le a \le 2V/3$. The determinant strip through a fixed point meets $O(K)$ prescribed cell pairs, as in the covering after (B.20); hence it contains $O(\mathfrak{m}_i(\ell)/K)$ eligible second points. Summation gives the thin-pair bound, without a further probability union. □

Rational numerators for one top parent and for a whole band. For a complete private band $J \subset (P, 2P]$, define

$$\mathcal{A}_p = [x_0/(2p), x_0/p] \cap \mathcal{S}_B, \quad \rho_p = p|\mathcal{A}_p|/x_0,$$
$$\mathcal{A}_J = \coprod_{p \in J} p\mathcal{A}_p, \quad \rho_J = \sum_{p \in J} \rho_p/p \asymp \rho_P/\log P.$$

Within a band $\rho_p \asymp \rho_P$. The aggregate product convolution has multiplicity one: its primes above B are the denominator $b < Y_{\min}$ and parent $p > Y_{\min}$, with distinguished roles. All these points are primitive. For a child $\ell \le Y_1$, expand the positive sum by parents and apply (G.8) to their smooth cofactors. Summing $p^{-1/2-\epsilon}$ gives relative cell error

$$\ll_\epsilon \frac{K\ell^{1+\epsilon}}{\rho_P (x_0/P)^{1/2-\epsilon}}. \tag{G.17}$$

Thus Lemma G.3 applies with V_i, ρ_i replaced by V_0, ρ_J.

An individual top parent uses $pu \equiv rb \pmod{Q}$. Its convolution is ub, of length $O(MV_0^2/p)$, and its main term is $|\mathcal{A}_p||\mathcal{B}|/(Q-1)$. The same variance calculation, including cells, gives

$$\mathbb{P}_\mu(\mathcal{E}_{0,p}) \le CK^4 \operatorname{polylog}(N) \rho_p^{-1} \left(\frac{p}{V_0^2} + \frac{1}{M} \right). \tag{G.18}$$

There is no coefficient-root condition at p. Averaging (G.18) over the parents in a band bounds the expected fraction of exceptional parents.

G.4 Gaussian distributions of polynomial values

Lemma G.4 (Gaussian images and conditioning)**.** *For one or two distinct primitive points $P_r = (a_r, b_r)$ in $V \le V_0$, let E have rows $v(P_r) = (a_r^{d-1}, a_r^{d-2}b_r, \dots, b_r^{d-1})$. Put $\Gamma = E\mathbb{Z}^d$, $B_E = 2\,\mathrm{diag}(G(P_r))E$, and $\Sigma = B_E B_E^T$. The ideal value law is pointwise within a multiplicative $1 + O(e^{-n^{10}})$ of*

$$\mathcal{Z}^{-1} \exp\Big(-\frac{\pi}{\sigma^2}(z - F_0(P))^T \Sigma^{-1}(z - F_0(P))\Big) \tag{G.19}$$

on the exact support $F_0(P) + 2\,\mathrm{diag}(G(P_r))\Gamma$. The same relative comparison holds after restriction to any nonempty event determined by this image. Under a common root $R_{q,s}$, divided values have center $F_0(P)/q$, shape Σ/q^2, and the actual divided affine support. The center need not belong to that support.

Proof. The full integer two-point kernel has the basis of coefficient vectors

$$(b_1X - a_1Y)(b_2X - a_2Y)X^{d-3-\nu}Y^{\nu}, \quad 0 \le \nu \le d-3.$$

The primitive linear factors divide every kernel polynomial over $\mathbb{Q}$; Gauss's lemma makes the quotient integral. Thus no index is missing. These basis vectors have norm at most $\sqrt{6}V^2 < 3V^2$. The one-point kernel uses shifted multiples of $bX - aY$, of norms at most $\sqrt{2}V$. For a rank-r lattice with basis norms at most B_0, every nonzero dual vector has length at least $1/B_0$. Disjoint balls imply at most $(2h+3)^r$ dual vectors of norm less than $(h+1)/B_0$. Hence

$$\sum_{u \in L^* \setminus \{0\}} e^{-\pi\sigma^2\|u\|^2} \le 2e^{-\pi(\sigma/B_0)^2 + 2r} \le e^{-n^{10}}.$$

Writing $\rho_\sigma(S) = \sum_{u \in S} e^{-\pi\|u\|^2/\sigma^2}$, Poisson summation gives, uniformly in the shift,

$$\rho_\sigma(L + a) = \sigma^r (\det L)^{-1}(1 + O(e^{-n^{10}})).$$

The minimum-norm real preimage of z is $B_E^T\Sigma^{-1}(z - F_0(P))$, orthogonal to the kernel. Its squared-norm exponential factors from the shifted-kernel mass. Normalization proves the result. Restriction to a subset changes only the normalizer. A root-compatible integral lift describes the support but does not move the Gaussian center. □

Root probabilities and finite implementation. Each admissible root equation has a unit coefficient in c_{d-1}. At most three linear root/parent equations are used together. Their primes are all at most Y_1: top private $p \mid G(P)$ does not condition $\mathcal{F}$; top incoming uses a child $q \le Y_1$; internal equations occur only at $i \ge 1$. If m_{root} is the product of their distinct primes, Poisson summation modulo m_{root} makes coefficient residues pointwise $1 + o(1)$ times uniform, since $\sigma/m_{\mathrm{root}} \to \infty$. A compatible system of total local rank at most three has probability at least $(2Y_1^3)^{-1}$. In particular

$$\nu_g(R_{q,s}) = (1 + o(1))/q, \qquad c \le (q-1)\nu_g(R_{q,s}) \le C \tag{G.20}$$

on admissible roots. The excluded root has probability zero. Conditioning the finite sampler costs $O(Y_1^3 e^{-n^{10}})$ in total variation distance. This error remains negligible after normalizing by the densities and taking the m-th roots below. The separate relative comparison for the ideal Gaussian image requires no conditioning loss.

Lemma G.5 (Separate periods and common content)**.** *For two rationally eligible points with $(N, G_1G_2) = 1$ in a common admissible q-root lattice, let $\Delta = |\det(P_1, P_2)| \neq 0$, $G_r = G(P_r)$, and $k = \Delta/q$. At the top put $q = 1$. The divided affine support has the form*

$$z_r = c_r^* + 2G_r u_r, \qquad u_1 \equiv \omega u_2 + b^* \pmod{k}, \qquad (\omega, k) = 1.$$

Its separate periods are

$$m_r = 2|G_r|k, \tag{G.21}$$

with exactly k phase pairs modulo $m_1 \times m_2$. Delete from k all prime powers whose prime divides $2G_1G_2$, obtaining $k_\perp$. Under uniform phases,

$$(z_1, m_1) = (z_2, m_2) = \delta \mid k_\perp, \qquad \mathbb{P}(\delta) = \varphi(k_\perp/\delta)/k_\perp. \tag{G.22}$$

Higher unit digits at shared G_r-primes need not be independent.

Proof. A unimodular variable change sends the points to $(1, 0)$ and (u, Δ), $(u, \Delta) = 1$, and induces a unimodular coefficient change. Its evaluation image is $\{(v_1, v_2) : v_2 \equiv u^{d-1}v_1 \pmod{\Delta}\}$. This identifies the lattice, not a Gaussian invariance. Choose a root-compatible coefficient lift before dividing by q; the support and periods follow. At a prime dividing G_r, the point is proportional to (r, Q), and $F_0(r, Q) = N$ is a unit. Thus every supported value is a unit there. Also $G_1 \equiv u'G_2 \pmod{\Delta}$ for a unit u'. A prime dividing k and one G_r divides both. Unequal G_r-valuations force $v_p(\Delta) = \min(v_p(G_1), v_p(G_2))$. At the other odd primes the value-coordinate relation is a homogeneous unit-ratio congruence modulo the full power in k. Its common capped valuation is geometric. Chinese remaindering the uniform phase parameter proves (G.22). □

Ellipses for the joint value distribution. The boxes imply $1/6 \leq a/b \leq 11/12$. Writing $v(P) = b^{d-1}(s^{d-1}, \ldots, s, 1)$, both the final vector and its derivative have bounded norms; its last two coordinates show that the angle between rows has sine $\asymp \Delta/(b_1b_2)$. Put $\vartheta = b_1b_2/\Delta$ and $T_r = 2\sigma|G_r|\|v(P_r)\|/q$. The marginal Gaussian parameters are T_r, the conditional ones are $\asymp T_r/\vartheta$. More precisely,

$$\frac{T_{r|3-r}}{m_r} \geq \frac{\sigma V^{d-3}}{\sqrt{2}\, 2^{d-2}}, \tag{G.23}$$

because projection to the last two columns gives

$$\operatorname{dist}(v(P_1), \mathbb{R}v(P_2)) \geq |b_1|^{d-2}\Delta/\sqrt{a_2^2 + b_2^2}.$$

The determinant, G_r, and root divisor cancel.

For the ellipse $\mathcal{E}_R$ of Mahalanobis radius R, vertical progression summation along the first-coordinate progression gives

$$|\mathcal{E}_R \cap \mathcal{L}| = \frac{\operatorname{area}(\mathcal{E}_R)}{(2|G_1|)m_2} + O\left(\frac{RT_2}{\vartheta m_2} + \frac{RT_1}{2|G_1|} + 1\right). \tag{G.24}$$

The chord-length function has total variation at most twice its maximum. The error is relatively small uniformly for $y^{-C_R} \leq R \leq n$, for every fixed C_R, by (G.23). Using

$$e^{-\pi r^2} = \int_r^\infty 2\pi R e^{-\pi R^2}\, dR$$

expresses the Gaussian as a mixture of uniform laws on the lattice points of the ellipses. Each mixing weight includes the number of lattice points in its ellipse. Its normalizer is $(1 + o(1))\sqrt{\det(\sigma^2\Sigma/q^2)}/\operatorname{covol}(\mathcal{L})$. The masses of radii below y^{-C_R} and above n are $O(y^{-2C_R})$ and $e^{-\Omega(n^2)}$, with negligible progression endpoints. The ellipse retains the dependence between both values.

G.5 Local densities and concentration of the relation counts

Here $V = V_i$, $y = y_i$, $Y = Y_i$; for the top density notation set $y_0 = Y_1$ and use root divisor $q = 1$. At level i define the following height and reference density:

$$\mathcal{X}_i = 16(n+2)^{12}\sigma M V_i^d, \qquad c_i(u) = \widehat{\rho}(2\mathcal{X}_i/u, y_i). \tag{G.25}$$

At the top write $c_0 = c_0(1)$. This reference is fixed before the environment, form, and root are drawn. The larger bound $2^{T_H}V_i^d$ controls integer size, but its slack can change the reference density by more than a polynomial factor. For a particular point let

$$T_P = 2\sigma|G(P)|\|v(P)\|/q, \qquad X_P = 8n^{11}T_P.$$

The logarithmic ratio of X_P to $\mathcal{X}_i/q$ is $O(d + \log n)$. All reference endpoints, including root and parent divisions, satisfy

$$cd \le u = \frac{\log x}{\log y} \le Cd, \qquad 0 \le \gamma = 1 - \beta(x,y) \le C/d, \qquad \gamma \log y = O(\log D). \tag{G.26}$$

Equations (B.27) and (B.28) therefore make these densities comparable by fixed factors. The endpoints themselves need not have a bounded ratio. The same classical estimates give $\widehat{\rho}(x,y) \ge e^{-C_\rho \Lambda_d}$ throughout this range, before any coefficient-law comparison.

Lemma G.6 (Unit restrictions and actual point means)**.** *For a rationally eligible point put $G_P = |G(P)|$. On odd integers coprime to G_P, the smooth reference density is $c_o(x,y)R_{G_P}(\beta)$, where*

$$c_o(x,y) = 2(1 - 2^{-\beta(x,y)})\widehat{\rho}(x,y), \quad R_{G_P}(\beta) = \prod_{\substack{p|G_P \\ 3\le p\le y}} \frac{1-p^{-\beta}}{1-p^{-1}} \prod_{\substack{p|G_P \\ p>y}} (1-p^{-1})^{-1}.$$

Uniformly,

$$|\log R_{G_P}(\beta)| \le C\gamma(1 + \log\log(3G_P)) + \frac{C\log G_P}{y^\beta \log y} = D^{-1/3+o(1)}. \tag{G.27}$$

Under the actual finite single-root law, for a parent $y < \ell \le Y$, $\ell \nmid G(P)$,

$$\mathbb{E}_{\nu_g}\left[\mathbf{1}_{\{0<|\mathcal{F}(P)|/q\in\ell\mathcal{S}_y\}} \mid R_{q,s}\right] \asymp \ell^{-1}c_i(q\ell). \tag{G.28}$$

The case $\ell = 1$ is the ordinary test. This is pointwise in g, P, without a lower bound on the number of points. If $\ell \mid G(P)$, the corresponding summand is zero.

Proof. Differentiating $\log(1-p^{-t})$ between β and 1 bounds the first product's absolute logarithm by $C\gamma\sum_{p|G_P}(\log p)p^{-\beta}$. Split at $z_0 = \max(3, (\log(2G_P))^2)$. The lower part is $O(\log z_0)$, since $\gamma\log\log(3G_P) = o(1)$; the upper part is at most $z_0^{-\beta}\log G_P$. Primes above y cost $C\log G_P/(y^\beta\log y)$.

For the finite formula, in Lemma C.4 put all primes of G_P among the avoidance primes, not in the forced divisor. Truncate avoidance divisors at E_{cut}, as in (C.35). Each small prime occurs once, so the absolute/signed Euler ratio is at most $C\log^2(2z)$. Larger avoidance primes contribute at most $\omega(G_Pk)z^{-\beta+1/\log z} = o(1)$. Rankin's tail is still

$$C(L/x)^{-1}\log^2(2z)\exp(-\log E_{\mathrm{cut}}/\log z).$$

Thus (C.19)–(C.20) apply with the same forced-divisor condition $\log(S_0E_{\mathrm{cut}}) = o(\log x)$. Substituting G_P for S_0 would be incorrect.

For the point mean, first apply (G.8) on the actual progression modulo $2G_P$, dividing by ℓ when a parent is prescribed. The preceding finite density formula applies to its principal term. On a central interval of length comparable to T_P the Gaussian weight is comparable to $1/T_P$, and

$$|F_0(P)/q|/T_P \le \mathrm{poly}(d)2^d/\sigma = o(1).$$

The direct interval estimate (B.26) gives the lower mean. Dyadic annuli or layer cake give the upper mean. The ratio estimate compares both with (G.28). The separate-period budgets below make the GRH and endpoint errors negligible even relative to $\ell^{-1}c_i(q\ell)$. The finite conditional TV cost follows (G.20). Positive summation proves the complete weighted-band version. □

The conditional smoothness density. Use the capped-prime modulus Q_o and multiplicative grid of Definition C.2, with the present A, d, Θ, so that $z = R = D^{8A}$ and $\eta = D^{-10A}/(C_g d\Theta)$. Put $S(u) = (u, Q_o)$, and choose the remaining cutoffs by

$$\begin{aligned} &\log U = C_U \log(2z)(\Lambda_d + A\log D), \quad \log E_{\mathrm{cut}} = C_{\mathrm{cut}} \log(2z)(\Lambda_d + A\log D), \\ &\nu = \left\lceil \frac{\log Y}{\lambda} A^4 (\log D)^4 \right\rceil, \quad E_\nu = \frac{(\sum_{3\le p\le z} p^{-1})^\nu}{\nu!}, \quad B_\nu = (1+z)^{\nu+1}. \end{aligned} \tag{G.29}$$

The fixed constants $C_U, C_{\mathrm{cut}}, C_X$ are chosen after the fixed density lower bound. They are proof parameters, not instructions to enumerate Q_o. Eventually $z < y$. For a grid cell I, let P_{I,G_P} be uniform on its odd integers coprime to G_P, and put

$$D_{I,G_P}(s) = \mathbb{E}_{P_{I,G_P}}[f(u) \mid S(u) = s], \tag{G.30}$$

where f is ordinary or normalized semismooth. The retained D_P^U is this function in the cell containing the value, and zero if $S > U$ or $|u| < X_P y^{-C_X}$. Cutoff boundaries are rounded to grid boundaries. It is one function of one value, fixed before a companion point, using the global G_P-unit law, not a fixed residue modulo $2G_P$. Null fibres are assigned value zero. Write $c_P = c_o(X_P, y)R_{G_P}(\beta(X_P, y))$ for the ordinary point reference; for a semismooth test use the corresponding positive sum with weights w_ℓ/ℓ and arguments X_P/ℓ. Put $c_{\min} = \inf_P c_P$ within the current test. Then $c_{\min} \ge e^{-C\Lambda_d}$, also for a complete band, since its additional $1/\log P$ costs only $O(\log D)$.

Lemma G.7 (Conditional densities in arithmetic progressions)**.** *Fix a supported residue $a \bmod m_0$ and write $\delta = (a, m_0)$. Then*

$$\begin{aligned} &\Bigg| \sum_{\substack{u\in I \\ u\equiv a \ (m_0)}} D^U_{I,G_P}(u) \\ &\qquad - \frac{1}{\varphi(m_0/\delta)} \sum_{\substack{u\in I \\ (u,m_0)=\delta}} D^U_{I,G_P}(u) \Bigg| \\ &\qquad\qquad \le C\left\{ UB_\nu + \frac{|I|}{m_0}(1+\log U)E_\nu \right\}. \end{aligned} \tag{G.31}$$

For each retained pattern, the arithmetic progression and the global unit distribution conditional on the specified gcd have the same ideal main term. For two points we condition on the common gcd in (G.22)*. The relative errors are bounded by*

$$\eta_{\mathrm{ar,G}} \le Cz\left\{E_\nu + \frac{U(2|G_P|)(2k)^2 B_\nu}{|I|}\right\}. \tag{G.32}$$

The full-content marginal has relative error $O(m_0/|I|)$. For a one-point use of (G.32)*, take $k = 1$.*

Proof. Put $u = \delta v$, $k' = m_0/\delta$. For a compatible pattern s, put $s_0 = (\delta, Q_o)$. Then $s = s_0 a_s$, $(a_s, k') = 1$, injectively in s. After division by a_s, remaining conditions avoid primes coprime to k'. Each Bonferroni divisor e has count $|I|/(m_0 a_s e) + O(1)$ in the progression. Its main term depends on s, δ, not the unit part of a. Average the same formula over all unit residues modulo k', multiply by $0 \le D_{I,G_P}(s) \le 1$, and sum. There are at most U patterns, and $\sum 1/a_s \le 1 + \log U$. This proves (G.31) with its exact finite principal average; no $m_0 E_\nu$ term occurs. For relative counts, absorb the base progression, full content and their avoidances into compatible CRT classes. Their number cancels between main term and error before division; the remaining Euler product is $\gg 1/\log(2z)$. The proof of (C.14) then gives the conservative (G.32), including rare patterns. Direct class counting gives the content marginal. □

Proposition G.8 (Higher moments of the conditional density)**.** *For each allowed point, condition the finite coefficient law on its prescribed root and put $a_P = \mathbb{E}[f(\mathcal{F}(P)/q) \mid R_{q,s}]$. Then*

$$\|D_P^U(\mathcal{F}(P)/q) - a_P\|_m \le D^{-1/4} a_P. \tag{G.33}$$

This is uniform for ordinary and normalized complete-band semismooth tests, and for unconditioned top points.

Proof. In the ideal G_P-unit pattern law, primes dividing G_P have valuation zero; the others have independent capped geometric valuations. Lemma C.5 applies to this subset of factors. The unit multiplier is $1 + o(1)$ by (G.27). The finite ordinary density product is (C.30) with error (G.32).

For a semismooth test, multiplication by $\ell > y > z$ preserves the recorded pattern. Its exact numerator in (G.30) is

$$\sum_{\substack{\ell \in J \\ \ell \nmid G_P}} w_\ell \#\{v \in I/\ell : v \text{ odd},\ (v, G_P) = 1,\ v \in \mathcal{S}_y,\ S(v) = s\}.$$

Finite counts on the divided intervals give

$$D_{I,G_P}(s) = \sum_{\substack{\ell \in J \\ \ell \nmid G_P}} \frac{w_\ell}{\ell} c_o(x_I/\ell, y) R_{G_P}(\beta_\ell) \mathcal{T}_s (1 - \beta_\ell)(1 + O(\epsilon_{\mathrm{loc}})). \tag{G.34}$$

The endpoint errors use the actual divided widths $|I|/\ell$; their factor $\ell \le Y$ is retained. Here $\mathcal{T}_s$ is the capped-valuation product of (C.15), and $\epsilon_{\mathrm{loc}} \ll d^{-1} + \eta_{\mathrm{FO},U} + \eta_{\mathrm{ar,G}}$, evaluated at those smallest divided widths and endpoints. Positivity prevents a $|J|$ penalty in the relative error. The reference sum is $\asymp \widehat{\rho}(x_I/P, y)/\log P$. Deleting parents dividing G_P costs $O(\log G_P/P) = o(1)$ of this sum. The pointwise score bound is simultaneous for all $0 \le 1 - \beta_\ell \le C/d$, so it bounds the positive mixture.

The one-dimensional Gaussian mass bound gives, with negligible progression endpoints,

$$\mathbb{P}\{\log(X_P/|u|) > v\} \le \min(1, \mathrm{poly}(n)e^{-v}) + e^{-c(\log M + d\log Y)}.$$

The center remains the original center. Thus (C.28)–(C.31) hold with $\log(Cd)$ replaced by $O(\log n)$. First restrict to $|u - F_0(P)/q| \le nT_P$; its complement has Gaussian mass $e^{-\Omega(n^2)}$. Within a retained cell in this region the Gaussian weight varies relatively by $D^{-10A+O(1)}$. Relative pattern comparison multiplies a nonnegative moment by $1 + o(1)$, rather than adding an error times a_P^{-m}. Hölder and the score bound now give

$$D^{o(1)}\{m\gamma_{\max}\log(2z) + \gamma_{\max}\log n + d^{-1} + \eta_{\mathrm{FO},U} + \eta_{\mathrm{ar,G}}\}$$

plus m-th roots of the deleted masses. The choices (G.29) imply

$$m\gamma_{\max}\log(2z),\ \eta_{\mathrm{FO},U} = D^{-1/3+o(1)},$$
$$\log B_\nu = o(d\log Y), \qquad -\log E_\nu \gg m\Lambda_d + A\log D.$$

Also $\log(UE_{\mathrm{cut}}) = o(\log x)$, $\log \eta^{-1} = o(\log y)$, and each direct short interval remains in (B.26). The fixed tail constants make the deleted masses negligible. The conditional finite-sampler error remains negligible after its m-th root and inverse density. On a discarded pattern $D_P^U/c_P - 1 = -1$; it costs no further inverse density.

Exact cancellation holds on each retained pattern under the global-unit law of a complete cell. Equations (G.8) and (G.31) transfer it to the fixed actual AP. A nearly constant Gaussian weight changes cancellation by its relative variation times the cell sum of $f + D$, which retains its density factor. Endpoint cells and deleted sets have the preceding bounds. The resulting relative centering error is $D^{-9A+O(1)} + e^{-c\Lambda_d}$. This centers at the true finite a_P. All other moment errors are $D^{-1/3+o(1)}$, proving the weaker (G.33). □

A fourth moment conditional on the gcd for semismooth tests. The ordinary full-gcd conditional density has bounded normalized fourth moment by Lemma C.6; removal of fixed G_P-unit primes removes factors. Normalize the weights in f_J by their maximum (the bounds are homogeneous). For f_J, a parent $\ell \nmid k_\perp$ contributes an ordinary conditional density with positive weight proportional to $w_\ell c_o(x/\ell, y)/\ell$. Minkowski bounds their sum. If $\ell^h \| k_\perp$, its summand requires $v_\ell(u) = 1$. The associated capped-content event has probability at most $1/\ell$, also when $h = 1$; the extra requirement that u/ℓ not be divisible by ℓ only decreases that probability. Its normalized fourth norm is at most $C\log P\,\ell^{-1/4}$. There are at most $\log k/\log P$ such parents, so the exceptional sum is at most $C\log k\, B^{-1/4} = o(1)$. Finite divided-interval counts are relative, without an inverse semismooth-density loss. The retained proxy also has bounded normalized fourth moment by (G.34) and the score bound. Conditional Jensen preserves it for full-gcd conditional means. These are bounds for conditional means, not for a raw Bernoulli indicator divided by its small mean.

Lemma G.9 (An upper bound for two eligible points)**.** *Under the actual single-root law, two distinct rationally eligible points with $(N, G(P_1)G(P_2)) = 1$ satisfy*

$$\mathbb{E}[f_1(\mathcal{F}(P_1)/q)f_2(\mathcal{F}(P_2)/q) \mid R_{q,s}] \le D^C c_1 c_2 + e^{-c\Lambda_d},$$

where c_r is its point reference density. Either test may be ordinary or a normalized complete-band semismooth test. The additive error can be made negligible relative to c_1c_2 by the fixed deletion constants. The bound is uniform in the angle between the points.

Proof. We work first with the ideal Gaussian law. Write $m_r = 2|G_r|k$ for the periods in (G.21), and retain the original center of the ellipse $\mathcal{E}_R$. We shall use $y^{-C_R} \le R \le n$, where C_R is a fixed constant to be chosen. In the coordinates scaled by T_1, T_2, the ellipse has diameter $O(R)$, area $\asymp R^2/\vartheta$, and perimeter $O(R)$. Cover it by the rectangles of a rectangular grid of side lengths $L_r = RT_r/(C_0\vartheta)$, with a fixed $C_0 \ge 1$. The union of the rectangles meeting the ellipse has area $O(R^2T_1T_2/\vartheta)$: after scaling, it is contained in the $O(R/\vartheta)$-neighborhood of the ellipse. This argument applies also when the two evaluation vectors are nearly parallel. The centers need not be in the support of the coefficient law.

The periods satisfy $m_r\vartheta = 2|G_r|b_1b_2/q$. Consequently $L_r/m_r \to \infty$ uniformly, by (G.23) and the parameter bounds in Proposition G.13. Moreover, each interval of length L_r lies in $[-X_r, X_r]$, where we may take $X_r = 8n^{11}T_r$. Every divided interval used below has length at least $L_r/(m_rY)$ when the divisor is at most m_rY; the explicit scales imply $L_r/(m_rY) > 2y$ for sufficiently large inputs.

We claim that on every supported phase $a_r \bmod m_r$, the average of either of the tests in the statement, over the progression in the corresponding grid interval, is $O(D^Cc_r)$. To see this, write $\delta = (a_r, m_r)$, so $\delta \mid k_\perp$ by (G.22). We first note an elementary consequence of Lemma B.7. For an interval $I \subset [-X, X]$ of length $L \ge 2y$, with $X \ge 2L$, the number of its nonzero integers whose absolute values are y-smooth is at most

$$CL\log(2y)\widehat{\rho}(X,y)(CXy/L)^{1-\beta(X,y)}.$$

Indeed, after reflection an interval away from zero is covered by (B.37) with modulus one. If it meets $[-L, L]$, its smooth count is at most $2\Psi(2L, y)$, and (B.29) gives the same bound. The constants can also absorb the endpoints of the intervals. We use this consequence with $X = X_r$ even after division of an interval by a positive integer; the dilation estimate applies at the common larger endpoint.

For the ordinary smoothness test, only smooth δ can contribute. By (G.8), its principal term is bounded by $1/\varphi(m_r/\delta)$ times the smooth count in the divided interval I_r/δ, with the remaining coprimality conditions omitted. After division by L_r/m_r, the preceding estimate bounds this by

$$C\frac{m_r/\delta}{\varphi(m_r/\delta)}\log(2y)\widehat{\rho}(X_r,y)(CX_ry\delta/L_r)^{1-\beta(X_r,y)}. \tag{G.35}$$

Here $\log(X_r\delta/L_r) = O(\log Y + \log n + C_R\log y)$, since $\delta \le k \le V^2$ and $\vartheta \le V^2$. The saddle-point bound (G.26) and the bounded ratio $\log Y/\log y$ make the last factor a fixed power of D. The totient ratio is also bounded by a fixed power of D. By the definition of c_r and the unit-factor estimate in Lemma G.6, the ordinary reference density c_r is comparable to $\widehat{\rho}(X_r, y)$.

For the semismooth test, let $J \subset (P, 2P]$ denote the complete parent band, with maximum weight one. If δ is smooth, expand the test as in (G.9). Only parents coprime to m_r can contribute; the main term for parent ℓ has the same factor $1/\varphi(m_r/\delta)$ and the smooth count in $I_r/(\delta\ell)$. The estimate above now gives an additional factor w_ℓ/ℓ and replaces δ in the dilation factor by $\delta\ell$. This changes the bound by at most a fixed power of D. If δ has a prime above y, it must contain exactly one such prime, a parent $\ell \in J$ to the first power, or the whole term vanishes. For this parent the divided interval is I_r/δ, and the same estimate applies without $1/\ell$. In both cases the normalized principal term is $O(D^C\widehat{\rho}(X_r, y))$. The semismooth reference density is bounded below by a constant multiple of $\widehat{\rho}(X_r, y)/\log P$. Indeed, the weights have maximum one and bounded ratio, the band contains $\asymp P/\log P$ primes, and the uniform Dickman estimate gives $\widehat{\rho}(X_r/\ell, y) \gg \widehat{\rho}(X_r, y)$. Omitting the parents dividing G_r removes only a negligible fraction of the primes in the band. The possible exceptional parent therefore costs only a further factor $O(\log P)$, which is polynomial in D.

For either test, the error in the progression count, normalized by L_r/m_r, is bounded by

$$C_\epsilon \frac{m_r}{L_r} X_r^{1/2+\epsilon} m_r^{\epsilon} \max(1, Y^{1/2-\epsilon}).$$

The same estimate holds for intervals crossing zero, by applying the prefix estimate (G.7) to both signs (and summing over parents when necessary). Since $m_r\vartheta \le CMV^3$, $L_r = RT_r/(C_0\vartheta)$ and $X_r/T_r = 8n^{11}$, this has precisely the separate-period upper bound verified in Proposition G.13, up to fixed powers of Y, y and n. Taking $\epsilon = 1/32$, the error divided by c_r is exponentially small in $d^2 m_b$. The elementary progression endpoint error is smaller. This proves the claim for every supported phase, without requiring a lower bound on the density of any particular content class.

Within a grid rectangle the support is the disjoint union of the k phase pairs described in (G.21). For any one pair its lattice points form a Cartesian product of two progressions. The preceding one-dimensional bounds therefore give a bound $D^C c_1 c_2$ on the average of the product on that rectangle. The number of lattice points in the rectangle is

$$\frac{kL_1L_2}{m_1m_2}\big(1 + O(m_1/L_1 + m_2/L_2 + m_1m_2/(L_1L_2))\big).$$

Sum over the covering rectangles. Their total area is at most a constant times the ellipse area, and (G.24) gives the corresponding lower bound on the number of points of the ellipse. Positivity allows the entire covering rectangles to be used. The resulting bound on the mean over the ellipse is $D^C c_1 c_2$, uniformly in R in the stated range. Integrating the ellipse counts with their Gaussian weights proves the same bound for the retained radii. The total omitted mass is $O(y^{-2C_R}) + e^{-\Omega(n^2)}$. Choose C_R after the fixed lower bounds on c_1, c_2; then this mass is negligible relative to c_1c_2 and is at most $e^{-c\Lambda_d}$. Finally the relative Gaussian image estimate and the finite-sampler total variation bound of Lemma G.4 transfer the result to the actual coefficient law. □

Proposition G.10 (Concentration under the coefficient distribution)**.** *Fix an environment satisfying the rational-count condition for the target in its search region. Let $Z = \sum_{P\in\mathcal{R}} f_P(\mathcal{F}(P)/q)$, $\mu = \mathbb{E}[Z \mid R_{q,s}]$, and suppose $\mu \ge e^{\Lambda_d/50}$. For ordinary tests or a normalized complete weighted parent band, the fixed proxy sum $\widetilde{Z} = \sum_{P\in\mathcal{R}} D_P^U(\mathcal{F}(P)/q)$ satisfies*

$$\|\widetilde{Z} - \mu\|_m \le D^{-1/4}\mu, \qquad \mathbb{E}[(Z-\widetilde{Z})^2 \mid R_{q,s}] \le D^{-4A+O(1)}\mu^2. \tag{G.36}$$

Hence $\mathbb{P}(|Z-\mu| > \mu/2 \mid R_{q,s}) \le D^{-3A}$, and $\mathbb{E}[(Z/\mu)^2 \mid R_{q,s}] \le D^C$. For an individual initial prime, set $q = 1$, without a coefficient-root condition at its rational parent.

Proof. Minkowski and Proposition G.8 give the first bound. For the residual take a complete grid rectangle of a nonthin pair. Expand each coordinate on its separate progression, using (G.8) and (G.31). The common-content principal expression is exactly, up to the stated relative finite counting errors,

$$\sum_{\delta\mid k_\perp} \frac{\varphi(k_\perp/\delta)}{k_\perp} \mathbb{E}_{P_{I_1,G_1}}[f_1 - D_1^U \mid (u,k_\perp) = \delta]\, \mathbb{E}_{P_{I_2,G_2}}[f_2 - D_2^U \mid (u,k_\perp) = \delta]. \tag{G.37}$$

Both margins use the same content. Higher unit digits at shared G-primes are accounted for by progression errors, not made independent.

Lemma C.8 applies to the independent geometric valuations after deleting fixed unit primes. Its parameters are

$$\eta_0 \ll \frac{\log k}{z\log z} + \frac{\log\log(3z)}{R}, \qquad \pi_0 \gg 1/\log(2z).$$

Use the fourth moment conditional on the gcd above. The relative finite counts (G.32), followed by the conditional-stability argument of Lemma C.9, bound the normalized principal residual by

$$C\sqrt{\log(2z)\eta_0} + \frac{C\{\eta_{\mathrm{ar,G}} + \mathcal{B}_U(L) + m_0/L\}}{c_{\min}^2}.$$

Here $L = \min(L_1, L_2)$, $m_0 = \max(m_1, m_2)$, and $\eta_{\mathrm{ar,G}}$ is the larger of the two errors in (G.32). The quantity $\mathcal{B}_U(L)$ denotes the larger of the bounds in (C.21) for the two coordinates, using this common lower bound on the cell lengths and the corresponding base-progression endpoint factor $2|G_r|$. This is $D^{-4A+O(1)}$.

If $L_r \asymp \eta x_r$, all normalized nonprincipal and proxy remainders are bounded by $D^C(e_1 + e_2 + e_1e_2)$, with

$$\begin{aligned} e_r = {} & C_\epsilon c_r^{-1} \frac{m_r^{1+\epsilon}}{\eta x_r^{1/2-\epsilon}} \max(1, Y^{1/2-\epsilon}) \\ & + Cc_r^{-1} \left\{ \frac{Um_rB_\nu}{L_r} + (1+\log U)E_\nu \right\}. \end{aligned} \tag{G.38}$$

The $Y^{1/2-\epsilon}$ is needed only for a semismooth test. Summing k phase pairs and dividing by their actual number of points

$$\frac{kL_1L_2}{m_1m_2}\left(1 + O(m_1/L_1 + m_2/L_2 + m_1m_2/(L_1L_2))\right)$$

cancels k. The character count has already canceled. The rational density cancels as well: $|\mathcal{R}|^2/\mu^2 \ll c_{\min}^{-2}$ uses this same fixed eligible point set. Proposition G.13 makes (G.38) exponentially small.

For a nonthin pair, $\vartheta \ll \Theta$. At Mahalanobis radius at most n, the logarithmic weight variation on a grid rectangle is

$$C\eta n^2\vartheta + C\eta^2n^2\vartheta^2 \le D^{-10A+O(1)}.$$

Multiply this relative error by the positive majorant $ff + fD + Df + DD$. Its uniform-cell principal expectation divided by c_1c_2 is D^C, by the fourth moments of full-content conditional means and conditional Jensen. Its progression errors were bounded above. Thus no absolute-TV loss times $c_{\min}^{-2}$ is inserted here. Sum with actual cell weights, without a factor equal to the number of grid cells. Boundary rectangles use the same majorant and their relative mass.

The retained proxy satisfies $D_P^U \le c_P \exp(CA(\log D)^2)$. For thin pairs, Lemma G.9 bounds ff. The uniform bound on the conditional density, together with the pointwise means, bounds the mixed terms and the product of two conditional densities. Point densities have bounded ratios in one search region. The rational thin-pair count therefore gives normalized error

$$[D^C + \exp(CA(\log D)^2)]/\Theta + e^{-c\Lambda_d}.$$

The diagonal uses $f^2 \le f$, the same uniform conditional-density bound, and the mean lower bound. Deleted strips, radius tails and sampler error were bounded above. Since A is a fixed power of $\log D$, every term except the principal residual is $o(D^{-4A})$. Markov at $m = 1000A$, then Chebyshev, gives the tail. The joint smoothness bound and the diagonal contribution give the asserted second-moment estimate. □

G.6 Propagation of exceptional sets

We bound the effect of a bad child by summing over all parents that can introduce it. The weight we give a parent depends only on its prime and its level, so the same weight can be used across roots and fields.

Write $T_i = V_i^2 \rho_i / \log V_i$ as above and put

$$\bar{\tau}_i(\ell) = T_i c_i(\ell)/\ell. \tag{G.39}$$

For an active parent satisfying the rational-count condition, the bound on the number of points and Lemma G.6 give $\tau_l(g) \asymp \bar{\tau}_i(\ell)$, uniformly in its root. Here τ_l is the conditional mean under the actual sampling law, whereas $\bar{\tau}_i$ is deterministic. Both use the cutoff y_i and the reference $\mathcal{X}_i$ of this same parent search region, including at a clamped transition. No such comparison is asserted on a parent that fails the rational-count condition.

Set

$$B_1(q) = q^{\gamma_0 - 1}, \qquad \gamma_0 = 1 - \beta(2\mathcal{X}_0, Y_1).$$

For each active complete band $J_e = \{p \text{ prime} : 2^e < p \le 2^{e+1},\ p \nmid N\}$ inside $(y_i, Y_i]$, put

$$\gamma_{i,e} = 1 - \beta(2\mathcal{X}_i/2^e, y_i)$$

and define

$$B_{i+1}(q) = B_i(q) + \sum_{e=e_{i+1}}^{e_i - 1} q^{\gamma_{i,e}-1} \sum_{\ell \in J_e} B_i(\ell), \qquad B < q \le y_i. \tag{G.40}$$

The first term is pass-through. Each profile is a positive sum of powers of exponent in $[-1, 0]$, so its within-band ratio is at most two. Let

$$W_i = \sum_{\substack{B < \ell \le Y_i \\ \ell \text{ prime},\ \ell \nmid N}} B_i(\ell).$$

This sum makes no field-dependent deletion. The saddle bound is $\gamma_{i,e} \log y_i = O(\log D)$. Hence

$$W_{i+1}/W_i \le D^C, \qquad W_1 \le D^C. \tag{G.41}$$

Indeed $\sum_{q \le y} q^{\gamma - 1} \le y^\gamma (1 + \log y)$; the parent bands partition their weights. Zero-mass domains are omitted.

For an event $\mathcal{A}$ on $(g, \mathcal{F}, l)$, define the finite measure

$$\mathfrak{B}_i(\mathcal{A}) = \frac{1}{W_i} \mathbb{E}_{g \sim \mu} \sum_{\ell, s} B_i(\ell) \mathbb{E}_{\mathcal{F} \sim \nu_g}[\mathbf{1}_{R_{\ell,s}} \mathbf{1}_{\mathcal{A}}]. \tag{G.42}$$

Its total mass is between fixed positive constants, by (G.20). We do not normalize by an unknown incoming mean, and do not claim this measure is an adaptive visitation law. Root flatness and $\nu_g(R_{\ell,s}) \le C/\ell$ give

$$\mathfrak{B}_i(\mathcal{E}_i) \le C \max_{\ell \le Y_i} \bar{\epsilon}_i(\ell). \tag{G.43}$$

The root sum cancels $1/\ell$; no label union occurs.

Proposition G.11 (Bad-parent mass in terms of bad-child mass)**.** *Declare a nonterminal active label bad if it is rational-bad, or if it fails the recursive goodness condition of Theorem G.1. Terminal labels have bad indicator zero, and pass-through copies have identical bad indicators. Put* $\delta_i = \mathfrak{B}_i(\mathrm{Bad}_i)$*. Let* $\bar{\epsilon}_i^{\mathrm{ord}}$ *be the largest error in* (G.15) *for own-box active queries, and* $\bar{\epsilon}_i^{\mathrm{in}}$ *the largest such error for child-root queries in that same parent search region. Then*

$$\delta_i \le D^C \delta_{i+1} + D^C\{\bar{\epsilon}_i^{\mathrm{ord}} + \sqrt{\bar{\epsilon}_i^{\mathrm{in}}} + D^{-3A/2}\}, \qquad \delta_j = 0. \tag{G.44}$$

If these rational errors are at most $e^{-c\Lambda_d}$*, uniformly, then*

$$\delta_1 \le D^{-A} + e^{-c'\Lambda_d} \tag{G.45}$$

for a fixed $c' > 0$*.*

Proof. Let $X_{l,(q,s)}$ be the formal incidence count of (A.4). It is zero off both its parent and child root events. Define the positive surrogate

$$H_i^{\mathrm{det}}(g, \mathcal{F}, q, s) = \sum_{\text{active } l=(\ell,u)} \frac{B_i(\ell)}{\bar{\tau}_i(\ell)} X_{l,(q,s)}. \tag{G.46}$$

On a parent satisfying the rational-count condition, $B_i(\ell)/\tau_l(g) \le CB_i(\ell)/\bar{\tau}_i(\ell)$. Use that inequality before dropping any restriction to good parents. The deletion implication (4.2), the ordinary tail, and positivity then give

$$\delta_i \le C\bar{\epsilon}_i^{\mathrm{ord}} + CD^{-3A} + \frac{C}{W_i}\mathbb{E}_{g,\mathcal{F}} \sum_{q,s} H_i^{\mathrm{det}} \mathbf{1}_{\mathrm{Bad}_{i+1}(q,s)} + \text{pass-through mass}. \tag{G.47}$$

The denominators in this bound are the deterministic quantities $\bar{\tau}_i(\ell)$.

For a parent band J_e, put

$$v_\ell = \frac{\ell B_i(\ell)}{T_i c_i(\ell)}.$$

The ratio between any two such weights is bounded throughout the band. Indeed, B_i is a positive sum of powers and the fixed-cutoff density ratio is bounded there. Divide by the maximum weight to apply the normalized semismooth estimates, and restore that common factor afterwards. Reorder the formal incidences in this band under the one child root. The exact result is the finite-parent sum

$$H_{i,J_e}^{\mathrm{det}} = \sum_{P \in \mathcal{R}_i(g,q,s)} \sum_{\substack{\ell \in J_e \\ \ell \nmid b(P)}} v_\ell \mathbf{1}_{\{0 < |\mathcal{F}(P)|/(q\ell) \in \mathcal{S}_{y_i}\}}.$$

Divisibility is part of the indicator, and the expression is zero off $R_{q,s}$. All numerical parents, terminal children, even valuations and raw content retain the formal convention. When a prime denominator equals a parent, the restriction matters. For upper estimates use the complete positive sum S_{i,J_e} obtained by dropping that mask:

$$H_{i,J_e}^{\mathrm{det}} \le S_{i,J_e}. \tag{G.48}$$

No lower mean for this masked incoming sum is required.

Under the prescribed child-root condition, if the rational count in the parent region satisfies the required bounds, the conditional mean of the complete sum is comparable to

$$\frac{T_i}{q}\sum_{\ell\in J_e}\frac{v_\ell}{\ell}c_i(q\ell)=\frac{1}{q}\sum_{\ell\in J_e}B_i(\ell)\frac{c_i(q\ell)}{c_i(\ell)}.$$

The same-parent-box saddle ratio is

$$c_i(q\ell)/c_i(\ell)\asymp q^{\gamma_{i,e}},\qquad q\le y_i. \tag{G.49}$$

Apply (B.27) with cutoff y_i. The change in the saddle point, multiplied by $\log q$, is bounded by

$$O\Big(\frac{(\log q)^2}{\log x\log y_i}\Big)=O(1/d).$$

Varying ℓ within its band changes the ratio by a bounded factor. Both densities use the parent scales y_i and $\mathcal{X}_i$, also when $y_i=B$.

Write $B_i^{\mathrm{act}}(q)=B_{i+1}(q)-B_i(q)$. On an incoming rational-good query, positive band summation and Proposition G.10 yield

$$\begin{aligned}\mathbb{P}(H_i^{\mathrm{det}}>CB_i^{\mathrm{act}}(q)\mid g,R_{q,s})&\le D^{-3A+O(1)},\\ \mathbb{E}[(H_i^{\mathrm{det}}/B_i^{\mathrm{act}}(q))^2\mid g,R_{q,s}]&\le D^C.\end{aligned} \tag{G.50}$$

There are polynomially many bands; their positive means sum to the corresponding profile. Zero active mass contributes zero. Increasing the threshold to $D^CB_i^{\mathrm{act}}(q)$, if desired, does not change the argument.

The exceptional rational-count event in the parent region $\mathcal{E}_i^{\mathrm{in}}(g,q,s)$ refers to root q,s in box V_i, not the child's own search region V_{i+1}. At every admissible g,q,s, even on this event, the one-point upper bound and (G.49) give

$$\mathbb{E}_{\mathcal{F}}[H_i^{\mathrm{det}}\mid g,R_{q,s}]\le CB_i^{\mathrm{act}}(q)\frac{R_i(g,q,s)}{\mathfrak{m}_i(q)}. \tag{G.51}$$

Use $\nu_g(R_{q,s})\le C/q$, then apply one Cauchy–Schwarz inequality in the joint measure of g and uniform s mod q:

$$\frac{1}{q}\sum_s\mathbb{E}_g\left[\frac{R_i}{\mathfrak{m}_i(q)}\mathbf{1}_{\mathcal{E}_i^{\mathrm{in}}}\right]\le C\sqrt{\bar\epsilon_i(q)}.$$

This uses exactly the matching second moment (G.14). Summing deterministic prime weights gives

$$\mathbb{E}_{g,\mathcal{F}}\sum_{q,s}H_i^{\mathrm{det}}\mathbf{1}_{\mathcal{E}_i^{\mathrm{in}}}\le CW_{i+1}\sqrt{\bar\epsilon_i^{\mathrm{in}}}. \tag{G.52}$$

An arbitrarily correlated child-bad indicator can be inserted, since it is at most one.

On the incoming-good part, truncate H_i^{det} at the threshold of (G.50). The bounded part of its child-bad integral is at most $D^CW_{i+1}\delta_{i+1}$. Cauchy–Schwarz in the actual conditional coefficient law, with the second moment and tail in (G.50), bounds the exceptional part by $D^{-3A/2+O(1)}W_{i+1}$. Pass-through costs at most $W_{i+1}\delta_{i+1}/W_i$, since $B_{i+1}\ge B_i$ on its domain. Substitute these estimates in (G.47) and use (G.41). This proves (G.44). Iterating through $j=O(\log D)$ costs $D^{O(j)}$ (even $D^{O(j^2)}$ would suffice). As $A=(j+10)^3$, this proves (G.45). □

G.7 Initial prime ideals in a common number field

Proposition G.12 (The fraction of bad parents in a top band)**.** *Let $J \subset (P, 2P]$ be a complete private band in $(Y_{\min}, Y_0]$. Use the original environment law, with no conditioning on a simultaneous top rational-count event. Write*

$$\epsilon_{0,J} = CK^4 \operatorname{polylog}(N)\rho_{\min,J}^{-1}(Y_0/V_0^2 + 1/M),$$

and

$$\bar{\epsilon}_{0,J}^{\text{in}} = CK^4 \operatorname{polylog}(N)\rho_J^{-1}(Y_1/V_0^2 + Y_1/M).$$

Then the average bad top-parent fraction satisfies

$$\delta_{0,J} := \frac{1}{|J|}\sum_{p\in J}\mathbb{P}_{\mu,\nu_g}(\mathrm{Bad}_{0,p}) \le \epsilon_{0,J} + CD^{-3A} + D^C\{\sqrt{\delta_1} + \sqrt{\bar{\epsilon}_{0,J}^{\text{in}}}\}. \tag{G.53}$$

If the rational errors are $e^{-\Omega(\Lambda_d)}$, then with probability $1 - o(1)$ a single field has at least $3/4$ good parents in every prescribed band.

Proof. Put $T_0 = V_0^2/\log V_0$ and $\bar{\tau}_p = T_0\rho_p c_0/p$, where c_0 is the reference density in (G.25). On $\mathcal{E}_{0,p}^c$, $\tau_p \asymp \bar{\tau}_p$. Average (G.18) over $p \in J$; its factor $|J|$ cancels, giving $\epsilon_{0,J}$. Declare these exceptional parents on the rational side bad without inverting their means. On the good part use $1/\tau_p \le C/\bar{\tau}_p$, the ordinary pointwise tail, and the deletion inequality. Only then drop the restriction to good parents by positivity. The resulting surrogate is

$$H_{0,J}^{\det}(q,s) = \sum_{p\in J}\frac{X_{p,(q,s)}}{|J|\bar{\tau}_p}. \tag{G.54}$$

It is zero off the child root and retains all formal incidences.

Let $R_J(g,q,s)$ be the rational relation count summed over the initial primes and $\mathfrak{m}_J(q) = T_0\rho_J/q$. The exact incidence reordering is

$$H_{0,J}^{\det}(q,s) = \sum_{P'\in\mathcal{R}_J(g,q,s)}\frac{1}{|J|\bar{\tau}_{p(P')}}\mathbf{1}_{\{0<|\mathcal{F}(P')|/q\in\mathcal{S}_{Y_1}\}}.$$

The rational parent $p(P')$ is unique. Within the band its weight is comparable to $P/(|J|T_0\rho_P c_0)$. Lemma G.6, without a lower bound on the number of points, gives

$$\mathbb{E}_{\mathcal{F}}[H_{0,J}^{\det}(q,s) \mid g, R_{q,s}] \le C\frac{c_0(q)}{qc_0}\frac{R_J(g,q,s)}{\mathfrak{m}_J(q)} \asymp B_1(q)\frac{R_J(g,q,s)}{\mathfrak{m}_J(q)}. \tag{G.55}$$

Here $|J| \asymp P/\log P$ and $\rho_J \asymp \rho_P/\log P$ cancel within this same band. The ratio $c_0(q)/c_0 \asymp q^{\gamma_0}$ uses this same top box and cutoff, not any comparison between bands.

For incoming-rational failure $\mathcal{E}_J^{\text{in}}$, use the aggregate version of Lemma G.3 and $\nu_g(R_{q,s}) \le C/q$. One joint Cauchy–Schwarz inequality in $(g, s \bmod q)$, exactly as in (G.52), gives

$$\mathbb{E}_{g,\mathcal{F}}\sum_{q,s}H_{0,J}^{\det}\mathbf{1}_{\mathcal{E}_J^{\text{in}}} \le CW_1\sqrt{\bar{\epsilon}_{0,J}^{\text{in}}}.$$

This is a bound on the weighted bad mass, not a conversion of a fixed-root bad probability into an incoming law.

On the incoming-good part, Lemma G.9 and the mean of the sum over the initial primes give

$$\mathbb{E}_{\mathcal{F}}[(H^{\mathrm{det}}_{0,J}/B_1(q))^2 \mid g, R_{q,s}] \le D^C.$$

Cauchy–Schwarz in the actual coefficient law and then in the positive deterministic root/profile measure bounds its correlated child-bad integral by $D^C\sqrt{\delta_1}$, using $W_1 \le D^C$. Thus the top uses one square root; internal levels used truncation and the linear recurrence. This proves (G.53).

The child's bad-state definition does not depend on the band weights. All these bounds therefore concern the same $g, \mathcal{F}$ and the same bad children. There are $O(e_0) = D^{O(1)}$ prescribed bands. Markov with threshold $1/4$, followed by their polynomial union, and (G.45), proves the final assertion. It does not combine different good fields from different bands. □

G.8 Verification of the parameter conditions

Proposition G.13 (Verification of the parameter conditions)**.** *The scales* (G.1)*, the finite coefficient schedule, and the conditions of the GRH program satisfy all hypotheses used above for one sufficiently large input threshold. In particular, all rational bad-mass bounds are* $e^{-\Omega(\Lambda_d)}$*, the normalized remainders in the polynomial-value and progression estimates are exponentially smaller than* D^{-4A}*, and* (G.6) *holds.*

Proof. Write $t = dm_b$. We first bound the number of descent levels. The unclamped fixed point of the cutoff recurrence is $(247e_B + 1125t)/1137$, and

$$e_B - \frac{247e_B + 1125t}{1137} = \frac{890e_B - 1125t}{1137} \ge \frac{7}{3411}t.$$

Thus the exact ceilings eventually permit strict decrease until the clamp, and

$$j = \frac{\log D}{3\log(430/51)} + O(1).$$

There is no definition of H or a later resource cap inside this recurrence. Next,

$$\log_2 M = \left(\frac{64}{27} + o(1)\right) d^2 m_b, \quad \log_2(\sigma M) = \left(\frac{1504}{405} + o(1)\right) d^2 m_b. \tag{G.56}$$

The same leading expression holds with H instead of σ. Let $T_H = \lfloor 15d^2 m_b/4 \rfloor$. The finite reference condition is

$$256(n+2)^2(d+2)hM2^d \le 2^{T_H}. \tag{G.57}$$

The difference between the leading coefficients is $15/4 - 1504/405 = 59/1620 > 0$. It covers both the actual coefficient bound and (G.25), including geometric rounding. In particular

$$\log_2(2\mathcal{X}_i) \le T_H + \frac{51}{100}de_i + \frac{247}{300}de_B.$$

The cutoff recurrence gives exactly

$$\frac{43}{10}e_{i+1} \ge \frac{51}{100}e_i + \frac{247}{300}e_B + \frac{15}{4}t.$$

Hence every algebraic smoothness argument is at most $43d/10$, including a clamped final transition. The lower endpoint is also a fixed positive multiple of d in smoothness units. This proves (G.26).

Geometry, finite conditioning, and numerical size. At the bottom,

$$\log_2(V_i/(2B)) \geq e_B/3 - 1, \qquad \log_2(V_i^2/Y_i) \geq 5e_B/3.$$

If the next cutoff is clamped, $\log_2(V_i/Y_{i+1}) \geq e_B/3$; otherwise the ceilings give

$$\log_2(V_i/Y_{i+1}) \geq 547t/1290 - 1.$$

At the top,

$$\frac{\log V_0}{\log M} \longrightarrow \frac{17}{60}, \quad \frac{\log Y_0}{\log M} \longrightarrow \frac{5}{9}, \quad \frac{127}{128} > \max\left(\frac{51}{100}, \frac{51}{430}\right).$$

These inequalities prove (H.2). They permit parents larger than their own search region but put all next-level children below the denominator interval. For Gaussian congruences the maximum product of three primes is Y_1^3, not Y_0^3, by the scope following (G.20). The exact leading gap is

$$2(51/100) - 3(51/430) = 714/1075 > 0.$$

It implies the inequality $\sigma \geq 2^{20A^2t}Y_1^3$, since $A^2 = o(d)$. Thus conditional finite TV errors and probability of rejection during initialization without a factor $O(Y_1^3/\sigma)$ are negligible. The additional initialization conditions $\sigma \geq 2^{d+3}(d+1)(d2^d+2)$, $h < N$, and $8MH < N$ also hold eventually.

The explicit bound on the integer values is

$$300T_H + 153de_0 + 247de_B \leq 150(n-1). \tag{G.58}$$

Its leading raw ratio is $17/60 < 1/2$; all other terms are lower order in n. Together with (G.57) it gives $|\mathcal{F}(P)| \leq \sqrt{N}$. Rational values are $< MV_0/8 < M^2 < N$. Zero and unit values are direct cases, and all numerical calls check $2 \leq x < N$. The raw value has fewer than n distinct prime factors, and an identity edge contributes one, proving the $n+1$ incidence bound.

Bounds on densities and means. The uniform Dickman comparison from [29], with $-\log \rho_D(u) = (1+o(1))u\log u$, gives

$$-\log c_A \leq (43/30 + o(1))\Lambda_d.$$

Since $36/25 - 43/30 = 1/150$, the pointwise lower bound p_A follows from Lemma G.6. For the rational densities the factor $B^{247/300}$ adds a constant to the smoothness argument, hence $o(\Lambda_d)$ to its logarithmic cost. There is no ordinary internal call at $i = 0$. The largest actual internal rational scale is MV_1; the top cofactor is MV_0/p, with $p \geq Y_{\min}$. Because $127/128 > 51/100$, the former has the larger leading smoothness cost. In units $\Lambda = D^{1/3}\lambda$, these costs are

$$s_A = 43/40, \qquad s_R = \frac{1 + (51/100)/(2(9/10)(430/51))}{3(3/4)(19/20)} = 5926/12255.$$

As $\Lambda_d \sim (3/4)\Lambda$, the combined loss is

$$\frac{s_A + s_R}{3/4} = \frac{152801}{73530}, \qquad \frac{25}{12} - \frac{152801}{73530} = \frac{773}{147060} > 0.$$

This strict margin allows for the density of prime denominators, fixed geometric constants, complete-band logarithms and rounding. On a state satisfying the rational-count condition a

primitive eligible point exists. Lemma H.2 gives $|\Omega_l| \le 32V_i^2/\ell$. The rational relation count and point mean lower bounds therefore give $\tau_l/|\Omega_l| \ge p_{\mathrm{den}}$. The same argument applies at the top.

For a normalized semismooth band, its mean gains only the factor $1/\log P$, not $1/P$, after summing the parents. With child $q \le y_i \le Y_i$, both ordinary and such semismooth geometric factors are at least a fixed polynomial factor times $V_i^2/Y_i \ge B^{5/3}$. The actual mean margin is

$$\frac{(2(4/3)-1)(19/20) - s_A - s_R}{3/4} = \frac{2429}{73530} > \frac{1}{50}.$$

Even the explicit density exponent leaves $(5/3)(19/15) - 25/12 = 1/36 > 1/50$. Thus all required internal means exceed $e^{\Lambda_d/50}$. An individual top mean has the additional positive margin

$$\log(V_0^2/Y_0) = (1/90 + o(1)) \log M;$$

the sum over the initial primes, with $q \le Y_1$, has a still larger geometric margin. Both dominate every $O(\Lambda_d)$ density loss.

Smooth-numerator congruences. For $\epsilon = 1/32$,

$$\frac{\log Y_1}{\log M} \to \frac{17}{258}, \qquad \frac{\log V_1}{\log M} \to \frac{289}{8600}.$$

The worst coefficient of $\log M$ in the internal error (G.12), including the cell normalization, is

$$\frac{33}{32}\frac{17}{258} - \frac{15}{32}\left(1 + \frac{289}{8600}\right) = -\frac{22927}{55040}.$$

In the sum over the initial primes, the smallest smooth cofactor corresponds to the largest parent Y_0; the cofactor scale satisfies $MV_0/Y_0 = M^{131/180+o(1)}$. Thus (G.17) has coefficient

$$\frac{33}{32}\frac{17}{258} - \frac{15}{32}\frac{131}{180} = -\frac{4511}{16512}.$$

Both are strictly below $-1/4$. Mesh and inverse density have logarithm $o(\log M)$. The individual initial-prime calculation uses no smooth progression theorem modulo its private parent.

Joint-root errors and top averages. For the internal level retain the level-specific bound

$$-\log \rho_i \le (320/513 + o(1))\Lambda_d + (51/(380d) + o(1/d)) \log Y_i.$$

Since $Y_i/V_i^2 \le Y_i^{-1/50} B^{-247/150}$, the first root-average error decreases with e_i eventually. Its bottom decay margin is

$$(5/3)(19/15) - 320/513 = 763/513 > 0.$$

The second internal term is $Y_1/M = M^{-241/258+o(1)}$. At the top the averaged individual powers are

$$Y_0/V_0^2 = M^{-1/90+o(1)}, \qquad M^{-1};$$

the aggregate root-average powers are

$$Y_1/V_0^2 = M^{-323/645+o(1)}, \qquad Y_1/M = M^{-241/258+o(1)}.$$

All inverse rational densities and $K^4 \operatorname{polylog}(N)$ are $e^{O(\Lambda_d)+o(\Lambda_d)}$. They are lower order at every top $\log M$ margin. The internal bottom margin is also strict, because $\log K = o(\Lambda_d)$. Therefore all errors used in the weighted recursion are $e^{-\Omega(\Lambda_d)}$.

Error terms for the separate periods. After root/parent division and the fixed near-zero, angular and radius padding, conservative retained endpoint/period bounds are

$$\log_2 X_* \ge \frac{37}{10} d^2 m_b + (51d/100 - C_0)e_i + (247d/300 - C_1)e_B,$$
$$\log_2 m_{\mathrm{coord}} \le \frac{12}{5} d^2 m_b + 153e_i/100 + 247e_B/100 + C_2.$$

They follow from $1504/405 > 37/10$ and $64/27 < 12/5$; all padding and parent factors cost only $O(e_i + e_B)$. For $\epsilon = 1/32$, the logarithm of the first term in (G.38), before lower-order factors, is at most

$$\frac{237}{320} d^2 m_b - \frac{153}{640} de_i - \frac{247}{640} de_B + O(e_i + e_B) + o(d^2 m_b + d(e_i + e_B)).$$

Its maximum is at the bottom. As $e_i \ge e_B \ge 19t/15$, this is at most

$$(-49/960 + o(1))d^2 m_b \le -d^2 m_b/25.$$

This margin is sufficient to absorb every fixed-power padding factor, inverse smooth density and parent factor, as well as η^{-1}. The same bound applies with $m_r\vartheta$ in Lemma G.9, since its logarithm has the same upper bound on the periods. The errors from the finite pattern counts and from the conditional densities are smaller. At $P = P_r$, the endpoint factor in (G.32) is $(2|G_r|)(2k)^2 = 4km_r$. Since $k = \Delta/q \ll V_i^2$, the scale definition gives

$$\log_2(4k) \le \frac{51}{50} e_i + \frac{247}{150} e_B + O(1).$$

The displayed endpoint and period bounds therefore imply

$$\log_2 \frac{4km_r}{X_*} \le -\frac{13}{10} d^2 m_b - \frac{51}{100} de_i - \frac{247}{300} de_B + O(e_i + e_B).$$

Thus the extra $4k$ costs only $O(e_i + e_B)$, within the endpoint margin of order $d^2 m_b + d(e_i + e_B)$. The logarithms of the remaining fixed-power padding and inverse-density factors are $o(\log X_*)$, as are $\log U, \log E_{\mathrm{cut}}, \log B_\nu$. Together with $-\log E_\nu \gg m\Lambda_d + A \log D$, these estimates control the relative counting errors, including after taking the m-th root. No product modulus G_1G_2k or factor $m_{\mathrm{coord}}E_\nu$ is used. Since the E_ν-term has no period multiplier, no comparison between $-\log E_\nu$ and $\log m_{\mathrm{coord}}$ is required at the bottom.

These bounds apply to ordinary internal counts for $i \ge 1$, incoming counts in the larger parent search region, counts for a single top parent without root conditioning, and top-band sums conditioned on a child root. □

Proof of Theorem G.1. Proposition G.13 supplies the finite ranges, density and mean margins. Lemma G.3 defines the local rational-good events, and Proposition G.10 supplies their conditional ordinary and semismooth tails. Proposition G.11 gives $\delta_1 \le D^{-A} + e^{-c\Lambda_d}$. Proposition G.12 then gives the simultaneous dense good-parent sets in a single field. The number of private bands is polynomial. Their construction requires

$$J_b + 1 \le \lfloor e_0/128 \rfloor, \qquad \lceil 127e_0/128 \rceil > e_1.$$

These hold eventually in the private-shape construction of the GRH program and do not change the preceding bandwise argument. The density, deletion and value-size conclusions give the remaining claims. Once the field is fixed, these sets precede all Search decisions. Their bounds therefore apply whenever Search visits one of their members.

□

H The factorization algorithm under GRH

We describe the algorithm G and prove its running-time bound. All the rows constructed in a given iteration use the field sampled in Appendix G; the existence of sufficiently many usable relations is established in Theorem G.1. At each node of the descent, the algorithm samples a point in the special-prime lattice and tests smoothness on the rational side. Only a successful rational test is followed by an algebraic test. Both the total number of sampled points and the number of algebraic tests are bounded. A completed row is treated as in the unconditional algorithm, with its rational correction and its own private square root. The estimates for the success probability use GRH. Every returned factor is verified independently of this hypothesis.

H.1 Parameters, preprocessing and polynomial selection

Use the shared data and notation of Section G.1. For the current input first perform exact primality and perfect-power tests and remove even factors. A prime is returned after deterministic verification; a nontrivial power gives a verified smaller factor. Thereafter N is odd and has at least two distinct prime divisors. Compute d, t, k_0, e_B, e_0 first. Reject the fast epoch if $e_0 \le e_B$, if an active step of (G.1) does not strictly decrease, or if e_B is not reached within $n+10$ steps. Only then compute j and (G.2). Empty required sampling intervals also fail the epoch. Every new outer input recomputes these data.

Test divisibility by every prime at most B, returning any proper factor found. If all tests pass, N is B-rough. The program next checks

$$\begin{aligned}
&4H \ge d2^d + 2, \quad h < N, \quad C_F < N, \quad \sigma \ge 2^{d+3}(d+1)(d2^d+2),\\
&\sigma \ge 2^{20A^2t}Y_1^3, \qquad 256(n+2)^2(d+2)hM2^d \le 2^{T_H},\\
&300T_H + 153de_0 + 247de_B \le 150(n-1), \qquad \text{(H.1)}\\
&V_i \le V_0, \quad 4V_0 < M, \quad V_i > 2B, \quad V_i^2 \ge Y_i, \quad V_i \ge 33Y_{i+1} \quad (0 \le i < j),\\
&4Y_{\min} > 3V_0, \qquad Y_{\min} > Y_1, \qquad 4Y_0 < 3M. \qquad \text{(H.2)}
\end{aligned}$$

If any condition fails, this iteration is unsuccessful. Proposition G.13 proves eventual validity; the persistent backup below covers all finite inputs. In particular, we do not require $V_i \ge 33Y_i$. The bounds below for the certificate, private shapes, records, and sampling depend only on these integer data. Compute them in the order given before drawing the environment, with $R_{\mathrm{env}} = n^3$.

To implement the environment law, try at most R_{env} uniform integer candidates in $(3M/4, M]$, testing primality exactly, and take the first prime Q. Exhaustion fails the epoch. Use (H.31) for scalar draws and then draw r in the prescribed interval. On the coupled ideal law the accepted prime is exactly uniform; the prime number theorem makes exhaustion probability $o(1)$. Return a proper $\gcd(Q, N)$ if found. Otherwise set $m_N = rQ^{-1} \pmod N$ and $G(X,Y) = QX - rY$. The coordinates r, Q are coprime since $r < Q$ and Q is prime.

The deterministic centered-base-Q and parity construction of Section A gives

$$F_0(r,Q) = N, \qquad \|F_0\|_\infty \le (d2^d+2)M, \qquad F_0 \bmod 2 \text{ irreducible of degree } d.$$

Its binary irreducible polynomial is found by exhaustive enumeration and Frobenius tests in $2^{O(d)}\,\mathrm{poly}(d)$ bit operations. Draw the coefficients once by Lemma H.1 and form $\mathcal{F}$ from (G.4). Then $\mathcal{F}(r,Q) = N$, $\|\mathcal{F}\|_\infty \le C_F$, and its primitive part F is irreducible of degree d. Its content e divides N; because $e \le C_F < N$, $e > 1$ is a proper factor. Thus a continuing path has $e = 1$ and $F = \mathcal{F}$.

Put $f(T) = F(T, 1)$ and gcd-test $A_f = \mathrm{lc}(f)$ and $f'(m_N)$ against N. Proper gcds exit; gcd N rejects the epoch. The coefficient and content bounds exclude a rejection without finding a factor from A_f or e. Conditional on the other coefficients, changing c_{d-1} changes the derivative of the sampled polynomial at m_N by the unit $2Q$ modulo N. Since $2H + 1 < N$, at most one supported coordinate value makes that derivative zero modulo N. Its ideal Gaussian mass is $O(1/\sigma)$, and the finite-sampler error is negligible. Thus the probability that setup rejects without finding a factor is $o(1)$ under the original distribution used in the theorem on relation counts.

Lemma H.1 (Finite Gaussian coefficients)**.** *There is a bounded procedure returning $c \in [-H, H]^d$ or a single failure outcome for the epoch such that its law has total variation distance at most $\exp(-n^{10})$, for sufficiently large n, from the product law on $\mathbb{Z}^d$ proportional to $\exp(-\pi\|c\|^2/\sigma^2)$. Every tape uses polynomially many bit operations in $n + \log \sigma$.*

Proof. For each coordinate use at most $R_G = \lceil n^{24} \rceil$ proposals. Each proposal draws a fixed-width uniform integer in a dyadic interval containing $[-H, H]$; an out-of-range draw consumes the proposal. On an in-range value a, use an independent P_G-bit threshold, $P_G = \lceil 16n^{12} \rceil$, to accept a dyadic probability $w_a \in [0, 1]$ with

$$|w_a - \exp(-\pi a^2/\sigma^2)| \le 2^{-P_G+2}.$$

If a coordinate exhausts its cap, declare coefficient sampling to have failed for the whole field. The dyadic probabilities are computed by rational interval arithmetic: the alternating arctangent series in Machin's formula compute π, and Taylor series with their factorial remainder bounds, followed by argument reduction and squaring, compute the exponential. Arguments whose elementary upper bound is below 2^{-P_G} can be assigned zero. Only polynomially many terms and bits in $P_G + \log \sigma$ are needed, so no transcendental oracle or unbounded precision loop is used.

The Gaussian mass in $[-\sigma, \sigma]$ is at least a constant times σ, whereas the proposal interval has length $O(\sigma n^{10})$. Each proposal therefore accepts with probability at least cn^{-10} for an absolute $c > 0$. Its failure due to the trial limit is at most $\exp(-cn^{14})$. The discarded Gaussian tail beyond $H = \sigma n^{10}$ is $O(\exp(-\pi n^{20}))$; the total error in dyadic weights, after normalizing by a mass $\gg \sigma$, is $O(n^{10}2^{-P_G})$. Sum these bounds over $d \le n$ coordinates to obtain the asserted total variation bound. This is a global sampling error. The pointwise relative image and conditioned Gaussian estimates used in supply are the separate statements proved in Appendix G. □

H.2 Uniform sampling in the special-prime lattices

Every active request uses the domain (G.5) and predicates of Section G.1. An internal label (i, ℓ, s) is admissible only if $B < \ell \le Y_i$, $\ell \nmid NQ$, $Qs - r \not\equiv 0 \pmod{\ell}$, and $\mathcal{F}(s, 1) \equiv 0 \pmod{\ell}$. If $\ell \le Y_{i+1}$, it passes unchanged to the next level; otherwise it is active. The top request has $Y_{\min} < p \le Y_0$, $p \nmid N$, and $s = rQ^{-1} \pmod{p}$, with no coefficient-root conditioning. The domain implies $V_i/8 < a < 2V_i/3$. All discovered proper gcds are absorbing successes.

Lemma H.2 (Counting and sampling the points in a box)**.** *The cardinality of* (G.5)*, its prefix counts in b, and the inverse of its lexicographic rank map can be computed in polynomial bit time in n. A trial using one $\lceil \log_2 |\Omega| \rceil$-bit block assigns exactly the same mass to every point of Ω and has geometric acceptance probability at least $1/2$. At a rational-good request,*

$$|\Omega| \le 32V_i^2/\ell. \tag{H.3}$$

Proof. Write $a = sb + \ell z$, $C_r = r - Qs$, $D_r = Q\ell$, and put

$$b_- = \lfloor V_i/2 \rfloor + 1, \quad b_+ = \lfloor 3V_i/4 \rfloor, \quad L = \lceil MV_i/16 \rceil, \quad U = \lfloor MV_i/8 \rfloor.$$

For a fixed b, the number of allowable z is

$$F_b = \left\lfloor \frac{C_r b - L}{D_r} \right\rfloor - \left\lfloor \frac{C_r b - U - 1}{D_r} \right\rfloor. \tag{H.4}$$

Prefix sums, with ℓ-divisible denominators subtracted, are four signed Euclidean floor sums. Here is an explicit finite recurrence for them. For $S(m, a, c, h) = \sum_{0 \le u < h} \lfloor (au + c)/m \rfloor$, divide a and c by the positive m using mathematical floor quotients and add $\lfloor a/m \rfloor h(h-1)/2 + \lfloor c/m \rfloor h$. Replace a, c by their residues. If $a = 0$, stop. If $y = ah + c < m$, stop. Otherwise replace (h, c, m, a) by $(\lfloor y/m \rfloor, y \bmod m, a, m)$ and repeat. Counting lattice points below the line proves this reciprocity; each normalization is exact and the successive positive moduli follow Euclid's algorithm. It applies to negative C_r as well. All arguments have $O(n)$ bits and there are polynomially many operations.

If $T = |\Omega| = 0$, Search returns local FAIL. Otherwise draw one rank from $\{0, \dots, 2^{\lceil \log_2 T \rceil} - 1\}$; a rank at least T consumes the trial. Binary search on the prefix counts locates b, and the interval offset in (H.4) locates z. The map is a bijection; for $T = 1$ it uses no bits. No list of points is constructed.

At a rational-good request there is a primitive point P_0 in Ω. It is a primitive vector in the index-ℓ lattice $a \equiv sb \pmod{\ell}$ and extends to a lattice basis. Its length is between $V_i/2$ and $2V_i$. Lines parallel to it have spacing $\ell/\|P_0\|$. The displayed rectangle has diameter below $2V_i$, so at most $1 + 4V_i^2/\ell$ such lines meet it, and each contains at most five of its lattice points. Thus $|\Omega| \le 5 + 20V_i^2/\ell \le 25V_i^2/\ell$, using $V_i^2 \ge Y_i \ge \ell$. This proves the weaker stated constant 32. Exact uniformity alone would not have supplied this necessary cardinality comparison. □

Smoothness tests and the resulting children. For every candidate we have $-G(P) = rb - Qa > 0$. The rational predicate is that $(a, b) = 1$, b is prime with $b > B$, and $-G(P)$ is B-smooth. The algebraic predicate is

$$0 < |\mathcal{F}(P)|/\ell \in \mathcal{S}_{Y_{i+1}}.$$

At the top the rational predicate instead requires $-G(P)/p$ to be B-smooth, and the algebraic predicate is $0 < |\mathcal{F}(P)| \in \mathcal{S}_{Y_1}$. The denominator primality test is exact. The parameter inequalities imply $b < Y_{\min} < p$ at the top and $b > Y_{i+1}$ at every internal next-level child comparison. The incoming active valuation is one because its parent exceeds the next cutoff.

Before either fallible smoothness test, compute the cheap gcds with N of both relevant numerical values, including rational and algebraic values; test also all denominators which will be inverted. If the gcd is nontrivial, return it before a rejection. Values zero and one are direct cases. Enforce $2 \le x < N$ for each numerical call, with $|\mathcal{F}(P)| \le \lceil \sqrt{N} \rceil$ and the bound on the initial coordinates $\lceil \sqrt{2N} \rceil$. The height bounds imply the first bound; rational values satisfy $0 < -G(P) \le MV_0/8 < M^2 < N$.

Use Proposition H.3 for the rational test first, and invoke the more expensive algebraic test only after rational acceptance. A failed test consumes this trial. On acceptance save the exact verified prime lists. A good prime $q > B$, $q \nmid eA_f \operatorname{disc} f$, in the primitive algebraic value creates the child $(q, ab^{-1} \bmod q)$ in the same number field precisely when its ideal valuation is odd. Small primes and bad primes are terminal; a prime at or below the next active cutoff passes unchanged to the appropriate next level without a new atom. The inverse of b is valid

by primitiveness and the leading coefficient condition. A large child cannot divide $G(P)$: at an internal point its rational part is B-smooth, and at the top its only possible large rational prime is $p > Y_1$. The child is therefore an admissible finite root.

The formal incidence count in the theorem on relation counts may include all distinct finite prime incidences, even valuations and bad terminal primes included. Terminal and factor-exit states have bad indicator zero. A pass-through has a single identity edge and no new atom. The actual child graph is a subgraph of this nonnegative overcount, so it has at most $M_c = n + 1$ incidences per candidate. This also follows from the enforced bound $\log_2 |\mathcal{F}(P)| < n$. In the mixed geometry a denominator can equal a formal parent prime. The finite parent mask is retained; the incoming proof uses only $H_{\mathrm{det}} = S_{\mathrm{fin}} \le S_{\mathrm{complete}}$, never an equality or a lower bound for the masked mean.

H.3 A bounded smoothness test under RH

Given a candidate value x and a smoothness bound y, the descent needs the prime factors of x only when they are all at most y. We use genus-two Jacobians to test this within a fixed budget. The genus-two method is due to Lenstra, Pila and Pomerance [42, 43], who predicted the limiting constant 2. See also the manuscript of Zhou [68]. Our probability estimate uses RH for the Riemann zeta function alone.

Proposition H.3 (Bounded smoothness recognition)**.** *Let $x, y \ge 2$, let $b(x) \le n$, and put*

$$k_y = \left\lceil \sqrt{b(y)b(b(y))} \right\rceil, \qquad z_y = 2^{k_y}.$$

For an integer $L_c \ge 1$, there is an explicitly bounded procedure $\mathrm{Smooth}(x, y, L_c)$ *with the following properties. It returns either the verified factorization of x into primes at most y, a verified prime factor exceeding y, or* FAIL*. Its accepting outputs are always correct, without any hypothesis. On every random tape its bit cost is*

$$z_y \, 2^{\lceil 101 k_y/100 \rceil} \, \mathrm{poly}(n + b(y)) L_c = L_y[1/2, 201/100] \, \mathrm{poly}(n) L_c \tag{H.5}$$

when $b(y) \le n$. Correctness and the cost bound hold for every finite y. Under RH, for every sufficiently large y, the probability of FAIL *is at most $2n \exp(-64 L_c)$ uniformly over y-smooth inputs x with $b(x) \le n$.*

The earlier RH analysis includes checked Cantor arithmetic and proper-factor extraction by the Chinese remainder theorem [33, `genus2rh`, Theorem R, Corollary R2, Lemma 3 and Propositions A–B]. We give the details of our bounded implementation. We sample a curve and a divisor from eight residues, then use an odd scalar to distinguish two prime components.

Smooth orders at every sufficiently large prime. Write $t_p = \sqrt{\log p \log\log p}$ and $w_p = \exp(t_p)$. For every sufficiently large prime p, set

$$x_p = p^2 - \tfrac{1}{2} p^{3/2}, \qquad h_p = p^{3/2}, \qquad I_p = (x_p, x_p + h_p].$$

The RH short-interval theorem of [67, Theorem 1.3], used with the fixed exponent $7/10$, gives

$$\begin{aligned} \frac{\Psi(x+h, w) - \Psi(x, w)}{h} &= \alpha(x, w) \frac{\Psi(x, w)}{x} (1 + O(\mathcal{E}_{\mathrm{sm}})), \\ \mathcal{E}_{\mathrm{sm}} &= \min\left\{ \frac{1}{u} + \frac{h \log(u+1)}{x \log w}, \frac{\log(u+1)}{\log w} \right\} \end{aligned} \tag{H.6}$$

for $x^{7/10} \le h \le x$ and $(\log x)^{K_0} \le w \le 2x$, where K_0 is a fixed constant and $u = \log x/\log w$. The saddle point satisfies

$$\alpha(x,w) = 1 - \frac{\log(u\log(u+1)) + O(1)}{\log w}.$$

All these ranges hold at $(x_p/2^j, h_p/2^j, w_p)$, $j = 0,1,2$. The classical smooth-number estimate [29, Theorem 1.1 and Corollary 2.3] gives

$$r_p := \frac{\Psi(x_p,w_p)}{x_p} = \exp(-(1+o(1))t_p). \tag{H.7}$$

This estimate holds for every sufficiently large prime p.

We need orders divisible by two exactly once. Applying (H.6) with $h = x$ gives $\Psi(2x,w_p) = (2+o(1))\Psi(x,w_p)$ at the two adjacent scales. Consequently

$$S_j := \Psi((x_p+h_p)/2^j, w_p) - \Psi(x_p/2^j, w_p) = (1+o(1))2^{-j}h_p r_p.$$

The number of w_p-smooth integers in I_p with two-adic valuation exactly one is therefore

$$S_1 - S_2 = (\tfrac{1}{4} + o(1))p^{3/2}r_p = p^{3/2-o(1)}. \tag{H.8}$$

Indeed, S_1 counts the even smooth orders and S_2 counts those divisible by four.

By [43, Theorem 1.1], for $p \ge 8100$, all but at most p integers in $[p^2 - p^{3/2}/2, p^2 + p^{3/2}/2]$ occur as the order of the Jacobian of at least

$$\frac{p^{9/2}}{24000(\log p)^2(\log\log p)^2}$$

nonsingular quintic polynomials over $\mathbb{F}_p$, allowing all nonzero leading coefficients. Subtracting the exceptional orders in (H.8) and then summing the disjoint order fibres proves that at least

$$p^6\exp(-(1+o(1))t_p) \tag{H.9}$$

coefficient vectors define nonsingular quintics f such that $\#J_f(\mathbb{F}_p)$ is w_p-smooth and $v_2(\#J_f(\mathbb{F}_p)) = 1$.

Sampling a curve and a divisor without extracting roots. Let a be an odd composite cofactor which is not a prime power. Draw eight independent $b(a)$-bit integers. If one is at least a, discard this entire trial. Otherwise use the eight resulting uniform residues in $\mathbb{Z}/a\mathbb{Z}$ to form

$$u = X^2 + u_1X + u_0, \qquad v = v_1X + v_0,$$
$$h = h_3X^3 + h_2X^2 + h_1X + h_0, \qquad f = uh + v^2.$$

The range event has probability at least 2^{-8}, and, conditional on it, the residues are exactly uniform. Compute the gcd with a of

$$h_3, \qquad \operatorname{disc}(f), \qquad \operatorname{disc}(u), \qquad \operatorname{Res}(u,f). \tag{H.10}$$

A proper gcd is returned immediately. A gcd equal to a discards the trial; if all gcds are one, (u,v) is a generic reduced degree-two Mumford representative at every prime divisor of a. This construction imposes $u \mid f - v^2$ identically and performs no square-root operation modulo a.

Fix a prime $p \mid a$ and reduce the eight residues modulo p. For a fixed nonsingular quintic f, a reduced degree-two Mumford pair (u, v) determines the unique cubic $h = (f - v^2)/u$. Hence, conditional on f and the stated checks modulo p, every allowed pair has the same probability. The degree-zero and degree-one representatives account for exactly $\#C_f(\mathbb{F}_p)$ elements of $J_f(\mathbb{F}_p)$. Repeated-root quadratics $u = (X - b)^2$ exclude at most $2p$ further pairs: the equations $v(b)^2 = f(b)$ and $2v(b)v'(b) = f'(b)$ have at most two solutions when $f(b) \neq 0$, and none when $f(b) = 0$. A common root of u and f excludes at most $10p + 2$ pairs (at most five rational roots, and at most two irreducible quadratic factors of f). Hence the omitted set has size at most

$$E_p \leq 13p + 4\sqrt{p} + 3. \tag{H.11}$$

We used the Weil bounds $\#C_f(\mathbb{F}_p) \leq p + 1 + 4\sqrt{p}$ and $(\sqrt{p} - 1)^4 \leq \#J_f(\mathbb{F}_p) \leq (\sqrt{p} + 1)^4$. For $p \geq 8100$ at least $p^2/2$ generic pairs remain. Multiplying this count by (H.9) and dividing by p^8 proves that the original eight-coordinate law has a useful p-component with probability at least $\exp(-(1 + o(1))t_p)$.

The standard monic Cantor formulas can be used without losing the nonmonic curve count. With the checked unit $A = h_3$, put

$$F(T) = A^4 f(T/A), \quad U(T) = A^2 u(T/A), \quad V(T) = A^2 v(T/A).$$

The curve isomorphism $(X, Y) \mapsto (AX, A^2 Y)$ makes F monic of degree five and preserves all group orders and point laws. Every displayed coefficient uses only nonnegative powers of A.

Computing in the Jacobian and detecting a factor. Use complete Cantor composition and reduction [18]. For two reduced pairs $(u_1, v_1), (u_2, v_2)$ compute

$$d_0 = \gcd(u_1, u_2, v_1 + v_2) = s_1 u_1 + s_2 u_2 + s_3(v_1 + v_2),$$
$$U = u_1 u_2 / d_0^2, \qquad V = \{s_1 u_1 v_2 + s_2 u_2 v_1 + s_3(v_1 v_2 + F)\}/d_0 \pmod{U}. \tag{H.12}$$

If $\deg U > 2$, replace it by $U' = (F - V^2)/U$, made monic, and $V' = -V \pmod{U'}$. The composed degree is at most four, and a single such reduction gives degree at most two. The identity is $(1, 0)$; zero remainders, inverses, doubling, and degree-one representatives use the complete Euclidean branches of these formulas.

Before any scalar inversion or zero/degree decision, compute the gcd of that scalar with a. A proper gcd is a factor. A scalar zero modulo a is handled by the zero branch and is never inverted; leading zero coefficients are removed. A unit is inverted, and nonzero divisor polynomials are made monic. Exact divisions and the Mumford identities are checked. If no factor is found, each operation and branch reduces to the correct field computation at every prime divisor of a. A purported exact division with a nonzero remainder must also yield a proper gcd: all prime reductions of that remainder vanish, so any nonzero coefficient has a proper gcd with a. This includes repeated prime factors of a. Thus no singular chart or nonunit failure is silently discarded on a useful component.

In particular, if a computed multiple is the identity at one prime and not the identity at another, a proper factor must have occurred. Otherwise its final monic reduced U would have degree zero at the first prime and positive degree at the second, which is impossible for a monic polynomial. The argument applies to each residue field separately, including when a has repeated prime factors.

The scalar used at a cofactor a is the odd integer

$$K_a = \prod_{\substack{3 \leq l \leq z_y \\ l \text{ prime}}} l^{\lfloor \log(16a^2)/\log l \rfloor}. \tag{H.13}$$

It is applied sequentially, prime power by prime power, without constructing a growing product by naive integer multiplication. For each l, compute the exponent as the largest integer v with $l^v \le 16a^2$, using integer arithmetic. Each prime power has $O(b(a))$ bits. If a useful prime $p \mid a$ has $\#J_f(\mathbb{F}_p) = 2m$ with odd, z_y-smooth m, then $m \mid K_a$, whereas K_a is odd. Multiplication by K_a maps a uniform Jacobian point onto its subgroup of order two, with two equally sized fibres: the group is the direct product of its subgroup of order two and its subgroup of odd order, and the scalar is odd and annihilates the latter subgroup. By (H.11), on the generic sampled subset each of the identity and nonidentity outcomes still has probability at least 1/3.

Fix any other prime $q \mid a$ and fix all eight residues modulo q. CRT makes the p-residues independent of them, even when a has prime-power divisors. On a useful p-tuple, a failed initial q-guard in (H.10) already gives a proper gcd. Otherwise the identity status of $[K_a]D$ modulo q is fixed. The p-status is opposite with probability at least 1/3, giving a proper factor by the preceding paragraph. This proves a splitting probability at least $\exp(-(1+o(1))t_p)$ per original trial. The argument works regardless of the order of the group at q, since only the identity status of its particular multiple is used.

The trial limit and the failure probability. First handle even factors, exact primality, and perfect powers. Trial-divide by all primes at most z_y. For each unresolved cofactor which is neither prime nor a prime power, perform at most

$$R_{\mathrm{curve}} = 2^{\lceil 101k_y/100 \rceil} L_c \tag{H.14}$$

independent curve trials, using the operations just specified. At a proper split, continue on both cofactors. An exhausted node returns FAIL. Primes are certified by a deterministic polynomial-time test [3, 2]; multiplicities and the product are verified. A factor tree has fewer than $2n$ nodes after perfect-power multiplicities are compressed, and all its arithmetic and random-bit loops have explicit finite bounds.

If the input is y-smooth, any unresolved target prime satisfies $z_y < p \le y$. Apply the preceding smooth-order count first at its own $w_p = \exp(t_p)$, subtracting the at most p exceptional orders there. Only then use $w_p \le z_y$ and monotonicity in the smoothness cutoff. Since the least remaining p tends to infinity with y, the error is uniform over this range. Therefore, for all sufficiently large y, the proper-split probability at every such node is at least $z_y^{-201/200}$. The cap in (H.14) makes its failure probability at most

$$\exp(-z_y^{1/200} L_c) \le \exp(-64 L_c).$$

Summing over the fewer than $2n$ nodes proves the failure bound. The search keeps the caller's cutoff y even for small x, since trial division already handles factors at most z_y.

There are at most z_y prime-power stages per curve, each using $O(b(a))$ fixed-genus additions and polynomial bit work. Generating these primes by a sieve or by successive deterministic primality tests has the same upper bound. Thus a trial costs $z_y \operatorname{poly}(n + b(y))$; the whole bounded tree satisfies (H.5). Finally $k_y \log 2 = (1+o(1))\sqrt{\log y \log\log y}$, proving the asserted L-constant. This completes the proof of Proposition H.3.

H.4 Bounds on the number of smoothness tests

Use the density exponents a_A, a_{den} and bounds p_A, p_{den} from (G.3). Lemma G.6 and Theorem G.1 give these lower bounds on the means used by Search. The comparison of incoming counts uses the means themselves.

Put $K = 2k_0$ and define

$$\begin{aligned} M_c &= n+1, & L_s &= b(64KM_c), & \varepsilon_s &= (16KM_c)^{-1}, \\ q_s &= 128L_s, & r_s &= 4096\,2^{a_{\mathrm{den}}}(q_s+L_s), & r_A &= 4096\,2^{a_A}L_s. \end{aligned} \tag{H.15}$$

A Search has at most r_s trials and at most q_s verified candidates which launch child descent. It also has at most r_A algebraic recognizer invocations. Increment the last counter after rational acceptance and before the algebraic invocation, irrespective of its answer. On overflow return local FAIL. There is no common limit on the searches at a level and no early termination of the run merely because of a bad child's local overflow. A failed child causes rejection of its parent candidate and a further attempt within that parent's own caps. Terminal requests return immediately; pass-throughs follow their identity edge. Good-state sets and conditional means are proof objects and are not computed by Search.

Fix an accepted field and a good request. Use $\mathcal{R}, \Omega, a_P, \tau, Z, D$ from Theorem G.1. Here a_P and τ still refer to the original coefficient law with this fixed environment and its prescribed root condition (none at the top); Z, D are counts in the already chosen field. The theorem implies

$$D/Z \ge 1/8, \qquad Z/|\Omega| \ge p_{\mathrm{den}}/4, \qquad D/|\mathcal{R}| \ge p_A/4. \tag{H.16}$$

The last inequality follows from $\tau = \sum_{P\in\mathcal{R}} a_P \ge p_A|\mathcal{R}|$. Among points passing the rational test it thus gives a lower bound on the fraction with a complete descent.

Lemma H.4 (Search with a local algebraic quota)**.** *In the exact-predicate program, every good request returns a complete descent or a proper factor with conditional probability at least* $1-\varepsilon_s$ *given any public history before the request. The bounds in* (H.15) *apply to every request, including bad ones, on every random tape.*

Proof. Induct backwards over the finite levels. Terminal states and identity transitions satisfy the claim directly. At a good candidate there are at most M_c required children. Their conditional failure probabilities, without any independence assumption, sum to at most $M_c\varepsilon_s \le 1/8$. Such a candidate therefore completes or finds a factor with probability at least $7/8$.

Each point in Ω has the same raw mass λ by Lemma H.2. Include all immediate terminations with a nontrivial factor as successes; do not condition on their absence. A rational opportunity is an eligible rational point or an immediate termination with a factor outside that set. If S_{out} is the latter raw mass, its conditional completion-or-factor probability is at least

$$\frac{(7/8)\lambda D + S_{\mathrm{out}}}{\lambda|\mathcal{R}| + S_{\mathrm{out}}} \ge \frac{7D}{8|\mathcal{R}|} > \frac{p_A}{8}. \tag{H.17}$$

Factor exits inside $\mathcal{R}$ only improve the numerator. Every actual algebraic invocation corresponds to a rational opportunity. A cached factor list can save an invocation. Exhausting the algebraic test limit implies at least r_A unsuccessful opportunities and has probability at most $(1-p_A/8)^{r_A} \le \exp(-512L_s)$.

For the probability estimate, extend the independent sequence of trials beyond the actual trial limit, and after any absorbing success or a different cap insert virtual successes. The preceding bad event is contained in the event of r_A unsuccessful marks in that extended stream. The bound thus applies even if the search would stop before another opportunity arises.

Among verified candidates, $D/Z \ge 1/8$ and the child bound give a completion probability exceeding $1/32$, including termination with a factor in the same way. Exhausting q_s such candidates therefore costs at most $\exp(-4L_s)$. A trial supplies a verified candidate or factor

with probability at least $p_{\mathrm{den}}/512$: the rank sampler has geometric mass at least $1/2$ and $Z/|\Omega| \geq p_{\mathrm{den}}/4$, which in fact give the stronger bound $p_{\mathrm{den}}/8$. The cap r_s gives mean at least $8(q_s + L_s)$; a conditional Chernoff bound, or the exponential supermartingale for the stopped Bernoulli trials, bounds shortage of q_s opportunities by $\exp(-129L_s)$. The sum of these three errors is below ε_s. This closes the backward induction, including pass-throughs. The same limits on the numbers of trials and calls apply to every bad state. □

For the comparison with exact tests, associate independent blocks of public random bits with each occurrence of a Search call, trial number, and predicate type. A block is sampled when needed. In particular, skipping a test does not change the distribution of the bits used by later trials or by the selection of the private roots. The ideal test accepts precisely the smooth inputs and has no hypothetical extra factor discoveries. The actual recognizer never falsely accepts. Thus, until the first one-sided false FAIL on an invoked smooth input, the actual and ideal executions coincide, including all labels, quotas, retained rows and child decisions, unless the actual routine has already returned a verified proper factor, which ends the search successfully. This coupling looks only at random choices made up to the current call.

H.5 Counting the prescribed initial factorizations

For each possible row attempt read a fresh block of exactly n private bits, interpreted as $h_i \in [0, 2^n)$. Out-of-range values are rejected; for $1 \leq h_i < N$ a nonunit gives a proper gcd. A unit releases only $R_i = h_i^2 \pmod N$ and remains in a separate fixed-width location for the private root. As in Theorem E.5, range, gcd, squaring and clearing use a padded polynomial schedule which does not reveal which unit root in a fixed fibre was used.

On public input R_i, two-dimensional Gauss reduction computes a canonical shortest nonzero vector in

$$\mathcal{L}_{R_i} = \{(u, v) \in \mathbb{Z}^2 : u \equiv R_i v \pmod N\}.$$

Use the fixed tie and sign rules of Lemma 5.1 and primitive normalization, checking every divisor used in normalization against N. If the divisor is nontrivial, return it. Reject zero coordinates or nonunit coordinates, and recognize both absolute coordinates with cutoff Y_0. A retained shape requires each coordinate to be squarefree, to have at most $K = 2k_0$ prime factors, and to have all its prime factors in $(Y_{\min}, Y_0]$. The bound on the shortest vector is $\lceil\sqrt{2N}\rceil$. Thus at most $2K$ top Search requests are launched for a shape, on every tape. Signs enter the row's rational factor q_i.

To bound the probability of these shapes, put

$$\begin{aligned} &m_k = b(k_0), && k = k_0 + \lceil k_0/m_k \rceil, && L_x = \lfloor (n-1)/2 \rfloor - 2, \\ &S_b = ke_0 - L_x, && D_0 = S_b + k, && J_b = \lfloor 2D_0/k \rfloor, && D_c = \lfloor kJ_b/2 \rfloor. \end{aligned} \tag{H.18}$$

Require $S_b \geq 0$, $k \leq K$, $S_b \leq D_c \leq S_b + k$, $2m_k(J_b + 1) \geq e_0$, $J_b + 1 \leq \lfloor e_0/128 \rfloor$, and $e_{\min} > e_1$. These are additional conditions for beginning an iteration. Set

$$P_{\mathrm{sh}} = 2^{16}(n+2)^4(128km_k)^{2k}, \qquad p_{\mathrm{new}} = P_{\mathrm{sh}}^{-1}. \tag{H.19}$$

Lemma H.5 (Probability of the initial factorizations)**.** *Fix the public environment and field. Suppose at least three quarters of the unit primes in each band*

$$(2^{e_0-h-1}, 2^{e_0-h}], \qquad 0 \leq h \leq J_b,$$

are good top parents. For sufficiently large B-rough inputs, a fresh n-bit private block gives a counted accepted shape all of whose parents are good with probability at least p_{new}. Moreover $\log P_{\mathrm{sh}} = (9/20 + o(1))\Lambda$.

Proof. Put $X = \sqrt{N}/4$ and $x_0 = 2^{L_x}$, so $X/2 < x_0 \le X$. All displayed bands lie above $Y_{\min}$ by the last condition on the prime intervals. The coefficient of z^{D_c} in $(1 + z + \cdots + z^{J_b})^k$ is at least $(J_b + 1)^k/(kJ_b + 1)$. Indeed these coefficients are symmetric and unimodal. To verify unimodality inductively, the difference of adjacent coefficients after convolution is the difference of two coefficients of the previous symmetric unimodal sequence, at distances ordered on the increasing half. The central coefficient is therefore at least their average.

In a band with upper endpoint w, the prime number theorem, the three-quarter assumption, and deletion of the at most $b(N)$ primes dividing N and at most k previously chosen primes leave at least $w/(32 \log Y_0)$ choices. This estimate holds uniformly over these bands since their least endpoint tends to infinity. For every ordered band tuple summing to D_c, choose distinct good primes in its positions. The product is at most $Y_0^k 2^{-D_c} \le x_0 \le X$. The product of the band upper endpoints is at least $2^{-k} x_0$. Divide the number of choices by $k!$; unique factorization bounds the overcount by that number. The set $\mathcal{A}$ of obtained squarefree products satisfies

$$\begin{aligned} |\mathcal{A}| &\ge \frac{(J_b + 1)^k Y_0^k 2^{-D_c}}{(kJ_b + 1)k!(32 \log Y_0)^k} \\ &\ge \frac{X}{4(n+2)^2 (128 k m_k)^k}. \end{aligned} \tag{H.20}$$

The last inequality uses the conditions on the central coefficient and the bands, $k! \le k^k$, and $kJ_b + 1 \le (n+2)^2$. Thus no smooth-number theorem for an arbitrary restricted set of good primes is being assumed.

The quotient of the unit group modulo its squares has $2^{\omega(N)}$ elements for odd N, also when N is not squarefree. At least $|\mathcal{A}|^2/2^{\omega(N)}$ ordered pairs of elements lie in the same quadratic class. For $\mathcal{A}_p = \{a \in \mathcal{A} : p \mid a\}$, $|\mathcal{A}_p| \le X/p$ and $\sum_p |\mathcal{A}_p| = k|\mathcal{A}|$. The number of noncoprime pairs is at most

$$\sum_p |\mathcal{A}_p|^2 \le kX|\mathcal{A}|/Y_{\min}.$$

Relative to the same-class bound this tends to zero. Indeed, B-roughness and the product count give

$$\omega(N) \log 2 \le \frac{\log N \log 2}{\log B} = o(\log Y_0),$$
$$\log(X/|\mathcal{A}|) = O(\Lambda) = o(\log Y_0),$$

whereas $\log Y_{\min} \ge (127/128) \log Y_0$. Hence at least $|\mathcal{A}|^2/2^{\omega(N)+1}$ pairs remain coprime.

Their positive reduced quotients $u/v \pmod N$ are distinct: equality makes the determinant a multiple of N of absolute value below $X^2 < N$, so it is zero. Their vectors have squared length at most $2X^2 = N/8$. Any independent lattice vector w has $\|(u,v)\|\|w\| \ge N$ and is strictly longer, and any collinear integer vector is an integer multiple. Thus (u,v) is the unique shortest vector up to sign and is preserved by the canonical normalization. Each quotient has exactly $2^{\omega(N)}$ unit private roots. The count therefore gives probability at least

$$\frac{|\mathcal{A}|^2}{2^{n+1}} \ge \frac{|\mathcal{A}|^2}{4N} \ge p_{\mathrm{new}}.$$

Range and unit filtering of the fixed private block are already included. Finally $k \sim (9/10)d$, $\log k = (1/3 + o(1)) \log \log N$, and $d \sim (3/4) D^{1/3}$. Taking logarithms of (H.19) gives $2(9/10)(3/4)/3 = 9/20$ as claimed. □

Theorem G.1 supplies the premise of this lemma simultaneously for all selected bands, with probability $1 - o(1)$ over the original environment and a single finite Gaussian draw, including the possibility of sampling failure. The theorem averages over the top parents within each band. Combining this estimate with the bound on the probability of rejection during initialization under the same distribution, we find that initialization returns a proper factor or accepts a field with the required good parents with probability at least $1/2$ for sufficiently large inputs. By Lemma H.4 and a union bound on at most $2K$ top calls, a counted shape completes a row or finds a factor with conditional probability at least $1 - 2K\varepsilon_s$. In particular a fresh shape attempt in the exact-predicate process has probability at least $p_{\mathrm{new}}/8$ of one of these outcomes.

H.6 Square certification with bounded running time

The square certificate of Appendix E applies to the single field used here. We bound the work needed to choose its prime probes, evaluate their characters on the completed rows, and recover the square root in the field. All three steps are unconditional.

Verifying the hypotheses of the square certificate. Specialize the field-data symbols J, A, m, V of Section E.1 to $1, A_f, m_N, V_0$, respectively. Write $D_f = eA_f \operatorname{disc} f$. The setup gives primitive irreducible f of odd degree, with $e, A_f, f'(m_N)$ checked units; a continuing path has $e = 1$. Every accepted numerical factor list is verified by Proposition H.3. For an internal atom its known incoming prime is adjoined to the cofactor list, giving the factorization of $|F(P)|$ into primes; at the top no prime needs to be adjoined. Recorded norm primes and all evaluation denominators are gcd-tested against N. Proper gcds exit, and a nonunit without a proper factor rejects the object. Consequently (E.2) and (E.6) give the required evaluation maps; no evaluation on the whole field is used.

The row construction of Appendix E.6, with all atoms in this field and top sign -1, gives the exact checked identity

$$R_i = q_i \prod_P \psi(\xi_P)^{c_{iP}} \pmod N. \tag{H.21}$$

Its rational record uses only signs, Q, and primes at most B. All four column families of (E.27) are retained: rational signs and prime parities, occurrence and good-small-ideal parities, shared norm signs and prime parities, and characters. In particular $E_i = \sum_P c_{iP}$ accounts for the exponent of Q in q_i. Lemma E.1 applies to every atom. Completed descents cancel odd valuations at good large ideals. The group H_{D_f} allows the remaining bad primes.

On discovery, norm columns are zero-extended using the retained complete records and the row parity invariant, as in Appendix E.6. No factorization of D_f is required. Duplicate completed rows are retained, and collection stops at $\mathcal{C} + 1$ rows without testing for a rank increase.

Bounds on columns and probe sizes. Use the integer upper bounds preceding (F.3), with $J = 1$ and the present d, C_F, V_0, n, B. These deterministically bound the field discriminants and every permitted norm of an atom, independently of the coefficient distribution. In particular set

$$\begin{aligned} S^* &= (2d)!(T^*)^4, \qquad t^* = 20d\{h^* + b(64) + 1\}, \\ \mathcal{C} &= 2B + 2 + dB + 1 + t^* + L_D^*, \qquad x = 2^{S^*}. \end{aligned} \tag{H.22}$$

Compute $T^*, h^*, L_D^*, \mathscr{D}_*, \mathscr{N}_*$ from the same integer formulas before drawing any rows. Their asymptotic costs at the new scales are computed in Section H.9.

Proposition E.3 now gives the same detecting fraction $1/(20d)$ for nonrational classes in H_{D_f}. All probes exceed $2, |D_f|, \mathscr{N}_*$, and an ideal integer proposal with a uniform index among d root positions accepts with probability at least $c/(dS^*)$, for an absolute $c > 0$. The class and discriminant bounds are those of Lemma E.2; no number-field GRH is used. For the simultaneous event, the integer upper bounds give $|H_{D_f}| \le H_* \le 2^{h^*}$: $X_0 \le 2^{L_g^*}$ and $1 + \log X_0 \le L_g^* + 3$ bound the corresponding factors of H_*. Thus

$$H_* e^{-t^*/(20d)} \le \exp\{-(1 - \log 2)h^* - b(64) - 1\} = o(1). \tag{H.23}$$

The exceptional probability in (H.23) covers the whole group, including any row products chosen later.

Proposition H.6 (Bounded character construction)**.** *Put $k_c = n^2$ and*

$$\begin{aligned} R_\chi &= n^2 dS^* b(64t^* + 2), & Q_{\mathrm{pr}} &= t^* R_\chi, \\ K_{\mathrm{sc}} &= Q_{\mathrm{pr}}\{k_c + (d+1)^2 k_c + 2\}. \end{aligned} \tag{H.24}$$

There is a finite implementation using at most R_χ candidates per required probe, whose work on every tape is $\mathrm{poly}(n, d)(S^* + 1)^4$*. Except with probability*

$$Q_{\mathrm{pr}} 4^{-k_c} + d Q_{\mathrm{pr}} (7/9)^{k_c} + K_{\mathrm{sc}} 2^{-k_c} = o(1), \tag{H.25}$$

it agrees with the true-prime, full-root proposal law of Proposition E.3. Its search exhaustion probability is also $o(1)$. No error in this implementation can invalidate a subsequently verified proper integer factor.

Proof. A scalar in an interval of length $L > 1$ is drawn using $\lceil \log_2 L \rceil$ bits, accepting a block below L. Allow at most k_c blocks; exhaustion discards the current candidate. Conditional on return the scalar is exactly uniform, and its probability of exceeding the prescribed limit is at most 2^{-k_c}. Length-one intervals use no bits. A candidate in $(2^{S^*-1}, 2^{S^*}]$ has a direct dyadic draw. Even candidates and candidates below three are rejected.

Apply k_c independent strong probable-prime tests with uniform bases in $\{1, \dots, p-1\}$. A true prime always passes; for each fixed composite the chance of passing all bases is at most 4^{-k_c} [57, Theorems 1 and 2]. This bound is conditional on the candidate being a fixed composite.

Suppose that p is prime. Compute

$$h_p = \gcd(f, X^p - X) \quad \text{in } \mathbb{F}_p[X].$$

The power is computed modulo f. Since $p > |D_f|$ and $p > 2$, f retains its degree and is squarefree. Thus h_p is exactly the product of its distinct linear factors. Split a fully split monic factor h_0 of degree $r \ge 2$ as follows: draw a uniform polynomial A of degree below r, compute

$$A^{(p-1)/2} \pmod{h_0}, \qquad \gcd(h_0, A^{(p-1)/2} - 1),$$

and keep a proper split. Cap attempts at k_c per node and process nodes in a fixed order; exhaustion discards the candidate. Evaluation at the r distinct roots is a bijection from these polynomials to $\mathbb{F}_p^r$. For any two roots, their evaluations are nonzero with opposite quadratic signs with probability $2((p-1)/(2p))^2 \ge 2/9$. Such an attempt splits them. There are at most $d - 1$ splitting nodes, and the total true-prime splitting probability of exceeding the prescribed limit is at most $d(7/9)^{k_c}$, also for adaptive nodes. Sort all roots and verify distinctness, vanishing, and that the product of the returned linear factors equals h_p.

Draw a uniform index in $\{1, \dots, d\}$; reject if it exceeds the number of returned roots. In the ideal procedure a particular pair (p, u) has probability exactly

$$\frac{2}{x}\frac{1}{d} = \frac{2}{dx} \tag{H.26}$$

per integer proposal. Drawing the index among all d positions gives the probability in (H.26) to each root, regardless of how many roots f has at p.

For a composite surviving the probable-prime test, all arithmetic is still bounded. Each inverse is checked by an integer gcd and a nonunit rejects the candidate; each polynomial division must decrease degree, and every quotient and loop is checked. These checks ensure termination even when a composite passes the primality test.

Index public random blocks by probe, candidate, base, splitting node, attempt, and coefficient. Couple to unlimited exact sampling from integer intervals and true-prime/full-root arithmetic using these same blocks. The processes agree until the first scalar timeout, composite false pass, or true-prime split timeout. Union bounds give (H.25); the scalar count is bounded by K_{sc}. The ideal acceptance probability $c/(dS^*)$ and the cap R_χ give a further $o(1)$ exhaustion probability. Here $\log Q_{\mathrm{pr}} = O(\Lambda) = o(n^2)$ and $\log t^* = O(\log n)$. Each error estimate applies conditional on the preceding draws, up to the first disagreement.

For $s = b(p) \le S^* + 1$, modular powering uses $O(s)$ multiplications of s-bit integers, hence $O(s^3)$ schoolbook bit operations. The bounded polynomial degrees and the n^2 repetitions multiply this by $\mathrm{poly}(n, d)$. There are $Q_{\mathrm{pr}} = \mathrm{poly}(n, d)S^*$ candidate positions, proving the fourth-power bound. Current-input primality and all returned factors are still certified exactly on polynomial-bit integers. The final modular-square and divisor tests below remain exact even on an erroneous auxiliary-prime branch. □

Character evaluation on completed rows. Store the t^* probe pairs once in a read-only table. Each Search state stores a reference to the table. Search, child creation, caching, and all acceptance predicates use no character column. Only after a whole row and its successful child trees have been completed, evaluate

$$\chi_{p,u}(\xi_P) = (a - bu)^{(p-1)/2} \pmod p$$

for its retained occurrences. At a true probe it is $+1$ or -1, since every permitted atom is a unit there; reject the epoch if a computed value is not a sign. If z_P is used instead, the A_f-symbol is cancelled by the E_i column. For the upper bound on the number of retained occurrences A_{occ} defined below, the total charge is

$$\mathrm{poly}(n, d)A_{\mathrm{occ}}t^*(S^* + 1)^3. \tag{H.27}$$

This mixed term is kept separately from the fourth-power setup term. Failed candidates and failed child trees use sparse local records; they do not copy the probe table, scan a dense retained matrix, or allocate a character vector. Clearing local scratch touches only those local records. This explicit memory discipline is why no $P_{\mathrm{raw}}S^*$ table-copy charge occurs.

H.7 Computation of the public square root

Lemma H.7 (Computing a square root in a number field)**.** *Let $g \in \mathbb{Z}[T]$ be monic irreducible of degree d and let $\Gamma \in \mathbb{Z}[T]/(g)$ be explicitly expanded. There is an unconditional deterministic*

algorithm deciding whether Γ is a square in $\mathbb{Q}[T]/(g)$ and returning an exactly verified root when it is. If h_g, h_Γ bound the coefficient bit lengths, its bit cost is

$$\mathrm{poly}(d)(2 + h_g + h_\Gamma)^3.$$

In particular, no integer factorization, maximal-order basis, unit group, or class group is required.

Proof. Let $K = \mathbb{Q}[T]/(g)$ and let β be the residue class of T. If $\Gamma = 0$, return zero. Otherwise form, for each integer $0 \le s \le 2d(d-1)$,

$$P_s(X) = \mathrm{Res}_T(g(T), (X - sT)^2 - \Gamma(T)).$$

Writing $\beta_1, \dots, \beta_d$ for the conjugates of β, the $2d$ roots of P_s are $s\beta_i \pm \sqrt{\Gamma(\beta_i)}$. Two roots belonging to the same embedding are distinct. Each pair from different embeddings excludes at most one value of s; there are at most $2d(d-1)$ such signed pairs. Thus one displayed shift has squarefree P_s. The algorithm finds it by exact resultants and a polynomial gcd, without computing embeddings.

The chosen P_s is monic, has degree $2d$, and has coefficient bit height $O(d(h_g+h_\Gamma+\log(2d)))$, by its Sylvester determinant. Factor it over $\mathbb{Q}$ by the deterministic LLL algorithm [39, Theorem 3.6]. If Γ is a square, squarefreeness forces the norms of its two linear factors over K to be distinct irreducible factors of degree d. If it is not a square, the norm of the irreducible quadratic is a power of one irreducible rational polynomial, and squarefreeness forces that power to be one. Trager's norm/gcd reduction [64, Theorem 2.2 and Corollary 2.5] therefore decides the alternatives and recovers a linear factor when it exists. More explicitly, reduction of a suitable rational factor modulo $(X - s\beta)^2 - \Gamma$ gives $uX + v$ with $u \neq 0$; its root is $-v/u$, and subtracting $s\beta$ gives the required square root. One field inverse is obtained by extended Euclid modulo g.

Fraction-free resultants, gcds, and the final inverse have cost $\mathrm{poly}(d)(2 + h_g + h_\Gamma)^2$; the rational LLL factorization has cubic dependence on the coefficient bit height. These are the claimed costs. Exact squaring verifies the answer. This explicit quadratic specialization also follows from the general polynomial factorization algorithm over number fields [41]. □

For a nonempty binary dependence I, use the selected-product and signed-squarefree conventions (E.4)–(E.5) with one field. The columns above give even E, $\gamma \in H_{D_f}$, and positive rational squares q_I and δ. On the simultaneous separation event, the odd-degree norm argument in Theorem E.4 proves that $\delta\gamma$ is a field square. This uses the norm and sign columns as well as ideal parity.

Construct the same integral target as (E.16):

$$\gamma^- = \prod_{c_P<0} z_P^{|c_P|}, \qquad \Gamma = \delta\, g'(\beta)^2 \prod_P z_P^{|c_P|}. \tag{H.28}$$

The identities in Theorem E.4 make Γ a square; its trace-dual step (E.17) places every root in $\mathbb{Z}[\beta]$. We obtain an exact root η of Γ by applying Lemma H.7 to its expanded coefficients. Reduce the returned coefficients over $\mathbb{Q}$ to their integer power-basis values before evaluating them modulo N; intermediate rational denominators of the factorization algorithm are not inverted modulo N.

The one-field specialization of (E.19) is

$$Y = \psi_{\mathrm{rat}}\big(\sqrt{q_I/\delta}\big) A_f^{-E/2} \frac{\psi(\eta)}{\psi(\gamma^-)\psi(g'(\beta))} \pmod N. \tag{H.29}$$

Compute the rational root from the saved factor exponents. The localization conditions above ensure that each denominator in (H.29) is a checked unit. Verify $\eta^2 = \Gamma$ exactly, that Y is a unit, and that $Y^2 = \prod_{i \in I} R_i \pmod N$. Failure rejects the epoch. Acceptance is correct even outside the separation and true-probe events.

The coefficient bound (E.18) applies with $R \le 1 + C_F^d$, $Z \le 1 + V_0(C_F + R)$, and $\log_2(2 + |\delta|) \le O(nA_{\mathrm{occ}})$. The same conjugate and Vandermonde bounds applied to Γ and partial products are covered by the explicit expanded coefficient height bound $\mathscr{L}$ in (H.30). Sequential product construction costs $\mathrm{poly}(n,d)A_{\mathrm{occ}}\mathscr{L}^2$; Lemma H.7 and exact verification cost $\mathrm{poly}(n,d)(A_{\mathrm{occ}} + \mathscr{L} + \mathcal{C} + 2)^3$. These charges contain no S^* and are polynomial in the *expanded* height, rather than in an unexpanded product description.

H.8 Storage, random sampling and factor extraction

First compute the certificate parameters in (H.22). In particular, $\mathcal{C}$ already includes the character columns. The following bounds now depend only on the input and the parameters of the program:

$$\begin{aligned}
R_{\mathrm{shape}} &= 32(\mathcal{C}+1)P_{\mathrm{sh}},\\
T_{\mathrm{Search}} &= 2KR_{\mathrm{shape}}(q_s M_c + 1)^{j+2},\\
P_{\mathrm{raw}} &= T_{\mathrm{Search}} r_s,\\
A_{\mathrm{occ}} &= 2K(\mathcal{C}+1)(M_c+1)^{j+2},\\
R_* &= 1 + C_F^d, \qquad Z_* = 1 + V_0(C_F + R_*),\\
L_0 &= 10 + b(d!) + (d-1)^2 b(R_*) + nA_{\mathrm{occ}}\\
&\qquad + (d-1)b(2R_*) + A_{\mathrm{occ}} b(Z_*),\\
\mathscr{L} &= (n+2)^{10}(L_0+1).
\end{aligned} \tag{H.30}$$

The epoch tries at most R_{shape} fresh private blocks, retains the next completed row whenever one occurs, and stops after $\mathcal{C}+1$ rows. A whole failed tree is discarded; retained rows keep only their successful trees. Each verified candidate launches at most M_c children, and only q_s candidates per Search do so. Summing the finite search tree therefore gives the stated T_{Search} and P_{raw}. In a final successful tree the multiplier is M_c, not $q_s M_c$, which gives A_{occ}. An occurrence is counted with multiplicity even when a cached factor record or relation is used again. No failed-attempt count enters the retained height $\mathscr{L}$. The term nA_{occ} bounds the size of the rational correction.

Evaluate characters only on retained rows and store the character table as specified in Section H.6. Failed trees keep only sparse local records; dense row registration and matrix access begin on retention. Thus trials incur neither $P_{\mathrm{raw}}\mathcal{C}^2$ matrix work nor $P_{\mathrm{raw}}S^*$ table copying. The rule retaining duplicate completed rows excludes an additional rank-filter cost.

For the outer scalar choices put

$$K_U = 16(1 + R_{\mathrm{env}} + d + P_{\mathrm{raw}} + Q_{\mathrm{pr}}), \qquad u_U = b(128K_U) + n. \tag{H.31}$$

Use at most u_U fixed-width rejection blocks for any such uniform interval choice; exhaustion fails the epoch. This agrees with unlimited exact sampling from integer intervals except with probability at most $K_U 2^{-u_U} \le 2^{-n}/128$. The exact Ω rank uses just its one direct dyadic block; private roots use exactly n bits; the Gaussian and genus-two routines have their own caps. Internal scalar choices of the large certificate use k_c and K_{sc} from Proposition H.6. They are not assigned an extra Markov time multiplier. If desired, the explicit rational expression

in (H.25) can be checked to be at most $1/4096$ as an additional admissibility condition; the check is polynomial in n and eventually passes.

The number of invoked smoothness predicates is bounded by

$$Q_{\mathrm{rec}} = 1 + 2R_{\mathrm{shape}} + 2P_{\mathrm{raw}}, \qquad L_{\mathrm{conf}} = b(2048Q_{\mathrm{rec}}(n+2)). \tag{H.32}$$

Use $L_c = L_{\mathrm{conf}}$ in every recognizer. Its actual caller cutoff is retained: Y_0 for a private coordinate, B for a rational cofactor, and Y_{i+1} for an algebraic cofactor at level i (including Y_1 at the top). All these cutoffs are at least B. Proposition H.3 is uniform in the numerical input, so, conditional on every public past, each invoked smooth query has error at most $2n\exp(-64L_{\mathrm{conf}})$ for sufficiently large N. Padding unused query positions by zero error gives an adaptive first-error union bound below $1/1024$. This is an error for the entire epoch, not a fixed absolute loss subtracted at each rare successful raw point. On the coupled event the ideal and actual processes have identical accepted shapes and child trees, unless an earlier verified proper factor has already been returned. Every tape has at most $T_{\mathrm{Search}}r_A$ expensive algebraic recognitions, regardless of whether that event occurs.

After a nonempty dependence and the verified public root Y have been determined, the first private operation involving retained roots is

$$d_N = \gcd\left(\prod_{i\in I} h_i - Y, N\right). \tag{H.33}$$

A nontrivial factor is verified by integer division and returned. A trivial result discards the stored private roots and fails the epoch. Only one gcd involving the retained private roots is computed in each run. All choices of fields, probes, Gaussian coefficients, geometric points, child retries, row retention, kernel vectors and public square roots depend only on public data and independent public randomness. Private records are not sorted, compared, hashed, compressed or deduplicated; a new row always has a separate record.

These are exactly the hypotheses of Theorem E.5: the private initialization is padded, every retained row has a separate private record, all intervening computation is public, and the nonempty selected product and verified public square precede the first gcd using the private roots. Its conditional proper-factor probability is therefore

$$1 - 2^{1-\omega(N)} \geq \tfrac{1}{2}. \tag{H.34}$$

This applies also to odd nonsquarefree N, and to a public transcript on which an auxiliary composite passed a probable-prime test: the theorem requires the final checked square identity, not correctness of every auxiliary predicate.

Proposition H.8 (Probability of finding a factor in one epoch)**.** *Assume GRH. For every sufficiently large odd, B-rough current input which is not a prime power, a fresh finite epoch of* G *returns a verified proper factor with probability at least* $1/512$*, conditional on every earlier public history of the factoring algorithm.*

Proof. The theorem on relation counts and the setup analysis give probability at least $1/2$ of either an immediate factor or an accepted field having the dense good-prime sets in every selected band. This statement is made on the original environment law; no future acceptance conditioning is imposed. Fix such a field. Lemmas H.5 and H.4 give conditional row-or-factor probability at least $p_{\mathrm{new}}/8$ in each fresh attempt of the ideal exact-predicate process. The cap R_{shape} gives at least $\mathcal{C}+1$ completed rows except with probability $\exp(-c(\mathcal{C}+1))$, by the

conditional Chernoff bound. An earlier factor is absorbing; inserting virtual completed rows after that outcome makes the same proof valid.

Character search exhaustion and simultaneous class-separation failure are $o(1)$ uniformly over accepted fields, by (H.23) and Proposition H.6. Outside these events, the matrix dependence and exact square-root computation succeed. All these pre-extraction events are public. Equation (H.34) supplies the final factor with conditional probability at least $1/2$.

Finally couple the finite environment scalars, bounded probe arithmetic and actual smoothness tests to this ideal process. Their total error is $o(1)+1/1024$. Subtracting this error from $1/4-o(1)$ leaves at least $1/512$ for sufficiently large inputs. Every new epoch uses fresh public and private random bits, so the same bound holds after any previous public history.

□

H.9 Expected bit complexity

Proposition G.13 verifies the specified parameters and every explicit parameter condition. We now bound the running time of the bounded program at those scales. In particular, $\Lambda_d = (\log 2)dm_b = (3/4 + o(1))\Lambda$ and $j = O(\log D)$.

Equations (H.22), (H.15) and (H.30) give

$$\begin{aligned}
\frac{\log \mathcal{C}}{\Lambda} = \frac{\log A_{\mathrm{occ}}}{\Lambda} = \frac{\log \mathscr{L}}{\Lambda} = \frac{19}{20} + o(1), \qquad \frac{\log S^*}{\Lambda} = \frac{1}{2} + o(1),\\
\frac{\log T_{\mathrm{Search}}}{\Lambda} = \frac{7}{5} + o(1),\\
\frac{\log P_{\mathrm{raw}}}{\Lambda} = \frac{237}{80} + o(1),\\
\frac{\log(T_{\mathrm{Search}} r_A)}{\Lambda} = \frac{62}{25} + o(1).
\end{aligned} \tag{H.35}$$

The factors arising from the depth of the descent have logarithm $O((\log n)^2) = o(\Lambda)$. The factorial in S^* contributes $\log(2d)! = (1/2 + o(1))\Lambda$; its other factors have polynomial bit length. Consequently, arithmetic modulo a character prime must be counted separately from arithmetic on integers of polynomial bit length in n.

The largest smoothness bound for the coordinates of the initial lattice vectors is Y_0, and the largest smoothness bound used on the algebraic side is Y_1. Proposition H.3 gives the following leading constants for a single call:

$$h_0 = \frac{201}{100}\sqrt{\frac{40}{81}}, \qquad h_1 = h_0\sqrt{\frac{51}{430}}.$$

The following inequalities suffice for the running-time bound:

$$h_0^2 = \frac{4489}{2250} < \left(\frac{3}{2}\right)^2, \qquad h_1^2 = \frac{76313}{322500}, \qquad \left(\frac{487}{1000}\right)^2 - h_1^2 = \frac{69601}{129000000} > 0. \tag{H.36}$$

Recognition at cutoff B has logarithmic cost $o(\Lambda)$. The confidence factor L_{conf} and all explicit polynomial factors also have logarithm $o(\Lambda)$.

The first two rows include all deterministic primality tests on data of polynomial bit length, exact gcds and verified local factor-list handling. The $2^{O(d)}$ initial parity-polynomial search is $\exp(o(\Lambda))$. The shape coordinates require at most $2R_{\mathrm{shape}}$ tests. The limit on the number of tests in each search bounds the total number of algebraic tests by $T_{\mathrm{Search}} r_A$. Each recognizer's bounded

Table 2: Bit-complexity bounds for one iteration of G, valid for every sequence of random bits. An entry c denotes an upper bound $\exp((c+o(1))\Lambda)$.

Operation	Leading upper bound
Base trial division and polynomial-size setup	$19/20$
Sampling lattice points, rational-side smoothness tests, and storage of unsuccessful searches	$237/80 = 2.9625$
Smoothness tests on the initial lattice vectors	$7/5 + h_0 < 29/10$
Algebraic-side smoothness tests	$62/25 + h_1 < 2967/1000$
Selection of character primes and computation of their roots	2
Characters on retained occurrences	$49/20 = 2.45$
Products, norm correction, Gaussian elimination, and algebraic square-root computation	$57/20 = 2.85$

internal factor tree is already included in its cost. Constructing the probes and evaluating the retained characters cost $(S^*)^4\,\mathrm{poly}(n)$ and $A_{\mathrm{occ}}(S^*)^3\,\mathrm{poly}(n)$, respectively. The last row includes $\mathcal{C}^3$ dense binary elimination, $A_{\mathrm{occ}}\mathscr{L}^2$ product expansion, and $\mathscr{L}^3\,\mathrm{poly}(n,d)$ root reconstruction and verification. Registered norm and character columns, rational correction, integer output heights and legal denominators are included in these bounds. The bounds require neither additional integer factorizations nor unbounded searches. We use dense Gaussian elimination and the algebraic square-root algorithm proved above.

Theorem H.9 (Running-time bound under GRH)**.** *The program specified in this appendix, with the trial-division procedure below, factors every integer $N \geq 2$ correctly and terminates on every random tape. Under GRH its expected bit complexity is at most $L_N[1/3, 297/100]$. Under the same hypothesis, each fast epoch finds a proper factor with probability at least $1/512$ on every sufficiently large nonterminal input. No extension of GRH to general number fields or Hecke characters, smoothness heuristic, or integer-factorization oracle is required.*

Proof. First consider correctness and termination, without GRH. All loops within a fast epoch have the displayed finite limits; all arithmetic accepts only checked identities. Prime factors are deterministically certified and every returned divisor is verified. At a fixed composite input, maintain a trial divisor t_0, initially two. After each failed epoch test the next integer t_0, increment it, and never reset this counter while splitting this input. By $t_0 = \lfloor\sqrt{N}\rfloor$ a proper divisor must have been found if no earlier fast epoch succeeded. Thus each splitting call terminates on every tape. The factoring algorithm recurses only on verified proper factors, hence terminates by induction on the input. This is the trial-division procedure of Section 6, with no auxiliary recursive factorization inside an epoch.

Assume now GRH. Proposition H.8 gives a conditional success probability at least $1/512$ on every sufficiently large nonterminal input, after every past. The number of fresh epochs has expectation at most 512; the trial divisions can only shorten it. The cost bound holds on every execution of an epoch, including a failed one. By (H.36), its largest leading coefficient is strictly below $2967/1000 < 297/100$. The $3/1000$ strict margin absorbs all fixed polynomial and growing-depth factors. In particular the expected local split work is at most $L_N[1/3, 297/100]$ without any assumed independence between cost and success.

There are at most $2\lfloor\log_2 N\rfloor - 1$ nodes in the verified outer factor tree, counting prime multiplicities. Monotonicity of the asymptotic upper bound and the strict margin absorb this polynomial factor. Inputs below the fixed eventual threshold have a uniform finite worst-case bound from the explicit limits and trial division; enlarging the global prefactor covers this finite

range. This proves the expected-time bound. The relation counts use GRH, whereas the analysis of the smoothness test requires only RH. The square certificate, exact field root, extraction symmetry, and correctness of every output are unconditional, as proved above. □