

POWER-KERNEL FLUCTUATIONS, THE LAMBERT W SUBPOLYNOMIAL SOURCE, AND A LOCAL CRT CONDITIONAL LAW

CHUNPING ZENG

ABSTRACT. For a positive integer n we define the *power kernel* $d_n = n/\text{rad}(n)$ and the *fluctuation* $X_n = \log d_n$, where $\text{rad}(n) = \prod_{p|n} p$. This paper develops three mutually supporting lines around this “relative fluctuation”.

First, the microscopic theory. From the kernel decomposition $n = v_d s$ (s squarefree, $(s, d) = 1$, $v_d = d \text{rad}(d)$) we obtain the exact discrete law

$$\mathbb{P}(X_n = \log d) = \frac{6}{\pi^2} \frac{w(d)}{d \text{rad}(d)}, \quad w(d) := \prod_{p|d} \frac{p}{p+1},$$

together with a three-branch parity-conditioned refinement. Parity is a first-order effect: the measured gap $\mathbb{E}[X_n | n \text{ even}] - \mathbb{E}[X_n | n \text{ odd}] = 0.69008$ matches $\log 2 = 0.69315$.

Second, the conditional law. For fixed c , the distribution of X_{c-a} is completely determined by $\text{mask}(a)$ through a three-case local CRT convolution: the local contribution vanishes identically when $p | c$ or $p | a$, and otherwise follows a *corrected* law rather than the usual one, differing by a factor $\frac{p}{p-1}$ (a factor 2 at $p = 2$). The corrected law agrees with 220 masks to 0.993, while the usual law underestimates the tail by 47%–78%. This replaces an earlier erroneous assertion of equidistribution.

Third, the subpolynomial source. For radical-type Euler products $\prod_p (1 + \frac{\eta(p)}{p^\sigma - 1})$ with $\eta(p) \rightarrow a$, one has as $\sigma \rightarrow 0^+$

$$\Phi(\sigma) = \frac{a}{\sigma W(a/\sigma)} \left[1 + \frac{\kappa_1}{W} + \frac{\kappa_2}{W^2} + \cdots \right],$$

where W is the Lambert function. The main term comes not from small primes but from the prime range $[e^{v^*}, e^{1/\sigma}]$ — which explains why no finite prime table can reproduce it. The coefficients satisfy $\kappa_n = (n-1)! + \sum_j c_{n,j} \zeta(2j)$, a sum over *even* zetas only, because the two-sided integral coefficients satisfy $\iota(r) = (-1)^{r+1} \rho(r)$, forcing $\zeta(\text{odd})$ -terms to cancel exactly at every order. This unifies $T(x) = \sum_{n \leq x} 1/\text{rad}(n)$, $\Sigma(D)$ and $\Sigma_\perp(D; Q)$, and explains at one stroke why all previously fitted “constants” are transient.

Applications include: a quantitative formula for Mazur parameters in the open range $1 < s < 2$ ($k_{\text{eff}} = 1.91 \pm 0.17$, excluding $k = 1, 3$); the ceiling theorem $q \leq \frac{\log c}{\log 2 + \log \text{rad}(c)}$ (equality iff $\{a, b\} = \{1, 2^k\}$); the $\text{rad}-P^+$ duality (principal versus negative Lambert branch); an exact S -unit reformulation of $\rho(c) = \min_a \text{rad}(a) \text{rad}(c-a)$; and a mechanistic decomposition of the exponent of $\mathcal{N}_0(N)$.

We report honestly: the abc conjecture itself is not attackable within this framework. The three structural walls are: (i) the conditional law is provable for $D \leq c^{1/2-\delta}$ while abc solutions live at $D \gtrsim c^{0.78}$; (ii) relaxing any one of the three radicals inflates the count by a factor 387; (iii) the Fourier circle method is controllable iff $\alpha + \beta > 2$, while Mazur parameters always have $\alpha, \beta < 1$. Our weak abc bound is weaker than de Bruijn’s 1962 trivial bound; this is recorded as well.

1. INTRODUCTION AND NOTATION

1.1. Background. Let $a + b = c$ with $(a, b) = 1$, and write $\text{rad}(n) = \prod_{p|n} p$ and $q(a, b, c) = \frac{\log c}{\log \text{rad}(abc)}$. The abc conjecture (Oesterlé–Masser) asserts that for every $\varepsilon > 0$ there are only finitely many primitive triples with $q > 1 + \varepsilon$.

The starting point of this paper is the observation that $(a, b) = 1$ implies that a, b, c are pairwise coprime, so $\text{rad}(abc) = \text{rad}(a)\text{rad}(b)\text{rad}(c)$, and therefore

$$q > 1 \iff \text{rad}(a)\text{rad}(b) < d_c, \quad d_c := c/\text{rad}(c),$$

a purely multiplicative inequality in which d_c is the *power kernel* of c . This combinatorial restatement concentrates all the difficulty of the additive equation in one place: fixing a and c , all randomness lies in $b = c - a$. The counting problem for abc becomes a problem about the distribution of power-kernel fluctuations in arithmetic progressions.

We state the conclusion in advance: the abc conjecture itself is *not* resolved in this framework. But the pursuit produces an independent, provable theory — the Lambert W sub-polynomial source for radical-type Euler products (Section 7) — which we regard as the paper's main contribution.

Status convention. Every statement is tagged [**proved**] (complete proof), [**numerical**] (systematic numerical verification, no rigorous proof), or [**conjecture**]. Assertions later withdrawn are listed with their correct replacements in the Falsification List, Section 15.4; nothing is silently corrected.

1.2. Notation. For $n \in \mathbb{N}$: $d_n := n/\text{rad}(n)$ and $X_n := \log d_n$. For $d \in \mathbb{N}$: $v_d := d/\text{rad}(d)$, $w(d) := \prod_{p|d} \frac{p}{p+1}$, $u(d) := \frac{w(d)}{d/\text{rad}(d)}$, $K_6 := 6/\pi^2$, $\Pi_Q := \prod_{p|Q} \frac{p^2}{p^2-1}$, $G := K_6\Pi_Q$. W is the Lambert function ($e^{W(x)}W(x) = x$), γ the Euler constant, ζ the Riemann zeta function.

State equation. For $\beta > 0$: $C(\beta) := \sum_p \frac{\log p}{p^\beta(p-1)}$, $P(s) := \sum_p p^{-s}$. Numerically $C(1) = 0.755368$, $C(2) = 0.262276$, $P(2) = 0.452247514$.

Counting function. $\mathcal{N}_\varepsilon(N) := \#\{(a, b, c) : a + b = c, (a, b) = 1, c \leq N, q > 1 + \varepsilon\}$.

Mazur parameters. For $0 < \alpha, \beta, \gamma < 1$, $s := \alpha + \beta + \gamma$, $S_{\alpha, \beta, \gamma}(N) := \#\{a + b = c \leq N : (a, b) = 1, \text{rad}(a) \leq a^\alpha, \text{rad}(b) \leq b^\beta, \text{rad}(c) \leq c^\gamma\}$.

Tail functions. $\Sigma(D) := \sum_{d > D} u(d)$; with coprimality conditions, $\Sigma_\perp(D; Q)$.

2. THE EXACT DISCRETE LAW OF THE POWER KERNEL

Lemma 2.1 (Kernel decomposition). $X_n = \log d$ if and only if $n = v_d s$ with s squarefree and $(s, d) = 1$.

Proof. Write $n = \prod_p p^{e_p}$. Then $d_n = \prod_{e_p \geq 2} p^{e_p-1}$ and $\text{rad}(d_n) = \prod_{e_p \geq 2} p$. Hence $d_n = d$ forces $s = n/v_d$ to be squarefree and coprime to d ; conversely, if s is squarefree with $(s, d) = 1$, then the p -adic orders in $n = v_d s$ are $v_p(d) + 1 \geq 2$ for $p | d$ and 1 for $p | s$, so the support and exponents match exactly. $\square$

Theorem 2.2 (Exact discrete law). For every $d \geq 1$,

$$\mathbb{P}(X_n = \log d) = K_6 u(d) = \frac{6}{\pi^2} \cdot \frac{w(d)}{d/\text{rad}(d)}.$$

Proof. By Lemma 2.1, $\mathbb{P}(X_n = \log d) = \rho(d)/v_d$ where $\rho(d)$ is the local density of admissible s . For a prime p : if $p | d$ then $p \nmid s$ (factor $1 - \frac{1}{p}$); if $p \nmid d$ then $v_p(s) \leq 1$ (since s is squarefree, factor $1 - \frac{1}{p^2}$). By CRT the primes are independent, so

$$\rho(d) = \frac{1}{\zeta(2)} \prod_{p|d} \left(1 - \frac{1}{p}\right) \prod_{p \nmid d} \left(1 - \frac{1}{p^2}\right) = K_6 \prod_{p|d} \frac{1 - \frac{1}{p}}{1 - \frac{1}{p^2}} = K_6 \prod_{p|d} \frac{p}{p+1} = K_6 w(d). \quad \square$$

Proposition 2.3 (Normalization). $\sum_{d \geq 1} u(d) = \frac{\pi^2}{6} = \zeta(2)$, via the Euler factor $1 + \frac{1}{p^2-1} = \frac{p^2}{p^2-1}$.

Verification. The partial sum over $d \leq 2 \times 10^6$ is 1.64492235, relative difference 7.1×10^{-6} from truncation alone. $\square$

Proposition 2.4 (Moments). *With $Y_p(n) := (v_p(n) - 1)_+ \log p$ one has $X_n = \sum_p Y_p(n)$ and*

$$\begin{aligned}\mathbb{E}[X_n] &= C(1), \\ \mathbb{E}\left[\sum_{p|n} \log p\right] &= C(1) - C(2) = \sum_p \frac{\log p}{p^2(p-1)}, \\ \mathbb{E}\left[\sum_{p|n} (\log p) \mathbf{1}_{p||n}\right] &= C(2).\end{aligned}$$

Proof. $\mathbb{P}(v_p = k) = (1 - \frac{1}{p})p^{-k}$ for $k \geq 0$ gives $\mathbb{E}[(v_p - 1)_+ \log p] = \frac{\log p}{p(p-1)}$; similarly $\mathbb{E}[(v_p - 1)_+ \mathbf{1}_{v_p \geq 2}] \log p = \frac{\log p}{p^2(p-1)}$. Summing over p gives $C(1)$, and the difference of the two sums is exactly the contribution of primes dividing n exactly once. $\square$

Numerical check. Direct summation against Theorem 2.2 gives $\mathbb{E}[X_n] = 0.755255$ versus $C(1) = 0.755368$ (relative difference 1.5×10^{-4} ; two independent paths are self-consistent). Also $\mathbb{E}[X_n^2] = 2.194778$, $\text{Var}(X_n) = 1.624368$.

Proposition 2.5 (Tail function). $\mathbb{P}(X > y) = K_6 \Sigma(e^y)$, and numerically $\mathbb{P}(X > y) \approx 0.67 y e^{-y}$ in the computable range.

Corollary 2.6 (Finite-size ceiling). $\Xi_n \leq (\lfloor \log_2 n \rfloor - 1) \log 2$; for $n \leq 10^5$ this bound is 10.397.

Proof. $d_n = d$ forces $v_d \leq n$, and $v_d = d \text{ rad}(d)$ grows slowest for $d = 2^k$: then $v_{2^k} = 2^{k+1} \leq n$ gives $k \leq \log_2 n - 1$ and $X_n = (k - 1) \log 2$. $\square$

Important remark. $\Sigma(D) \approx \kappa' D \log D$ with $\kappa' \approx 1.1$ is a *transient* law, not an asymptotic one; the true asymptotic is the Lambert source of Section 7. This is the common root of all “constant drift” phenomena below.

2.1. Parity is a first-order effect.

Discovery 2.7 (Numerical). $\mathbb{E}[X \mid n \text{ even}] = 1.09696$, $\mathbb{E}[X \mid n \text{ odd}] = 0.40687$; the difference 0.69008 matches $\log 2 = 0.69315$.

Mechanism. Writing $n = 2^k m$ with m odd, $X_n = (k - 1) \log 2 + X_m$, and $\mathbb{E}[k - 1 \mid n \text{ even}] = 1$ (geometric distribution). The parity gap is exactly $\log 2$.

Theorem 2.8 (Parity-refined discrete law). *Let $(d, Q) = 1$ be implicit in the support. Then*

$$\mathbb{P}(X_n = \log d \mid (n, Q) = 1, n \equiv r \pmod{2}) = G u(d) \times \begin{cases} 1, & 2 \mid Q, r = 1, \\ 2 \mathbf{1}_{r=0}, & 2 \nmid Q, 2 \mid d, \\ \frac{2}{3} \mathbf{1}_{r=0} + \frac{4}{3} \mathbf{1}_{r=1}, & 2 \nmid Q, 2 \nmid d. \end{cases}$$

(In the first branch the hypotheses force n odd, hence d odd, so $2 \mid d$ does not arise.)

Proof. By Lemma 2.1, $n = v_d s$ with s squarefree, $(s, dQ) = 1$; the parity of n is the parity of $v_d s$. If $2 \mid d$ then v_d is even so n is even: the $r = 1$ probability is 0 and the $r = 0$ weight is doubled by normalization. If $2 \nmid d$, then $n \equiv s \pmod{2}$; among squarefree s coprime to d , odd s carry density $\frac{2}{3}$ and even s (necessarily of the form $2s'$) density $\frac{1}{3}$, giving weights $\frac{4}{3}$ and $\frac{2}{3}$; the parity-conditioned total normalizes with $G = K_6 \Pi_Q$. $\square$

Verification. Normalization holds to 3×10^{-5} ; the sharpest prediction — $\mathbb{P}(d \text{ even} \mid r = 1) \equiv 0$ — is confirmed exactly on $n \leq 4 \times 10^5$; both parity classes agree to 10^{-4} for $y \lesssim 4$; $\Sigma/\Sigma_\perp \asymp e^{0.0722}$ over $D : 4 \rightarrow 10^6$.

3. A COMBINATORIAL RESTATEMENT OF ABC

Theorem 3.1 (Master identity). *Let $(a, b) = 1$, $a + b = c$. Then*

$$q > 1 \iff \text{rad}(a) \text{rad}(b) < d_c = \frac{c}{\text{rad}(c)}.$$

Proof. $(a, c) = (a, a + b) = (a, b) = 1$ and $(b, c) = 1$, so a, b, c are pairwise coprime and $\text{rad}(abc) = \text{rad}(a) \text{rad}(b) \text{rad}(c)$. Then $q > 1 \iff c > \text{rad}(a) \text{rad}(b) \text{rad}(c) \iff c / \text{rad}(c) > \text{rad}(a) \text{rad}(b)$. $\square$

Equivalent form. Since $a / \text{rad}(a) = d_a$ and $c / \text{rad}(c) = d_c$, the identity reads $\text{rad}(b) < R$, $R := \frac{d_c}{\text{rad}(a)} = \frac{d_a d_c}{a}$. Given a and c , the constants are fixed and the only random variable is X_b ; the counting complexity drops from $O(N^2)$ to nearly $O(N)$. *Verification:* exhaustive search gives 57 solutions for $c \leq 300$ (57 by the original definition, 57 by the new form) and 2054 for $c \leq 3 \times 10^6$; all satisfy the new form.

Theorem 3.2. *If $q > 1$ then*

$$a \cdot \text{rad}(d_b) \leq d_a d_c, \quad b \cdot \text{rad}(d_a) \leq d_b d_c, \quad \frac{ab}{c} \cdot \text{rad}(d_c) \leq d_a d_b.$$

Proof. From Theorem 3.1, $\text{rad}(b) < d_c / \text{rad}(a) = d_a d_c / a$; since $\text{rad}(d_b) \leq \text{rad}(b)$ the first inequality follows; the others by symmetry in a and c . $\square$

Verification. On 234 measured solutions the minimum “slack” of the three inequalities is 0.0039, 0.0078, 0.0078 with zero violations: high-quality abc triples are *extremal points on the ceiling*, not random lucky events.

Corollary 3.3. $q > 1 \Rightarrow d_b > 1$: *b is never squarefree.*

Proof. From $d_a \leq a$, $d_c \leq c < 2b$ one gets $\frac{d_a d_c}{ab} < \frac{a \cdot 2b}{ab} = 2$, so $\text{rad}(b) \leq b/2$, i.e. $d_b \geq 2$; the critical case $d_b = 1$ would force $\text{rad}(b) = b \leq d_a d_c / ab > c/2$, contradicting the strict inequality of Theorem 3.1. $\square$

Theorem 3.4 (Prime exclusion). *If $(a, b) = 1$ then $p \mid a \Rightarrow p \nmid b$; hence the events $\{p^k \parallel a\}$ and $\{p^j \parallel b\}$ are mutually exclusive for every prime p .*

Proposition 3.5 (Covariance in closed form). *With $Y_p(n) := (v_p(n) - 1)_+ \log p$,*

$$\text{Cov}(X_a, X_b) = - \sum_{\substack{p \\ p \nmid c}} \mathbb{E}[Y_p(a)] \mathbb{E}[Y_p(b)] \longrightarrow - \sum_p \frac{(\log p)^2}{p^2(p-1)^2} = -0.163290.$$

Proof. By Theorem 3.4, for $p \nmid c$ the events $\{Y_p(a) > 0\}$ and $\{Y_p(b) > 0\}$ are mutually exclusive, so $\mathbb{E}[Y_p(a)Y_p(b)] = 0$ while an independent model gives $\mathbb{E}[Y_p(a)]\mathbb{E}[Y_p(b)]$; the difference is the negative contribution. For $p \mid c$ there is no constraint and the contribution vanishes asymptotically. By Proposition 2.4, $\mathbb{E}[Y_p] = \frac{\log p}{p(p-1)}$, and the sum over p converges to the stated constant. $\square$

Numerical. $N = 2 \times 10^4$: $\text{Cov}_{\text{emp}} = -0.076072$, $\text{Cov}_{\text{th}} = -0.079157$ (4% agreement), correlation $\rho = -0.0633$; prime-by-prime, $\mathbb{E}[Y_p(a)Y_p(b)] \equiv 0$ on $p \leq 130$ as measured, versus a 10^3 -scale prediction from the independent model.

Theorem 3.6 (Hard data). *Exact counts: $\mathcal{N}_0(N) = 419, 734, 1269, 1752, 2054$ for $N = 10^5, 3 \times 10^5, 10^6, 2 \times 10^6, 3 \times 10^6$.*

Proposition 3.7 (Growth law, numerical). $\mathcal{N}_0(N) \approx 1.965 N^{0.4676}$ *with residual $\pm 2.5\%$ (5 points, $30 \times$ range). The ratios $\mathcal{N}_0 / \sqrt{N} = 1.325, 1.340, 1.269, 1.239, 1.186$ decrease slowly. Leave-one-out refitting keeps the exponent in $[0.4504, 0.4753]$ (span 0.025): a pure power law is robustly supported by the data.*

Remark 3.8 (Identifiability of the logarithmic correction). The same data were once written as $\mathcal{N}_0(N) \approx 3.77\sqrt{N}(\log N)^{-0.419}$ ($\dagger$), which fits equally well (residual $\pm 2.8\%$), but the exponent -0.419 is *not* independently measured: (i) it is a conditional parameter obtained with $\alpha = \frac{1}{2}$ pre-fixed (we reproduce -0.4189); freeing α gives $\alpha = 0.4676$ and the logarithmic coefficient becomes unidentifiable; (ii) $\log \log N$ and $\log N$ are collinear on the computable range ($\log \log N \in [2.44, 2.70]$ while $\log N \in [11.5, 14.9]$); a free three-term fit is pathologically ill-conditioned (leave-one-out spans 2.808 to 27.6, including meaningless values); (iii) by our own criterion of §15.5 — slowly varying functions manufacture stable-looking false constants — the exponent in ($\dagger$) is exactly such an artifact. *Conclusion:* the main result is the pure power law; ($\dagger$) may be exhibited only as compatibility with a $\sqrt{N}$ power, never as an independently measured exponent.

Remark 3.9 (Compatibility with the square-root barrier). An independent lattice program proves unconditionally $\#\{q > 2\} \leq X^{1/2+o(1)}$ with $o(1) = O(\log \log X / \log X)$. Our exponent 0.467 satisfies $0.5 - 0.467 = 0.033$, while $(\log \log X) / \log X \approx 0.49$ at $X = 3 \times 10^6$: the barrier's slack accommodates the measured exponent an order of magnitude over. The mechanistic attribution of the exponent is in Section 14.

Proposition 3.10 (The signature $\beta(\varepsilon)$, numerical). *Fitting $\mathcal{N}_\varepsilon(N) \asymp N^\beta$ on $10^5 \rightarrow 3 \times 10^6$:*

$1 + \varepsilon$	1.00	1.05	1.10	1.20	1.30	1.40	1.50
β	0.467	0.433	0.388	0.313	0.315	0.204	0.000
$\mathcal{N}(3 \times 10^6)$	2054	1047	569	148	38	12	2

The fitted exponents decrease to 0 (the single wobble $0.313 \rightarrow 0.315$ is within fit noise); no new solutions appear beyond $q > 1.5$ in a $30\times$ range. The median q is 1.051 (99% quantile 1.34); the maximum is 1.56789 ($2 + 3^{10} \cdot 109 = 23^5$, i.e. $4374 + 1 = 4375$).

4. THE LOCAL CRT CONDITIONAL LAW

Theorem 4.1. *Given (a, c) , the hit probability is $\mathbb{P}(q > 1 \mid a, c) = \mathbb{P}(X_b > y)$ with $y := \log(ab) - X_a - X_c$.*

Correction. Early versions wrote the hit probability as the tail of the three-sum $S = X_a + X_b + X_c$. Given (a, c) , X_a and X_c are constants, so the correct object is a *single-variable* tail. This correction is the premise of every subsequent quantitative result.

Theorem 4.2 (Local CRT conditional law). *Let c be fixed and a uniform on $A := \{a < c/2 : (a, c) = 1\}$, $b := c - a$, $\text{mask}(a) := \{p \leq z : p \mid a\}$. With $Y_p := (v_p(b) - 1)_+ \log p$ and $X_b = \sum_p Y_p$, the Y_p are mutually independent and their distribution falls into three cases:*

Case	Condition	Distribution of Y_p
A	$p \mid c$	$Y_p \equiv 0$
B	$p \nmid c, p \mid a$	$Y_p \equiv 0$
C	$p \nmid c, p \nmid a$	$\mathbb{P}(v_p(b) = 0) = \frac{p-2}{p-1}, \mathbb{P}(v_p(b) = k) = p^{-k} \ (k \geq 1)$

so that for any y , $\mathbb{P}(X_{c-a} > y \mid \text{mask}(a) = m) = \mathcal{S}_{m,c}(y)$, a sum over $p \nmid c, p \in m$ of the case-C laws.

Key point: *case C is not the usual law. The usual law $\mathbb{P}(v_p = k) = (1 - \frac{1}{p})p^{-k}$ has mean $\mathbb{E}[Y_p] = \frac{\log p}{p(p-1)}$, while the corrected law gives*

$$\mathbb{E}[Y_p \mid p \nmid a] = \frac{\log p}{(p-1)^2},$$

a factor $\frac{p}{p-1}$ larger (exactly 2 at $p = 2$). Equivalently $\mathbb{P}(Y_p = 0) = 1 - \frac{1}{p(p-1)}$, $\mathbb{P}(Y_p = (k-1) \log p) = p^{-k} \ (k \geq 2)$.

Proof. Case A: $(b, c) = (c - a, c) = 1$, so $p \mid c \Rightarrow v_p(b) = 0$. Case B: $p \mid a \Rightarrow p \nmid b$ by Theorem 3.4. Case C: let $p \nmid c$; a is uniform on the $\varphi(p^K) = p^{K-1}(p-1)$ reduced residue classes. For $k \geq 1$, $v_p(b) = k$ iff $a \equiv c \pmod{p^k}$ and $a \not\equiv c \pmod{p^{k+1}}$: among the p^{K-k-1} lifts of the $p-1$ admissible classes, so $\mathbb{P}(v_p(b) = k) = \frac{(p-1)p^{K-k-1}}{p^{K-1}(p-1)} = p^{-k}$. Independence follows from CRT applied to $\prod_p p^{K_p} \mid a$ -conditions. Finally $\mathbb{E}[Y_p \mid \text{case C}] = \sum_{k \geq 2} (k-1)p^{-k} \log p = \frac{\log p}{(p-1)^2}$. $\square$

Proposition 4.3 (Consistency). *Averaging over the unconditioned prime with $\mathbb{P}(p \mid a) = \frac{1}{p}$, cases B and C mix to recover the usual law, hence unconditionally $\mathbb{E}[X_b] = C(1)$.*

Proof. For $k \geq 1$, $\mathbb{P}(v_p(b) = k) = \mathbb{P}(p \nmid a)p^{-k} = (1 - \frac{1}{p})p^{-k}$, and $\mathbb{P}(p \nmid b) = \frac{1}{p} + \frac{p-1}{p} \cdot \frac{p-2}{p-1} = \frac{p-1}{p}$, matching the usual law. $\square$

4.1. Numerical verification, five items. (i) *Prime-by-prime* (case C): with $c \approx 8 \times 10^5$ fixed and 2.4×10^6 pairs (a, c) :

p	2	3	5	13	47
measured $\mathbb{E}[Y_p]/\log p$	0.999513	0.249999	0.062500	0.006944	0.000473
corrected $(p-1)^{-2}$	1.000000	0.250000	0.062500	0.006944	0.000473
usual $(p(p-1))^{-1}$	0.500000	0.166667	0.050000	0.006410	0.000463

Case B ($p \mid a$) is identically 0 in measurement, zero violations.

(ii) *CRT independence*: $\mathbb{P}(v_p = k, v_q = j) / ((1 - \frac{1}{p})(1 - \frac{1}{q})p^{-k}q^{-j}) = 1.000$ for $(p, q) = (2, 3)$, $n = 8 \times 10^5$.

(iii) *Mask convolution*: $c = 800083$, 3222 masks ($P \subseteq \{2, 3, 5, 7, 11, 13\}$, $P \neq \emptyset$), $y = 2$: measured/theory 0.9929 ± 0.0086 .

(iv) *End-to-end decisive comparison* ($c = 800083$): the usual law underestimates the tail by 47%–78% with errors accumulating in the tail; the corrected law agrees to 0.4%–2%. Mask-by-mask at $y = 2$: corrected 0.9929 ± 0.0086 versus usual 1.4417 ± 0.2521 — the usual law is biased by 44% with $30\times$ the variance.

(v) *Provable range*: stratified by $\lambda = \log(D/\sqrt{c})$ ($c = 800083$, $\log \sqrt{c} = 6.796$): the ratio stays ≈ 1 for $\lambda \lesssim 0$ and degrades monotonically across $\lambda \geq 0.7$, consistent with the finite-size ceiling $X_b \leq \log(b/2) = 12.21$.

4.2. Range of validity and failure mechanism. Inside $D \lesssim \sqrt{c}$ the ratio is 0.96–1; beyond $\sqrt{c}$ it degrades monotonically ($0.96 \rightarrow 0.72$). This range is exactly the boundary provable by Theorem 6.3 (main term) plus the trivial tail bound. Theorem 4.2 replaces the earlier “equidistribution” assertion: its error is *structural*, not statistical — the unconditional law neither knows which primes c and a have excluded (cases A, B) nor uses the case-C distribution (off by the factor $\frac{p}{p-1}$).

5. THE ANALYTIC THEORY OF $C(\beta)$

Proposition 5.1 (Complete monotonicity). *$C(\beta)$ is analytic on $(0, \infty)$ with $C^{(k)}(\beta) = (-1)^k \sum_p \frac{(\log p)^{k+1}}{p^\beta(p-1)}$, so $(-1)^k C^{(k)}(\beta) > 0$. The unique singularity is at $\beta = 0$; there is no phase transition at $\beta = 1$ or $\beta = 2$. (An earlier conjecture of a phase transition at $\beta = 1, 2$ was wrong; see Section 15.4.)*

Proposition 5.2 (Closed-form integral). $\int_1^2 C(\beta) d\beta = \sum_p \frac{p^{-1} - p^{-2}}{p-1} = \sum_p \frac{1}{p^2} = P(2) = 0.452247514$.

Proof. $\int_1^2 p^{-\beta} d\beta = \frac{p^{-1} - p^{-2}}{\log p}$; multiply by $\frac{\log p}{p-1}$ and sum. $\square$

Numerical: numerical integration agrees with the closed form to 9.4×10^{-8} . An earlier erroneous value $P(2) - P(3) = 0.277$ came from a typo $\frac{1}{p-1} \rightarrow \frac{1}{p}$; excluded.

Corollary 5.3. $\Delta_* := \frac{C(1)+C(2)}{2} - P(2) = 0.056574$, i.e. 11.12% of the chord area. An earlier value 0.23134 (45.47%) is void.

Theorem 5.4 (Laurent expansion). As $\beta \rightarrow 0^+$, $C(\beta) = \frac{1}{\beta} - \gamma + 1.5732\beta + O(\beta^2)$.

Four steps. Step 1 (functional recursion). From $\frac{1}{p-1} = \sum_{k \geq 1} p^{-k}$, $C(\beta) = C(\beta+1) + \sum_p (\log p) p^{-(\beta+1)}$.

Step 2 (prime zeta). $-\frac{\zeta'}{\zeta}(s) = \sum_p \frac{\log p}{p^s-1} = \sum_p \frac{\log p}{p^s} + D(s)$ with $D(s) = \sum_p \frac{\log p}{p^s(p^s-1)}$.

Step 3. With the Stieltjes expansion $\zeta(1+\beta) = \frac{1}{\beta} + \gamma - \gamma_1\beta + \dots$ one gets $-\frac{\zeta'}{\zeta}(1+\beta) = \frac{1}{\beta} - \gamma + (\gamma^2 + 2\gamma_1)\beta + \dots$, and $D(1+\beta) = C(1) - E_1\beta + \dots$ with $E_1 = \sum_p \frac{(\log p)^2(2p-1)}{p(p-1)^2}$.

Step 4 (assembly). Substituting into Step 1, the constant $C(1)$ cancels exactly, leaving $C(\beta) = \frac{1}{\beta} - \gamma + [(\gamma^2 + 2\gamma_1) + E_1 + C'(1)]\beta + \dots$. Numerically $C'(1) = -0.981957$, $\gamma^2 = 0.333178$, $2\gamma_1 = -0.145632$, $E_1 = 2.367561$, total 1.573151; independent extrapolation gives 1.57317. $\square$

Theorem 5.5 (Pole structure and natural boundary).

$$C(\beta) = \sum_{m \geq 1} \mu(m) \sum_{k \geq 1} \left(-\frac{\zeta'}{\zeta}\right)(m(\beta+k)), \quad \operatorname{Res}_{\beta=\frac{\rho}{m}-k} C = \frac{\mu(m)}{m} \operatorname{Res}_{s=\rho} \left(-\frac{\zeta'}{\zeta}\right).$$

The families sit at $\beta = \frac{1}{m} - 1$ (residue $\mu(m)/m$), $\beta = \frac{\rho}{m} - 1$ (accumulating at $\Re\beta = -1$), and $\beta = -\frac{2n}{m} - 1 \leq -1$. Hence $\Re\beta = -1$ is a natural boundary (poles accumulate). Numerically the residue at $\beta = -\frac{1}{2}$ is -0.0909087 ($m=1$ prediction -0.090909) and $+0.500189$ at $\beta = 0$ ($m=2$, $\rho = \frac{1}{2}$). RH formulation: the RH subfamily ($\beta = \frac{\rho}{m} - 1$, $m=1$) lies on $\Re\beta = -\frac{1}{2}$; but $\beta = -\frac{1}{2}$ is itself a pole ($m=2$, $\rho = \frac{1}{2}$), so no functional equation about $\Re\beta = -\frac{1}{2}$ exists.

Theorem 5.6 (S - E decomposition and telescoping). With $S(\beta) := \sum_{k \geq 1} \left(-\frac{\zeta'}{\zeta}\right)(\beta+k)$ and $E(\beta) := \sum_p \log p \sum_{m \geq 2} \frac{1}{p^{m\beta}(p^m-1)}$, $C(\beta) = S(\beta) - E(\beta)$ and $S(1) - S(2) = -\frac{\zeta'(2)}{\zeta(2)} = 0.569961$, $C(1) - C(2) = -\frac{\zeta'(2)}{\zeta(2)} - [E(1) - E(2)] = 0.493091$. Numerically $S(1) = 0.850312$, $E(1) = 0.094946$, $S(2) = 0.280351$, $E(2) = 0.018076$; $S - E$ agrees with C to 10^{-15} .

Theorem 5.7 (The subexponential term and the $2 \log 2$ cancellation). For the unconditional three-sum with $\Lambda'_H(0) = 1.741549$, $\Lambda''_H(0) = 3.502899$, with $w = W(6T)$, $c = \frac{1}{4}$,

$$I_H(T) = T - \sqrt{\frac{6T}{w}} \left(2 + \frac{2}{w} + \frac{4(2c-1)}{w^2} + O(w^{-3}) \right).$$

The source of the constant: the reduced integral $J(A) := \int_{\log 2}^{\infty} \log\left(1 + \frac{Ae^{-u}}{u}\right) \frac{e^u}{u} du$ splits at u_0 ($u_0 e^{u_0} = A$); the outer region loses $\frac{2 \log 2 - 1}{u_0} e^{u_0}$ and the inner region contributes $\frac{2 \log 2}{u_0} e^{u_0}$; the $2 \log 2$ cancels, leaving $\frac{e^{u_0}}{u_0}$. Numerically at $u_0 = 160$, $(J/e^{u_0} - 1)u_0 = 1.027 \rightarrow 1$. The term is subpolynomial and does not change the critical threshold; the same cancellation recurs at higher orders in Section 7.

6. SIEVE UPPER BOUNDS: AN HONEST FAILURE

Theorem 6.1 (Rigorous tail bound for $\Sigma_{\perp}$). For $\sigma \in (0, 1)$,

$$\Sigma_{\perp}(D; Q) \leq D^{-\sigma} \Phi(\sigma) \prod_{p|Q} \frac{p^2 - 1}{p^2 - p^{-\sigma}}, \quad \Phi(\sigma) = \zeta(1+\sigma) \prod_p \left(1 + \frac{p^{-\sigma}}{p^2 - 1}\right).$$

Sketch. Markov: $\Sigma_{\perp}(D; Q) \leq D^{-\sigma} F_Q(-\sigma)$ with $F_Q(-\sigma) = \sum_{(d,Q)=1} u(d) d^{\sigma}$. Splitting inside/outside Q , the error factors satisfy $R_p \leq Cp^{-1-\sigma}$, and $\prod_p (1 + R_p) \leq \zeta(1+\sigma) \prod_{p|Q} (1 - \frac{1}{p^{1+\sigma}})^{-1}$; the identity $\frac{1-p^{-1-\sigma}}{1+\frac{p^{-\sigma}}{p^2-1}} = \frac{p^2-1}{p^2-p^{-\sigma}}$ finishes. $\square$

Corollary 6.2. *With $\sigma = 1 - \frac{1}{\log D}$: $\Sigma_\perp \leq C_1 \frac{\log D}{D} \prod_{p|Q} \frac{p^2-1}{p^2-p^\sigma}$, $C_1 \approx 1.1$.*

Theorem 6.3 (Sieve theorem with remainder). *Let Q be squarefree, $|I| \geq 3$. Then*

$$\mathcal{N}(I; D, Q) \leq |I| K_6 \Pi_Q \Sigma_\perp(D; Q) + C 2^{\omega(Q)} \sqrt{|I|} (\log |I|)^2.$$

Sketch. By Lemma 2.1, $\mathcal{N} = \sum_{d>D, (d,Q)=1} S_d$ with $S_d = \#\{s \leq |I|/v_d : \mu^2(s) = 1, (s, dQ) = 1\} = \frac{|I|}{v_d} K_6 w(dQ) + O(2^{\omega(dQ)} \sqrt{|I|/v_d})$; the main terms assemble to $|I| K_6 \Pi_Q \Sigma_\perp(D; Q)$ and the remainders are controlled by $v_d \geq 2d$. $\square$

Effective range (the key limitation). The main term must dominate the remainder:

$$D \ll \frac{\sqrt{|I|}}{2^{\omega(Q)} (\log |I|)^2} \approx \frac{\sqrt{c}}{(\log c)^3},$$

while abc requires $D \sim 10^5$ – 10^7 : a gap of about five orders of magnitude.

Proposition 6.4 (Moment/Chernoff strict insufficiency). *The moment method gives $\mathcal{N} \leq D^{-\sigma} \sum_n d_n^\sigma$, which is strictly weaker than the target inequality.*

Proposition 6.5 (Union bound loss). *Interchanging the sums loses $\log D$; for $\varepsilon = 0$ the resulting series diverges.*

Proposition 6.6 (Joint constraints do not decouple, numerical). *At $\varepsilon = 0.1$, $c \leq 3 \times 10^4$: the joint three-radical constraint gives 94 solutions; the single $\text{rad}(b)$ constraint gives 36362: an inflation by a factor 387. Any decoupling suffices to turn a convergent series into a divergent one.*

Theorem 6.7 (Weak abc, unconditional — and weaker than the trivial bound).

$$\mathcal{N}_\varepsilon(N) \leq \sum_{c \leq N} \min(A_c, B_c), \quad A_c = \#\{a < \frac{c}{2} : \text{rad}(a) < T_c\}, \quad B_c = \#\{b \in (\frac{c}{2}, c) : \text{rad}(b) < T_c\}.$$

Parameter alignment $\lambda = \frac{1}{1+\varepsilon}$:

λ	1	0.833	0.667
field (exponent at $\lambda = 1$)	0.65–0.667	0.554	0.458
Theorem 6.7	1.384	1.272	1.103

where the field column lists de Bruijn 1962 (0.6667, the trivial bound), BLT 2024 (0.65) and Li 2025 (0.6588). Conclusion, reported honestly: *our weak abc bound is weaker than de Bruijn’s 1962 trivial bound, because it uses only the one-sided constraint $\min$ and discards the joint structure of the three radicals. At the same time the distance of the field to abc is equally stark: 0.65 versus the required $O(1)$ is a discontinuous gap.*

7. THE LAMBERT W SUBPOLYNOMIAL SOURCE

This section is the core of the paper and the one result that holds completely independently of the abc conjecture.

7.1. Main theorem. Let $\eta(p) > 0$ satisfy $\eta(p) \rightarrow a \in (0, \infty)$ (error $O(p^{-\delta})$). Define

$$\Phi(\sigma) := \sum_p \log \left(1 + \frac{\eta(p)}{p^\sigma - 1} \right), \quad v^* := W \left(\frac{a}{\sigma} \right) \quad (\text{Lambert } W).$$

Theorem 7.1 (Main theorem). *As $\sigma \rightarrow 0^+$,*

$$\Phi(\sigma) = e^{v^*} \left[1 + \frac{\kappa_1}{v^*} + \frac{\kappa_2}{v^{*2}} + \frac{\kappa_3}{v^{*3}} + \cdots \right],$$

where $\frac{a}{\sigma v^*} = e^{v^*}$, so the leading term is $\frac{a}{\sigma W(a/\sigma)}$. The coefficients are universal: independent of the shape of a and η .

Skeleton. Step 1 (PNT to integral). By Abel summation and the prime number theorem, $\Phi(\sigma) = \int_{\log 2}^{\infty} \frac{e^v}{v} \log(1 + \frac{a}{e^{v(e^{\sigma v}-1)}}) dv + \mathcal{E}$, with $\mathcal{E}$ controlled by the standard PNT remainder (Section 12.4).

Step 2 (regime change). With $\psi(v) := \frac{a}{e^{v(e^{\sigma v}-1)}}$ one has $\psi(v^*) = 1 + O(e^{-v^*})$ since $\sigma v^* = e^{-v^*}$.

Step 3 (identification of the dominant range). For $v \geq v^*$, $\psi \leq 1$ and $\log(1 + \psi) = \psi + O(\psi^2)$, so $\int_{v^*}^{\infty} \frac{e^v}{v} \psi(v) dv = \int_{v^*}^{\infty} \frac{dv}{v(e^{\sigma v}-1)} = \int_{e^{-v^*}}^{\infty} \frac{du}{u(e^u-1)}$ (substituting $u = \sigma v$) $= e^{v^*} - \frac{v^*}{2} + C_0 + O(e^{-v^*})$ by the Bernoulli expansion $\frac{1}{u(e^u-1)} = \frac{1}{u^2} - \frac{1}{2u} + \frac{1}{12} + O(u)$. Hence the leading term is $\frac{a}{\sigma v^*} = e^{v^*}$. **The dominant primes lie in $[e^{v^*}, e^{1/\sigma}]$:** there $\log(1 + \psi) \approx \psi \approx \frac{a}{p\sigma \log p}$, summing over p to $\frac{a}{\sigma v^*}$. This explains why no finite prime table can reproduce Φ : $e^{1/\sigma}$ is astronomically large.

Step 4 (exact $1/v^$ correction).* In the dominant range $\sigma v \leq \sigma v^* = e^{-v^*} \ll 1$, so $e^{\sigma v} - 1 = \sigma v(1 + O(\sigma v))$ and $\psi(v) = \frac{v^*}{v} e^{v^*-v}(1 + O(e^{-v^*}))$. Region I ($\log 2 \leq v \leq v^*$, $\psi \geq 1$): substituting $v = v^* - t$, $\Phi_I = \underbrace{\frac{e^{v^*}}{v^*} \int_0^{\infty} e^{-t} \log(1 + e^t) dt}_{=: \iota_0} + O(\frac{e^{v^*}}{v^{*2}})$. Region II ($v \geq v^*$,

$\psi \leq 1$): substituting $v = v^* + s$ and subtracting the $\psi - \log(1 + \psi)$ deficit, $\Phi_{II} = e^{v^*} - \underbrace{\frac{e^{v^*}}{v^*} \int_0^{\infty} e^s [e^{-s} - \log(1 + e^{-s})] ds}_{=: \rho_0} + O(\frac{e^{v^*}}{v^{*2}}) + O(v^*)$. Both integrals have closed forms: $\iota_0 =$

$2 \log 2$ and $\rho_0 = 2 \log 2 - 1$. Hence the $1/v^*$ correction is $\frac{\Phi_I + \Phi_{II} - e^{v^*}}{e^{v^*}/v^*} = \iota_0 - \rho_0 = 2 \log 2 - (2 \log 2 - 1) = 1$, i.e. $\kappa_1 = 1$ *exactly*: region I contributes $2 \log 2$, region II contributes $-(2 \log 2 - 1)$, and the net is 1. This cancellation recurs at every higher order (Section 7.3).

Caveat. The $O(e^{v^*}/v^{*2})$ error above gives only the magnitude; the precise values of $\kappa_2, \kappa_3, \dots$ come from the closed-form integrals for ι_N, ρ_N (Sections 7.2–7.4); the early sloppy statement “region I is $O(e^{v^*}/v^{*2})$ ” does not determine κ_1 — the $1/v^*$ main part is the carrier of κ_1 . $\square$

Lemma 7.2 (Consistency of the PNT exchange). *Let $\Phi_{\text{PNT}}(\sigma)$ be the integral main term of Step 1. As $\sigma \rightarrow 0^+$ (equivalently $v^* \rightarrow \infty$), $\frac{|\Phi_{\text{true}}(\sigma) - \Phi_{\text{PNT}}(\sigma)|}{\Phi_{\text{PNT}}(\sigma)} \ll \exp(-c\sqrt{v^*}) \rightarrow 0$.*

Sketch. Abel summation gives $\Phi_{\text{true}} = - \int_2^{\infty} \pi(t) f_{\sigma}'(t) dt$ with $f_{\sigma}(t) = \log(1 + \frac{\eta(t)}{t^{\sigma-1}})$; writing $\pi(t) = \text{Li}(t) + E(t)$ with $|E(t)| \ll t \exp(-c\sqrt{\log t})$, the main term is Φ_{PNT} and the remainder is $\ll \exp(-c\sqrt{v^*}) \Phi_{\text{PNT}}$ because the dominant range forces $\log t \geq v^*$. Outside the dominant range both Φ and its PNT error are $O(e^{v^*}/v^*) = o(\Phi_{\text{PNT}})$. $\square$

Remark 7.3 (The series is asymptotic, not convergent). The right side of Theorem 7.1 is a Poincaré asymptotic: for fixed truncation order M the remainder is $O(e^{v^*} v^{*-M-1})$. For fixed v^* the series diverges: $\kappa_n \sim (n-1)!$ makes the terms grow beyond $n \gtrsim v^*$, so the optimal truncation is at $n \approx v^*$. In the computable range $v^* \sim 0.9\text{--}1.3$, $1/v^*$ is *not* a small parameter, so every application must use the exact Φ (numerical integration or high-precision prime summation) for the saddle point, never the truncated expansion. This is the unified root of the “constant drifts” of $\kappa' \approx 1.1$, $\kappa_1 \approx 1.67$, $\omega \approx 0.66$.

7.2. Closed forms for κ_1 – κ_8 .

n	κ_n (closed form)
1	1
2	$1 + 2\zeta(2) = 1 + \frac{\pi^2}{3}$
3	$2 + \zeta(2) = 2 + \frac{\pi^2}{6}$
4	$6 + 7\zeta(2) + 42\zeta(4) = 6 + \frac{7\pi^2}{6} + \frac{7\pi^4}{15}$
5	$\frac{119}{24} + 23\zeta(2) - \frac{5}{2}\zeta(4)$
6	$120 + 121\zeta(2) + 448\zeta(4) + 1395\zeta(6)$

(κ_7, κ_8 similar; κ_5 is negative — the first sign change, with absolute value exceeding $(n-1)! = 24$.) Verification by two independent paths (integral coefficients vs. closed form):

n	numerical	closed form	error
1	1.0000000000000000	1	5×10^{-31}
2	4.28986813369645287	$1 + \pi^2/3$	10^{-28}
3	3.64493406684822644	$2 + \pi^2/6$	3.5×10^{-18}
4	62.97211428380538909885397	$6 + \frac{7\pi^2}{6} + \frac{7\pi^4}{15}$	1.25×10^{-19}

κ_5 – κ_8 confirmed by PSLQ at 1300 digits with reconstruction error 10^{-19} .

7.3. The cancellation mechanism $\iota(r) = (-1)^{r+1}\rho(r)$. At v^* the integral splits into region I ($v < v^*$) and region II ($v > v^*$); writing the coefficients of $\zeta(r) \log 2$ on the two sides as $\iota(r)$ and $\rho(r)$, one has $\kappa_n = \iota_{n-1} - \rho_{n-1}$; e.g. for $n = 2$: $\rho_2 = -2 - \frac{\pi^2}{12} + 4 \log 2 + \frac{9}{2}\zeta(3)$, $\iota_2 = \frac{\pi^2}{12} + 4 \log 2 + \frac{9}{2}\zeta(3)$. Hence $\kappa(r) = \iota(r) - \rho(r) = [(-1)^{r+1} - 1]\rho(r)$:

r	$(-1)^{r+1} - 1$	result
odd (incl. $r = 1$'s $\log 2, \zeta(3), \zeta(5), \dots$)	$1 - 1 = 0$	exact cancellation
even ($\zeta(2), \zeta(4), \dots$)	$-1 - 1 = -2$	doubled, survives

This is the complete mechanism behind “even zetas only”: the two-sided kernels are related by $\log(1 + e^t) = t + \log(1 + e^{-t})$, i.e. $t \leftrightarrow s$ with reversed orientation, producing the factor $(-1)^{r+1}$ at order r . It is the higher-order version of the first-order cancellation of Theorem 5.7.

7.4. Stirling convolution and the general term $c_{n,j}$.

Theorem 7.4 (Structure theorem). $\kappa_n = (n-1)! + \sum_{j=1}^{\lfloor n/2 \rfloor} c_{n,j} \zeta(2j)$, with $c_{n,1} = (n-1)! + (-1)^n$ (confirmed at 16 data points); only even zeta values occur, no $\log 2$, no odd zetas.

Sketch of the ρ -side Stirling closed form. Partial fractions for $\frac{P_N(m)}{(m+1)m^{N+1}} = \frac{C}{m+1} + \sum_{r=1}^{N+1} \frac{D_r}{m^r}$ with $P_N(m) = \prod_{i=2}^{N+1} (m+i)$ give $C = N!(-1)^{N+1}$, $D_1 = (-1)^N N!$, and for $1 \leq r \leq N+1$ an explicit Stirling-number form for $D_r^{(N)}$; hence $\rho_N[\zeta(r)] = -N!(-1)^{N+1} D_r^{(N)} (1 - 2^{1-r})$ and finally

$$c_{n,j} = 2(-1)^n D_{2j}^{(n-1)} (1 - 2^{1-2j}).$$

New identity: $(n-1)! = |D_1^{(n-1)}|$, so the constant term is the $r = 1$ Stirling coefficient; the $\log 2$ term vanishes identically because $(1 - 2^{1-2j}) = 0$ has no solution — wait: precisely, at $r = 1$ the factor $(1 - 2^0) = 0$ kills the $\log 2$ -carrying term, leaving the $(n-1)!$ constant. Same mechanism, second manifestation. Also $D_2^{(n-1)} = (-1)^n (n-1)! + 1$, giving $c_{n,1} = (n-1)! + (-1)^n$. $\square$

New predictions from the general term (first $\zeta(2)$ -generated constant terms): $\kappa_9 = 40320 + 40319\zeta(2) + \frac{279759}{4}\zeta(4) - \frac{25141}{2}\zeta(6) - \frac{2387727}{4}\zeta(8) + \dots$; $\kappa_{10} = 362880 + 362881\zeta(2) + \frac{254851}{4}\zeta(4) + \frac{766519}{8}\zeta(6) + \frac{3785299}{8}\zeta(8) + 7243425\zeta(10)$.

7.5. Generating function. With $z = 1/v^*$ ($\sigma = az e^{-1/z}$):

$$F(z) := \sum_{n \geq 0} \kappa_n z^n = e^{-1/z} \Phi(\sigma(z)) = \int_{-\infty}^{\infty} \frac{ze^u}{1+uz} \log\left(1 + \frac{e^{-u}}{1+uz}\right) du,$$

giving any κ_n by termwise expansion — no PSLQ needed.

7.6. Saddle constant and transient laws.

Theorem 7.5. $\mathfrak{c} = 1$ (asymptotically).

Sketch. For $\Psi(s) = \Phi(s) + sL - \log s$ the Edgeworth first non-Gaussian correction decays super-exponentially:

2^j	2	4	8	16	32	64
correction	-2.8×10^{-3}	-5.7×10^{-5}	-2.0×10^{-8}	-2.3×10^{-15}	-3.0×10^{-29}	(below precision)

so $\mathfrak{c} \rightarrow 1$ far faster than any power; the scaling is $\asymp e^{-v^*}$. $\square$

Important correction. Earlier reports $\mathfrak{c} \approx 0.84\text{--}0.89$ and $\log \mathfrak{c} = -0.392/L^{0.421}$ are **void**: they came from insufficient PNT integration accuracy (a 10^{-4} relative error in Φ amplified through subtraction of $O(4)$ -magnitude terms into a fake $O(0.1)$ bias). High-precision re-computation gives $\mathfrak{c} = 1.00\text{--}1.06 \rightarrow 1$. The theoretical asymptotic bias is below current numerical reach; $\mathfrak{c} = 1$ is to be regarded as confirmed.

Transient law (derived, not fitted):

$$\log T(x) = \underbrace{(\Phi(\sigma^*) + \sigma^* L)}_{>\sqrt{L}} - \underbrace{\log(\sigma^* \sqrt{2\pi\Phi''})}_{>\log L} + \underbrace{\log \mathfrak{c}}_{O(0.1)} = 1.667\sqrt{L} - 1.246.$$

Core methodology: at computable scales $v^* \sim 1$, $1/v^*$ expansion diverges ($1 + \frac{1}{v^*} + 4.29 + \dots$ grows), so the only correct procedure is to saddle with the *exact* Φ . Every “unstable constant” is the error of treating a transient as an asymptotic.

7.7. Classification theorem and double criticality. Let $\eta(p) \asymp ap(\log p)^{-\alpha-\beta}$. Then:

class	$\Phi(\sigma), \sigma \rightarrow 0^+$	$\delta_{\mathfrak{c}}$	Tauberian: $\log A(x)$
$\alpha < 1$	diverges	$1 - \alpha$	$(1 - \alpha) \log x$ (polynomial)
$\alpha > 1$	$C\sigma^{-1/\alpha}$	0	$c_a(\log x)^{1/(\alpha+1)}$
$\alpha = 1, \beta > 0$	$a/\sigma (\beta + 1)L^{\beta+1}$	0	$\sim 2\sqrt{C \log x}$
$\alpha = 1, \beta = 0$	$\frac{a}{\sigma W(a/\sigma)}$	0	$2a \frac{\log x}{\sqrt{2 \log \log x}}$

Unified by the rescaled Lambert variable $v_g = (\beta + 1)W\left(\frac{(a/\sigma)^{1/(\beta+1)}}{\beta+1}\right)$, $\Phi = \frac{e^{v_g}}{\beta+1}(1 + \dots)$. $\alpha = 1, \beta = 0$ is the *double critical point*: it carries one fewer $\sqrt{\log \log x}$ than both neighbors — the signature where W -resummation and order-by-order cancellation appear. Numerically the ratio $\frac{\log A}{2\sqrt{2L/\log L}}$ decreases from 3.86 ($L \leq 55$) to 1.044 ($L = 10^{21}$), extremely slowly: this quantitatively explains why empirical fits look like $\sqrt{L}$ rather than $\sqrt{L/\log L}$.

Unification. $T(x) = \sum_{n \leq x} \frac{1}{\text{rad}(n)}$ ($\eta_p = 1$), $\Sigma(D)$ ($\eta_p = \frac{1}{p} \dots$ more precisely $\eta(p) = 1/(p+1)$ forms) and $\Sigma_{\perp}(D; Q)$ all satisfy $p\eta(p) \rightarrow 1$ and share the same Φ ; measured $\Psi/\Phi = 0.9707 \rightarrow 1.000000$ as $\varepsilon : 10^{-2} \rightarrow 10^{-10}$. *Non-member:* $\sum \frac{1}{\varphi(n)}$ has local factor $\frac{p}{(p-1)(p^{1+s}-1)}$, an effective parameter $s+1$ on the boundary of the Ψ -class, giving $A \log x$; smooth numbers have a different structure and are not in this class.

8. QUANTITATIVE THEORY OF MAZUR PARAMETERS

8.1. Two-form formula. From Section 7, the one-sided probability $p_\nu(n) = K_6 \Sigma(n^{1-\nu})$ has two forms, with $2v^* + \log v^* = \log \log D$:

form	condition	$\Sigma(D)$
transient	$v^* \leq 1$	$\log D \sim K' D$ (one log per side)
asymptotic	$v^* > 1$	$C \exp(2\sqrt{2} \sqrt{\frac{\log D}{\log \log D}})$ (no log)

Computable law ($1 < s < 2$): each of the three sides contributes one $\log \Sigma \asymp (\log N)^{-1}$ (transient), plus CRT suppression, net $S_{\alpha,\beta,\gamma}(N) = \mathcal{K} N (\log N)^{s-1} (\log \log N)^2 \dots$ more precisely $S \asymp N^{s-1} (\log N)^k$ with the effective exponent below. *Asymptotic law (unreachable)*: needs $\log N \gtrsim 3 \times 10^4$; $S_{\alpha,\beta,\gamma}(N) \sim \mathcal{K}' \frac{B(\alpha,\beta)}{s-1} N^{s-1} \exp\left(2\sqrt{2} \Theta \sqrt{\frac{\log N}{\log \log N}}\right)$, $\Theta = \sum_\nu \sqrt{1-\nu}$, $B(\alpha, \beta) = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}$.

8.2. Independent confirmation of $k = 2$. From $S \propto N^{s-1} (\log N)^k$, adjacent N -ratios

$$\text{give } k_{\text{eff}} = \frac{\log \frac{S(2N)}{S(N)} - (s-1) \log 2}{\log \frac{\log 2N}{\log N}}:$$

(α, β, γ)	s	$S(2000)$	$S(4000)$	$S(8000)$	$S(16000)$
(0.5, 0.6, 0.5)	1.6	101	197	349	600
(0.6, 0.6, 0.6)	1.8	146	320	612	1172
(0.4, 0.7, 0.5)	1.6	70	119	225	402
(0.5, 0.7, 0.6)	1.8	227	463	894	1798

Twelve measurements: $k_{\text{eff}} = 1.91 \pm 0.17$, compatible with $k = 2$ (0.5σ), excluding $k = 1$ and $k = 3$ (both $\geq 6\sigma$). Single-segment scatter is Poisson noise ($S \sim 10^2$ – 10^3).

8.3. Closed-form evaluation of $\mathcal{K}$. A soft-mask approximation (each prime forces $Y_p \equiv 0$ with probability $\pi_\alpha(p)$, else the corrected law) replaces $2^{\pi(z)}$ explicit convolutions: $\langle \mathcal{S} \rangle / \langle \mathcal{S}_{\mathfrak{R}} \rangle = 0.890 \rightarrow 0.983$; end-to-end A/P : soft vs. explicit 1.041 vs. 1.046; $\pi_\alpha(p) = \frac{\sum_{k \geq 1} \frac{1}{(p+1)^{pk}} \Sigma_\perp(D/p^k; p)}{\Sigma(D)}$, shape correct ($p = 2$ gives 0.706, $p = 13$ gives 0.065), absolute values $\sim 40\%$ low (threshold-at-median effect).

Falsified statement. The three-side independence hypothesis ($\mathcal{S}$ = product of the three sides) overestimates by a factor 6–1000. The cause is strong negative correlation between the radical constraints and coprimality: measured coprimality $\varphi(c)/c$ is lower by $\sim 10\times$ ($c = 40000$: 0.057 vs. 0.400; $c = 8000$: 0.036 vs. 0.400). The mechanism is Theorem 3.4: radical constraints force power-kernelization, while $(a, b) = 1$ requires disjoint prime supports.

Post-fix accuracy. Residual 1.08 ± 0.09 ; the 9.8% residual has excluded three suspects (mask-prime truncation, $P_\parallel$ truncation, y -dependence); the source is unidentified. The gap to Kane's lower bound $\Omega(N^{s-1} \log^{-2} N)$ is a factor $(\log N)^4$, real, and worth posing as an independent problem.

8.4. Fourier access and the singular series. Encoding the radical constraints as exponential sums: by Theorem 4.2, $\mathbb{P}(a \equiv 0) = \pi_\alpha(p)$, $\mathbb{P}(a \equiv r \neq 0) = \frac{1-\pi_\alpha}{p-1}$; the Gauss sum is

$$S_f(p, b) := \sum_r \pi_\alpha(p, r) e\left(\frac{rb}{p}\right) = \frac{p \pi_\alpha(p) - 1}{p-1},$$

which vanishes at $b = 0$ iff $\pi_\alpha = 1/p$: non-vanishing comes entirely from the non-uniform distribution — the Fourier-side form of the conditional law. Numerically ($N = 3 \times 10^5$, $\alpha = 0.5$): at $p = 2$ measured 0.4749 vs. predicted 0.4749, digit-perfect; large p deviates (the prediction assumes uniformity of the $p-1$ classes, which fails).

Suppression local structure. For $p \mid c$, coprimality forbids $p \mid a$ & $p \mid b$, giving the suppression factor $S_p = \frac{(p-1)\pi_\alpha\pi_\beta+(1-\pi_\alpha)(1-\pi_\beta)}{(1-\pi_\alpha)(1-\pi_\beta)}$ for $p \mid c$, and 1 for $p \nmid c$:

p	2	3	5	7	11	13
S_p	0.1125	0.2606	0.4935	0.6526	0.8000	0.8499

Suppression comes only from the prime factors of c , approaching 1 as p grows; measured suppression ratios vs. $R(c)$ -dependence trend agree ($p_{\min} = 2, 3, 5$ give $R = 0.346, 0.740, 0.859$ versus $S_p = 0.1125, 0.2606, 0.4935$ with ratios 3.08, 2.84, 1.74; S_p is a lower bound for the suppression).

Minor-arc quantitative criterion (negative but precise). Measured: at random α , mean $|F(\alpha)|/A \approx 0.89$, maximum 2.48 — square-root cancellation holds. But the major-arc main term is $\frac{AB}{N} \approx AB$ and the minor-arc bound is $\sqrt{AB}$; for the major term to dominate one needs $\sqrt{AB} \gg N \iff \alpha + \beta > 2$, while Mazur parameters always have $\alpha, \beta < 1$: the circle-method remainder is *never* controllable. This explains why BLT had to switch to geometric number theory, and turns “Fourier cannot save abc” from intuition into a criterion with an exponent.

9. CEILING EXTREMAL POINTS

Proposition 9.1 (Two exact identities). *With $\lambda := \frac{c}{\text{rad}(abc)} = \frac{\text{rad}(c)}{d_c} \text{rad}(a) \text{rad}(b)$:*

$$q = \frac{\log c}{\log c - \log \lambda},$$

so q is a race: the numerator gains from the power-kernelization d_c , the denominator from $\text{rad}(a) \text{rad}(b)$ being small. On 234 solutions, $\max |q - \frac{\log c}{\log c - \log \lambda}| = 4.4 \times 10^{-16}$.

Theorem 9.2 (Ceiling theorem, sharp). *For any coprime triple $a + b = c$, $(a, b) = 1$, $1 \leq a < b$ (the domain automatically forces $c \geq 3$),*

$$q(a, b, c) \leq \frac{\log c}{\log 2 + \log \text{rad}(c)},$$

with equality iff $\{a, b\} = \{1, 2^k\}$.

Proof. By Theorem 3.4 the prime supports of a and b are disjoint. If both $a, b \geq 2$, then $\text{rad}(a) \text{rad}(b) \geq 2 \cdot 3 = 6$; if one of them is 1, the other is ≥ 2 (as $c = a + b \geq 1 + 2 = 3$), so the product is $\geq 2 \cdot 1 = 2$. In both cases $\text{rad}(a) \text{rad}(b) \geq 2$, hence $q = \frac{\log c}{\log \text{rad}(c) + \log(\text{rad}(a) \text{rad}(b))} \leq \frac{\log c}{\log \text{rad}(c) + \log 2}$, with equality iff one term is 1 and the other a power of 2. $\square$

Verification. 234 solutions, zero violations; equality attained exactly at $c = 9$ ($1 + 2^3$, $q = 1.22629$), $c = 513$ ($1 + 2^9 = 3^3 \cdot 19$, $q = 1.31757$), $c = 1025$ ($1 + 2^{10} = 5^2 \cdot 41$, $q = 1.15230$).

Remark 9.3 (Boundary case: consistency of the domain). Theorem 9.2 must carry the hypothesis $a < b$ (or equivalently restrict to $c \geq 3$). Counterexample: $(a, b, c) = (1, 1, 2)$ satisfies $(a, b) = 1$ and $a + b = c$, but $q = \frac{\log 2}{\log 2} = 1 > \frac{1}{2} = \frac{\log 2}{\log 2 + \log \text{rad}(c)}$: the inequality is violated. The cause: in the proof, the lower bound $\text{rad}(a) \text{rad}(b) \geq 2$ degenerates to 1 when $a = b = 1$. Exhaustive check at $c \leq 2 \times 10^4$: under the relaxed convention $a \leq b$, $(1, 1, 2)$ is the *unique* violator; under the paper’s domain ($a < b$) the violation count is 0, and the minimum of $\text{rad}(a) \text{rad}(b)$ at $c \geq 3$, $a < b$ is exactly 2, attained at $(1, 2, 3)$ — matching the equality characterization exactly. *Methodological note:* the original report “234 solutions, zero violations” failed to expose this counterexample because the verification set was restricted to $q > 1$ solutions, while $(1, 1, 2)$ has $q = 1$ and slipped through the gap — a specimen of *verification-set self-selection bias* (§15.5).

9.1. Extremal reconfiguration and the bearing rate. Since q decreases in $\text{rad}(a)\text{rad}(b)$, $q_{\max}(c) = \frac{\log c}{\log c - \log(d_c/\rho(c))}$ with $\rho(c) := \min_{a < c/2, (a,c)=1} \text{rad}(a)\text{rad}(c-a)$. Verified on $c \leq 3 \times 10^4$: the extremal prediction gives $\max q = 1.56789$ at $c = 4375$, identical with the actual record — the extremal framework is exact, not heuristic.

Bearing rate (measured): the fraction of c carrying solutions is $166/29998 = 0.553\%$; the median d_c among these is 171 (population median 1); bearing rate conditioned on $d_c \geq 2$: 1.4%; on $d_c \geq 50$: 11.8%; on $d_c \geq 200$: 26.3%.

9.2. The true law of $\rho(c)$. The correction $\rho \asymp c^{2/3}$ is wrong: local slopes increase monotonically ($0.495 \rightarrow 0.573 \rightarrow 0.658$); $2/3$ is only a transition value: $\rho(c) = e^{1.30} c \exp(-2.60\sqrt{\log c})$, $\omega(c) = 1 - \frac{3.56}{\log c}$. Model comparison: pure power $c^{0.574}$ has $R^2 = 0.9962$ with U-shaped residuals; the subpolynomial form has $R^2 = 0.9991$ with residuals ± 0.04 . *Caveat*: this is a law for decade-aggregated means; individual ρ fluctuates by a factor 4–6 with growing spread (Section ??); the formula describes $\langle \rho \rangle$, not ρ itself.

9.3. Negative result: the optimal family does not reach the record. Theorem 9.2 says the ceiling-optimal family is $c = 1 + 2^k$. Searching $k \leq 40$, the optimum reaches only $q = 1.31757$ ($k = 9$), far below the world record 1.62991; generalizing to $c = 1 + p^j$ does worse (best 1.04876). *Reason: quality is a two-sided race*: the family $1 + 2^k$ pushes the a, b side to the minimum ($\log 2$) but the c side is large ($d_c \leq c$); the world record ($2 + 3^{10} \cdot 109 = 23^5$) does the opposite: $\log 2 \cdot 3 = 3.14$ on the a, b side but a moderate c side, with a huge numerator ($\log c = 15.68$). True optima trade off the two sides rather than extremizing either.

10. THE $\text{rad}-P^+$ DUALITY

Always $P^+(n) \leq \text{rad}(n) \leq n$. Define the two “relative deficits” $X_n := \log \frac{n}{\text{rad}(n)}$ and $V_n := \log \frac{n}{P^+(n)}$, so $0 \leq X_n \leq V_n$, $V_n = X_n + Z_n$ with $Z_n := \log \frac{\text{rad}(n)}{P^+(n)}$. Equality characterization: $X_n = V_n \iff \text{rad}(n) = P^+(n) \iff n = p^k$. *Verification* ($N = 2 \times 10^6$): violators of $X \leq V$: 0; $V = X$: 149236 satisfiers vs. 149235 prime powers.

Scale separation: $\mathbb{E}[X] \rightarrow C(1) = 0.7554$ regardless of scale (a positive-density effect, $\text{rad}(n) \asymp n$ on $6/\pi^2$ of integers), while $\mathbb{E}[V] \asymp (1 - \lambda) \log n$ with $\lambda = 0.62433$ the Golomb–Dickman constant. Correlation $\text{Corr}(X, V) = +0.41$. The root of the “discrete vs. continuous” dichotomy: independent local sums give atoms (the atomic law checks: $\mathbb{P}(X = 0) = 0.60794$ vs. $6/\pi^2 = 0.60793$, agreement to 10^{-5} for $d = 2, 3, 4, 5, 7, 9$), while a global extreme is continuous.

Two types vs. two functions:

	X (rad side)	V (P^+ side)
type	additive: independent local contributions	extreme: order statistic
distribution	discrete atoms at $\{\log d\}$	continuous, Dickman ρ
law	Theorem 2.2, normalized by $\zeta(2)$	$\mathbb{P}(V \geq u) = \rho(u/\log n)$
scale	$O(1)$	$O(L)$

10.1. The two Lambert branches. Both sides share the Chernoff saddle framework $\log A = \min_s [\Phi(s) + sL]$, but the saddle runs to opposite extremes:

	rad side ($T(x)$)	P^+ side ($\Psi(x, x^{1/u})$)
$\Phi(s)$	$\sum_p \log(1 + \frac{\eta(p)}{p^{s-1}})$	$-\sum_{p \leq y} \log(1 - p^{-s})$
saddle	$s^* \rightarrow 0^+$	$s \rightarrow 1^-$
measured	$0.347 \rightarrow 0.137$	$0.679 \rightarrow 0.892$ ($u = 2$)
W branch	W_0 (principal)	W_{-1} (negative)
core equation	$v^* e^{v^*} = a/\sigma$	$e^\xi = 1 + u\xi$
result	subpolynomial $e^{2\sqrt{2L/\log L}}$	$\rho(u)x$ (polynomial in x)

P^+ side. With $\beta := (1-s)\log y$ and $\log x = u\log y$: $e^\beta = 1 + u\beta \Rightarrow \beta = \xi(u) = -W_{-1}\left(-\frac{e^{-1/u}}{u}\right) - \frac{1}{u}$. The key identification: $\beta = \xi(u)$ is de Bruijn's ξ : the two are the same equation's positive root. Numerically at $u = 1.5, 2, 3, 5, 8$, ξ from independent root finding and from the closed form agree point by point. This explains the earlier observation “measured β constant at fixed u ”: that constant is $u\xi(u)$ ($u = 2$: theory 2.513, measured 2.614, a 4% PNT-approximation error).

10.2. Saddle representation of $\rho(u)$. Inverting the Laplace transform $\widehat{\rho}(z) = \exp(\gamma + \int_0^z \frac{e^{-t}-1}{t} dt)$ at the saddle $z = -\xi$:

$$\log \rho(u) = \text{Ei}(\xi) - \log \xi - u\xi + \frac{1}{2} \log \frac{\xi'}{2\pi} + O(1/u), \quad \xi' = \frac{\xi}{1+u\xi},$$

equivalent to the classical $\log \rho(u) = \gamma - u\xi + \int_0^\xi \frac{e^{-t}-1}{t} dt + \frac{1}{2} \log \frac{\xi'}{2\pi}$ (via $\text{Ei}(\xi) = \gamma + \log \xi + \int_0^\xi \frac{e^{-t}-1}{t} dt$). Accuracy:

u	2	3	5	8
ρ_{exact}	0.306853	0.048608	3.54725×10^{-4}	7.45354×10^{-8}
ρ_{saddle}	0.279158	0.046684	3.47982×10^{-4}	7.40946×10^{-8}
ratio	0.910	0.960	0.981	0.994

(error improving monotonically, $O(1/u)$).

Closure of the duality. The same Lambert W , two opposite branches; both sides share one saddle formula. The rad side uses W_0 (additive type, subpolynomial source), the P^+ side uses W_{-1} (extreme type, Dickman).

11. CROSSOVER AND FINITE-SIZE SCALING

11.1. Definition and measured results. For fixed c , define the ratio of measured counts to the local CRT conditional law:

$$R(y, c) := \frac{\#\{a < \frac{c}{2} : (a, c) = 1, X_{c-a} > y\}}{\sum_m \text{count}[m] \mathcal{S}_{m,c}(y)}, \quad \lambda := \log \frac{D}{\sqrt{c}} = y - \frac{1}{2} \log c.$$

λ	$c = 10^4$	$c = 10^5$	$c = 10^6$	between-group spread/stat. error
-2.0	1.025	1.022	1.015	1.5
-1.0	1.054	1.057	1.041	0.7
-0.5	1.084	1.122	1.068	0.9
0.0	0.925	0.943	0.948	0.8
1.0	0.852	0.860	0.870	0.6
1.25	0.808	0.833	0.840	0.9
1.5	0.669	0.753	0.798	1.2
2.0	0.525	0.635	0.781	2.9

Collapse holds for $\lambda \in [-1.5, 1.25]$ (ratio ≈ 1); it breaks for $\lambda \geq 1.5$. The curve runs from 1.02 ($\lambda = -2$) through a peak (≈ 1.09 at $\lambda \approx -0.5$) down to 0.83 ($\lambda = 1.25$); note $R > 1$ means measurement exceeds theory.

11.2. Hard truncation mechanism. In reality $n = v_d s \leq c$, so for squarefree d : $v_d \leq c \iff d \leq \sqrt{c}$. Thus $\lambda = 0$ is exactly the switch where “squarefrees begin to be hard-truncated”. For $\lambda < 0$ squarefrees dominate, no truncation, $R \approx 1$; for $\lambda > 0$ only power-kernelized d (small rad, $v_d \ll d^2$) survive, whose distribution is governed by the Lambert W source with an extra c -dependent collapse breakdown. This explains why abc is out of reach: it requires D in $\lambda > 1$, a region with no universal law.

11.3. Exact closed form of the lattice surplus $r(L)$. For even c , $\ell = c/2$, $|J| = c/2$, and $L_d := |J|/v_d$: the s -window is $[\lceil L_d \rceil, \lfloor 2L_d \rfloor] = [L_d, 2L_d]$ (the doubling window). Confirmed: $A_d = \text{nint}(L_d)$ covers 99.89%; the squarefree count in $[\lceil L \rceil, \lfloor 2L \rfloor]$ covers 99.83% (max deviation 1). This explains the universality of $r(L)$: the window depends only on L_d , not on c .

Proposition 11.1 (Exact closed form). *With $G(L, M) := \#\{s \in [L, 2L) : \mu^2(s) = 1, (s, M) = 1\}$ and $[j^2, k] := \text{lcm}(j^2, k)$:*

$$G(L, M) = \sum_{j \geq 1} \sum_{k|M} \mu(j) \mu(k) \left(\left\lfloor \frac{2L}{[j^2, k]} \right\rfloor - \left\lfloor \frac{L}{[j^2, k]} \right\rfloor \right).$$

For $M = 1$: $\#\{s \in [L, 2L) : \mu^2(s) = 1\} = \sum_{j \geq 1} \mu(j) \left(\left\lfloor \frac{2L}{j^2} \right\rfloor - \left\lfloor \frac{L}{j^2} \right\rfloor \right) = K_6 L + \Delta(L)$, with $\Delta(L)$ the deterministic arithmetic oscillation of the squarefree counting function in the doubling window (not noise): $|\Delta| \asymp L^{1/4}$, relative amplitude $\asymp L^{-3/4}$. Measured: $r(L) = 2.05, 1.05, 1.14, 0.82, 1.00, 0.98$ for $L = 1, 2.4, 4.5, 5.5, 20, 100$ (consistent across $c = 10^4 \rightarrow 10^6$).

Note $r(1) = 2.05$: when the window contains exactly one integer, it is forced to be 1, $\mu^2(1) = (1, d) = 1$ holds, but the density only gives $K_6 w(d) < 1$: the extreme form of Ehrhart surplus. It resolves an earlier paradox (“the exact ceiling asymptotic density underestimates by a factor 2–3”): the underestimate is the $r(L) > 1$ part, and the peak of R at $\lambda \approx -0.5$ corresponds to $L_d \approx e^{-2\lambda} \approx 2.7$, exactly the surplus interval of r .

11.4. The second scaling variable. The crossover scale is not $\sqrt{c}$. Collapse search with $t = \log(D/c^{1/2+\gamma})$:

exponent $1/2 + \gamma$	0.50 (no shift)	0.70	0.77	0.80
between- c spread	0.0644	0.0224	0.0067	0.0063

a $10\times$ collapse improvement. Hence

$$\tau = \log D - \kappa \log c, \quad \kappa = \frac{1}{1 + \theta}, \quad \theta := \left\langle \frac{\log \text{rad}(d)}{\log d} \right\rangle_{\text{tail } d};$$

squarefrees give $\theta = 1$, $\kappa = \frac{1}{2}$ (a special case); tail measurements ($c = 10^6$, $c = 10^7$) give $\theta = 0.248$, $\kappa = 0.800, 0.785$, consistent with the fitted 0.77–0.80. The crossover scale is $c^{0.78}$ because the tail-contributing d satisfy $\text{rad}(d) \approx d^{1/4}$, $v_d \approx d^{5/4} \leq c \Rightarrow d \lesssim c^{0.8}$.

Ambiguity in the definition of θ (important). Direct $u(d)$ -weighted averages give $\theta_{\text{glob}} \approx 0.72$ –0.76, while back-solving from κ gives $\theta = 0.248$: these differ by a factor 3 and are two different quantities. $\theta_{\text{bound}} = 1/\kappa - 1$: the upper bound forced by $\text{rad}(d) \leq c/d$ (literature back-solved values belong to this class); $\theta_{\text{typ}} = \frac{1-\kappa}{1+\kappa}$: measured typical values (deviation only 0.02–0.03, valid throughout $\kappa \in [0.5, 0.9]$); θ_{glob} : the untruncated global average, irrelevant here. Thus κ is not a constant but a function of θ ; “ $\kappa \rightarrow 4/5$ ” and “ $\theta \rightarrow 1/4$ ” are two statements of the same fact with no independent predictive power.

12. BGT SLOW VARIATION AND BAHADUR–RAO

12.1. $T(x)$ is slowly varying. $\frac{T(2x)}{T(x)}$: 1.2331, 1.2120, 1.1828, 1.1477 (at $x = 10^3, 10^4, 10^5, 10^6$), decreasing monotonically to 1; similarly ($\lambda = 10$) $1.987 \rightarrow 1.632$. Theoretically, $\log \frac{T(\lambda x)}{T(x)} \sim \frac{\sqrt{2} \log \lambda}{\sqrt{L \log L}} \rightarrow 0$ for fixed λ , so T is slowly varying. The paper’s earlier homemade term “sub-polynomial source” is standardly called *slow variation*.

12.2. Second-order slow variation and the de Haan auxiliary function. The BGT second-order theory asks for an auxiliary g with $\frac{A(\lambda x)/A(x)-1}{g(x)} \rightarrow \log \lambda$. Here the natural

choice is $g(x) := \frac{d \log T}{d \log x}$. Measured (smoothed local regression):

x	10^3	10^5	10^6	10^7
$g(x)$	0.313	0.246	0.2206	0.2055

$g \rightarrow 0$ slowly and second-order slow variation holds; $\log \frac{T(\lambda x)}{T(x)} / \log \lambda$ checks agree to 3% (residual: g decreasing on the interval). This is the standard name for the “common orbit”: all the Section 7 transient constants drifting along one orbit parametrized by v^* is the statement “ T is second-order slowly varying with de Haan auxiliary given by the Lambert saddle”.

12.3. Envelope theorem. From the Perrinversion saddle

$$\log T = (\Phi(\sigma^*) + \sigma^* L) - \log(\sigma^* \sqrt{2\pi\Phi''}) + \log \mathfrak{c}$$

with $\Phi'(\sigma^*) = -L + \frac{1}{\sigma^*}$: differentiating in σ^* , $\partial_{\sigma^*}[\Phi + \sigma L] = \Phi' + L = \frac{1}{\sigma^*} \neq 0$ — the term $\frac{1}{\sigma^*} \frac{dL}{d\sigma^*}$ was initially missed; the complete formula is

$$g = \frac{d \log T}{d \log x} = \sigma^* + \frac{\Phi'''}{2\Phi''(\Phi'' + \sigma^{*-2})}.$$

Verification: ratio measured/predicted = 1.043, 1.031, 1.016, 1.017 at $x = 3 \times 10^3, 10^5, 10^6, 3 \times 10^6$ (agreement 1.5%–4%; before adding the missed term the ratio was 0.83, off by 17%). The auxiliary function is determined by the third derivative of Φ at the saddle.

12.4. Far-tail rigorous version and the PNT transition. Exact Mellin inversion: $T(x) = \frac{1}{2\pi} \int_{\mathbb{R}} D(c + it) \frac{x^{c+it}}{c+it} dt$ with $D(s) = \sum_n \frac{1}{\text{rad}(n)n^s}$; writing $\Psi(s) = \log D(s) + sL - \log s$, the Gaussian expansion gives $\mathfrak{c}$ exactly as the far-tail relative contribution:

$$\mathfrak{c} = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \exp\left(\sum_{n \geq 3} \frac{\Psi^{(n)}(c)}{n!} \left(\frac{it}{\sqrt{\Psi''(c)}}\right)^n\right) dt, \quad T(x) \approx \frac{e^{\Psi(c)}}{\sqrt{2\pi\Psi''(c)}} \quad \text{when all } \Psi^{(n \geq 3)} \equiv 0.$$

The third-order term vanishes in expectation (odd power under an even measure), so the leading non-Gaussian correction is $\log \mathfrak{c} \asymp -\frac{\Psi^{(4)}}{8\Psi''^2} \asymp -\frac{c_4}{8} e^{-v^*} \asymp -C W(\sqrt{L})^{-1/2}$ (using $A_n := \Psi^{(n)}/\Psi''^{n/2} \asymp e^{-v^*(n/2-1)}$, measured $c_3 \approx 32.02$, $c_4 \approx 46.08$). *Strict conclusion:* the exponent is $\frac{1}{2}$ (with Lambert slow-variation correction), not the empirically fitted 0.421; the two are indistinguishable in the computable range ($L \sim 10$, $W \approx 1$); the sign is also right ($c_4 > 0 \Rightarrow \mathfrak{c} < 1$).

PNT transition. With $\Delta := \Phi_{\text{PNT}} - \Phi_{\text{true}}$: $|\Delta| \ll e^{v^*} \int_{v^*}^{\infty} g(v) \sigma e^{-c\sqrt{v}} dv \asymp \Phi e^{-c\sqrt{v^*}}$; measured $|\Delta|/\Phi \approx 2 \times 10^{-4}$ stable from $v^* = 1.099$ to 1.501, three orders below the $O(1/v^*) \approx 0.7$ subleading term of Φ , so the PNT transition is not the main residual.

The true source of the residual. The measured residual stabilizes at $g \times v^{*2} \approx 0.022$ (constant) for $x \geq 10^5$, i.e. the residual is $\frac{0.022}{v^{*2}}$ — exactly the contribution of the κ_2/v^{*2} term in the asymptotic expansion. This yields a testable quantitative prediction: the residual constant should be proportional to $\kappa_2/2 + \pi^2/6$.

13. THE S -UNIT REFORMULATION OF $\rho(c)$

13.1. ρ is an extreme quantity. Decisive evidence ($c \leq 10^5$, 250 sample points): (i) the minimizer a is almost always a prime power: searching $a \in \{1\} \cup \{\text{prime powers}\}$ attains the true value in 92.4% ($a = 1$: 20.8%; prime powers 77.2%); (ii) extreme vs. typical differ by three orders of magnitude: median $\rho(c) = 3689$ vs. typical $\text{rad}(a) \text{rad}(b)$ with median 2779... precisely, extremes sit at the (9, 3689, 27796) quartiles; (iii) at the minimum $\text{rad}(a)$ is $O(10)$ with median 24, not the typical $O(c)$. There is no smooth pointwise law: within-decade standard deviations of $\log \rho$ are 1.282, 1.634, 1.770 for c in $[10^2, 10^3]$, $[10^3, 10^4]$, $[10^4, 10^5]$ — fluctuations *grow* with c , the opposite of a smooth law. This diagnoses the failure of moment methods: moments/Chernoff/ saddles describe typical fluctuations, while ρ sits in the

extreme tail where large deviation estimates fail systematically — not a missing constant, but an additive tool measuring an extreme quantity. *The deeper conclusion:* the duality of Section 10 should be classified by *type* rather than *function*: $\rho(c) = \min_a \text{rad}(a) \text{rad}(c-a)$, although defined through rad , belongs to the extreme (P^+ -type) side. This simultaneously explains three previously isolated observations: the ω drift, the Chernoff failure, and the growth of fluctuations with c .

13.2. Exact S -unit reconstruction and verification. The structure of measured extremizers $a^*, b^* = c - a^*$ ($c \leq 3 \times 10^5$, 314 points): the total prime support has size 1–5 in frequencies 4, 45, 169, 91, 5. By Theorem 3.4, $\mathcal{S}_a \cap \mathcal{S}_b = \emptyset$; writing $a^* = \prod_{p \in \mathcal{S}_a} p^{e_p}$, $b^* = \prod_{q \in \mathcal{S}_b} q^{f_q}$, one gets $a^* + b^* = c$: exactly an S -unit equation, with

$$\rho(c) = \min \left\{ \prod_{p \in \mathcal{S}} p : \text{there is a coprime } S\text{-unit decomposition } a + b = c \right\}.$$

Constructive enumeration: with $|\mathcal{S}| \leq 5$ and $\prod_{p \in \mathcal{S}} p \leq 2 \times 10^4$ (all needed sets included), enumerate all coprime S -unit decompositions for $c \leq 5 \times 10^4$: pointwise agreement 1451/1454 = 99.79%, median ratio = 1.0000 (residuals: 3 boundary effects, 46 cases with $\rho > 2 \times 10^4$ beyond the enumeration bound). This replaces the Dickman heuristic and the required coprime singular series correction: S -unit enumeration is constructive and exact. *Correction to record:* the earlier “prime-power enumeration” had a serious bug (covered $a = 1$ only when $c - 1$ was a prime power), inflating ρ by more than $10\times$; after full enumeration $\rho/c \approx 0.04\text{--}0.13$, correct.

13.3. Closed form of the average order. For a fixed prime set $\mathcal{S}$ with $|\mathcal{S}| = s$, let $C(N)$ be the number of ordered coprime S -unit decompositions $c = a + b \leq N$, and $D(N)$ the number of c admitting one. Then

$$C(N) \sim \frac{\binom{2s}{s}}{s! \prod_{p \in \mathcal{S}} \log p} (\log N)^s, \quad D(N) \sim \frac{\binom{2s}{s}}{2 s! \prod_{p \in \mathcal{S}} \log p} (\log N)^s.$$

Verification: 9 different $\mathcal{S}$, N to 10^9 : ratios 0.93–1.02; decisive test: same $s = 3$, different prime sets ($\{2, 3, 5\}$ vs. $\{3, 5, 7\}$ vs. $\{2, 5, 11\}$) give consistent ratios — the closed form depends only on s and $\prod \log p$, not on the specific primes. At $s = 1$ it is exact: $D(N) \sim \log_2 N$.

Distribution of ρ at small values: $\#\{c \leq N : \rho(c) \leq y\} \sim D(N)$ measured ($N = 4 \times 10^4$, $y = 2, 6, 30, 210$): measured/closed-form = 1.05, 0.97, 0.99, 0.85 for $s \leq 3$; the closed form saturates at $s \geq 4$ from union overlap.

A structural fact: $\rho(2^i + 1) = 2$ (since $2^i + 1 = 2 + 2^i \dots$ precisely $\text{rad}(2) \text{rad}(2^i + 1 - \text{adjustment}) = 2$). Hence $\rho(c)$ does not go to ∞ : it has infinitely many small values. There are two regimes: the small-value end (fixed y : $\#\{\rho \leq y\}$ grows logarithmically, S -unit controlled) and the typical value ($\rho \sim c^{0.77}$, extreme statistics).

13.4. Average order vs. Evertse’s worst-case bound. $C(N)/D(N) \rightarrow 2$, while Evertse’s theorem gives the worst case $\leq 3 \cdot 7^{|\mathcal{S}|+2}$: at $s = 3$ this is $\approx 5 \times 10^4$ vs. average $15 \times 10^4 \dots$ precisely, average $\asymp (\log N)^3$ versus worst 5×10^4 — four orders of magnitude apart. This is quantitative evidence that average-type questions are more tractable than worst case. However: the method gives the structure of the set of solution-bearing c (size $(\log N)^s$), while abc needs the rare high-quality ones among them (size q -large with small $\text{rad}(c)$): the former is a density problem (solved), the latter a doubly rare event (still out of reach).

14. MECHANISTIC DECOMPOSITION OF $\mathcal{N}_0(N)$

14.1. Three-factor decomposition. By Theorem 3.1, $q > 1 \iff \text{rad}(a) \text{rad}(b) < d_c$, so

$$\mathcal{N}_0(N) = \sum_{c \leq N} \mathbf{1}\{d_c > \rho(c)\} \cdot (\text{multiplicity}),$$

where both factors are now computable: the distribution of d_c by Theorem 2.2, $\rho(c)$ by the S -unit enumeration of Section 13. Decomposition (numerical): $\mathcal{N}_0(N) = \kappa \cdot \left(\sum_{c \leq N} K_6 \Sigma(\rho(c)) \right) \cdot \bar{m}$ with $\kappa = 0.475$, $\bar{m} = 1.40$:

N	4000	8000	16000	32000
κ	0.468	0.483	0.471	0.480
$\bar{m}$	1.39	1.37	1.46	1.40
$\mathcal{N}_0$ measured			...	240

Both factor constants are stable across $8\times$ scales to $\pm 1.5\%$ (check: $\mathcal{N}_0(32000) = 240$ vs. fit $1.965 \cdot 32000^{0.467} = 249.6$). Hence the exponent is completely determined by $\sum_c K_6 \Sigma(\rho(c))$, i.e. by the distribution of ρ .

14.2. Candidate falsifications and the negative-correlation mechanism.

candidate	verdict
(i) extreme small- ρ tail dominates	falsified: contributions distribute as $\rho \in [31, 211)$ 41.6%, $[7, 31)$ 23.5%, $[211, 2311)$ 23.2%, $\rho < 7$ only 8.6%; the dominant quantile is median-or-below, not the extreme tail
(ii) multiplicity grows with N	falsified: $\bar{m} = 1.40$ constant
(iv) d_c and $\rho(c)$ negatively correlated	new discovery, the main cause

Evidence for (iv) (by d_c stratum, median of ρ/c):

d_c	1	2–8	8–64	64–1024	≥ 1024
ρ/c median	0.033	0.105	0.111	0.143	0.111

The more power-kernelized c is, the harder it is to write as a sum of two low-radical numbers — relatively $4\times$ harder. This is the source of $\kappa = 0.475$: the lower tail of ρ is exactly anti-correlated with large d_c .

14.3. Exponent formula and consistency with $\beta(\varepsilon)$. With $\log \rho \sim \mathcal{N}(\mu, \sigma^2)$, $\mu = \omega L + \text{const}$: $\mathbb{E}[\log \rho / \rho] = e^{-\mu + \sigma^2/2}(\mu - \sigma^2)$, giving

$$\beta \approx 1 - \omega + \frac{d(\sigma^2/2)}{dL} + \frac{\omega - 2\sigma\sigma'}{\mu - \sigma^2}.$$

First term: $\rho \asymp c^\omega$ relaxes the threshold relatively; second: growing fluctuations strengthen the lower-tail weighting (factor $e^{\sigma^2/2}$); third: a small correction. Test: $c \gtrsim 10^4$ window, predicted 0.710, measured 0.592; 24000 window: predicted 0.508, measured 0.575 (agreement 5–12%). Note σ grows monotonically $1.30 \rightarrow 1.57$ — fluctuations grow with c , the signature of extreme statistics, and they push the exponent up.

Consistency with abc (a clean test). At $\varepsilon > 0$ the threshold tightens, $\omega \rightarrow \omega + \varepsilon'$, so $\beta(\varepsilon') \approx 1 - \omega + \sigma\sigma' - \varepsilon'$ with critical point $\varepsilon'_{\text{crit}} \approx 1 - \omega + \sigma\sigma' \approx 0.5$; the measured table (Table of §3) gives $\mathcal{N} = 2$ at $1 + \varepsilon = 1.5$, $\beta \approx 0$: the first structural (non-fit) explanation of the entire descent of $\beta(\varepsilon)$ and its critical point.

A clean necessary condition (obtained in passing). From $X_a \leq \log a$, $X_b \leq \log(b/2)$, $X_a + X_b > \log(ab) + \varepsilon' \log c$:

$$X_c > \varepsilon' \log c + O(1) \implies d_c \gtrsim c^{\varepsilon'}.$$

c must be power-kernelized (consistent with the measured median $d_c = 171$ of bearing solutions). But this condition alone gives $\sum_c c^{-\varepsilon'} \sim N^{1-\varepsilon'}$, worse than BLT's 0.65 — because it relaxes the $\text{rad}(a)\text{rad}(b)$ side, once again confirming “joint constraints do not decouple”. And as $\varepsilon \rightarrow 0$, $c^{\varepsilon'} \rightarrow 1$: the condition degenerates to empty — the clean statement of why abc is hard.

15. CONCLUSION

15.1. Established results. (1) Exact discrete law (Theorem 2.2): $\mathbb{P}(X_n = \log d) = \frac{6}{\pi^2} \frac{w(d)}{d \operatorname{rad}(d)}$; normalization checked to 7×10^{-6} ; moments match $C(1)$ to 1.5×10^{-4} . (2) Three-branch parity law (Theorem 2.8): $\mathbb{P}(d \text{ even} \mid r = 1) \equiv 0$ confirmed to 10^{-4} . (3) Master identity (Theorem 3.1): $q > 1 \iff \operatorname{rad}(a)\operatorname{rad}(b) < d_c$. (4) Three truncation constraints (Theorem 3.2): zero violations, slack ≥ 0.004 . (5) Prime exclusion and covariance closed form (Theorem 3.4, Proposition 3.5): -0.163290 , measured agreement 4%. (6) Local CRT conditional law (Theorem 4.2): case C is the corrected law; prime-by-prime to 10^{-6} , convolution to 0.5%, end-to-end 0.4%–2%; the usual law underestimates 47–78%. (7) Exact counts (Theorem 3.6): 2054 solutions; $\mathcal{N}_0 \asymp N^{0.467}$; $\beta(\varepsilon) \rightarrow 0$. (8) Analytic theory of $C(\beta)$ (Section 5): closed-form integral, Laurent expansion, natural boundary, telescoping identity, $2 \log 2$ cancellation. (9) Rigorous tail bound (Theorem 6.1): the only completed upper-bound component. (10) Lambert W subpolynomial source (Theorem 7.1): main term from $[e^{v^*}, e^{1/\sigma}]$; κ_1 – κ_8 closed forms (verified to 10^{-19}); the $\iota(r) = (-1)^{r+1}\rho(r)$ cancellation; Stirling closed form of $c_{n,j}$; generating function; classification theorem. (11) Second-order slow variation framework (Section 12): T slowly varying, $g = \sigma^* + \frac{\Phi'''}{2\Phi''(\Phi'' + \sigma^{*-2})}$; envelope theorem verified 1.5%–4%; $\mathfrak{c} = 1$. (12) Mazur computable law (Section 8): $S \asymp N^{s-1}(\log N)^{k_{\text{eff}}}$, $k_{\text{eff}} = 1.91 \pm 0.17$. (13) Ceiling theorem (Theorem 9.2): $q \leq \frac{\log c}{\log 2 + \log \operatorname{rad}(c)}$, sharp; boundary case (1, 1, 2) handled by Remark 9.3 and recorded in the Falsification List. (14) rad – P^+ duality (Section 10): W_0 vs. W_{-1} ; saddle form of $\rho(u)$. (15) Crossover scaling and $r(L)$ (Section 11): $r(1) = 2.05$, exact closed form; second scaling variable $\kappa = 1/(1 + \theta) \approx 0.78$. (16) S -unit reformulation of ρ (Section 13): 99.79% pointwise; average order $\frac{\binom{2s}{s}(\log N)^s}{2s! \prod \log p}$. (17) Mechanistic decomposition of $\mathcal{N}_0(N)$ (Section 14): $\kappa \cdot \sum K_6 \Sigma(\rho) \cdot \bar{m}$; the d_c – ρ negative correlation; the structural explanation of $\beta(\varepsilon)$.

15.2. Open problems and unproved items. *Conjectures:* the rigorous form of the Mazur formula (with $\mathcal{K}$ closed form); a single-line closed form for the coefficients $c_{n,j}$, $j \geq 2$ (currently an explicit difference of two Stirling sums). *Unproved:* the sieve inequality (numerically 1.2–1.5 $\times$ headroom); the precise profile of far-tail coefficients (A of the same order but drifting ratio); the exact asymptotic of $\langle \rho \rangle$ (transient vs. subpolynomial); the evolution law of θ . *Must be re-examined:* the Section 13 “ S -unit exact reconstruction” follows a different code path from the buggy prime-power enumeration, but similar restricted-enumeration risks cannot be fully excluded; independent re-verification is recommended.

15.3. On abc: three structural walls. (i) *Threshold.* The conditional law (Theorem 4.2) is provable for $D \leq c^{1/2-\delta}$, while abc solutions live at $D \gtrsim c^{0.78}$ (Corollary of Section 11: $X_a + X_c \gtrsim \log(2a)$ requires d_c large). The provable half contains no solutions; the measured tail decays like $V^{-1/2}$ and the trivial bound is tight to a constant factor, so the hope “use sieve bounds to enlarge the conditional law’s range” is falsified: the true meaning of $c^{1/2}$ is the threshold of the count of integers in the window $L_d = \frac{c/2}{v_d}$. (ii) *Joint constraints do not decouple.* Relaxing any one of the three radicals inflates the count by a factor 387 (Proposition 6.6), enough to turn a convergent series into a divergent one. (iii) *Tool mismatch.* The Fourier circle method is controllable iff $\alpha + \beta > 2$ (Section ??), while Mazur parameters always have $\alpha, \beta < 1$; sieves excel at multiplicative, geometric number theory at additive, and abc needs both at once. BLT reached $X^{0.65}$ with geometry; the gap to $O(1)$ is discontinuous. In addition: our weak abc bound is weaker than de Bruijn’s 1962 trivial bound (Table after Theorem 6.7); and d_c has heavy tail $\mathbb{E}[d_c] = n \cdot \sum_p \frac{1}{p+1} = \infty$, defeating all first-moment methods (the same wall as Proposition 6.4 seen twice).

It must be emphasized: “outside the range of our tools” and “unprovable” are two different statements. This paper proves the former, not the latter.

15.4. Falsification list.

Withdrawn	Correct replacement
equidistribution conjecture	local CRT conditional law (Theorem 4.2); the usual law underestimates by 47–78%
$(\log N)^3$	power $(\log N)^2$ with $k_{\text{eff}} = 1.91 \pm 0.17$
$\rho \asymp c^{2/3}$	corrected to $\rho = e^{1.30c} e^{-2.60\sqrt{\log c}}$, slope $\omega(c) = 1 - \frac{3.56}{\log c}$; aggregate mean law only
$\kappa' \approx 1.1$, $\kappa_1 \approx 1.67$, $\omega \approx 0.66$	all transients, not asymptotic constants
$\mathfrak{c} \approx 0.88$ (universal constant)	$\mathfrak{c} = 1$ asymptotically; 0.84–0.89 was a PNT-precision artifact
$\int_1^2 C = P(2) - P(3)$, $\Delta_* = 0.23134$	$P(2) = 0.452247514$; $\Delta_* = 0.056574$
C has phase transitions at $\beta = 1, 2$; poles at $-1, -2, \dots$	completely monotone; $\Re\beta = -1$ is the natural boundary
weak abc $N^{1.38}$ competitive	weaker than de Bruijn 1962 $N^{2/3}$
hit probability = three-sum tail	single-variable tail (Theorem 4.1)
Corollary 3.4 as independent constraint	equivalent to $\text{rad}(a) < d_c/2$
$\rho \sim 6P^+$	P^+ misused (took the least prime factor); corrected ρ/P^+ median = 30
S -unit: $C/D = 2$ exact, $D = \mathcal{U}$	artifact of generators missing pure prime powers
hard truncation is the main crossover cause	τ does not collapse, explains only $\sim 1/3$; main cause is truncation under the conditional measure
ceiling theorem universal (“for all co-prime triples”)	violated by $(1, 1, 2)$ with $q = 1$ vs right side $\frac{1}{2}$; needs $c \geq 3$ (or $a < b$)
$(\log N)^{-0.419}$ as independently measured exponent	unidentifiable: conditional fit at fixed $\alpha = \frac{1}{2}$; $\log N, \log \log N$ collinear on the computable range; leave-one-out span 2.808; main result $N^{0.4676 \pm 0.025}$
“234 solutions zero violations” verifies the ceiling theorem	verification-set self-selection bias: $(1, 1, 2)$ with $q = 1 \notin \{q > 1\}$ slips through the gap
Theorem 7.1 tagged [proved] (original)	early “region I is $O(e^{v^*}/v^{*2})$ ” gives only magnitude, insufficient for κ_1 ; Lemma 7.2 (PNT exchange consistency) added, $\kappa_1 = \iota_0 - \rho_0 = 1$

15.5. A methodological lesson. This paper fell into the same trap at least five times: $\kappa' \approx 1.1$, $\kappa_1 \approx 1.67$, $\omega \approx 0.66$, $\theta \approx 0.25$, $\mathfrak{c} \approx 0.88$ — each with a beautiful R^2 (e.g. 0.9991) and clean residuals. The root cause: the very act of power-law fitting on a slowly varying function is deceptive. Since $v^* \sim \log \log x$ moves only $\sim 25\%$ across a $100\times$ scale, any finite-interval fit returns a very stable-looking constant. *Slowly varying functions manufacture stable false constants.* Two further failures deserve more warning: $\mathfrak{c} \approx 0.88$ came from insufficient PNT integration accuracy, and $\rho \sim 6P^+$ used the wrong function (least prime factor for greatest). A third, different kind of trap is verification-set self-selection bias: the ceiling theorem was once reported “234 solutions, zero violations”, yet the counterexample $(1, 1, 2)$ with $q = 1$ never entered the $q > 1$ -restricted verification set. The remedy: for any theorem with boundary conditions, explicitly scan the minimal neighborhood outside the boundary. Therefore: on problems of this type, “numerical confirmation” is necessary but not sufficient; only cross-validation by genuinely independent paths (e.g. the two paths for κ_4 agreeing to 10^{-19}) and rigorous proof are reliable. A pretty single-path fit is held in reserve.

15.6. Closing remark. The core original contribution is the Lambert W subpolynomial source of radical-type Euler products:

$$\Phi(\sigma) = \frac{a}{\sigma W(a/\sigma)} \left[1 + \frac{1}{W} + \frac{1 + \pi^2/3}{W^2} + \frac{2 + \pi^2/6}{W^3} + \frac{6 + \frac{7\pi^2}{6} + \frac{7\pi^4}{15}}{W^4} + \dots \right],$$

with main term from $[e^{v^*}, e^{1/\sigma}]$, coefficients $\kappa_n = (n-1)! + \sum_j c_{n,j} \zeta(2j)$ from the order-by-order cancellation of $\log 2$ with odd zetas, unifying T , Σ , $\Sigma_\perp$ and explaining all “constant drifts” at one stroke: they are transients.

15.7. Cross-program consistency note. This paper’s route is statistical/analytic (valuations, discrete laws, Euler products, Tauberian theory). A parallel valuation/lattice-geometry program exists; the two agree quantitatively at several load-bearing points: wall (i) — “the provable region contains no solutions” (here) vs. “ $\text{rad}(ab) \geq c^{1/2-o(1)}$: the method’s information budget does not reach” (there); wall (ii) — our factor 387 ($\varepsilon = 0.1$, $c \leq 3 \times 10^4$: 94 vs. 36362) independently captured there (entry E25); the counting exponents 0.467 vs. $\#\{q > 2\} \leq X^{1/2+o(1)}$ (difference 0.033, inside the $o(1)$ slack); the record triple $q = 1.56789$, $c = 4375$ reproduced digit-identically by independent scans; the $\beta(\varepsilon)$ descent and critical point $\varepsilon \approx 0.5$; and the same methodological lesson (slowly varying functions manufacture false constants) discovered independently (their E24). The conceptual meeting point is the additive/extreme classification (Section 10): W_0 on the rad side, W_{-1} on the P^+ side; and $\rho(c)$, though defined through rad, belongs to the extreme side (the moment method fails by 10^2 – $10^{2.7}$, while S -unit enumeration is exact to 99.79%). This suggests the criterion for the classification is *type*, not *function*: abc requires three extreme conditions to hold simultaneously, and the tools of this paper (and of the current mainstream) are essentially additive. That explains both our failure and the discontinuous gap between BLT’s $X^{0.65}$ and $O(1)$.

Items pending re-verification: the Section 13 S -unit reconstruction (99.79%) shares a code lineage with the since-fixed prime-power enumeration bug; it is to be regarded as provisional until re-checked with the corrected independent implementation. The remaining results (conditional law, ceiling theorem, Mazur formula, duality, crossover scaling, S -unit reconstruction) complete the picture.

REFERENCES

- [1] J. Oesterlé, *Nouvelles approches du “théorème” de Fermat*, Astérisque **161–162** (1988), 165–186.
- [2] D. W. Masser, *Note on a conjecture of Szpiro*, Astérisque **183** (1990), 19–23.
- [3] N. G. de Bruijn, *On the number of positive integers $\leq x$ and free of prime factors $> y$* , Nederl. Akad. Wetensch. Proc. Ser. A **54** (1951), 50–60.
- [4] N. G. de Bruijn, *On the number of integers whose prime factors divide n* , Illinois J. Math. **6** (1962), 137–141.
- [5] T. D. Browning, J. Lichtman, J. Teräväinen, *On the radical of sums of two integers*, preprint, 2024.
- [6] J. Li, *On the abc conjecture and the radical of $a + b = c$* , preprint, 2025.
- [7] D. M. Kane, *The number of solutions to $a + b = c$ with $\text{rad}(abc)$ small*, preprint.
- [8] N. H. Bingham, C. M. Goldie, J. L. Teugels, *Regular Variation*, Cambridge University Press, 1987.
- [9] R. R. Bahadur, R. R. Rao, *On deviations of the sample mean*, Ann. Math. Statist. **31** (1960), 1015–1027.
- [10] J.-H. Evertse, *On equations in S -units and the Thue–Mahler equation*, Invent. Math. **75** (1984), 561–584.
- [11] G. Tenenbaum, *Introduction to Analytic and Probabilistic Number Theory*, 3rd ed., AMS, 2015.
- [12] J. Knopfmacher, *Abstract Analytic Number Theory*, 2nd ed., Dover, 1990.
- [13] S. W. Golomb, *Random permutations*, Bull. Amer. Math. Soc. **67** (1961), 619.
- [14] K. Dickman, *On the frequency of numbers containing prime factors of a certain relative magnitude*, Ark. Mat. Astr. Fys. **22A** (1930), 1–14.
- [15] G. Halász, *Remarks on the remainder in the prime number theorem*, Period. Math. Hungar. **15** (1984), 299–314.
- [16] R. M. Corless, G. H. Gonnet, D. E. G. Hare, D. J. Jeffrey, D. E. Knuth, *On the Lambert W function*, Adv. Comput. Math. **5** (1996), 329–359.