hexagon:2611.01292v1

Second-Order Minimum Heights and Arithmetic Phase Transitions in Egyptian Fraction Hypergraphs

Pengcheng Chen

Abstract

We study the least maximum denominator in representations of a fixed rational by exactly K distinct unit fractions, and the minimum number of forbidden denominators that obstruct all such representations. For every fixed rational $0 < r \le 1$, the minimum height is

$$\frac{K}{1-e^{-r}} + \frac{1 - r/(e^r-1)}{1-e^{-r}} \frac{K \log\log K}{\log K} + o_r\left(\frac{K \log\log K}{\log K}\right).$$

For $r = 1$ the second-order coefficient is $e(e-2)/(e-1)^2$. At this onset height the limiting transversal cost has a zero plateau, a nonlinear part governed by $v \log v - v + 1$, and a linear part. For proper rational targets we also establish the macroscopic phase diagram, extremal stability, and second-order windows at fixed linear heights, for both at-most and exact length. Our approach combines synchronizing divisor carriers, Fourier inversion on the actual rational lattice with a cardinality coordinate, sparse modular repair, and a weighted deletion ledger. Complementary results give a uniform deficit at larger heights and a first-order law for at-most-length representations in fixed positive-density denominator families.

MSC 2020: 11D68, 11B30, 05C65, 60F05.
Keywords: Egyptian fractions, hypergraph transversals, adversarial deletion, local limits, modular absorption.

1 Phase diagram and main results

All logarithms are natural, and intervals of integers are intersected with $\mathbb{Z}$. Write $R(E) = \sum_{n \in E} 1/n$, and fix a rational $r = a/b \in (0, 1]$ in lowest terms. The macroscopic phase theorems below are restricted to $0 < r < 1$; the onset theorem and its minimum-height corollary also include $r = 1$. The vertices are the integers in $[2, M]$. The two edge families consist of sets E with $R(E) = r$ and respectively $|E| \le K$ or $|E| = K$. Their transversal numbers are denoted by $\tau_K(r; M)$ and $\tau_K^{=}(r; M)$; an empty family has transversal number zero. Put

$$A = (e^r - 1)^{-1}, \quad D_* = A + 1, \quad c = rA(A+1), \quad \delta_K = \frac{\log\log K}{\log K}.$$

For the proper-target macroscopic results only, define $q = \max_{p|b} p^{v_p(b)}$. This symbol is not used in the unit-target onset proof. For $0 < r < 1$, one has $q > e^r$: if $q = 2$ then $b = 2$ and $r = 1/2$, and otherwise $q \ge 3$. The question has several different regimes. The first is governed by the largest available reciprocal sum, another by divisibility of the common denominator, and a third (for exactly K terms) by the smallest available reciprocal sum.

The unit-target minimum-height question was posed by Erdős and Graham [1, p. 33]. Martin [2, Theorem 1 in the published version; Theorem 2 in arXiv v1] proved that the minimum maximum

denominator in an exact K-term representation of a fixed positive rational is $K/(1-e^{-r})+O_r(K\delta_K)$, with optimal error order. Croot [3] established a related sharp-order short-interval theorem. Bloom [4] and Liu–Sawhney [5] developed powerful Fourier and pruning techniques for reciprocal subset sums. The latter paper's Proposition 5.2 concerns arbitrary sets and general rational grid targets, and its denominator applications use two scales (we refer to the numbering of arXiv v1). Our carrier arguments are adaptations of those methods, not claims of invention of the circle method or of modular cancellation. The specific questions here are sharp adversarial costs, exact cardinality and the onset window.

Theorem 1.1 (Macroscopic heights and stability)**.** *Assume $0 < r < 1$. For each fixed $D > 0$ and $M = DK + o(K)$,*

$$\tau_K(r; M)/K \longrightarrow \min(A, D/q), \tag{1}$$

$$\tau_K^{=}(r; M)/K \longrightarrow [\min(A, D/q, D - D_*)]_+. \tag{2}$$

For the second assertion suppose $D > D_*$. Every transversal with size equal to the limiting minimum times K, plus $o(K)$, is at symmetric-difference distance $o(K)$ from one of the minimizing templates: an initial segment of the vertices, the set of multiples of q, or a terminal segment. The first assertion has the same stability conclusion with only the first two templates. Segment lengths may be taken to be the size of the transversal, up to rounding. At a tie the assertion is proximity to the union of the template classes, not to every class.

In particular, the critical height for saturation at AK is qA for at most K terms and $\max(qA, 2A+1)$ for exactly K terms.

Theorem 1.2 (Second-order height windows)**.** *Assume $0 < r < 1$. Fix $D > 0$ and $u \in \mathbb{R}$, and let $M = \lfloor DK + uK\delta_K \rfloor$. For at most K terms use the index set $I = \{1, 2\}$; for exactly K terms assume $D > D_*$ and use $I = \{1, 2, 3\}$. Define*

i	$h_i(D)$	$g_i(D, u)$
1	A	$-D(e^{c/D} - 1)$
2	D/q	$u/q - D(e^{1/q} - 1)$
3	$D - D_*$	$u - D(e^{1-c/D} - 1)$

Let $h_* = \min_{i \in I} h_i(D)$ and $I_* = \{i \in I : h_i(D) = h_*\}$. For the corresponding transversal number τ,

$$\tau = Kh_* + K\delta_K \min_{i \in I_*} g_i(D, u) + o(K\delta_K). \tag{3}$$

An asymptotically optimal transversal (with an $o(K\delta_K)$ error) must be $o(K)$-close to a template that minimizes both levels of this comparison.

The second-order comparison must be restricted to the first-order minimizers. When $q > 1 + e^r$, the exact-length phases occur in the order terminal, divisibility, initial, with junctions $qD_*/(q-1)$ and qA. When $q < 1 + e^r$, the divisibility phase is absent and the junction is $2A + 1$. Formulas (3) also cover ties without requiring a nondegeneracy assumption.

The synchronized-carrier and actual-lattice completion estimates are proved in Sections 2–4, with the general pruned grid in Section 12.2. The capped cost profile in this paper suffices for its active parameters; it must not be identified with the full Dickman profile beyond exponent two.

The boundary $D = D_*$ calls for a different argument, because a transversal then has only second-order size. In the following theorem and corollary, let $0 < r \le 1$ be any fixed rational, including the integer target $r = 1$. Put

$$\alpha = 1 - rA, \qquad J(v) = v \log v - v + 1, \qquad u_0 = D_*\alpha, \quad u_1 = D_*\alpha e^{\alpha}.$$

Here $0 < \alpha \le 1 - 1/(e-1) < \log 2$. Indeed $r/(e^r - 1)$ is decreasing for $r > 0$, and $1 - 1/(e-1) < 1/2 < \log 2$. In particular $1 < e^{\alpha} < 2$, including at the endpoint $r = 1$.

Theorem 1.3 (Nonlinear onset)**.** *For every fixed rational $0 < r \le 1$, for $M = \lfloor D_*K + uK\delta_K \rfloor$ and each fixed $u \in \mathbb{R}$,*

$$\lim_{K\to\infty} \frac{\tau^{=}_K(r; M)}{K\delta_K} = \Psi_r(u) = \begin{cases} 0, & u \le u_0, \\ D_*J(u/u_0), & u_0 < u < u_1, \\ u - D_*(e^{\alpha} - 1), & u \ge u_1. \end{cases} \tag{4}$$

If $u < u_0$, the edge family is empty for all sufficiently large K. No assertion of eventual emptiness at equality is made.

Corollary 1.4 (Minimum height)**.** *Let $0 < r \le 1$ be a fixed rational, and let $m_K(r)$ be the least M for which an exact K-term representation exists. Then*

$$m_K(r) = D_*K + D_*(1 - rA)K\delta_K + o_r(K\delta_K). \tag{5}$$

For the unit target, this gives the explicit refinement

$$m_K(1) = \frac{e}{e-1}K + \frac{e(e-2)}{(e-1)^2}\frac{K \log\log K}{\log K} + o(K\delta_K), \tag{6}$$

where the coefficient of $K\delta_K$ is $0.661303112661534\ldots$. For $K \ge 2$ a representation of 1 by positive distinct unit fractions cannot contain denominator 1, so this is the same minimum-height problem as Martin's for large K. More generally, if $h = \lfloor \beta K\delta_K \rfloor$ with fixed $\beta \ge 0$ and $m_K(r; h)$ is the least height with $\tau^{=}_K > h$, the second coefficient is $u(\beta)$, where

$$u(\beta) = \begin{cases} D_*\alpha v, & \beta \le D_*J(e^{\alpha}), \quad J(v) = \beta/D_*, \ 1 \le v \le e^{\alpha}, \\ \beta + D_*(e^{\alpha} - 1), & \beta \ge D_*J(e^{\alpha}). \end{cases}$$

In particular $u(\beta) = u_0 + \alpha\sqrt{2D_*\beta} + O_r(\beta)$ as $\beta \downarrow 0$, after taking the large-K limit.

The lower bound on the coefficient in (5), including (6), follows directly from the forbidden large primes, without Fourier analysis. The upper bound and nonlinear onset require the full repair argument. Neither a finer onset scale nor the eventual existence of representations at $u = u_0$ follows from the stated errors. Sections 11 and 12 establish the complementary growing-height and restricted-family results. The latter concerns at most K terms; exact length in restricted families requires additional lattice information.

2 Arithmetic estimates and two carrier regimes

We use the standard prime number estimates

$$\log \operatorname{lcm}(1, \ldots, \lfloor x \rfloor) \ll x, \qquad \sum_{p \le x} p^{-1} = \log\log x + c_M + O(1/\log x), \qquad \pi(x) \ll x/\log x. \tag{7}$$

We also use the standard upper-bound sieve: for primes from an arbitrary set $\mathcal{P}$ below $\exp(\log X/\sqrt{\log\log X})$, the count of integers in $[1, X]$ divisible by none of these primes is $\ll X\prod_{p\in\mathcal{P}}(1-1/p)$ (the fundamental lemma, in the form of [5, Lemma 2.4]). We only apply it at well-separated exponential scales. For later estimates,

$$\sum_{n\le X}(3/2)^{\Omega(n)} \ll X(\log(2X))^{1/2}, \qquad \sum_{p^j>Y,\ j\ge 2} p^{-j} \ll Y^{-1/2}. \tag{8}$$

The first follows by expanding the nonnegative divisor convolution with local coefficients $g(p^j) = \frac{1}{2}(3/2)^{j-1}$ and applying Mertens to the Euler product. The second follows by splitting at $p = \sqrt{Y}$ and summing geometric tails.

Write $L = \log N$, $\lambda = \log L$, $\delta = \lambda/L$. A power-smooth integer has every prime-power divisor at most Y. We work in two regimes for a pool $P \subseteq [m, N]$:

Fine. $m = N/L^3$, $Y = N/L^B$, with fixed B sufficiently large. Assume $|P| \ge c_0 N$ and $|P \cap s\mathbb{N}| \ge \eta N/s$ for every prime power s appearing in P, where $c_0, \eta > 0$ are fixed.

Coarse. $m = N/L^{1/100}$, $Y = Ne^{-12L^{9/10}}$. Assume $|P| \ge N/L^{3/50}$ and $|P \cap s\mathbb{N}| \ge L^{-7/100}N/s$ for every prime power s appearing in P.

In both regimes assume P is power-smooth and $\Omega(n) \le 20\lambda$ for $n \in P$. When $s = 1$, the local bound follows from the global size. Write $Q = \mathrm{lcm}(P)$; then $\log Q \ll Y$. The following lemma formalizes the common-carrier device.

Lemma 2.1 (Carriers and synchronizing primes)**.** *In either regime, for any appearing power s (or $s = 1$), even after deletion of $o(N/s)$ elements in the fine regime or $o(L^{-7/100}N/s)$ in the coarse regime, the surviving s-multiples contain a family d_sU_s, with $s \mid d_s$ and $U_s \subseteq [1, z_s]$, $z_s = N/d_s$. In the fine regime, for fixed constants $d, J, \nu, \gamma > 0$,*

$$L^{d\nu} \le z_s \le L^{3+dJ}, \qquad |U_s| \ge \gamma z_s.$$

In the coarse regime,

$$e^{L^{7/10}} \le z_s \le L^{1/100}e^{L^{9/10}}, \qquad |U_s| \gg L^{-27/100}z_s.$$

Any two obtained sets U_s have common prime divisors whose product exceeds N. A common divisor means a prime dividing some member of each set, not necessarily all members.

Proof. Write $X = N/s$ and let T be the surviving set of cofactors. For the construction of a single fiber only the local lower bound for this s is required, along with the common interval bounds. In particular, when $s = 1$ the same construction applies to a subset having the stipulated global density, with the fine density constant decreased if necessary; no health assumptions for its other powers are needed. Define $s_y(k) = \prod_{p\le y} p^{v_p(k)}$. Its logarithmic sum over $k \le X$ is $\ll X\log y$, by Chebyshev and summing the valuations. In the fine regime take $y = L^d$. For any desired constant small proportion, one may choose $\nu > 0$ sufficiently small so that the count of integers avoiding all primes in $[y^\nu, y]$ is less than that proportion of X. Indeed the prime reciprocal sum tends to $\log(1/\nu)$; if a subcollection of primes up to y has reciprocal sum v, expanding the square and counting multiples gives

$$\sum_{k\le X}\left(\sum_{p\in\mathcal{P}_0}\mathbf{1}_{p|k} - v\right)^2 \le Xv + 2v|\mathcal{P}_0| \le 2Xv.$$

The floor errors in the first moment cost at most $2v|\mathcal{P}_0|$; the pairwise intersections of different primes have counts at most $X/(pp')$. The final inequality holds since $X \geq y^3$. The number missing every prime is therefore at most $2X/v$. Markov's inequality for $\log s_y$ then allows a fixed large J such that a fixed positive fraction of X remains in T with $y^\nu \leq s_y(k) \leq y^J$.

Partition by the unique rough part $j = k/s_y(k)$, writing $z = X/j$ for each fiber. Since $k \geq (m/N)X$, a nonempty fiber satisfies $y^\nu \leq z \leq L^3 y^J$. For fixed U, the count of y-smooth integers up to z is $\gg_U z$ uniformly for $1 \ll z \leq y^U$. One can use the fixed-parameter smooth-number estimate [7, Theorem 1.1]; alternatively, for $z > y$ multiply smooth integers in $[1, z/p]$ by primes in $[y^{1/4}, y^{1/2}]$ and induct on U in steps of $1/4$. Each integer is counted at most $4U$ times, and the prime reciprocal sum tends to $\log 2$. The base case $z \leq y$ is immediate. Use $U = J + 3$ (with $d \geq 1$). Uniqueness of the rough parts implies $\sum_j z_j \ll_J X$, since the full smooth fibers give disjoint subsets of $[1, X]$. Thus one of the fibers has density at least $\gamma > 0$, depending only on the original fixed density, J, ν, not on d. It yields $d_s = sj$.

For the coarse regime use $y = e^{L^{4/5}}$. By the upper-bound sieve the integers lacking every prime from $[e^{L^{7/10}}, y]$ have count $O(XL^{-1/10})$: the product of the sieve factors has this order by Mertens, and $X \geq e^{12L^{9/10}}$ permits the sieve. The logarithmic moment eliminates $O(XL^{-1/10})$ integers with smooth part greater than $e^{L^{9/10}}$. These losses are $o(L^{-7/100}X)$. Partition the remaining set by rough parts. The sum of reciprocals of possible rough parts is bounded by the Euler product $\prod_{y<p\leq X}(1 - 1/p)^{-1} \ll L^{1/5}$. Hence $\sum_j z_j \ll XL^{1/5}$ and one fiber has density $\gg L^{-7/100-1/5} = L^{-27/100}$. Its asserted range follows from $k \geq L^{-1/100}X$.

For synchronization in the fine regime use primes in $[L^2, L^{d\nu/4}]$. If a set of density at least γ lacks every prime in a subcollection with reciprocal sum v, counting multiples yields

$$\sum_{n\leq z}\left(\sum_{p\in\mathcal{P}_0}\mathbf{1}_{p|n} - v\right)^2 \leq zv + 2v|\mathcal{P}_0| \leq 2zv.$$

Indeed $|\mathcal{P}_0| \leq z^{1/4}$, so this proves $v \leq 2/\gamma$. The total prime reciprocal sum is $\log(d\nu/8) + o(1)$. Choose d large after fixing γ, ν, J, so that this exceeds $4/\gamma + 2$; the fine bound $B > 4d$ ensures $X \geq y^3$ in the earlier variance argument. The common primes have reciprocal sum at least one and are all at least L^2, so their product exceeds N.

In the coarse regime use primes in $[e^{L^{1/20}}, e^{L^{3/5}}]$, with reciprocal sum $\frac{11}{20}\lambda + o(1)$. The upper-bound sieve, applicable since $z \geq e^{L^{7/10}}$, implies that the missing primes for any fiber have reciprocal sum at most $(27/100)\lambda + O(1)$. Thus the common primes have reciprocal sum at least $\lambda/100 - O(1)$, and their product exceeds N. In both regimes the construction and bounds are uniform after the stated small deletions. □

3 Weighted exact completion

Lemma 3.1 (One- and two-dimensional completion)**.** *In either regime above, give each $n \in P$ an independent Bernoulli variable X_n of parameter $p_n \in [L^{-4}, 1 - L^{-4}]$. Set $Z = \sum X_n/n$, $T = \sum X_n$. If $Q\mathbb{E}Z \in \mathbb{Z}$, then*

$$\mathbb{P}(Z = \mathbb{E}Z) \sim \frac{1}{Q\sqrt{2\pi \operatorname{Var} Z}}. \tag{9}$$

If moreover $|P| \geq (1/2 + \beta)N$ for a fixed $\beta > 0$ and $\mathbb{E}T \in \mathbb{Z}$, then, with $\Sigma = \operatorname{Cov}(Z, T)$,

$$\mathbb{P}(Z = \mathbb{E}Z, T = \mathbb{E}T) \sim \frac{1}{2\pi Q\sqrt{\det \Sigma}}. \tag{10}$$

The conclusions are uniform for the stated parameter regimes. In the fine case the fixed B may be chosen large enough in terms of the fixed densities (and, for the two-dimensional assertion, β), independently of the probabilities. In the coarse case the displayed fixed exponents suffice.

Proof. We give details for the two-dimensional case. For the one-dimensional assertion set $b' = 1$ and $\theta = 0$ throughout, omit the good-phase density test and rational approximation below, and begin directly with the bad-power classification. The size bound in either regime suffices for the one-dimensional variance estimate. Write $e(t) = e^{2\pi i t}$, $\|t\| = \operatorname{dist}(t, \mathbb{Z})$ and $\phi(\xi, \theta) = \mathbb{E}e(\xi(Z - \mathbb{E}Z) + \theta(T - \mathbb{E}T))$. From the Bernoulli factor bound,

$$|\phi(\xi, \theta)| \le \exp[-cL^{-4} \sum_{n \in P} \|\xi/n + \theta\|^2]. \tag{11}$$

Frequencies range over $[-Q/2, Q/2] \times [-1/2, 1/2]$. Choose, in the fine case, fixed $w > 3 + dJ + 2$, $\zeta > w + 3 + dJ + 2$ and $B > \max(4d, 2w + 12, 2\zeta + 6)$, increasing B if needed. Use

$$W = N/L^w, \quad \Delta = L^{-\zeta}, \quad T_* = L^{2w+10}.$$

In the coarse case use $W = Ne^{-2L^{9/10}}$, $\Delta = e^{-4L^{9/10}}$, $T_* = e^{5L^{9/10}}$. Round thresholds if necessary. In both regimes W is smaller than all carriers in Lemma 2.1; $T_* = o(\eta N/Y)$ with $\eta = L^{-7/100}$ in the coarse case. Also $NL^{-4}\Delta^2 \gg Y + L$ (with ratio to $Y + L$ tending to infinity).

If more than $\beta N/2$ of the denominators have $\|\xi/n + \theta\| \ge \Delta$, their entire integral contribution is negligible, by (11) and $\log Q \ll Y$. Otherwise the good-phase subset has cardinality at least $(1/2 + \beta/2)N$. Apply the carrier construction for $s = 1$ to this subset. In its fiber there are two cofactors $u < v$ with difference $O(1)$ in the fine case or $O(L^{27/100})$ in the coarse case, by ordering its elements and averaging gaps. The identity

$$v(\xi/(dv) + \theta) - u(\xi/(du) + \theta) = (v - u)\theta$$

shows that θ is near a reduced rational a'/b' with $b' \ll 1$ in the fine case, $b' \le L$ for large N in the coarse case, and

$$|b'\theta - a'| = o(W/N). \tag{12}$$

In the fine case this follows from $u, v \le L^{3+dJ}$ and the choice of ζ; in the coarse case the error is $O(L^{1/100}e^{-3L^{9/10}})$. There are $O(L^2)$ possible reduced rationals within the allowed frequency interval. We can integrate over their neighborhoods with overcounting.

Put $v' = b'\xi$. Call a prime power s bad at v' if at least T_* denominators divisible by s have distance from v' to their multiples at least $W/2$. For every such denominator, (12) implies $\|\xi/n + \theta\| \ge W/(4b'N)$ for sufficiently large N. If there are j bad powers, their witnesses count at least $jT_*/(20\lambda)$ distinct denominators. Hence

$$|\phi(\xi, \theta)| \le \exp\left(-\frac{cjT_*L^{-4}W^2}{\lambda(b')^2N^2}\right) \le N^{-20j}. \tag{13}$$

The fine-regime bound uses bounded b'; the coarse-regime exponent is exponentially larger than jL even with $b' \le L$. For a good power, remove its fewer than T_* witnesses and apply Lemma 2.1. Each $d_s u$ in its fiber has a multiple in $(v' - W/2, v' + W/2)$. Since $d_s > W$, all these multiples are the same integer x_s. The common primes from Lemma 2.1 divide $x_s - x_{s'}$; their product exceeds $N > W$, forcing equality. Thus, writing d' for the lcm of the good powers (or 1 if there are none), v' lies within $W/2$ of a multiple of d'. Modulo Q this has measure at most $(W + 1)Q/d'$, and

$Q/d' \le N^j$. The v' integration runs over b' periods and the change of variables cancels that factor. With at most N^j choices of bad powers, all contributions with $j \ge 1$, including the $O(L^2)$ rational choices, are $O(L^2 \sum_{j\ge1} N^{1-18j})$. This also proves the analogous one-dimensional bound by setting $b' = 1, \theta = 0$.

If $j = 0$, then $v' = hQ + O(W)$ for an integer h. For every good-phase denominator, (12) and $W/m = o(1)$ imply

$$hQ/n + a' \equiv 0 \pmod{b'}.$$

Indeed, after multiplying the distance by b', the error is bounded by $b'\Delta + O(W/m) + |b'\theta - a'| = o(1)$. The numerator $hQ/n + a'$ is an integer since $n \mid Q$, so the stated congruence follows. Since a', b' are coprime, h is invertible modulo b'. If a prime $\ell \mid b'$ also divides Q, then every good-phase denominator must have valuation $v_\ell(n) = v_\ell(Q)$, and there are at most $N/2$ of them. If $\ell \nmid Q$, multiplication by n shows instead that all such denominators lie in one nonzero residue class modulo ℓ, of size at most $N/2 + 1$. Both contradict their density. Thus $b' = 1$, $\theta = o(W/N)$, and ξ is within $W/2$ of zero in the chosen period. In the one-dimensional case this last conclusion follows immediately from $j = 0$.

On this central region all unwrapped phases are small. In two dimensions the normalized covariance $\operatorname{diag}(N, 1)\Sigma \operatorname{diag}(N, 1)$ is $\succeq cNL^{-4}I_2$: a positive-density set has two fixed positive-size subcollections in a fixed proportion of $[1, N]$ with separated denominators. For example, discard $n < cN$ for small fixed c, then take the lowest and highest fixed fractions of the remaining denominators. The corresponding vectors $(N/n, 1)$ have uniformly separated first coordinates and bounded norm. In one dimension $\operatorname{Var}(Z) \gg |P|/(N^2L^4)$. These bounds give Gaussian decay outside the box $|\xi| \le m^{3/5}$, $|\theta| \le m^{3/5}/N$, with superpolynomially small error. In that box the uniform centered Bernoulli expansion has total cubic error $O(Nm^{-6/5}) = o(1)$. Therefore ϕ equals the centered Gaussian characteristic function times $1 + o(1)$ uniformly there. The Gaussian integral is bounded below by a negative power of L in dimension two, using the upper covariance bound from $m \le n \le N$, and is likewise not small in dimension one. The preceding absolute errors are consequently negligible relative to the full Gaussian integral.

Finally, use continuous inversion for Z on its lattice $Q^{-1}\mathbb{Z}$ and the usual integer inversion for T. At the integral mean the exact probability is $1/Q$ times the integral of ϕ over the displayed rectangle. Gaussian integration proves (10) and its one-dimensional counterpart. This is inversion on the actual lattice, not a congruence modulo one followed by an unverified concentration argument. □

Remark 3.2. The density assumption in two dimensions is important. For an all-odd pool, $QR(S) \equiv |S| \pmod 2$. One-dimensional lcm divisibility alone cannot ensure arbitrary cardinalities. At density strictly above $1/2$, the rational resonance argument eliminates this extra obstruction. At near-full density one can alternatively use the many surviving $(2j, 3j)$ pairs, whose phases satisfy $3\phi_{3j} - 2\phi_{2j} = \theta$.

4 Sparse modular covering at a second scale

Lemma 4.1 (Sparse rational grid)**.** *Let $U \subseteq [1, X]$ have size at least ϵX for a fixed $\epsilon > 0$, and suppose every element has a prime divisor exceeding $z \ge 2$. For either fixed choice $\vartheta \in \{1/100, 1/20\}$ (with $1/100$ the default in this paper), there is a set $C \subseteq U \cap [c_\epsilon X, X]$ such that, with $\ell = \log X$ and $Q_C = \operatorname{lcm}(C)$,*

$$|C| \asymp X\ell^{-\vartheta}, \qquad R(C) \asymp_\epsilon \ell^{-\vartheta}, \qquad \log Q_C \gg z \log z\, \ell^{-\vartheta}, \tag{14}$$

$$[R(C)/3, 2R(C)/3] \cap Q_C^{-1}\mathbb{Z} \subseteq \{R(E) : E \subseteq C\}. \tag{15}$$

All conclusions hold uniformly for large X (depending on ϵ).

Proof. Discard $n < \epsilon X/4$, and choose $\left\lfloor X\ell^{-\vartheta} \right\rfloor$ of the remaining integers. Remove all integers with a prime-power divisor greater than $Xe^{-12\ell^{9/10}}$, or with $\Omega(n) > 20\log\ell$. The count removed is $O(X\ell^{-1/10}) + o(X\ell^{-1/10})$ by (7)–(8). Repeatedly remove all multiples of a power s whose surviving nonempty pool has size less than $\ell^{-7/100}X/s$. Each power triggers at most once, with total count $O(X\ell^{-7/100}\log\ell) = o(X\ell^{-\vartheta})$. For both choices of ϑ the loss is $o(X\ell^{-\vartheta})$, the retained size is at least $X\ell^{-3/50}$ for large X, and the fixed lower endpoint $\epsilon X/4$ exceeds $X/\ell^{1/100}$ eventually. This leaves a coarse-regime pool for Lemma 3.1 at scale X. Its size and reciprocal mass have the asserted orders. Each prime exceeding z divides at most X/z of its elements, so the lcm contains at least $\gg z\ell^{-\vartheta}$ such distinct primes. This proves (14). For any grid target in (15) use the uniform inclusion probability $R_{\text{target}}/R(C) \in [1/3, 2/3]$ and the one-dimensional part of Lemma 3.1. □

In particular, if $p > X$, $\log Q_C \gg \log p$ with ratio tending to infinity, and $\log X = O(\log\log p)$, the consecutive grid numerators in (15) span every residue class modulo p, and also any additional fixed coprime modulus. More generally we may prescribe a numerator modulo an integer d whenever $Q_C R(C)/d \to \infty$. We shall also use this with a composite power d; this is just counting consecutive integers, not a finite-field estimate for a composite modulus. This proves the precise sparse covering required below without invoking a cardinality-constrained version of [5, Proposition 5.2].

5 First-order phase transitions and stability

Let $N = M$ throughout, with $N \sim DK$ for a fixed $D > 0$. The initial-segment deletion $[2, h+1]$ leaves at most the mass of the first K available reciprocals, which is less than $\log(1 + K/(h+1))$. Taking $h = \lfloor AK \rfloor$ (intersecting with the vertex set if necessary) hits all edges. Deleting all multiples of q prevents the lcm from being divisible by b, and costs $\lfloor N/q \rfloor$. Finally, deleting the largest h denominators forces every exact K-term sum to be at least $\sum_{n=N-h-K+1}^{N-h} 1/n$, provided K terms remain. The cost for making this exceed r is $[N - D_*K]_+ + O_r(1)$; if fewer than K terms remain the obstruction is automatic. These give the three first-order upper bounds. If $D < D_*$ there are eventually no exact K-term edges, by the same capacity comparison without deletions.

For lower bounds we record a coarse forcing test. Put $\rho = L^{-1/100}$ and ignore $n < \rho N$. The deletion and iterative power-pruning procedure in the proof of Lemma 4.1, without its preliminary downsampling, at scale N, removes $O(NL^{-1/10} + NL^{-7/100}\lambda) = o(\rho^2 N)$ additional vertices, hence reciprocal mass $o(\rho)$. Here we used $1/n \le 1/(\rho N)$ after the initial cutoff. In particular any fixed positive fraction of surviving multiples of a target prime power can be retained through this pruning. The following two tests are useful (the margins in them are fixed positive constants, and the cutoff is implicit):

- If the largest at most K available reciprocals sum to more than r by a fixed margin, and each highest prime power of b has $\gg N$ surviving multiples, reserve $\lfloor \rho N \rfloor$ multiples for each such power and fill a pool of at most K elements with the largest remaining reciprocals. To justify the capacity loss near very small denominators, the reserves can replace at most $O_r(\rho N)$ of the *smallest weights* in the unconstrained top-K pool, so their discarded mass is $O_r(\rho)$ when the pool has $\gg N$ elements. If the full available pool has fewer than K elements, simply include it. The initial cutoff itself removes $o(N)$ points. Whenever invoking this test below the stated fixed capacity margin holds after that cutoff. The pool has size $\gg N$, its pruning loses $o(\rho)$ mass, and each reserved class survives. The lcm is divisible by b. Set the Bernoulli mean reciprocal sum equal to r by a common parameter; the mass is $O(\log L)$, so the probability lies

in $[L^{-4}, 1 - L^{-4}]$ for large N. Apply the one-dimensional coarse completion lemma. Capacity is guaranteed by the size of the pool.

- For exactly K terms, retain the whole available pool after the cutoff and prune it. Suppose the largest and the smallest K reciprocal sums lie on opposite sides of r with fixed margins, every target power has $\gg N$ surviving multiples, and the available density is at least $1/2 + \beta$ with fixed $\beta > 0$. Then the pruned pool still has these properties (up to harmless smaller margins), provided the cutoff and pruning do not change the endpoints by more than $o(1)$. In the applications, removing $o(N)$ points only increases the minimum and decreases the maximum by at most $o(1)$ after the initial cutoff, or the endpoints are checked directly by rank comparisons. Set $p_n = L^{-4} + (1 - 2L^{-4})z_n$ for $z_n \in [0, 1]$. The required fractional cardinality differs from K by $O(NL^{-4})$, and the required reciprocal sum from r by $O(L^{-4}R(P))$. The endpoint changes from this fractional cardinality shift are $O(L^{-4}/\rho) = o(1)$. The two target means lie within the interval of achievable weighted masses at that cardinality, by the extremal weight allocations and convexity. Apply the two-dimensional coarse completion lemma.

For clarity, these tests can always be applied with the top and bottom sums evaluated *after* pruning, using the uniform rank estimates below; no bound on the reciprocal mass of the initial cutoff is needed.

Write $h = |F|$. If h is less than $\min(A, D/q)K$ by a fixed multiple of K, then the top available reciprocal sum is larger than r by a fixed margin after the cutoff. If there are at least K available denominators this follows from comparison with deleting the first $h + O(\rho N)$ integers; otherwise the sum of the entire available set is at least $\log(N/(h + O(\rho N))) + o(1) \geq \log q - o(1) > r$. Every target power $s \mid b$ has at least $N/s - h - O(\rho N) \gg N$ multiples. This proves the first lower bound. For exact length and $D > D_*$, the strict inequality $h < \min(A, D/q, D - D_*)K - \epsilon K$ ensures that at least K denominators remain, and both capacities have fixed margins by the initial- and terminal-segment comparisons. The largest possible first-order cost is at most $(D-1)K/2$, since $\min(A, D - D_*) \leq (D-1)/2$. Thus the remaining density exceeds $1/2$ by a fixed amount. The second test applies. The same rank comparisons after pruning remain valid, since the total cutoff and pruning count is $o(N)$. This proves the limits in Theorem 1.1. At $D = D_*$ the upper bound gives $o(K)$ and hence the asserted limit.

Here is the stability argument in a form that also applies when D is approached along a sequence. Suppose a sequence of transversals has the first-order minimum cost plus $o(K)$. Pass to a subsequence on which each normalized distance to a template has a limit. If no minimizing template is at distance $o(N)$, every target power has $\gg N$ surviving multiples: failure would force $h \geq N/s - o(N)$, and consequently $s = q$ and the multiples template would be minimizing and at distance $o(N)$. For a pool with at least K survivors, a displacement of $2d$ from the initial-segment deletion of the same size increases its top-K mass by at least $\min(d, K)^2/N^2$. To see this, even the least favorable $\min(d, K)$ restored low denominators and displaced larger ones give a sum of reciprocal gaps at least $\sum_{j \leq \min(d,K)}(2j-1)/N^2$. The corresponding terminal-segment displacement decreases the bottom-K mass by the same amount. Both comparisons can be made after adding the initial cutoff and the coarse pruning to the deleted set, changing its size and template distances by $o(N)$. If fewer than K survivors remain in the at-most case, the earlier $\log q > r$ margin suffices. Thus a nonminimizing template has a fixed margin automatically, and a minimizing but distant template yields a fixed margin by the displacement estimate. Both forcing tests now contradict transversality. This proves stability and finishes Theorem 1.1.

6 Large-prime attacks and the cost of repairing them

This section gives a deletion ledger for both the onset and the macroscopic second-order analysis. In the onset application $0 < r \le 1$; all mentions of a divisibility template and its prime power q apply only to the proper-target macroscopic results. When $b = 1$, the target-divisibility requirement is automatic. Continue to write $L = \log N$, $\lambda = \log L$, $\delta = \lambda/L$. The parameters B and all positive auxiliary density constants in this section are fixed before N tends to infinity. Define

$$Y = N/L^B, \qquad v_p = \log(N/p)/\lambda \quad (p > Y).$$

All these large primes exceed $\sqrt{N}$ and none divides b when N is sufficiently large. Their multiples form disjoint pools. We shall repeatedly use the scale integration rule

$$\sum_{p>Y}^{p\le N} \frac{g(v_p)}{p} = \delta \int_0^B g(v)\,dv + o(\delta) \tag{16}$$

for bounded, fixed piecewise continuously differentiable functions g (and interval indicators). Indeed Mertens and partial summation give $\int_0^B \lambda g(v)/(L-\lambda v)\,dv$ with error $O_g(1/L)$; the replacement of the integrand denominator by L costs $O_B(\delta^2)$. Floors of N/p introduce total error $O(\pi(N)) = o(N\delta)$.

6.1 An explicit attack

Put $f(v) = 0$ for $v \le 1$, $f(v) = \log v$ for $1 < v \le 2$, and $f(v) = \log 2$ for $v > 2$. The last part will only be used as a lower bound on repair failure cost. For $1 < v < 2$, attack every prime $p > N/L^v$ by deleting its multiples $pm \le N$ for which m has a prime-power divisor exceeding $c_1 L$, with $c_1 > 0$ sufficiently small. Call the actual deleted set E_v and the union of the entire attacked multiple pools V_v. If an exact sum r avoiding E_v used a nonempty set of multiples of such p, their cofactor reciprocal sum would vanish modulo p, since $p^2 > N$ and $p \nmid b$. All its cofactors divide $Q_0 = \operatorname{lcm}(1, \dots, \lfloor c_1 L \rfloor)$. But the numerator obtained with denominator Q_0 is positive and less than p: indeed $Q_0(1 + \log L^v) < p$ by Chebyshev and small c_1. This is impossible. Thus every exact representation avoiding E_v avoids V_v.

For $x = N/p \le L^v$, distinct primes above $c_1 L$ have product greater than x, for large N. Counting their multiples, and separately bounding larger higher powers using (8), shows the cost in the p-pool is $xf(v_p) + O_v(x/\lambda)$, uniformly (it is zero for $x < c_1 L$). Here the prime count error $\pi(x)$ is $O(x/\lambda)$ for $x \ge c_1 L$ and the shift by the fixed c_1 in the Mertens sum has the same size. It follows from (16) that

$$|E_v| = N\delta J(v) + o(N\delta), \qquad |V_v| = vN\delta + o(N\delta). \tag{17}$$

Uniformly in $0 \le X \le N$, $|V_v \cap [1, X]| = v\delta X + o(N\delta)$, since the count of multiples of each prime differs from X/p by at most one. For fixed q, the same formula for the intersection with $q\mathbb{N}$ has density $v\delta/q$. Also the primes $p > N/(c_1 L)$ give a *free* attack of density $(1 + o(1))\delta$: their entire pools are unusable without any deletions. A bounded Lipschitz weight w has total weight over a p-pool equal to $(N/p)\int_0^1 w(t)dt + O_w(1)$, a direct Riemann sum estimate. Thus (17) applies to weighted benefits as well.

6.2 Marking damaged small powers

Exclude an exceptional set E_0 consisting of $n < N/L^3$, integers with a higher prime-power divisor $p^j > Y$ $(j \ge 2)$, and integers with $\Omega(n) > 20\lambda$. Then

$$|E_0| = o(N\delta) \tag{18}$$

by (8). Let G be the complement of E_0, all large-prime multiples and the actual deletion set F. In a divisibility-template argument only, also exclude $q\mathbb{N}$ from G, temporarily; call a power allowed if it is not divisible by q in that case, and otherwise allow all powers at most Y. Mark an allowed power $s \le Y$ as bad if $|G \cap s\mathbb{N}| < \eta N/s$ for a fixed sufficiently small $\eta > 0$. Include zero-size pools in this definition. For each base prime retain only its lowest marked power, forming the pairwise coprime set $\mathcal{B}$. Let

$$T_s = \sum_{s\in\mathcal{B}} 1/s, \qquad V_s = \bigcup_{s\in\mathcal{B}} (s\mathbb{N} \cap [1, N]), \qquad P = G \setminus V_s.$$

(The notation T_s here denotes a scalar total, not a separate quantity indexed by s.) We work either with $|F| = O(N\delta)$ (onset), or F at distance $o(N)$ from a minimizing macroscopic template F_0.

We claim $T_s = O(\delta)$ at onset and $T_s = o(1)$ for the macroscopic templates. For this purpose one must not assume uniform small exceptional losses for each power separately. Instead put $g(n) = \sum_{s\in\mathcal{B}} \mathbf{1}_{s|n}$. Coprimality gives

$$\sum_{n\le N} g(n)^2 \le N(T_s + T_s^2).$$

In the onset case set $F_0 = \varnothing$. Before excluding E_0 or the large primes, the surviving multiples of each allowed s in the complement of F_0 have count at least $c_0 N/s - O(1)$ with fixed $c_0 > 0$: this holds for both segment templates, and for a multiples template it follows by subtracting the multiples of $\mathrm{lcm}(s, q)$, with ratio at most $1/2$ if $q \nmid s$. At onset take $c_0 = 1$. In the multiples-template case, excluding $q\mathbb{N}$ is already part of this comparison. Since all bad powers have fewer than $\eta N/s$ actual survivors, summing the lost multiples gives

$$(c_0 - \eta - o(1))NT_s \le \sum_{n\in(F\setminus F_0)\cup E_0} g(n).$$

Here the total large-prime overlap is at most $NT_s \sum_{p>Y} 1/p = O_B(NT_s\delta)$ (the primes are coprime to all s); the floor errors are $O(|\mathcal{B}|) \le YT_s = o(NT_s)$. Cauchy–Schwarz and (18) now show $T_s \ll (|F \setminus F_0| + |E_0|)/N$, because this fraction is $o(1)$ and $\eta < c_0/2$. This proves the claim. This argument accounts even for powers all of whose multiples are exceptional.

Every allowed power s' not divisible by a bad power has at least $\eta N/s'$ multiples in G. Intersections with other-base bad powers remove at most $(N/s')T_s = o(N/s')$; the one possible larger same-base bad power removes fewer than $\eta N/(2s')$. Thus

$$|P \cap s'\mathbb{N}| \ge \eta N/(3s'). \tag{19}$$

In particular all allowed nonexcluded powers occur in the lcm $Q = \mathrm{lcm}(P)$, not merely the powers already known to occur. For every fixed power allowed in this comparison, there remain a positive proportion of its multiples by the template comparison, (18), and the $o(N)$ extra deletions. Consequently all fixed target powers occur in Q, except powers divisible by q in the multiples-template case. There $q = \ell^a$ implies $v_\ell(Q) = a - 1$, and the other highest target powers divide Q. At onset $|P| = N - O_B(N\delta)$ and in the macroscopic case it is within $o(N)$ of the template complement. Thus in all cases P satisfies the fine one-dimensional completion hypotheses; in an active exact-length phase it also satisfies the density condition of Lemma 3.1. For the multiples-template phase this uses $D/q \le (D-1)/2$, hence $1 - 1/q > 1/2$ by a fixed margin whenever that phase is active.

6.3 Large-prime failure versus sparse repair

Choose small fixed $\gamma, \epsilon > 0$, and for each $p > Y$, put $x = N/p$ and $z_p = \max(L^{1+\gamma}, \sqrt{x})$. Classify its pool as healthy if at least ϵx cofactors m have $pm \notin F \cup E_0 \cup V_s$ and have a prime divisor greater than z_p. In the multiples-template case use this same test, including surviving q-multiples, rather than the restricted pool G. Denote the set of failed large primes by $\mathcal{D}$, and put

$$V_p = \bigcup_{p \in \mathcal{D}} (p\mathbb{N} \cap [1, N]), \qquad V = V_s \cup V_p, \qquad \omega = T_s + \sum_{p \in \mathcal{D}} 1/p.$$

In a healthy pool Lemma 4.1 supplies a reserve C_p of size $O(x(\log x)^{-1/100})$, with all cofactors at least $c_\epsilon x$. Its lcm Q_p satisfies $\log Q_p \gg z_p \log z_p (\log x)^{-1/100} \gg L$ with ratio tending to infinity, as $x \le L^B$ and $z_p \ge L^{1+\gamma}$. Thus it covers all residues modulo p, including after excluding a specified single seed if necessary, by constructing the reserve from the remaining fixed-density set. It can correct any chosen base subset disjoint from that reserve, by adding a subset whose reciprocal sum cancels the residue modulo p. The resulting block has no p in its reduced denominator. In the nondivisibility cases all its other prime-power divisors occur in Q by (19): every cofactor of an available pm avoiding E_0, V_s has its prime powers at most Y and not divisible by any bad power. This includes the cofactor 1. No assumption of lcm equality between the local reserve and the global pool is made.

The reserves in distinct large-prime pools are disjoint from each other and from P. Since every healthy pool has $x \ge L^{1+\gamma}$, their total cardinality and reciprocal mass in the original scale are bounded by

$$\sum_{p \text{ healthy}} |C_p| \ll_{\epsilon, B} N\delta\lambda^{-1/100} = o(N\delta), \qquad \sum_{p \text{ healthy}} R(pC_p) = o(\delta). \tag{20}$$

The second estimate uses $pm \ge c_\epsilon N$ for each reserved denominator. This is why a bounded harmonic repair cost alone would not suffice for the present problem.

For a failed p-pool, count all cofactors with a prime factor greater than $z_p \ge \sqrt{x}$. There is no overlap between distinct such primes. Mertens and floor errors show this count is at least $x(f(v_p) - O(\gamma) - o(1))$ uniformly; if the lower bound is negative it is trivial. For $v_p > 2$, the bound $\log 2$ follows from primes above $\sqrt{x}$, or loses at most $O(\gamma)$ if the other cutoff is larger. The number remaining is less than ϵx. In charging actual deletions exclude V_s and charge exceptional losses separately. Using the disjointness of the large pools, (18), and $|V_s \cap \bigcup_{p>Y} p\mathbb{N}| \le O_B(NT_s\delta) = o(N\delta)$, we obtain

$$|F \cap (V_p \setminus V_s)| \ge \sum_{p \in \mathcal{D}} (N/p) f(v_p) - O_B((\gamma + \epsilon)N\delta) - o(N\delta). \tag{21}$$

Floors of counts, including when x is small, are harmless by $O(\pi(N)) = o(N\delta)$; all uses of Mertens for positive f take place at $x \gg L$. By pairwise coprimality of $\mathcal{B}$ and disjointness of the large pools, uniformly in $0 \le X \le N$,

$$|V \cap [1, X]| = \omega X + o(N(T_s + \delta)). \tag{22}$$

Indeed the errors from small-small overlaps are $O(NT_s^2)$, from large-small overlaps $O_B(NT_s\delta)$, and from floors $O(YT_s + \pi(N))$. In the multiples-template case each bad power is coprime to q: an allowed power of the base of q is fixed and has already been shown not bad. Thus $|V \cap q\mathbb{N}| \le (N/q)\omega$ in that case.

6.4 A common cost ledger

In either segment case or at onset, there are fewer than ηNT_s surviving nonexcluded small-power multiples in G. Summing over the marked powers, the union bound and (18) give

$$|F \cap V_s| \geq (1 - \eta - o(1))NT_s - o(N\delta). \tag{23}$$

Indeed $|V_s| = NT_s + O(NT_s^2 + YT_s)$, and the exceptional and large-prime exclusions in V_s have count at most $|E_0| + O_B(NT_s\delta)$. In the multiples-template case we shall prove below, for a transversal, the analogous inequality with 2η in place of η. Given either such inequality with $j\eta$ ($j = 1$ or 2), for any fixed $0 < \alpha' < \log 2$ one has

$$\begin{aligned} |F \cap V| - \alpha' N\omega \ &\geq (1 - j\eta - \alpha' - o(1))NT_s \\ &\quad - N\delta(e^{\alpha'} - 1) - O_B((\gamma + \epsilon)N\delta) - o(N\delta). \end{aligned} \tag{24}$$

To verify this, sum (23) (or its $j = 2$ version) and (21). Optimize the negative contribution by including only the positive values of $\alpha' - f(v_p)$, then apply (16). The integral is

$$\int_0^B (\alpha' - f(v))_+ dv = \alpha' + \int_1^{e^{\alpha'}} (\alpha' - \log v)dv = e^{\alpha'} - 1. \tag{25}$$

We always take $B > 2$. The term $(1 - j\eta - \alpha')NT_s$ is strictly positive if η is chosen small. In particular, the error $o(NT_s)$ in the ledger must be absorbed by this term, not treated as $o(N\delta)$ without justification.

7 Macroscopic second-order bounds

We prove Theorem 1.2. All quantities D, u are fixed in the limit. Write $N = \lfloor DK + uK\delta_K \rfloor$, so $N\delta = (D + o(1))K\delta_K$ and $\delta \sim \delta_K$.

7.1 Upper bounds and comparison of scales

After the attack E_v in Section 6.1, the effective vertices have density $d = 1 - v\delta$, with a uniform rank error $o(N\delta)$. On any interval bounded away from zero on the N scale, partial summation gives the corresponding reciprocal sums with error $o(\delta)$. In particular:

- If the initial-segment branch is active, delete the smallest remaining h vertices, with $h = K/(e^{r/d} - 1) + o(N\delta)$, increasing the lower-order adjustment if necessary to make the maximal capacity strictly less than r. The corresponding maximal sum is $d \log(1 + K/h) + o(\delta)$. There are enough denominators at that rank: $D \geq qA > A + 1$. Thus the total attack cost is at most

$$AK - vcK\delta + N\delta J(v) + o(N\delta).$$

- Delete the remaining multiples of q, at cost $(N/q)(1 - v\delta) + o(N\delta)$, in addition to E_v. This yields total cost at most $N/q - v(N/q)\delta + N\delta J(v) + o(N\delta)$.

- For exact length and active terminal branch, delete the largest remaining h vertices. The minimal K-term mass is $d \log((dN - h)/(dN - h - K)) + o(\delta)$, at $h = N - D_*K + O(N\delta)$. Its critical value occurs at $h = dN - K/(1 - e^{-r/d})$. This is positive to first order since $D > D_*$. The total cost is at most

$$N - D_*K - v(N - cK)\delta + N\delta J(v) + o(N\delta).$$

Here the rank and mass comparisons follow by locating the boundary ranks to accuracy $o(N\delta)$ and applying partial summation; increasing h by $\epsilon' N\delta$ produces a strict obstruction, and one may let $\epsilon' \downarrow 0$ after taking the large-N limit. This also avoids treating a capacity equality as a strict hitting condition.

The relevant benefit parameters are $\alpha_1 = c/D$, $\alpha_2 = 1/q$, and $\alpha_3 = 1 - c/D$. Each active parameter is in $(0, \log 2)$. For the first branch, $D \geq qA$ implies $c/D \leq r/(q(1-e^{-r}))$; for $q \geq 3$ this is at most $1/(3(1-e^{-1})) < \log 2$, and for $q = 2$ we use $r = 1/2$ directly. For the third branch, $D_* < D \leq 2A + 1$ and $c < D_*$, so

$$0 < 1 - c/D \leq 1 - r/(2\sinh r) < \log 2.$$

(The last inequality follows, for instance, from monotonicity of $r/\sinh r$ and its value at 1.) The second parameter is at most $1/2$. Therefore we may optimize over $1 < v < 2$, and the maximum of $\alpha_i v - J(v)$ is $e^{\alpha_i} - 1$, attained at $v = e^{\alpha_i}$. Substituting N proves the upper bound in (3).

7.2 Capacity forcing near the segment templates

By the first-order bounds and stability, a sequence of minimum transversals approaches, on subsequences, one of its minimizing templates. For a segment template use the unrestricted construction of Section 6; all target powers divide Q. Put $h = |F|$, $h' = h - |F \cap V|$, and $d = 1 - \omega$. We claim respectively for the initial and terminal templates that a transversal must satisfy

$$h \geq AK + |F \cap V| - cK\omega - o(N(T_s + \delta)), \tag{26}$$

$$h \geq N - D_*K + |F \cap V| - (N - cK)\omega - o(N(T_s + \delta)). \tag{27}$$

Here and below a necessary inequality with a little-oh error means that for each fixed positive error tolerance it holds for all sufficiently large N, with the fixed auxiliary constants as specified. We give the details of the forcing, including the cardinality constraint.

By (22), the effective rank function, before the h' additional deletions, differs from dX by $o(N(T_s + \delta))$. Excluding E_0 and the healthy large-prime reserves costs $o(N\delta)$ additional ranks. In the initial-template branch the top-K available mass is at least

$$d\log(1 + K/h') - o(T_s + \delta),$$

with the analogous bottom-K upper bound $d\log((dN - h')/(dN - h' - K)) + o(T_s + \delta)$ in the terminal branch. To justify these uniform errors, compare with deleting the smallest (respectively largest) h' effective vertices. The cutoff ranks at the two ends of the extremal interval have errors $o(N(T_s + \delta))$, and the associated endpoint denominators are bounded below by a positive multiple of N: $h' \sim AK$ in the initial case, and $dN - h' - K \sim AK$ in the terminal case. In the first comparison the actual top-K set may contain smaller denominators, which only helps the lower bound. Rank errors are converted to mass errors by partial summation on the corresponding comparison interval. The inversion thresholds expand as

$$K/(e^{r/d} - 1) = AK - cK\omega + O(N\omega^2),$$
$$dN - K/(1 - e^{-r/d}) = N - D_*K - (N - cK)\omega + O(N\omega^2).$$

Since $T_s = o(1)$, $\omega^2 = o(T_s + \delta)$. A failure of the asserted necessary inequality by a fixed multiple of $N(T_s + \delta)$ would leave a capacity margin of order $T_s + \delta$, in particular of order δ.

Here is the exact construction from such a margin. Choose a top-K (or bottom-K) set S_0 of remaining effective vertices, avoiding all reserves. For each large prime in S_0, keep its selected

multiples and add a subset of its reserve to cancel that prime. Write C_0 for the original large-prime part and C for its corrected union. By (20), the excess length is $o(N\delta)$ and the added mass $o(\delta)$. Also $|C| = O_B(N\delta)$, and $R(C) = o(1)$, even if S_0 uses some small denominators: the full harmonic mass of all large-prime multiples is bounded by

$$\sum_{p>Y} p^{-1}(1 + \log(N/p)) = O_B(\lambda^2/L) = o(1).$$

The corrected blocks are disjoint from P, and $QR(C) \in \mathbb{Z}$. The remaining target is $r' = r - R(C) > 0$, with $Qr' \in \mathbb{Z}$.

For at most K terms, first remove enough of the largest selected smooth denominators to compensate the added terms of C and leave a count slack of a sufficiently small fixed multiple of $N\delta$. More explicitly, the slack can be chosen to be $\epsilon_0 N\delta$ with ϵ_0 small in terms of the fixed capacity margin, plus $|C| - |C_0|$. Such removals cost $O_r(\epsilon_0\delta) + o(\delta)$ reciprocal mass: since the original top set has $K \asymp N$ elements, a positive fraction of its smooth elements have denominators $\gg N$. Denote the resulting smooth set by S_1. Give every denominator in $P \setminus S_1$ probability L^{-4}, and those of S_1 a common parameter in $[L^{-4}, 1 - L^{-4}]$. The mean reciprocal sum can be set to r' because the extreme means are below and above it (the floor-probability mass is $o(\delta)$). The mean number selected is at most $K - |C| - \frac{1}{2}\epsilon_0 N\delta$. The probability of exceeding $K - |C|$ is $e^{-\Omega(N\delta^2)}$ by Hoeffding's inequality [6], which is $o(Q^{-1})$ since $B > 2$. Lemma 3.1 gives exact reciprocal mass with probability at least a constant times Q^{-1}, so the two requirements intersect. This proves the at-most forcing.

For exactly K terms put $K' = K - |C|$. Remove precisely $|C| - |C_0|$ of the largest denominators from the original smooth part of S_0, leaving K' smooth terms. The removed mass is $o(\delta)$, preserving the required upper-end or lower-end margin for r'. In the initial-template case the opposite (minimum) endpoint has a fixed margin: the full upper tail after deleting the initial segment remains essentially intact, and $D \geq 2A + 1 > D_*$. Thus $\log(D/(D-1)) < r$. In the terminal-template case the opposite (maximum) endpoint has a fixed margin, as a sufficiently low fixed-proportion interval contains fewer than $K/2$ points and mass exceeding $r + 1$; the terminal template leaves it intact up to $o(N)$ holes. These same comparisons apply to the smooth pool P, which differs from the corresponding template complement by $o(N)$, and to $K' = K - o(N)$ and $r' = r - o(1)$. In both cases $K' < |P|$ with fixed linear room.

At fractional size k, denote the extreme sums in P by $L_P(k), U_P(k)$, using linear interpolation between integers. Set $p_n = L^{-4} + (1 - 2L^{-4})z_n$. The required size of z is $(K' - L^{-4}|P|)/(1 - 2L^{-4})$, a shift of $O(NL^{-4})$. Its effect on either endpoint is at most $O(NL^{-4}/m) = O(1/L) = o(\delta)$. The required mass differs from r' by $O(L^{-4}(R(P) + 1)) = o(\delta)$. The endpoints still straddle it, so convexity provides the required $z_n \in [0, 1]$. Apply the two-dimensional fine completion in Lemma 3.1. This produces exactly K' smooth terms, and together with C an exact K-term representation. This contradiction proves (26)–(27).

In (26) or (27), combine the cost ledger (24) with benefit $\alpha' = c/D$ or $1 - c/D$. The difference from the actual ratio cK/N is $O(\delta)$, which contributes only $o(N(T_s + \delta))$ because $T_s = o(1)$. Choose η small enough that the positive NT_s coefficient absorbs every $o(NT_s)$ error. For fixed η, fix the fine parameters and B; then take γ, ϵ sufficiently small, and finally take N large. The resulting segment-branch lower bounds agree with their upper bounds at the $K\delta_K$ scale.

8 The divisibility branch and target-adic completion

For the remaining branch assume F is a transversal at distance $o(N)$ from $q\mathbb{N} \cap [2, N]$ (where $q = \ell^a$ for a prime ℓ). Use the restricted construction in Section 6.2: P is the healthy smooth non-q pool, and $v_\ell(Q) = a - 1$ with all other target powers present. In particular Qr is a rational of reduced denominator ℓ. We need a correction S_* outside P, constructed using surviving q-multiple seeds, with

$$Q(r - R(S_*)) \in \mathbb{Z}, \qquad |S_*| = o(N), \qquad R(S_*) = o(1). \tag{28}$$

For each of the three resources described below, if it remains sufficiently abundant then such a correction exists. All size estimates in this section are for a fixed r, η and fixed other auxiliary constants. A positive fraction of $N\delta$ is more than sufficient for the abundance thresholds involving powers of L.

8.1 A correction in a bad small-power pool

Suppose for some $s = p^j \in \mathcal{B}$ there are at least $\eta N/s$ nonexceptional smooth q-multiples of s outside F. Here $p \neq \ell$ by the observation after (22). The other bad bases exclude at most $(N/s)T_s = o(N/s)$ such denominators. By discarding very high p-valuation relative to j and very high ℓ-valuation (each exclusion can be made smaller than $\eta N/(8s)$ by choosing a fixed exponent increment), and then using the pigeonhole principle, one obtains a subclass with exact valuations $v_p(n) = j'$ and $v_\ell(n) = k \geq a$, where $s' = p^{j'} = O_\eta(s)$ and $k = O_{r,\eta}(1)$. These denominators are $s'm$ with $p \nmid m$, $v_\ell(m) = k$. At scale $X = N/s' \geq L^B$ they yield a positive-density set of cofactors, with density depending only on r, η. There is a fixed $\nu > 0$ such that a positive proportion of these cofactors have a prime factor above X^ν. To see this without a density assertion about the adversary, first discard $m < c'X$ for a sufficiently small fixed $c' > 0$. The sum of $\log s_{X^\nu}(m)$ over $m \leq X$ is $O(\nu X \log X)$, so the count of X^ν-smooth $m \geq c'X$ is $O(\nu X)$ by Markov. Choose ν small in terms of the density. This leaves a positive-density eligible set.

Apply Lemma 4.1 to this set to obtain a reserve C. Fix B in advance sufficiently large that $B\nu > 2$. Then $\log Q_C/L \to \infty$ by (14). The cofactor lcm has the form $Q_C = \ell^k D_C$, with D_C coprime to ℓ and $D_C \mid Q_0$, where $Q = \ell^{a-1}Q_0$. Indeed the base p has been stripped entirely and no other bad base divides any cofactor to its marked exponent; all other powers are allowed and appear in Q. To cancel the extra s' and match the target ℓ-adic part, prescribe a grid numerator $j_0 = s'z_0$ satisfying

$$z_0 Q_0/D_C \equiv \ell^{k-a+1} Qr \pmod{\ell^{k-a+1}}. \tag{29}$$

The right side is an integer, and the coefficient on the left is a unit modulo ℓ. This restricts j_0 to a progression of modulus $s'\ell^{k-a+1} = O_{r,\eta}(N)$, whereas the grid interval contains $\gg Q_C(\log X)^{-1/100}$ consecutive numerators, much larger than this modulus. Pick one by (15), giving $R(E) = j_0/Q_C$ for $E \subseteq C$. For $S_* = s'E$ we then have $R(S_*) = z_0/Q_C$ and (29) implies (28). In addition $|S_*| = O(X(\log X)^{-1/100}) = o(N)$ and $R(S_*) = O((\log X)^{-1/100}/s') = o(1)$. It is disjoint from P because every member is divisible by q.

8.2 Corrections from nonexcluded target multiples

Any nonexceptional smooth q-multiple outside $F \cup V_s$ has $Q/n \bmod 1$ a nonzero element of an ℓ-power-order subgroup of $\mathbb{Q}/\mathbb{Z}$, with order bounded by a fixed power of L (since $\Omega(n) \leq 20\lambda$). Its non-ℓ powers all divide Q. If at least $\epsilon_1 N\delta$ such denominators survive for some fixed $\epsilon_1 > 0$, there are more than any fixed power of L of them. By pigeonholing among at most a fixed power of L

classes, one can choose a number of distinct equal-class denominators, itself at most a fixed power of L, whose class sum is Qr mod 1. This uses the fact that every nonzero cyclic ℓ-power subgroup contains the entire subgroup of order ℓ. Since $n \geq N/L^3$ their mass is $o(1)$, proving (28).

Suppose instead that at least $\epsilon_1 N\delta$ nonexcluded q-multiples belong to healthy large-prime pools. There are at least $\epsilon_1 N\delta/L^B$ distinct such pools. Choose one seed pm from each, and construct the reserve C_p excluding the seed (Section 6.3). Prescribe a numerator j for $R(E) = j/Q_p$ with

$$1/m + j/Q_p \equiv 0 \pmod{p}.$$

This is a valid condition since $p \nmid mQ_p$. We may also require that the ℓ-valuation of the reduced denominator of $1/m + j/Q_p$ be at least a. To check it, put $k = v_\ell(m) \geq a$, $k' = v_\ell(Q_p)$. If $k > k'$, cancellation of that leading denominator is impossible. If $k \leq k'$, prescribe j modulo ℓ to keep the numerator at common denominator $\operatorname{lcm}(m, Q_p)$ a unit modulo ℓ; exactly one residue is forbidden because the coefficient of j is an ℓ-unit. The CRT with the modulus p is valid and the grid interval spans all needed residues by (14). After dividing by p, the block $B_p = \{pm\} \cup pE$ has nonzero class $QR(B_p)$ mod 1 of bounded polylogarithmic ℓ-power order. All other denominator primes and powers divide Q: the p factor has been canceled, and other bad powers were excluded in both the seed and reserve. These blocks are pairwise disjoint, each of polylogarithmic length, and with denominators at least N/L^3. There are more than any fixed power of L of them. The same pigeonhole argument on their classes yields (28), with total length polylogarithmic and total mass $o(1)$. Unlike the first two constructions, some correction terms within these blocks need not be q-multiples; the entire blocks nevertheless remain outside P because they contain the factor $p > Y$. This is sufficient for disjointness.

8.3 Why the correction defeats a transversal

For a multiples-template transversal any successful S_* as above yields a contradiction. The remaining $r' = r - R(S_*) \to r$ is allowed by Q. The smooth non-q pool differs by $o(N)$ points from the fixed-density periodic non-q integers. Hence there is a fixed-proportion low interval with more than $r + 1$ reciprocal mass but fewer than $K/2$ denominators, obtained by taking its upper endpoint sufficiently small and then its lower-to-upper ratio sufficiently small. This is enough for at-most forcing using the one-dimensional lemma, floor probabilities L^{-4} and a linear cardinality slack, as in Section 7.2. For exact length let $d_q = 1 - 1/q$. Since the divisibility branch is active we have $d_q D \geq D_*$. For $K' = K - |S_*| = K - o(N)$ the minimum mass in the non-q template complement tends to

$$d_q \log \frac{d_q D}{d_q D - 1} \leq d_q r < r.$$

Uniform rank estimates on the upper-tail interval show that the $o(N)$ perturbation to P does not affect this fixed margin. The low interval gives the opposite fixed margin. The probability interpolation of Section 7.2 and Lemma 3.1 produce the remaining K' smooth terms. This uses the fact that in an active exact-length divisibility phase $q \geq 3$ and $|P| > N/2$ by a fixed margin. In either edge family we obtain a representation contradicting that F hits all edges.

Consequently the first resource cannot occur for a transversal: each bad small-power pool has fewer than $\eta N/s$ surviving smooth q-multiples. By definition it has fewer than $\eta N/s$ surviving smooth non-q multiples. The proof of (23) now gives its version with 2η in place of η, including the same exception and overlap errors. Also by the other two constructions

$$|(q\mathbb{N} \cap [1, N]) \setminus (F \cup V \cup E_0)| = o(N\delta). \tag{30}$$

If this were false along a subsequence then a fixed positive multiple of $N\delta$ would survive in one of the two resources. Thus, using coprimality of all bad powers with q,

$$h \geq N/q + |F \cap V| - (N/q)\omega - o(N\delta).$$

The ledger (24) with $j = 2$ and $\alpha' = 1/q$ applies. Its small-power term is positive for sufficiently small η, giving

$$h \geq N/q - N\delta(e^{1/q} - 1) - O_B((\gamma + \epsilon)N\delta) - o(N\delta).$$

Choose the fixed B after η and the carrier parameters, but before γ, ϵ. The small-power correction above only requires $B\nu > 2$ for ν depending on r, η; it introduces no dependence on γ, ϵ. Taking their limits after $N \to \infty$ gives the required matching lower bound. Together with both segment bounds this proves Theorem 1.2. Strict separation between minimizing second-order costs excludes the higher-cost template subsequences, proving the asserted second-order template selection.

9 The onset window: weighted loss and a budget constraint

We now prove Theorem 1.3 for each fixed rational $0 < r \leq 1$, without using either macroscopic phase theorem or the divisibility-template proof of Section 8. All affordable deletion sets have size $O(N\delta) = o(N)$. For a proper rational, the finitely many fixed powers of its reduced denominator survive the pruning of Section 6.2; for $r = 1$ the denominator is $b = 1$ and Qr is automatically integral. The large primes in Section 6.1 are coprime to b in both cases. Only the unrestricted marking and repair construction is used here. Write $D_* = A + 1$, $N = \lfloor D_*K + uK\delta_K \rfloor$, and set $s = e^{-r} = A/D_*$. In this section u and all budget multipliers are fixed when K tends to infinity.

9.1 Linear response of the minimum mass

Define a nonnegative Lipschitz weight on $[0, 1]$ by

$$w(t) = \begin{cases} 0, & 0 \leq t \leq s, \\ (1 - s/t)/(1 - s), & s < t \leq 1. \end{cases} \tag{31}$$

Then $w(1) = 1$, $w \leq 1$, and $\int_0^1 w(t)dt = 1 - rA = \alpha$. For a set E of removed vertices let $\mathcal{W}(E) = \sum_{n \in E} w(n/N)$.

Lemma 9.1 (Uniform capacity response)**.** *If $|E| = O(N\delta)$, then the minimal reciprocal sum among K surviving vertices is*

$$L_{[2,N]\setminus E}(K) = r + \frac{\delta}{A}\left(\frac{\mathcal{W}(E)}{N\delta} - \frac{u}{D_*}\right) + o(\delta), \tag{32}$$

uniformly over such E, for fixed constants in the cardinality bound.

Proof. The original largest K denominators have mass $\log(N/(N - K)) + O_r(1/N) = r - u\delta_K/(AD_*) + o(\delta_K)$, by differentiation at $N = D_*K$. Removing d of them requires d replacements in the interval just below $N - K$. Since $d \leq |E|$, all replacements are within $O(|E| + 1)$ of that boundary. The total replacement mass therefore differs from $d/(N - K)$ by $O((|E| + 1)^2/N^2)$. The net increase is

$$\sum_{n \in E \cap [N-K+1, N]} ((N - K)^{-1} - n^{-1}) + O((|E| + 1)^2/N^2).$$

Replacing $(N-K)/N$ by s changes each summand by $O_r(\delta/N)$; for points in the thin interval where the indicator changes, the summand itself is $O_r(\delta/N)$ since $w(s)=0$. Thus the net increase is $\mathcal{W}(E)/(AN)+O_r(\delta^2+1/N)$, as $(1-s)/s=1/A$. Use $\delta\sim\delta_K$ to conclude. □

The normalization in (32) is crucial: one unit of direct deletion near N contributes asymptotically one unit to $\mathcal{W}$, whereas an entire large-prime pool contributes the average α per vertex.

9.2 Budget optimization and upper bounds

Suppose a total budget is $HN\delta$, with $H\geq 0$. The continuum relaxation has a zero-cost failed density on $0\leq v\leq 1$; for $1<v<2$, the cost of failing a whole pool per vertex is $f(v)=\log v$ and its weight benefit per vertex is α. Direct deletion can yield asymptotic weight one per point near N. The relevant cost and weight are given by (17) and its weighted version. The exact value of the relaxation is

$$\Phi(H)=\inf_{\mu\geq 1}\mu(H+e^{\alpha/\mu}-1)=\begin{cases}\alpha v_H, & 0\leq H\leq J(e^\alpha),\quad J(v_H)=H,\ 1\leq v_H\leq e^\alpha,\\ H+e^\alpha-1, & H\geq J(e^\alpha).\end{cases} \tag{33}$$

Indeed the multiplier $\mu\geq 1$ bounds the payoff from direct deletions, and

$$\int_0^B(\alpha-\mu f(v))_+dv=\mu(e^{\alpha/\mu}-1).$$

Differentiating $\mu(H+e^{\alpha/\mu}-1)$ yields $H-J(e^{\alpha/\mu})$. When $0<H<J(e^\alpha)$ the minimizer satisfies $J(e^{\alpha/\mu})=H$; at $H=0$ take the infimum as $\mu\to\infty$. For larger H, the minimum is at $\mu=1$. The maximizing strategy for the primal relaxation, and asymptotically for the constructed attacks, attacks all pools up to v_H, or up to e^α followed by direct deletion. Thus this evaluation is not an assumption about the optimality of a particular deletion pattern; its applicability to arbitrary adversaries is established next.

Here $\alpha=1-r/(e^r-1)\in(0,1-1/(e-1)]$, so $e^\alpha<2$ with strict slack even at $r=1$. For $u<u_0$, the free obstruction in Section 6.1 already makes the weight loss $\alpha N\delta+o(N\delta)>(u/D_*)N\delta$, and (32) rules out all exact representations. For $u=u_0$ attack up to $v>1$ arbitrarily close to 1. For $u_0<u<u_1$ attack up to v slightly greater than u/u_0; its weight exceeds the threshold and the cost tends, as the fixed overshoot tends to zero, to $N\delta J(u/u_0)$. For $u\geq u_1$ attack up to e^α, then delete enough of the largest remaining vertices to surpass the threshold. They have weight $1-o(1)$ because only $O(N\delta)$ vertices have been excluded. The needed direct cost, normalized by $N\delta$, tends to $u/D_*-\alpha e^\alpha$, and the total cost to $u/D_*-(e^\alpha-1)$. Since $N\delta\sim D_*K\delta_K$, these constructions prove the upper bound in (4). Each strict mass obstruction follows from (32), since the union of failed and directly removed vertices has size $O(N\delta)$. There is no assertion that a limiting mass equality itself obstructs all representations.

9.3 Arbitrary-adversary bound

Let F be any deletion set of cardinality at most $HN\delta$. Use the unrestricted pool and marking of Section 6. In this regime $T_s=O_H(\delta)$, $|V|=O_B(N\delta)$, and $|P|=N-o(N)$. All fixed target powers survive, so $b\mid Q$ (vacuously when $b=1$). We need a weighted counterpart of (24) in which the multiplier may be greater than one. For fixed $\mu\geq 1$ split actual cost into $F\cap V_s$, $F\cap(V_p\setminus V_s)$ and

$F \setminus V$. By Riemann sums,

$$\mathcal{W}(V_s) \le \alpha N T_s + O_w(|\mathcal{B}|),$$
$$\mathcal{W}(V_p) = \alpha \sum_{p \in \mathcal{D}} N/p + O_w(\pi(N)).$$

The first error is $o(N\delta)$ since $|\mathcal{B}| \le YT_s$ and $T_s = O_H(\delta)$. For the second, $\pi(N) = O(N/L) = o(N\delta)$. By (23) the small-power gain less μ times the cost is bounded by $(\alpha - \mu(1-\eta))NT_s + o(N\delta) \le o(N\delta)$, with $\eta < 1-\alpha$. Direct deletions have weight at most one and hence nonpositive gain after subtracting μ times cost. Finally (21) bounds the large-pool gain less μ times its cost by

$$N \sum_{p>Y} \frac{(\alpha - \mu f(v_p))_+}{p} + O_B(\mu(\gamma+\epsilon)N\delta) + o(N\delta).$$

Use (16) and the integral in (33). Overlap of the two virtual sets only decreases the weight, so we have proved

$$\mathcal{W}(F \cup V) \le \mu|F| + N\delta\,\mu(e^{\alpha/\mu} - 1) + O_B(\mu(\gamma+\epsilon)N\delta) + o(N\delta). \tag{34}$$

All estimates are uniform for F in the budget. Given a fixed error tolerance, choose μ first to approximate the infimum in (33) (a large finite μ is sufficient at $H = 0$). Choose the fine carrier parameters and B for fixed η; these are independent of γ, ϵ. Then choose γ, ϵ sufficiently small and take N sufficiently large. It follows that the right side of (34) can be made at most $(\Phi(H) + \text{tolerance})N\delta$. This order of choices is essential; no uniformity in unbounded μ or vanishing local density is claimed. This proves the needed bound for the actual adversary, rather than only for whole-pool attacks in an abstract optimization.

9.4 Exact-length forcing below the budget boundary

Suppose $\Phi(H) < u/D_*$. Choose a fixed small margin relative to this gap and select constants as in Section 9.3. Remove $F \cup V \cup E_0$ and all large-prime repair reserves. By (20), these latter reserves and E_0 add only $o(N\delta)$ to the deleted ranks and weight. All removed sets together have size $O_B(N\delta)$, so the capacity response (32) applies to their union. Its K largest surviving vertices S_0 satisfy

$$R(S_0) < r - c_3\delta \tag{35}$$

for fixed $c_3 > 0$. All of S_0 lies above a fixed positive proportion of N (the rank of its smallest member is at least $N - K - O(N\delta) \sim sN$).

Split off the selected large-prime portion C_0 and correct it with the reserved cofactors, keeping the original selected terms and adding the terms that cancel each large prime. The corrected union C avoids F and is disjoint from P; $QR(C) \in \mathbb{Z}$, by the precise cofactor lcm implication in Section 6.3. The total number of terms of C is $O_B(N\delta)$, and $R(C) = O_B(\delta)$ here since both the selected terms and added terms lie on the scale N. Set

$$K' = K - |C|, \qquad r' = r - R(C).$$

Both K' and Qr' are integers. Indeed $QR(C) \in \mathbb{Z}$ by the cofactor-lcm implication and cancellation modulo each large prime; $Qr \in \mathbb{Z}$ by the preceding argument (for the unit target, $Qr = Q$). Moreover $K'/N \to 1/D_* \in (0,1)$ and $r' = r - O(\delta) > 0$, so both targets are in the stated range for large N. The extra terms in the correction have count $o(N\delta)$ and mass $o(\delta)$. Remove exactly that

many terms from the original smooth part of S_0 (there are $K - O(N\delta)$ such terms). The remaining smooth set has K' terms, and together with C its reciprocal sum is less than $r - c_3\delta + o(\delta)$. Thus

$$L_P(K') \le r' - \tfrac{1}{2}c_3\delta. \tag{36}$$

The upper endpoint has a fixed positive margin. To see this, fix a sufficiently small $c_2 > 0$ so that $c_2 N < K/3$ for large N, and then choose $c_1 \in (0, c_2)$ with $\log(c_2/c_1) > r + 2$. The interval $[c_1 N, c_2 N]$ has fewer than $K'/2$ points but reciprocal mass exceeding $r + 1$ in P for large N: only $o(N)$ vertices are missing from it. Complete to K' vertices using other members of P, possible because $K'/N \to 1/D_* < 1$ and $|P| = N - o(N)$. Therefore $U_P(K') > r' + c_4$ with fixed $c_4 > 0$.

Now impose the probability floor $p_* = L^{-4}$ and write $p_n = p_* + (1 - 2p_*)z_n$. The desired means $\mathbb{E}T = K'$, $\mathbb{E}Z = r'$ correspond to fractional cardinality and weight

$$k_z = (K' - p_*|P|)/(1 - 2p_*), \quad R_z = (r' - p_* R(P))/(1 - 2p_*).$$

The cardinality shift is $O(NL^{-4})$ and the reciprocal shift is $O(L^{-4}(1 + R(P))) = O(L^{-3})$. Since every denominator is at least N/L^3, the extremal masses at size k_z differ from those at K' by at most $O(L^{-1}) = o(\delta)$. Thus (36) and the fixed opposite margin allow a convex combination of the minimum and maximum fractional allocations to attain R_z. Lemma 3.1 now applies in the fine two-dimensional regime, with $|P| = N - o(N)$, $Qr' \in \mathbb{Z}$ and $K' \in \mathbb{Z}$, and gives positive probability of simultaneously attaining (r', K'). In particular, no assumption that each large-prime repair preserves its length has entered the proof. Appending C yields an exact K-term representation avoiding F.

We have shown that every deletion budget $HN\delta$ with $\Phi(H) < u/D_*$ fails to be a transversal for all sufficiently large N. For $u > u_0$, Φ is continuous and strictly increasing, so take any H below $\Phi^{-1}(u/D_*)$ and then approach that endpoint. This proves the matching lower bound $D_*\Phi^{-1}(u/D_*)$ on the $K\delta_K$ scale. For $u \le u_0$, nonnegativity and the upper bound already give the limit zero. This proves Theorem 1.3, including $r = 1$. In particular the unit-target extension uses neither a largest prime power of 1 nor a limiting argument through proper rational targets.

10 Minimum heights, regularity and remaining problems

Proof of Corollary 1.4. Fix $\epsilon > 0$. Theorem 1.3 gives eventual emptiness below $u_0 - \epsilon$ and a positive limiting normalized transversal number above $u_0 + \epsilon$. In particular an edge exists at the latter height. The edge family grows monotonically with height, so these two heights sandwich the minimum, and $\epsilon \downarrow 0$ proves (5). For a fixed positive β the function in (4) has a unique positive-level inverse. Apply the same strict sandwich to the level β; the integer rounding of the budget is $o(K\delta_K)$. At $\beta = 0$ use the just-established existence threshold, not an inverse of the entire zero plateau. The claimed inverse is then obtained from (33). Finally $J(1+t) = t^2/2 + O(t^3)$ proves the square-root expansion. At $r = 1$ we have $A = 1/(e-1)$, $D_* = e/(e-1)$ and $\alpha = (e-2)/(e-1)$; hence $D_*\alpha = e(e-2)/(e-1)^2$, as in (6). The two strict sandwiches apply to this fixed target directly. □

Some further consequences concern the shape of the limit, rather than a finer uniform estimate. The derivative of the middle piece of Ψ_r is $(D_*/u_0)\log(u/u_0)$, tending to zero at u_0 and to one at u_1. Thus the function is continuously differentiable and not twice continuously differentiable at either junction. Near the first junction,

$$\Psi_r(u_0 + \varepsilon) = \frac{D_*}{2u_0^2}\varepsilon^2 + O_r(\varepsilon^3), \quad \varepsilon \downarrow 0.$$

This is a statement about the limiting function. It does not justify predicting constant-sized transversals at distance $\sqrt{K\delta_K}$ from $D_*K + u_0K\delta_K$: the present remainder can be much larger and the true finite-scale center can drift.

For the macroscopic phase junctions, set $C_i(D) = D(e^{\alpha_i} - 1)$. Comparing the active second-order functions gives the following offsets from the respective first-order junctions, when those junctions are active:

$$\begin{aligned} D_{21} = qA, &\quad u_{21} = q(C_2(D_{21}) - C_1(D_{21})), \\ D_{32} = qD_*/(q-1), &\quad u_{32} = \tfrac{q}{q-1}(C_3(D_{32}) - C_2(D_{32})), \\ D_{31} = 2A + 1, &\quad u_{31} = C_3(D_{31}) - C_1(D_{31}). \end{aligned}$$

These are comparisons of limiting costs within the $K\delta_K$ windows. They do not assert finite-size uniqueness of the transition or the exact shape of the optimal deletion set. The active-template stability proved above is macroscopic; a description of the microscopic distribution of actual deletions inside each failed pool is still a further question.

11 Uniform deficit and quantitative stability at larger heights

The preceding macroscopic second-order formula assumes a fixed proper rational and a fixed linear height window; the exact-length onset result allows a fixed rational $0 < r \le 1$. The following complementary results apply to all fixed positive rationals, including integer targets, and allow heights growing subexponentially relative to K. In this section alone write $r = u/b > 0$ in lowest terms, $a_r = e^r - 1$, $\alpha_r = 1/a_r$, and $\Delta_r(K; M) = K/a_r - \tau_K(r; M)$. The definition of the at-most-K edge family is unchanged.

Theorem 11.1 (Transversal deficit)**.** *For every fixed positive rational r there exists an integer $D_r \ge 2$ such that, for every fixed $0 < \epsilon < 1$, uniformly over integers*

$$D_rK \le M \le K\exp((\log K)^{1-\epsilon}),$$

one has

$$\Delta_r(K; M) \asymp_{r,\epsilon} \frac{K \log\log K}{\log K} \tag{37}$$

as $K \to \infty$. In particular this holds at $M = D_rK$ and at $M = \lfloor CK\log K \rfloor$ for every fixed $C > 0$ (with constants allowed to depend on r, C).

Theorem 11.2 (Stability)**.** *For $M \ge D_rK$, every minimum transversal F satisfies*

$$|F \triangle [2, 1 + \lfloor K/a_r \rfloor]| \ll_r K\sqrt{\frac{\log\log K}{\log K}}.$$

11.1 A protected high reservoir

We only require a specialized form of reservoir pruning, rather than the more general weighted local limit theorem with arbitrary deterministic weights.

Lemma 11.3 (Uniform prefix loss and reservoir health)**.** *Fix $b_0 \ge 1$. There are constants $B > 0$, $\eta_0 > 0$, $C_1 > 0$ and an absolute $C_2 > 0$ such that the following holds for large N. Write $L = \log N$,*

$\lambda = \log L$, $\delta = \lambda/L$ and $Y = N/L^B$. If $F \subseteq [2, N]$ has $h_H = |F \cap (N/2, N]| \le \eta_0 N$, there is $G \subseteq [1, N]$ with

$$|[1, X] \setminus G| \le C_1 X\delta + C_2 X h_H/N \qquad (1 \le X \le N). \tag{38}$$

For $H = G \cap (N/2, N] \setminus F$ and $Q = \operatorname{lcm}(H)$, we have $|H| \gg N$, $b_0 \mid Q$, $n \mid Q$ for all $n \in G$, and $\log Q \ll Y$. For every $A \subseteq G \cap [N/L^2, N/2] \setminus F$, the one-dimensional conclusion of Lemma 3.1 holds on $P = A \cup H$ with parameters in $[L^{-4}, 1 - L^{-4}]$ and with this Q.

Proof. Choose a fixed small threshold $\eta = 1/16$. First exclude E, the integers with a prime-power divisor above Y or with $\Omega(n) > 20\lambda$. By Mertens' estimate and (8), uniformly for $1 \le X \le N$,

$$|E \cap [1, X]| \ll_B X\delta.$$

Indeed the first exclusion is bounded by $X \sum_{Y < s \le N,\ s \text{ prime power}} s^{-1} \ll_B X\delta$ and the second by $O(XL^{1/2 - 20\log(3/2)})$. Let $H_0 = (N/2, N] \cap \mathbb{N} \setminus (E \cup F)$. Mark powers $s \le Y$ for which $|H_0 \cap s\mathbb{N}| < \eta N/s$, retaining only the smallest marked power for each base prime in a set $\mathcal{B}$. Put $T_0 = \sum_{s \in \mathcal{B}} s^{-1}$. Since $(N/2, N]$ contains $N/(2s) + O(1)$ multiples of s, for sufficiently large N the excluded points in the upper half contribute at least $N/(3s)$ to each marked class. With $g(n) = \sum_{s \in \mathcal{B}} \mathbf{1}_{s \mid n}$, pairwise coprimality gives $\sum_{n \le N} g(n)^2 \le N(T_0 + T_0^2)$. Therefore Cauchy–Schwarz yields

$$T_0^2/9 \le (h_H/N + |E|/N)(T_0 + T_0^2).$$

If the factor in parentheses is below $1/18$, then $T_0 \le 18(h_H/N + |E|/N)$ (also valid for $T_0 = 0$). Define G by excluding E and all multiples of powers in $\mathcal{B}$. The count of multiples in a prefix is at most XT_0, so (38) follows with an absolute $C_2 = 18$ and a constant C_1 allowed to depend on B.

Choose η_0 sufficiently small that $T_0 < \eta/6$ for all large N and that each of the finitely many highest prime powers of b_0 has at least $\eta N/s$ multiples in H_0 (the latter follows from the half-interval count and $h_H + |E|$). The same property holds for their lower powers. If $s \le Y$ is not divisible by any marked power, it is not marked, and the deletion of other-base marked multiples costs at most $(N/s)T_0$. The one possible larger same-base marked power removes fewer than $\eta N/(2s)$ of its H_0 multiples. Thus

$$|H \cap s\mathbb{N}| \ge (\eta/2 - T_0)N/s \ge \eta N/(3s).$$

Every power appearing in G is such an s; hence $n \mid \operatorname{lcm}(H)$ for $n \in G$, and $b_0 \mid Q$. The density of H is bounded below by an absolute positive constant by (38) and the size of F in the upper half. On $A \cup H$ every appearing power consequently meets the fine health hypothesis with fixed lower density $\eta/3$. All denominators exceed N/L^3, and the imposed arithmetic restrictions are those of the fine regime. Fix B sufficiently large for Lemma 3.1 with these fixed density constants, before choosing the final sufficiently large N; its choice does not change C_2 or the small fixed threshold for h_H/N. This proves the stated one-dimensional inversion, with the actual lcm equal to Q. □

11.2 Linear-height construction

Write $r = u/b$ in lowest terms, and $\alpha_r = 1/a_r$. First observe that for any M, deleting $[2, h+1] \cap [2, M]$ with $h = \lfloor K/a_r \rfloor$ is a transversal, since even the sum of the first K possible remaining denominators is

$$\sum_{j=1}^{K} \frac{1}{h + 1 + j} < \log(1 + K/(h+1)) < r.$$

In particular $\tau_K(r; M) \le \lfloor K/a_r \rfloor$.

For the lower bound, use Lemma 11.3 with $b_0 = b$. Fix an integer $D_1 > 2 + \alpha_r$ and choose an integer D_r so large that

$$D_r > 2D_1, \qquad C_2 D_1/D_r \le 1/2, \qquad \alpha_r/D_r < \eta_0. \tag{39}$$

Take $N = D_r K$, and suppose $F \subseteq [2, N]$ has size $h \le \alpha_r K - C_r K\delta$ with a constant C_r to be specified. Put $h_H = |F \cap (N/2, N]|$. Lemma 11.3 applies. At scale N use $L = \log N$, $\lambda = \log L$ and $\delta = \lambda/L$. Let $X = D_1 K < N/2$ and let d_F be the number of forbidden or non-good denominators in $[2, X]$. Crucially, the original deletions here and the reservoir deletions are disjoint. Thus (38) and (39) give

$$d_F \le (h - h_H) + C_1 D_1 K\delta + C_2 D_1 K h_H/N \le h - h_H/2 + O_r(K\delta). \tag{40}$$

This is the cost-transfer step that permits a linear rather than logarithmically larger height.

Put $J_K = \lceil K\delta \rceil$, $k_0 = K - J_K$ and take A to be the smallest k_0 good, non-forbidden integers in $[J_K, X]$. There are enough because $J_K + d_F + k_0 - 1 < X$ for large K. The i-th chosen integer is at most $J_K + d_F + i - 1$, so

$$R(A) \ge \sum_{n=J_K+d_F}^{J_K+d_F+k_0-1} \frac{1}{n} \ge \log \frac{d_F + K}{d_F + J_K}. \tag{41}$$

For $d_F \le \alpha_r K - C' K\delta$ with a sufficiently large fixed C', the last expression is at least $r + \delta$. To see this, expand the expression at $d_F = \alpha_r K - C' K\delta$ and $J_K = K\delta + O(1)$:

$$r + \left(\frac{a_r^2 C'}{1 + a_r} - a_r \right) \delta + O_r(\delta^2 + K^{-1});$$

the expression decreases with d_F. Choose C_r so that (40) implies this bound on d_F.

As $J_K \ge N/L^2$ for large K, the main set is admissible for Lemma 11.3. On H take $p_n = L^{-2}$, and on A take the constant probability

$$\theta = \frac{r - L^{-2} R(H)}{R(A)} \in (0, 1).$$

Here $R(H) \le \log 2 + o(1)$ and $r + \delta \le R(A) \ll \log L$. These bounds ensure $L^{-4} \le \theta \le 1 - L^{-4}$ for large N, so all probabilities satisfy Lemma 11.3. The mean is r, which is on the Q^{-1} lattice since $b \mid Q$. For the number Z_H of chosen reservoir terms, $\mathbb{E} Z_H \le N/L^2$ and the elementary Chernoff bound gives

$$\mathbb{P}(Z_H > J_K) \le \left(\frac{e \mathbb{E} Z_H}{J_K} \right)^{J_K} \le e^{-J_K} = o(Q^{-1}).$$

Indeed $J_K/(N/L^2) \to \infty$ and $J_K \gg \log Q$ in the sense $J_K/(N/L^B) \to \infty$. Lemma 11.3 gives exact mass with probability asymptotic to $1/(Q\sqrt{2\pi \operatorname{Var} Z})$, where Z is the sampled reciprocal sum. Since $\operatorname{Var} Z \le N/(N/L^2)^2 = o(1)$, the probability is at least $1/Q$ for large N, while the failure probability above is $o(1/Q)$. Hence there is an outcome $S \subseteq A \cup H$ with $R(S) = r$ and $|S| \le k_0 + J_K = K$. It avoids F and its height is at most N. Consequently

$$\tau_K(r; D_r K) \ge \alpha_r K - O_r(K \log\log K / \log K). \tag{42}$$

By monotonicity in the height, this lower bound holds for every $M \ge D_r K$.

11.3 Matching obstruction throughout the growing-height range

We prove the upper bound on $\tau_K(r;M)$ in Theorem 11.1 throughout its stated height range. Write $\ell = \log K$ and fix $0 < \epsilon < 1$. Consider primes

$$\mathcal{P} = \{p : K/\ell^{\epsilon/2} \le p \le K/\ell^{\epsilon/4}\}.$$

For large K, they do not divide b, and $pq > M$ for every $p, q \in \mathcal{P}$, including $p = q$, because $M \le K\exp(\ell^{1-\epsilon}) = K^{1+o(1)}$.

If an exact representation S contains $s > 0$ multiples $pm_1, \dots, pm_s$ of p, then $p \nmid m_i$ and multiplication of $R(S) = r$ by p, followed by reduction modulo p in the localization of $\mathbb{Z}$ at p, gives $\sum_i m_i^{-1} \equiv 0 \pmod p$. Hence p divides the positive integer $U = \sum_i \prod_{j\neq i} m_j$. Since $m_i \le M/p$, for $s \le \ell^\epsilon/4$ we have

$$\log U \le \log s + (s-1)\log(M/p) \le (1/4 + o(1))\ell < \log p.$$

This is impossible, so $s > \ell^\epsilon/4$ whenever $s > 0$. Among integers at most $h + 2K + 1$ with $h \asymp_r K$, only $O_r(\ell^{\epsilon/2})$ are multiples of p; hence a representation using p must contain $\gg \ell^\epsilon$ larger multiples of p.

For $h \asymp_r K$, delete $F_h = [2, h+1]$ and let $B_h = [h+2, h+K+1]$, $v = (h+K+1)^{-1}$, $I = [h+2, h+1+\lfloor K/2\rfloor]$. For any representation avoiding F_h with $|S| \le K$, the identity

$$\begin{aligned} R(B_h) - R(S) = &\sum_{n\in B_h\setminus S} (1/n - v) + \sum_{n\in S\setminus B_h} (v - 1/n) \\ &+ (K - |S|)v \end{aligned} \tag{43}$$

has nonnegative terms. For each $p \in \mathcal{P}$ there are $t_p = |I \cap p\mathbb{N}| \asymp K/p$ integers of I divisible by p. If S omits all multiples of p, their contribution in the first sum of (43) is $\gg_r t_p/K$. Otherwise the second sum contains $\gg \ell^\epsilon$ multiples of p above $h + 2K + 1$, each contributing $\gg_r 1/K$. As $t_p = O(\ell^{\epsilon/2})$, the latter cost is also $\gg_r t_p/K$. No denominator is counted twice across these primes. It follows that

$$R(B_h) - r \gg_r \sum_{p\in\mathcal{P}} p^{-1} \asymp_\epsilon \frac{\log\ell}{\ell}.$$

The last estimate follows from (7), with main term $(\epsilon/4)\log\ell/\ell$. Now take $h = \lfloor \alpha_r K - c_{r,\epsilon} K\log\ell/\ell\rfloor$. Then $R(B_h) - r = O_r(c_{r,\epsilon}\log\ell/\ell + 1/K)$ by harmonic-sum comparison. Choosing $c_{r,\epsilon} > 0$ small enough contradicts the preceding lower bound if S exists. Thus F_h is a transversal and the matching upper bound follows. Together with (42), this proves Theorem 11.1.

11.4 Quantitative stability

Let $M \ge D_r K$, let F be a minimum transversal of size h, and put $N = D_r K$, $X = D_1 K$. Apply the same construction to $F \cap [2, N]$, using $L = \log N$ and $\delta = (\log L)/L$. From the basic upper bound and (42), $h = \alpha_r K + O_r(K\delta)$. Let $h_{\text{out}} = |F \setminus [2, N]|$, and define h_H as above. Equation (40) now gives $d_F \le h - h_{\text{out}} - h_H/2 + O_r(K\delta)$. If $d_F \le \alpha_r K - C'K\delta$ with C' as in the linear-height construction, the construction produces a representation avoiding F, impossible. Consequently $h_{\text{out}} + h_H = O_r(K\delta)$ and the extra exclusions in $[1, X]$ from (38) number $O_r(K\delta)$.

Choose an integer $h' = h + O_r(K\delta)$ with $h' \ge d_F$ and $h' \ge h$. Put $d = |[2, h+1] \setminus F|$, and let d' be the number of good, non-forbidden integers in $[J_K, J_K + h' - 1]$. Translating endpoints and removing the extra excluded integers costs $O_r(K\delta)$, so $d' \ge d - O_r(K\delta)$. Take $t = \min\{d', k_0\}$.

Since at most h' integers of $[2, X]$ are excluded, the k_0 smallest available integers in $[J_K, X]$ have reciprocal mass at least

$$R(A) \geq \sum_{n=J_K+h'}^{J_K+h'+k_0-1} \frac{1}{n} + \sum_{j=1}^{t} \left(\frac{1}{J_K + h' - j} - \frac{1}{J_K + h' + j - 1} \right)$$
$$\geq \log \frac{h' + K}{h' + J_K} + \frac{t^2}{X^2} \geq r - O_r(\delta) + \frac{t^2}{X^2}.$$

For clarity, if $d' < k_0$, at most d' exclusions remain above $J_K + h' - 1$, so at least $k_0 - d'$ of the baseline k_0 integers can be chosen, and the worst case removes the first d' of them and uses the largest d' possible holes below the boundary. If $d' \geq k_0$, the k_0 largest possible holes already give the displayed bound. All involved endpoints lie below X for large K. The summand differences are at least $(2j-1)/X^2$.

If d were larger than a sufficiently large constant times $K\sqrt{\delta}$, the last lower bound would be at least $r + \delta$. The same reservoir and sampling construction would contradict the transversal property. Hence $d = O_r(K\sqrt{\delta})$. Since $|F| = |[2, h+1]|$, their symmetric difference has size $2d$. Replacing h by $\lfloor \alpha_r K \rfloor$ changes the interval by only $O_r(K\delta)$ points. This proves Theorem 11.2.

12 Restricted denominator families

12.1 First-order law and scope

Let $\mathcal{A} \subseteq \mathbb{N}$ be fixed, with natural density $d > 0$. For $r = a/b > 0$ in lowest terms let $\mathcal{Q}_b$ denote the highest prime power $p^{v_p(b)}$ for each prime $p \mid b$; when $b = 1$ this set is empty. Assume for every $q \in \mathcal{Q}_b$ that the density

$$d_q = \lim_{X \to \infty} \frac{|\mathcal{A} \cap q\mathbb{N} \cap [1, X]|}{X}$$

exists. For $M \geq 2$ define $\tau_{K,\mathcal{A}}(r; M)$ as the smallest number of deleted vertices in $\mathcal{A} \cap [2, M]$ hitting every subset S of that family with $|S| \leq K$ and $R(S) = r$, where $R(S) = \sum_{n \in S} 1/n$. The transversal number of an empty edge family is zero.

Theorem 12.1 (First-order height formula)**.** *Fix $D > 0$. With the convention $\min_{q \in \varnothing} d_q = +\infty$, we have*

$$\lim_{K \to \infty} \frac{\tau_{K,\mathcal{A}}(r; \lfloor DK \rfloor)}{K} = \min \left\{ \frac{1}{e^{r/d} - 1},\ dDe^{-r/d},\ D \min_{q \in \mathcal{Q}_b} d_q \right\}. \tag{44}$$

The middle branch is essential for general targets. In the full integer family, for $0 < r < 1$ with reduced denominator b, it is dominated by the prime-power branch because $q_*(b) > e^r$. For arbitrary $r > 0$, and in particular for integer targets, the finite total reciprocal mass must be retained in the formula.

This first-order statement uses natural densities only, not a specific distribution in short intervals. It is complementary to Martin's prescribed-length theorem [2] and the positive-density reciprocal sum tools of Bloom and Liu–Sawhney [4, 5]. It does not follow merely by replacing each integer with its average-density analogue in an exact-cardinality result.

12.2 A pruned exact grid

The coarse completion lemma above provides the precise grid input needed for arbitrary fixed families, without invoking a stronger external proposition about exact rational sums.

Lemma 12.2 (Pruned rational grid)**.** *Fix $c, c' > 0$. Let $X \to \infty$, $l = \log X$, and $A_0 \subseteq [cX, X] \cap \mathbb{N}$ have size at least $c' X l^{-1/20}$. One can remove $o(Xl^{-1/20})$ elements to obtain A such that, writing $Q = \mathrm{lcm}(A)$, every $u \in Q^{-1}\mathbb{Z}$ with*

$$R(A)/\log X \le u \le R(A)/(1 + 1/\log X)$$

is a reciprocal subset sum of A. In particular the central third of the reciprocal-mass interval is filled at its grid points.

Proof. Apply the arithmetic exclusions and iterative power pruning from the proof of Lemma 4.1 directly to A_0, without downsampling or a large-prime-factor condition. Those steps use neither condition. Their total loss is $O(Xl^{-1/10} + Xl^{-7/100} \log l) = o(Xl^{-1/20})$. Thus the remaining size is at least $(c'/2)Xl^{-1/20} \ge Xl^{-3/50}$, and the fixed lower endpoint cX exceeds $X/l^{1/100}$ eventually. The arithmetic restrictions and all local density conditions of the coarse regime hold. For an admissible u set $p_n = u/R(A)$ on every member; then $l^{-4} \le p_n \le 1 - l^{-4}$ for large X, and $Q\mathbb{E}Z = Qu \in \mathbb{Z}$. The scalar conclusion of Lemma 3.1 gives positive probability of the exact target. □

12.3 Proof of the height formula

Let $N = \lfloor DK \rfloor$, and denote the right side of (44) by C_*.

12.3.1 Necessary bounds

Order the elements of $\mathcal{A} \cap [2, \infty)$ as $a_1 < a_2 < \cdots$. Natural density gives $a_j = j/d + o(j)$ and $|\mathcal{A} \cap [2, N]| = dDK + o(K)$. For fixed $h_0 > 0$ and $k_0 > 0$, if there are enough elements, the reciprocal sum of ranks $\lfloor h_0 K \rfloor + 1$ through $\lfloor (h_0 + k_0)K \rfloor$ tends to $d \log(1 + k_0/h_0)$, by uniformity of $a_j/(j/d) \to 1$ for $j \asymp K$ and harmonic-sum estimates.

For the first branch, delete the smallest $(1/(e^{r/d} - 1) + \epsilon)K$ available denominators, or all vertices if fewer exist. The largest possible K-term reciprocal sum of survivors is less than r for large K: if there are fewer than K survivors it is bounded by the sum of the next K ranks in the infinite family. This proves the first upper bound after $\epsilon \downarrow 0$. For the second branch, delete the smallest $(dDe^{-r/d} + \epsilon)K$ vertices (take ϵ small enough that this is below dDK). The total reciprocal mass remaining tends to at most $d \log(dD/(dDe^{-r/d} + \epsilon)) < r$. This proves the second upper bound. For any $q = p^{v_p(b)} \in \mathcal{Q}_b$, deleting all multiples of q in the family removes $d_q DK + o(K)$ vertices. The surviving lcm then has p-valuation less than $v_p(b)$, making r impossible. This proves the third upper bound, including when a d_q vanishes. If $C_* = 0$, nonnegativity proves the claimed limit, so henceforth assume $C_* > 0$.

12.3.2 Constructing a small candidate set with the right lcm

Fix $0 < \epsilon < C_*$ and suppose an arbitrary $F \subseteq \mathcal{A} \cap [2, N]$ satisfies $|F| \le (C_* - \epsilon)K$. We show it cannot hit all representations. Put $h_0 = C_* - \epsilon/2 > 0$. Choose sufficiently small $s > 0$ so that excluding all $n < sK$ adds fewer than $\epsilon K/4$ effective deletions; it suffices that $s < \epsilon/4$, since there are at most sK such integers. Write $U = (\mathcal{A} \cap [sK, N]) \setminus F$. Its complement in the finite family has size at most $h_0 K$ for large K.

The total supply of q-multiples in U, for every $q \in \mathcal{Q}_b$, is at least $d_q DK - (C_* - \epsilon)K - sK - o(K) \ge \epsilon K/2$. Choose $\rho > 0$ sufficiently small and reserve $\lceil \rho K \rceil$ such multiples for each q, taking their union B_0 (the choices need not be disjoint). Let $t_b = |\mathcal{Q}_b|$. Thus $|B_0| \le t_b \lceil \rho K \rceil$; each reserved class contains at least ρK elements. If $b = 1$ reserve nothing. Complete this union with the smallest

available elements of $U \setminus B_0$ until its size reaches K or U is exhausted; call the result A_0. The set A_0 contains all of the first k ranks of U where $k = \min(\lfloor (1-2t_b\rho)K \rfloor, |U|)$ (for $b = 1$ use $k = \min(K, |U|)$). Indeed each such rank is either reserved or among the first $K - |B_0|$ unreserved ones. The reciprocal mass is consequently at least that of the corresponding worst-case rank interval after deleting $\lfloor h_0 K \rfloor + 1$ smallest members of the finite family. More explicitly, let $j_0 = \lceil h_0 K \rceil$. The interval of ranks

$$j_0 + 1, \ldots, \min\{|\mathcal{A} \cap [2, N]|, j_0 + \lfloor (1-2t_b\rho)K \rfloor\}$$

provides a lower bound by rearrangement: the survivor rank j is at most $j_0 + j$, and the total number of survivors is at least the size of that rank interval. Hence

$$\liminf R(A_0) \geq d \log \frac{\min(dD, h_0 + 1 - 2t_b\rho)}{h_0}. \tag{45}$$

At $\rho = 0$, the right side strictly exceeds r, because $h_0 < C_* \leq 1/(e^{r/d} - 1)$ and $h_0 < C_* \leq dDe^{-r/d}$. By continuity the same is true for sufficiently small fixed ρ. All candidate denominators lie in $[sK, N]$, a fixed proportional interval. Moreover (45) implies $|A_0| \gg K$ since all denominators are at least sK. The construction has $|A_0| \leq K$, with a fixed positive mass surplus and a positive number of representatives of every necessary highest prime power.

12.3.3 Exact extraction

Apply Lemma 12.2 with $X = N$ and the proportional lower cutoff sK/N. Its pruning removes $o(N(\log N)^{-1/20}) = o(K)$ vertices. Thus $R(A) = R(A_0) - o(1) > r$ with a fixed surplus, and each reserved class still contains some member. Consequently $b \mid Q = \mathrm{lcm}(A)$ (all highest prime powers are present). The upper bound $R(A) = O_{s,D}(1)$ follows from the proportional denominator interval. For sufficiently large K,

$$R(A)/\log N \leq r \leq R(A)/(1 + 1/\log N).$$

The target is a legal rational grid point, since Qr is integral. Lemma 12.2 gives $S \subseteq A$ with $R(S) = r$. It automatically has at most K members, avoids F, belongs to $\mathcal{A}$, and has height at most N. This proves the lower bound for every fixed $\epsilon > 0$, and hence (44).

12.4 Examples and critical heights

When all relevant $d_q > 0$, set $C_1 = (e^{r/d} - 1)^{-1}$. The smallest fixed height coefficient at which the right side of (44) reaches C_1 is

$$D_{\mathcal{A},*} = C_1 \max\left\{ \frac{e^{r/d}}{d}, \frac{1}{\min_{q \in \mathcal{Q}_b} d_q} \right\},$$

with the second term taken as zero when $b = 1$. If any $d_q = 0$ there is no finite D attaining the positive plateau on this linear scale.

Squarefree denominators. Their density is $d = 6/\pi^2$ by the identity $\mu^2(n) = \sum_{k^2 \mid n} \mu(k)$. For a prime p, the density of squarefree multiples of p is $d/(p+1)$, by the Euler product (the local factor at p changes from $1 - 1/p^2$ to $1/p - 1/p^2$). A nonsquarefree target denominator makes representation impossible, not merely difficult. For squarefree $b > 1$ with largest prime factor $P^+(b)$ we obtain

$$\lim_{K \to \infty} \frac{\tau_{K,\mathrm{sf}}(r; \lfloor DK \rfloor)}{K} = \min\left\{ \frac{1}{e^{r/d} - 1}, dDe^{-r/d}, \frac{dD}{P^+(b) + 1} \right\},$$

with critical height $\max(e^{r/d}, P^+(b) + 1)/(d(e^{r/d} - 1))$.

One residue class. For $\mathcal{A} = \{n : n \equiv a_0 \pmod m\}$ with fixed $m \geq 1$, $d = 1/m$. The density d_q is $1/\operatorname{lcm}(q, m)$ if $\gcd(q, m) \mid a_0$, and zero otherwise, by the Chinese remainder theorem. In particular an odd denominator family cannot represent a rational with an even reduced denominator. For odd b the odd family has $d = 1/2$ and $d_q = 1/(2q)$. These give explicit piecewise-linear formulas from Theorem 12.1.

Full integers. For $d = 1$ and $0 < r < 1$, $q_*(b) \geq 2$ and $q_*(b) > e^r$: if $q_* = 2$ then $b = 2$ and $r = 1/2$, otherwise $q_* \geq 3 > e^r$. Thus the formula reduces to $\min((e^r - 1)^{-1}, D/q_*(b))$. For $r = 1$, $b = 1$, and $D < e/(e - 1)$ the total mass branch D/e is the smaller one. This explicitly shows why the denominator-mass branch cannot be omitted outside the original proper-fraction regime.

12.5 Scope: exact length and varying families

The restricted-family argument does not handle exact cardinality. For a finite pool P of at least two distinct denominators put $Q = \operatorname{lcm}(P)$ and $g = \gcd_{n \in P}(Q/n - Q/n_0)$. Then the integer lattice generated by $(Q/n, 1)$ is $\{(u, k) : u \equiv (Q/n_0)k \pmod g\}$, by subtracting generators and applying Bezout. For an odd pool this imposes $QR(S) \equiv |S| \pmod 2$, a genuine extra obstruction. Moreover algebraic lattice membership does not ensure quantitative equidistribution. A restricted-family prescribed-length theorem needs to treat these additional congruences and near-resonances. The present first-order result also does not cover density-zero families such as shifted primes, nor families depending on N (for example N^θ-smooth numbers) without suitable uniform rank and local-density inputs.

13 Scope and remaining questions

The macroscopic results at fixed proper targets, the exact-length onset at fixed rational $0 < r \leq 1$, the growing-height deficit, and the restricted-family first-order law have different hypotheses. In particular, the last is not a second-order formula for arbitrary families. Neither the limiting onset function nor the macroscopic stability theorem describes the microscopic location or uniqueness of optimal forbidden sets. A quantitative remainder at the onset, an exact-cardinality theory for restricted families with nontrivial lattice congruences, and sharp coefficients for the full growing-height range are natural further problems.

For the main second-order results the delicate inputs are the common-carrier synchronization, removal of extra rational cardinality resonances, disjoint sparse repairs with both length and mass cost controlled, target-adic correction, and the uniform capacity and cost comparisons. We retain the actual-lcm inversion and the explicit order of auxiliary constants; in particular no fixed-target congruence or exact-cardinality version of an external grid theorem is assumed. The refined Martin coefficient and nonlinear onset are offered for independent mathematical review; no claim of formal certification or established priority is made.

References

[1] P. Erdős and R. L. Graham, *Old and New Problems and Results in Combinatorial Number Theory*, Monographies de L'Enseignement Mathématique, no. 28, Université de Genève, Geneva, 1980. `https://mathweb.ucsd.edu/~ronspubs/80_11_number_theory.pdf`.

[2] G. Martin, Denser Egyptian fractions, *Acta Arith.* **95** (2000), 231–260. doi:10.4064/aa-95-3-231-260; arXiv:math/9811112.

[3] E. S. Croot III, On unit fractions with denominators in short intervals, *Acta Arith.* **99** (2001), 99–114. arXiv:math/9904181.

[4] T. F. Bloom, with an appendix by T. F. Bloom and B. Mehta, On a density conjecture about unit fractions, *J. Eur. Math. Soc.* **27** (2025), no. 11, 4563–4589. doi:10.4171/JEMS/1456; arXiv:2112.03726.

[5] Y. P. Liu and M. Sawhney, On further questions regarding unit fractions, *Int. Math. Res. Not.* 2026, no. 2, rnaf382. doi:10.1093/imrn/rnaf382; arXiv:2404.07113v1 (the numbered results cited here refer to arXiv v1).

[6] W. Hoeffding, Probability inequalities for sums of bounded random variables, *J. Amer. Statist. Assoc.* **58** (1963), 13–30.

[7] A. Hildebrand and G. Tenenbaum, Integers without large prime factors, *J. Théor. Nombres Bordeaux* **5** (1993), no. 2, 411–484. doi:10.5802/jtnb.101.