hexagon:2611.01130v1

Local Capacity and Thresholds for Simultaneous Egyptian Fraction Packings

pengcheng chen

26 November 2026

Abstract

We study pairwise disjoint sets of denominators in a common interval whose reciprocal sums are prescribed rational numbers. Put $L = \log N$ and $U = \log L$. For targets in a fixed compact subinterval of $(1/d, \infty)$ satisfying a common denominator condition, our sufficient and necessary bounds give the same first two terms in the loss from the harmonic mass of $[d, N]$, namely $U^2/(2L) + U \log U/L$. The construction also bounds the number of terms in each group. For unit targets, let $\mathfrak{C}_N$ be the maximum reciprocal mass compatible with the zero-residue conditions at the primes above $\sqrt{N}$. For all but $O(X^{11/12})$ integers $N \le X$, both the maximum number of disjoint unit sums and the largest integer representable by a single subset are equal to $\lfloor \mathfrak{C}_N \rfloor$. If X_j denotes either the least height permitting j disjoint unit sums or the least height permitting one representation of j, then

$$\log X_j = j - \gamma + \frac{(\log j)^2}{2j} + \frac{\log j \, \log\log j}{j} + O\Big(\frac{\log j \, \log\log\log j}{j}\Big).$$

Keywords: Egyptian fractions, reciprocal sums, disjoint packings, smooth numbers, modular subset sums.
MSC 2020: 11D68, 11B75, 05D99.

1 The packing problem and a uniform construction

Write $[d, N] = \{d, d+1, \dots, N\}$, $\mathcal{H}(S) = \sum_{n \in S} 1/n$, and $H_N = \mathcal{H}([1, N])$, with $H_0 = 0$. All logarithms are natural. For large N put

$$L = \log N, \quad U = \log L, \quad E(N) = U^2/L, \quad y = N^{1/20},$$

and define the common denominator

$$D_N = \operatorname{lcm}(1, \dots, \lfloor y \rfloor) \prod_{\substack{y < p \le \sqrt{N} \\ p \text{ prime}}} p. \tag{1}$$

For a rational s, let $\operatorname{den}(s)$ denote its positive reduced denominator. An integer is y-powersmooth if all its prime-power divisors are at most y. Thus a divisor of D_N may also contain primes in $(y, \sqrt{N}]$, each to exponent at most one.

Erdős and Graham [5] asked how many disjoint subsets of $[1, N]$ can each have reciprocal sum 1. Let this maximum be $k(N)$. As recorded in Problem 296 of Bloom's database [2], Hunter and Sawhney observed that a greedy application of Bloom's density theorem [1] gives $k(N) = (1 - o(1)) \log N$. For the related problem of representing a single integer, Croot [3] proved capacity bounds with coefficients $9/2$ and $1/2$ on the scale $E(N)$, and proposed sharpness of the latter (see the discussion after his main theorem). In the inverse-height formulation, Yokota [8, Theorem 1 and Corollary 1] obtained $\log T_j \le j - \gamma + (\pi^2/3 + o(1))(\log j)^2/j$, where

T_j is the least height permitting a representation of the integer j. This records the improvement without suppressing the integer rounding in the counting formulation. The distinction between representing an integer and partitioning a representation into prescribed sums is important. A union of disjoint unit-sum sets has integer reciprocal sum, but we do not assume any converse.

Here we establish a more general packing theorem, for a list of rational targets which may depend on N. It simultaneously controls total reciprocal mass and individual cardinalities.

Theorem 1 (Simultaneous rational packing)**.** *Fix constants $0 < a \le A < \infty$ and a positive integer d with $1/d < a$. There are constants $C > 0$ and N_0 depending only on a, A, d with the following property. For $N \ge N_0$, let $r_1, \dots, r_m \in [a, A] \cap \mathbb{Q}$ be a finite nonempty list with $\mathrm{den}(r_i) \mid D_N$. Set $R = \sum_i r_i$ and $W = H_N - H_{d-1}$.*

(i) *If*

$$R \le W - \frac{U^2}{2L} - C\frac{U \log U}{L}, \tag{2}$$

then there exist pairwise disjoint $S_1, \dots, S_m \subseteq [d, N]$ with $\mathcal{H}(S_i) = r_i$ and

$$|S_i| \le \frac{N r_i}{\max\{R, L/2\}} + C\frac{N}{L^2} \qquad (1 \le i \le m). \tag{3}$$

(ii) *Conversely, if pairwise disjoint sets with these reciprocal sums exist (without any cardinality restriction), then*

$$R \le W - \frac{U^2}{2L} + C\frac{U}{L}. \tag{4}$$

All constants are uniform over the list of targets, including their numerators and denominators.

In particular, for each fixed $\epsilon > 0$, mass at most $W - (1/2 + \epsilon)E(N)$ suffices, while mass greater than $W - (1/2 - \epsilon)E(N)$ is impossible once N is sufficiently large. The coefficient $1/2$ is sharp in this threshold sense; integer rounding must be retained in specializations to equal targets.

Proof strategy. For the initial uniform bound, the height bound of Yokota [7], refined by Liu and Sawhney [6], supplies short modular covers. We use one cover for each prime above $\sqrt{N}$, cancelling that prime from the sum of its entire available block of multiples. This is related to the deletion strategy of Croot [3]; the additional packing step is to retain these corrected blocks as indivisible objects. The blocks are then ordered by cardinality divided by reciprocal mass and allocated by minimum normalized load. This controls the cardinality of every group via a tail-mass estimate. Disjoint medium-prime repair pools and prime-marked smooth closures handle the individual residuals. We give all these steps explicitly below.

Further results. Section 8 applies the construction to fixed rational targets, unit sums, integer sums and packings with a prescribed bound on the number of terms. Theorem 12 determines the next term in the capacity loss. Its proof combines short modular covers for almost all primes with an upper bound on the number of distinct rational reciprocal sums. For unit targets, Theorem 16 uses the actual minimum deletion in each large-prime block and retains individual length bounds. Corollary 17 then locates the heights at which integer rounding can affect the answer.

2 Arithmetic preliminaries

The denominator of a sum or difference of rationals divides the least common multiple of their denominators. We use this both for powersmoothness and for the stronger condition of dividing D_N. Congruences of rationals modulo a prime p are taken in $\mathbb{Z}_{(p)} = \{a/b \in \mathbb{Q} : p \nmid b\}$, with reduction into $\mathbb{F}_p$.

Lemma 2 (Modular coverage)**.** *For every sufficiently large prime p and every $a \in \mathbb{F}_p$, some subset $J \subseteq [1, q(p)]$ satisfies $\sum_{b\in J} b^{-1} = a$ in $\mathbb{F}_p$, where*

$$q(p) = \left\lceil \log p \, (\log\log p)^4 \right\rceil .$$

Proof. The case $a = 0$ is immediate. For the representative $1 \le a < p$, Theorem 1.5 of Liu and Sawhney [6] gives a distinct unit-fraction representation of a/p with largest denominator at most $p \log p (\log\log p)^3 (\log\log\log p)^{O(1)} \le pq(p)$ for all sufficiently large p, uniformly in a. Since $q(p) < p$, no denominator is divisible by p^2. Multiply by p and reduce modulo p: denominators coprime to p contribute zero, and denominators of the form pb contribute b^{-1}. The corresponding distinct cofactors $b \le q(p)$ give the assertion. Alternatively, the bound of Yokota [7], as also recorded in [6, Section 1], gives a distinct unit-fraction representation of a/p with largest denominator $\ll p \log p (\log\log p)^4 (\log\log\log p)^2$, uniformly in a. Since $(\log\log\log p)^2 = o(\log\log p)$, the implied constant is absorbed for sufficiently large p by taking the denominator bound $pq_Y(p)$, where $q_Y(p) = \lceil \log p (\log\log p)^5 \rceil < p$. The identical reduction therefore proves the modular coverage statement with $q(p)$ replaced by $q_Y(p)$. This alternative suffices for all conclusions of Theorem 1; the necessary parameter changes are recorded after (14). □

We record elementary estimates used in the smooth closure and in the construction. The prime number theorem gives $\log \operatorname{lcm}(1, \dots, \lfloor x \rfloor) = (1 + o(1))x$. Also

$$\sum_{\substack{p^b > z \\ b \ge 2}} p^{-b} \ll z^{-1/2}. \tag{5}$$

Indeed, for $p \le \sqrt{z}$ the geometric tail starting at the first $p^b > z$, $b \ge 2$, is at most $2/z$, and for $p > \sqrt{z}$ it is at most $2/p^2$. Summing proves the estimate. For each fixed integer $u \ge 2$ and all sufficiently large T, the interval $(T, 2T]$ contains $\gg_u T$ integers whose prime-power divisors are at most $T^{1/u}$. Indeed the fixed-parameter smooth-number estimate [4, Section II] first gives this lower bound with a restriction on prime factors. The integers excluded by a prime power $p^b > T^{1/u}$, $b \ge 2$, number at most $2T$ times the sum in (5), which is $o(T)$. Taking $u = 20$ and $T = N/2$ supplies $\gg N$ y-powersmooth integers in $[N/2, N]$.

We also need a fixed $C_0 > 1$ for which the $X^{1/5}$-powersmooth integers in $[X, C_0X]$ have total reciprocal mass at least $3/2$ for all large X. For $T = 2^jX$ the estimate with $u = 6$ supplies at least cT integers in $(T, 2T]$, each of weight at least $1/(2T)$, with an absolute $c > 0$. Choose an integer J with $Jc/2 \ge 3/2$. For $0 \le j < J$ one has $(2^jX)^{1/6} \le X^{1/5}$ for all sufficiently large X. The intervals are disjoint, so $C_0 = 2^J$ suffices.

Lemma 3 (Uniform smooth closure)**.** *There are absolute constants $C_0 > 1$ and X_0 such that for every integer $X \ge X_0$ and every rational $s \in [1/2, 3/4]$ whose reduced denominator is $X^{1/5}$-powersmooth, there exists a finite set $T \subseteq [X, C_0X]$ of $X^{1/5}$-powersmooth integers with $\mathcal{H}(T) = s$.*

Proof. This is a uniform specialization of Croot's Proposition 2 [4, pp. 103–104], with parameter $\varepsilon = 1/20$. We give the needed special case of his Fourier argument, including a short-interval

separation estimate, so that the uniformity and the divisor multiplicity bound are explicit. Put $P = \operatorname{lcm}(1, \dots, \lfloor X^{1/5} \rfloor)$, so that $Ps \in \mathbb{Z}$, and let

$$Y_0 = \{n \in [X, (1 + 1/\log X)X] : n \mid P\}.$$

Here and below $\|t\|$ denotes the distance of a real number t to the nearest integer.

Short-interval separation. We first prove that, uniformly for integers h with $X/2 < |h| \le P/2$, there are $\gg X^{3/4}$ elements $n \in Y_0$ with $\|h/n\| \ge X^{-1/4}/2$. This adapts Croot's Lemma 5 and its proof [4, pp. 112–114]. Let $M = X^{3/4}(\log X)^3$. There are $\gg M$ integers $v \in [M, 2M]$ dividing P: the $M^{1/4}$-powersmooth integers in this interval alone number $\gg M$ by the preceding powersmooth estimate, and $M^{1/4} < X^{1/5}$ for large X. For each such v, set

$$\mathcal{M}(v) = \{n \in Y_0 : v \mid n,\ \omega(n/v) \le 3\},$$

where ω counts distinct prime factors. We claim $\operatorname{lcm} \mathcal{M}(v) = P$.

To prove this identity, fix a maximal prime power $p^\alpha \le X^{1/5}$ for a prime $p \mid P$, and let p^e be the exact power of p dividing v (with $e = 0$ if $p \nmid v$). Then $e \le \alpha$. Put $z = (X/(vp^{\alpha-e}))^{1/2}$ and choose distinct primes ℓ_1, ℓ_2 not dividing vp in the interval

$$[z, z\sqrt{1 + 1/\log X}].$$

Such primes exist uniformly in v and p. Indeed

$$\frac{X^{1/40}}{\sqrt{2}(\log X)^{3/2}} \le z \le \frac{X^{1/8}}{(\log X)^{3/2}}.$$

The interval has length $\asymp z/\log X$ and its endpoints have logarithms comparable to $\log X$. The classical quantitative PNT, $\pi(t) = \operatorname{Li}(t) + O(t \exp(-c\sqrt{\log t}))$ for some $c > 0$, gives $\gg z/(\log X)^2$ primes in this interval, uniformly. This number exceeds $\omega(vp) + 2 = O(\log X)$ for all sufficiently large X. In particular the primes can be chosen outside the factors of both v and p. They are less than $X^{1/5}$. The integer $vp^{\alpha-e}\ell_1\ell_2$ belongs to Y_0 (its prime-power factors do not exceed $X^{1/5}$), is divisible by p^α, and belongs to $\mathcal{M}(v)$. Every maximal prime power in P thus divides some member of $\mathcal{M}(v)$, and all members divide P. The claimed identity follows.

Consider the open interval $I = (h - X^{3/4}, h + X^{3/4})$. Since $X/2 < |h| \le P/2$ and $X^{3/4} < X/2$ eventually, I contains no multiple of P. If every member of $\mathcal{M}(v)$ divided an integer in I, all these integers would have to coincide: they are multiples of v and $v > 2X^{3/4}$. The lcm identity would then force a multiple of P in I. Consequently for each v there is a witness $n \in \mathcal{M}(v)$ dividing no integer in I.

For any fixed witness n, its possible cofactors n/v are divisors q of n with $\omega(q) \le 3$. The total number of nontrivial prime powers dividing n is $\Omega(n) \le \log n/\log 2 = O(\log X)$, where Ω counts prime factors with multiplicity. Each such q is a product of at most three of these prime powers. Hence

$$\#\{q : q \mid n,\ \omega(q) \le 3\} \le \sum_{k=0}^{3} \Omega(n)^k = O((\log X)^3).$$

This is the multiplicity bound needed here in place of the stronger little-o estimate displayed on p. 113 of [4]. Counting the $\gg M$ carriers with this multiplicity bound gives $\gg M/(\log X)^3 =\gg X^{3/4}$ distinct witnesses. For each, the distance from h to the nearest multiple of n is at least $X^{3/4}$, so $n \le 2X$ implies $\|h/n\| \ge X^{-1/4}/2$. This proves the separation estimate.

Fourier closure. Take C_0 as in the dyadic mass bound above. Let Y consist of the $X^{1/5}$-powersmooth integers in the shortest initial interval of $[X, C_0X]$ whose mass is at least $2s$. The preceding smooth estimate guarantees its existence for all large X, uniformly in s. Then

$$0 \le \mathcal{H}(Y) - 2s \le 1/X, \qquad |Y| \ge X.$$

Moreover Y includes every $X^{1/5}$-powersmooth integer in $[X,(1+1/\log X)X]$, since even the unrestricted mass of this short interval is $o(1)$, whereas $2s \geq 1$.

Write $e(t)=\exp(2\pi it)$ and $F(h)=\prod_{n\in Y}(1+e(h/n))$. The number of subsets of Y with reciprocal sum congruent to s modulo 1 is

$$P^{-1}\sum_{-P/2<h\leq P/2} e(-hs)F(h).$$

For $|h|\leq X/2$, the real part of $e(-hs)F(h)$ is nonnegative: the factors $\cos(\pi h/n)$ are nonnegative, and the remaining phase is $\pi h(\mathcal{H}(Y)-2s)$ of absolute value at most $\pi/2$. At $h=0$ the contribution is $2^{|Y|}$. For $X/2<|h|\leq P/2$, the separation estimate gives $\gg X^{3/4}$ integers $n\in Y_0\subseteq Y$ with $\|h/n\|\geq X^{-1/4}/2$. Using $|1+e(h/n)|=2|\cos(\pi h/n)|$ and $|\cos(\pi t)|\leq \exp(-c\|t\|^2)$ for an absolute $c>0$, together with $\log P=(1+o(1))X^{1/5}$, we obtain

$$|F(h)|\leq 2^{|Y|}\exp(-c''X^{1/4})=o(2^{|Y|}/P).$$

The sum of the absolute values of all high-frequency terms is $o(2^{|Y|})$, as there are at most P frequencies. The nonnegative real parts at low frequencies, including $2^{|Y|}$ at zero, therefore imply positivity of the Fourier sum for all large X, uniformly in s. Finally every subset sum is nonnegative and at most $2s+1/X<s+1$, while $s-1<0$. Therefore a congruence to s modulo 1 is an exact equality. This proves the claimed uniform form without any dependence on the varying numerator or denominator. □

Lemma 4 (Large-prime mass)**.** *If $w=U+O(\log U)$ and $w>0$, then*

$$\sum_{Ne^{-w}<p\leq N}\frac{H_{\lfloor N/p\rfloor}}{p}=\frac{w^2}{2L}+O(U/L). \tag{6}$$

This is exactly the mass of all integers in $[1,N]$ divisible by at least one of the indicated primes.

Proof. These primes exceed $\sqrt{N}$, so their sets of multiples are disjoint. Uniformly for $p\leq N$, $H_{\lfloor N/p\rfloor}=\log(N/p)+O(1)$. By the prime number theorem and partial summation, the error contributes $O(w/L)$. To make the main term quantitative use the classical PNT error $\pi(t)=\mathrm{Li}(t)+O(t\exp(-c\sqrt{\log t}))$ (any error $O(t/(\log t)^3)$ suffices here). For $g(t)=\log(N/t)/t$, partial summation on $[Ne^{-w},N]$ bounds the resulting error by $O((w+w^2)e^{-c'\sqrt{L}})$. Consequently the main term equals

$$\int_{Ne^{-w}}^{N}\frac{\log(N/t)}{t\log t}\,dt=\int_0^w\frac{v}{L-v}\,dv=\frac{w^2}{2L}+O(w^3/L^2).$$

Since $w\asymp U$, the stated bound follows. □

3 The arithmetic obstruction

We prove Theorem 1(ii) first. Let r be any rational whose reduced denominator divides D_N, and suppose $\mathcal{H}(S)=r$ for $S\subseteq[d,N]$. If a prime $p>10N/L$ divides some element of S, set $q=\lfloor N/p\rfloor$ and $J=\{b:pb\in S\}$. For large N, $p>\sqrt{N}$ and $p\nmid \mathrm{den}(r)$. Multiplying by p and reducing modulo p gives

$$\sum_{b\in J}b^{-1}=0\quad\text{in }\mathbb{F}_p.$$

The set $J\subseteq[1,q]$ is nonempty. Put $D=\mathrm{lcm}(1,\dots,q)$; the positive integer $D\sum_{b\in J}1/b$ must be divisible by p. But $q<L/10$, $p\nmid D$, and the PNT gives $\log D\leq 2q+O(1)$. Hence this integer is at most $DH_q\ll N^{1/5}(1+\log L)<p$, a contradiction.

None of the multiples of primes above $10N/L$ can therefore occur in any of the representations. For large N all these multiples lie in $[d,N]$, and their total mass by Lemma 4 with $w=U-\log 10$ is $U^2/(2L)+O(U/L)$. Summing the representations proves (4). The argument also applies to any single integer reciprocal sum, even one depending on N and not bounded above.

4 Common repair pools

For the rest of the proof assume (2). We use

$$Q = \left\lceil LU^4 \right\rceil, \qquad Z = N/Q, \qquad Q_0 = \left\lceil L^{12} \right\rceil. \tag{7}$$

Thus $2Q_0 < y < \sqrt{N} < Z < N$ for sufficiently large N. If $R < L/2$, append copies of a fixed positive integer $b \geq a$ until the total target mass $\widetilde{R}$ first reaches $L/2$. The resulting targets are bounded by $\max(A, b)$ and their denominators still divide D_N. The total is less than $L/2 + b$, so (2) still holds for the enlarged list for large N. If $R \geq L/2$, take $\widetilde{R} = R$. Relabel the enlarged list as $r_1, \ldots, r_m$, keeping the original targets first; the auxiliary groups will be discarded at the end. Thus throughout the construction

$$L/2 \leq \widetilde{R} \leq L + O_d(1), \qquad m = O_a(L), \qquad \widetilde{R} \geq \max\{R, L/2\}. \tag{8}$$

All implicit constants in the construction may depend on the fixed a, A, d and the fixed choice of b.

Lemma 5 (Powersmooth dilation)**.** *Let $A \in \mathbb{N}$ satisfy $1 \leq A \leq 2Q_0$ and let $p > y$ be a prime with $pA \leq N$. There is a positive integer t such that*

$$\frac{N}{B_0 pA} \leq t \leq \frac{N}{pA}, \qquad ta \text{ is } y\text{-powersmooth for every } 1 \leq a \leq A,$$

where B_0 is an absolute constant. In particular $p \nmid ta$.

Proof. Let B_0 be the largest of the first 40 primes. For each such prime ℓ allow powers up to the largest $\ell^e \leq y/A$. The product of the 40 maximal powers is $\gg (y/A)^{40} > N$ for all sufficiently large N. Starting at 1, multiply by the allowed prime factors one at a time until the next multiplication would exceed $N/(pA)$. The last product t is within a factor B_0 of this target. Every prime-power component of t is at most y/A. For $a \leq A$, multiplying by a increases any one prime-power component by at most A, including primes not dividing t. Thus ta is y-powersmooth. □

Lemma 6 (Disjoint covers for medium primes)**.** *There exist pairwise disjoint sets $\mathcal{L}_1, \ldots, \mathcal{L}_m$ of primes in $(Q_0, 2Q_0]$, each of size $\lfloor L^{10} \rfloor$, with the following property. For every prime $y < p \leq \sqrt{N}$, every i and every $a \in \mathbb{F}_p$ there is $J \subseteq \mathcal{L}_i$ with $\sum_{\ell \in J} \ell^{-1} = a$ in $\mathbb{F}_p$.*

Proof. By the PNT there are $\gg Q_0/\log Q_0$ primes in $(Q_0, 2Q_0]$. By (8), $m \lfloor L^{10} \rfloor \ll_a L^{11} = o(Q_0/\log Q_0)$, so we can choose disjoint sets of the indicated size. Choose each $\ell \in \mathcal{L}_i$ independently with probability $1/2$. At frequency $h \in \{1, \ldots, p-1\}$ the modulus of the Fourier coefficient of the resulting subset sum is

$$\prod_{\ell \in \mathcal{L}_i} |\cos(\pi z_\ell / p)|,$$

where z_ℓ is the centered integer representative of $h\ell^{-1} \pmod p$. If $|z_\ell| < p/L^3$, then $\ell z_\ell = h + jp$ with $|j| \leq 2Q_0/L^3 + 1$. For each such j the integer $h + jp$ is nonzero and of absolute size $O(NQ_0)$, so at most $O(L/\log Q_0)$ distinct primes $\ell \geq Q_0$ divide it. The number of exceptional ℓ is therefore $O(Q_0/(L^2 \log Q_0)) = o(|\mathcal{L}_i|)$, uniformly in p, i, h. All other factors are at most $\exp(-c/L^6)$ for an absolute $c > 0$. Hence every nonzero Fourier coefficient has modulus at most

$$\exp(-c'L^4) < N^{-2}$$

for some $c' > 0$ depending only on a. Fourier inversion shows that the probability of every specified residue is at least $p^{-1}(1 - (p-1)N^{-2}) > 0$. □

4.1 Reserving the repair pools

Choose distinct marker primes $P_1, \dots, P_m$ in $(\sqrt{N}, 2\sqrt{N}]$, possible by the PNT. For each medium prime $y < p \le \sqrt{N}$ fix t_p from Lemma 5 with $A = 2Q_0$, and reserve the sets

$$\mathcal{C}_{i,p} = \{pt_p\ell : \ell \in \mathcal{L}_i\}. \tag{9}$$

Each element lies in $[N/(2B_0), N]$. Its only prime factor greater than y is p, occurring to the first power; the remaining cofactor is y-powersmooth. Hence all these sets are mutually disjoint, first by their unique large prime, then by disjointness of the sets $\mathcal{L}_i$ for a fixed p. None contains a marker prime as a factor. Their total mass, summed over all i, p, is

$$\mathcal{H}(\bigcup_{i,p} \mathcal{C}_{i,p}) \ll Q_0/\sqrt{N}. \tag{10}$$

Indeed there are at most $O(\sqrt{N}Q_0)$ candidates, each of weight $O(1/N)$.

Independently choose a set $\mathcal{R}$ of $\lceil N/L^4 \rceil$ y-powersmooth integers in $[N/2, N]$, using the smooth-number estimate of Section 2. It has mass between L^{-4} and $O(L^{-4})$, and is disjoint from every set in (9) by powersmoothness.

4.2 The available base set

Let V be obtained from $[d, N]$ by deleting all integers having a prime factor greater than Z, all integers divisible by $p^b > y$ with $b \ge 2$, all multiples of any marker P_i, and both reserved pools $\bigcup_{i,p} \mathcal{C}_{i,p}$ and $\mathcal{R}$. The first restriction deletes mass

$$\frac{U^2}{2L} + O(U \log U/L),$$

by Lemma 4 with $w = \log Q = U + 4 \log U + o(1)$. By (5) the total mass deleted by all other restrictions is at most

$$\begin{aligned} &\sum_{\substack{p^b > y \\ b \ge 2}} p^{-b} H_{\lfloor N/p^b \rfloor} + \sum_{i=1}^{m} P_i^{-1} H_{\lfloor N/P_i \rfloor} + O(Q_0/\sqrt{N} + L^{-4}) \\ &\qquad \ll Ly^{-1/2} + mL/\sqrt{N} + Q_0/\sqrt{N} + L^{-4} = o(U/L). \end{aligned}$$

Overlaps only reduce the deletion cost. Consequently

$$\mathcal{H}(V) = W - \frac{U^2}{2L} + O_{a,A,d}(U \log U/L). \tag{11}$$

5 Large-prime block cancellation

This step differs from repairing the groups individually. For each prime $\sqrt{N} < p \le Z$ not equal to a marker, put

$$V_p = \{n \in V : p \mid n\}.$$

The blocks V_p are disjoint, since a number at most N cannot be a multiple of two distinct primes greater than $\sqrt{N}$. Write $\sigma_p = \mathcal{H}(V_p)$. The reduced denominator of σ_p has p-adic valuation at most one, so $p\sigma_p \in \mathbb{Z}_{(p)}$.

Lemma 7 (Block cancellation)**.** *For each such p there is a subset $D_p \subseteq V_p$ such that the reduced denominator of $\mathcal{H}(V_p \setminus D_p)$ is coprime to p, and*

$$\mathcal{H}(D_p) \le B_0 Q H_Q/N. \tag{12}$$

Moreover, every prime power dividing that reduced denominator is either at most y or is a prime at most $\sqrt{N}$.

Proof. Apply Lemma 5 with $A = Q$, obtaining a multiplier t_p. Its hypotheses hold because $pQ \le ZQ = N$. We first check the crucial availability condition

$$\{pt_p a : 1 \le a \le Q\} \subseteq V_p. \tag{13}$$

Indeed each candidate is between $p > d$ and N for large N. Its only prime factor greater than y is $p \le Z$, which occurs to the first power. Its cofactor $t_p a$ is y-powersmooth, so it passes the higher prime-power restriction. It is not divisible by any marker, because p is not a marker and all other primes are at most y. It cannot belong to any medium repair pool, whose unique prime factor greater than y is at most $\sqrt{N}$, or to the powersmooth pool $\mathcal{R}$. This verifies all the restrictions defining V and proves (13).

Since $p > \sqrt{N}$, Lemma 2 applies and $q(p) \le Q$ for large N (the defining function is increasing). Choose $J_p \subseteq [1, Q]$ with

$$\sum_{a \in J_p} a^{-1} = t_p(p\sigma_p) \quad \text{in } \mathbb{F}_p.$$

Set $D_p = \{pt_p a : a \in J_p\}$. The denominators in this set are distinct, and by (13) are all available. In the local ring $\mathbb{Z}_{(p)}$,

$$p\big(\sigma_p - \mathcal{H}(D_p)\big) = p\sigma_p - t_p^{-1} \sum_{a \in J_p} a^{-1} \in p\mathbb{Z}_{(p)}.$$

It follows that $\sigma_p - \mathcal{H}(D_p) \in \mathbb{Z}_{(p)}$, which proves cancellation. The cost is at most $(pt_p)^{-1} H_Q \le B_0 Q H_Q / N$. The remaining sum has denominator dividing the least common multiple of the remaining denominators. None of those denominators has a second prime greater than $\sqrt{N}$; all their higher prime powers are at most y. Together with cancellation, this proves the final assertion. □

Put $B_p = V_p \setminus D_p$. The total deletion cost is

$$\sum_p \mathcal{H}(D_p) \ll \frac{Z}{L} \frac{QH_Q}{N} = O(U/L), \tag{14}$$

using $Z = N/Q$, $\pi(Z) \ll Z/L$ and $H_Q \ll U$.

Alternative height-bound input (Yokota). All the preceding estimates, and hence the rest of the proof, remain valid without invoking Theorem 1.5 of [6]. Use the modular coverage variant $q_Y(p) = \lceil \log p (\log\log p)^5 \rceil$ from [7] in Lemma 2, replace Q in (7) by $\lceil LU^5 \rceil$, and retain $Z = N/Q$, $Q_0 = \lceil L^{12} \rceil$ and every other parameter. We still have $q_Y(p) \le Q \le 2Q_0$ for $p \le N$, $\sqrt{N} < Z$ for large N, and $pQ \le N$ for $p \le Z$, as required in Lemma 7. In the base-mass calculation only the value of $w = \log Q$ changes: it is now $U + 5\log U + o(1)$. Thus the first deletion has mass

$$\frac{(\log Q)^2}{2L} + O(U/L) = \frac{U^2}{2L} + O(U \log U/L),$$

with all other deletions unchanged in order. The total block-cancellation cost in (14) remains $O(U/L)$, since $ZQ = N$ and $H_Q = O(U)$. The subsequent allocation and exact-closure estimates depend only on these bounds and the unchanged medium-prime and marker pools. This gives the same main theorem and all its corollaries, with possibly different implicit constants. The main route below retains the original choice $Q = \lceil LU^4 \rceil$.

Retain each nonempty B_p as an indivisible object. Every element of V not divisible by a prime greater than $\sqrt{N}$ is retained as a singleton object. Write $\mathcal{B}$ for this family of mutually disjoint blocks. By (11) and (14),

$$\sum_{B \in \mathcal{B}} \mathcal{H}(B) = W - \frac{U^2}{2L} + O_{a,A,d}(U \log U/L). \tag{15}$$

In particular, if the constant in (2) is chosen large enough, the mass is at least $\widetilde{R}$, including in the padded case. Each block has reciprocal sum with reduced denominator dividing D_N. Also, every block originating from a large prime p has mass at most $H_N/\sqrt{N}$ and all its denominators exceed $\sqrt{N}$. These properties also hold if such a block contains just one element.

Only one modular cover per large prime is required, regardless of how many target representations are requested. Medium primes instead have separate repair covers for every group.

6 Allocation with cardinality control

For each nonempty block $B \in \mathcal{B}$ put

$$w(B) = \mathcal{H}(B), \qquad c(B) = |B|, \qquad x(B) = c(B)/w(B).$$

The effective denominator $x(B)$ is the harmonic mean of the elements of B, and $x(B) \le N$. Order the blocks by nondecreasing $x(B)$ (ties broken arbitrarily). At each step allocate the next block to a group having least normalized load, defined as its accumulated mass divided by its target r_i.

6.1 Stopping with small positive residuals

Set $\delta = L^{-6}$ and $\eta = a^{-1}L^{-10}$. All blocks with $x(B) \le L^{10}$ are singletons, and their total mass is $O(U)$. Before each addition the least normalized load is at most $O(U/\widetilde{R}) = o(1)$, since the weighted average uses weights r_i. The normalized mass of an added singleton is at most $1/(da)$. Consequently, during and after this initial stage, every normalized load is at most

$$\beta + o(1), \qquad \beta = 1/(da) < 1. \tag{16}$$

All later blocks have normalized mass at most η: this is immediate for singletons of denominator greater than L^{10} and follows from $w(B) \le H_N/\sqrt{N}$ for large-prime blocks. Continue in the prescribed order until every normalized load is at least $1 - \delta - \eta$. An addition before stopping increases the least load by at most η, and no load can exceed $1 - \delta$ for large N. Exhausting the blocks before stopping would force their total mass to be less than $\widetilde{R}$, contrary to (15). Thus the process stops. Let A_i be the union of blocks allocated to group i, and let $\Delta_i = r_i - \mathcal{H}(A_i)$. We have

$$r_i\delta \le \Delta_i \le r_i(\delta + \eta), \qquad \operatorname{den}(\Delta_i) \mid D_N. \tag{17}$$

The denominator assertion follows from the same property of the blocks and targets.

6.2 The tail estimate

We give a length estimate for this allocation. Denote the ordered blocks by $B_1, B_2, \ldots$, and let T be the final allocated index. Write $u_i(t)$ for group i's normalized load after the first t allocations and $v(t) = \min_i u_i(t)$; v is nondecreasing.

Choose a fixed integer $D \ge d$ such that $1/(aD) < 1/16$. Call a group special if it receives a denominator at most D; there are at most D such groups. All such denominators occur as singleton blocks in the initial stage. The normalized load of every nonspecial group at the end of the initial stage is at most $1/16 + o(1) < 1/8$, by the same minimum-load argument as in (16).

Let t_0 be the number of blocks in the full ordered list with $x(B) \le N^{1/3}$. All such blocks are singletons. If $T \le t_0$, then $|A_i| \le r_i N^{1/3} = o(N/L^2)$, which suffices. Suppose $T > t_0$. By (15) and the construction, the total mass of the first t_0 blocks is $L/3 + O_d(1)$: the mass removed from the entire interval $[d, N]$ before forming the blocks is $O(U^2/L) = o(1)$, and none of the

large-prime blocks has a denominator at most $N^{1/3}$. We claim that $v(t_0) > 1/8$. Otherwise every nonspecial group has load at most $1/8+\eta$ at t_0: if it has had no addition after the initial stage, this follows from its initial bound; otherwise apply the minimum-load rule at its last addition and monotonicity of v. The special groups have combined mass $O_{a,A,d}(1)$, since their loads are at most 1 and there are at most D of them. The total mass at t_0 would therefore be at most $(1/8+\eta)\widetilde{R} + O_{a,A,d}(1) \le L/8 + O_{a,A,d}(1)$, a contradiction.

It follows that every nonspecial group has received an addition since the initial stage by time t_0. For every $t_0 \le t \le T$ the minimum-load rule at its last addition gives

$$v(t) \le u_i(t) \le v(t) + \eta \qquad \text{for every nonspecial group } i. \tag{18}$$

For $t_0 \le t \le T$ let $W_t = \sum_{t<k\le T} w(B_k)$, and let $W_{i,t}$ be the part of this tail mass allocated to group i. Let R_0 denote the sum of the nonspecial targets, so $R_0 = \widetilde{R} - O_{a,A,d}(1) \asymp L$. Summing (18) at the two endpoints gives

$$W_t \ge R_0(v(T) - v(t) - \eta).$$

For every group i, including the special ones, if it receives a block after time t, its final load is at most $v(T)+\eta$ by the rule at its last addition. Its load at t is at least $v(t)$. If it receives no such block its tail mass is zero. Thus

$$W_{i,t} \le r_i(v(T) - v(t) + \eta).$$

In both cases we obtain the uniform tail inequality

$$W_{i,t} \le \frac{r_i}{R_0} W_t + 2r_i\eta. \tag{19}$$

To pass from mass to cardinality, write $x_k = x(B_k)$. Discrete summation by parts gives

$$\begin{aligned}\sum_{\substack{k>t_0 \\ B_k \text{ allocated to } i}} c(B_k) &= x_{t_0+1}W_{i,t_0} + \sum_{t=t_0+1}^{T-1} (x_{t+1}-x_t)W_{i,t} \\ &\le \frac{r_i}{R_0}\sum_{k>t_0}^{T} c(B_k) + 2r_i\eta x_T \\ &\le \frac{r_i N}{R_0} + 2r_i\eta N.\end{aligned}$$

Here the coefficients are nonnegative because the x_k are nondecreasing, and the blocks are disjoint subsets of $[d,N]$. The prefix contributes at most $r_i N^{1/3}$. Since $R_0 = \widetilde{R} - O_{a,A,d}(1)$ and $\widetilde{R} \asymp L$, we have proved

$$|A_i| \le \frac{N r_i}{\widetilde{R}} + O_{a,A,d}(N/L^2) \tag{20}$$

for every group. This is a cardinality bound for the actual denominators, not merely a bound on the number of allocated blocks.

7 Simultaneous exact closure

7.1 Medium-prime cancellation within each group

Fix a group i. For a prime $y < p \le \sqrt{N}$ dividing the reduced denominator of its current residual Δ, use Lemma 6 to choose $J \subseteq \mathcal{L}_i$ with

$$\sum_{\ell\in J} \ell^{-1} = t_p(p\Delta) \quad \text{in } \mathbb{F}_p.$$

Subtract $\sum_{\ell\in J}(pt_p\ell)^{-1}$ from the residual, and add the corresponding elements of $\mathcal{C}_{i,p}$ to group i. As in Lemma 7, the new residual is in $\mathbb{Z}_{(p)}$. Its denominator has not acquired any other prime above y, since the cofactors are y-powersmooth. No higher prime power above y can be introduced either. Thus each medium prime needs at most one cancellation for each group, in any order. All these additions are mutually disjoint by (9), and are disjoint from the base sets by construction.

The mass of all additions in this stage, even over all groups, is bounded by (10), which is $o(a\delta)$. In particular no residual becomes nonpositive. Denote the selected denominators for group i by F_i and its new residual by Δ_i'. It has y-powersmooth reduced denominator, and

$$a\delta/2 \le \Delta_i' \le 2r_i\delta. \tag{21}$$

7.2 Shared smooth pool and prime-marked final repairs

For each group i in turn, add unused denominators from $\mathcal{R}$ until the residual θ_i satisfies

$$\frac{1}{2P_i} \le \theta_i \le \frac{1}{2P_i} + \frac{2}{N}. \tag{22}$$

The initial residual exceeds this interval for large N by (21). Add a term only while the residual exceeds the upper endpoint. Every term in $\mathcal{R}$ is at most $2/N$, so no addition can cross the lower endpoint. The sum of the residuals before this stage is at most $2\widetilde{R}\delta = O(L^{-5})$, whereas $\mathcal{H}(\mathcal{R}) \ge L^{-4}$. The pool cannot be exhausted before all groups reach their target intervals. Let E_i denote the terms assigned to group i in this stage. The sets E_i are disjoint, and each θ_i still has y-powersmooth reduced denominator.

Finally, take

$$X_i = \lfloor N/(C_0P_i)\rfloor, \qquad s_i = P_i\theta_i.$$

Then $X_i \asymp \sqrt{N}$, $y < X_i^{1/5}$, and (22) gives $1/2 \le s_i \le 1/2 + 2P_i/N \le 3/4$ for large N. Multiplication by P_i cannot enlarge the reduced denominator. Lemma 3 therefore supplies a set $T_i \subseteq [X_i, C_0X_i]\cap\mathbb{N}$ of $X_i^{1/5}$-powersmooth integers with $\mathcal{H}(T_i) = s_i$. Put

$$G_i = \{P_it : t \in T_i\}.$$

All its elements are at most $P_iC_0X_i \le N$, and $\mathcal{H}(G_i) = \theta_i$. Every element has exactly one marker prime as a divisor, namely P_i, since $X_i^{1/5} \ll N^{1/10} < P_j$ for every j. The sets G_i are thus disjoint from one another. Every previous pool and base object avoids all marker primes, so the G_i are also disjoint from everything previously used. This argument does not require their numerical intervals to be disjoint.

All new denominators are also at least d for large N. The medium-prime stage uses at most $O(\sqrt{N}Q_0)$ terms in total. For group i, the smooth-pool stage uses at most $N\Delta_i' = O_{a,A,d}(N/L^6)$ terms, because each term is at least $1/N$, and the marked closure uses at most $C_0X_i = O(\sqrt{N})$ terms. These three quantities are $o(N/L^2)$. Consequently

$$S_i = A_i \cup F_i \cup E_i \cup G_i$$

are pairwise disjoint subsets of $[d, N]$ with $\mathcal{H}(S_i) = r_i$ and $|S_i| \le Nr_i/\widetilde{R} + O_{a,A,d}(N/L^2)$. Discarding the auxiliary groups and using (8) proves (3). Enlarging C to accommodate the finitely many estimates completes the proof of Theorem 1. □

8 Applications

It is convenient to use the abbreviations

$$B_+(N;C) = \frac{U^2}{2L} + C\frac{U\log U}{L}, \qquad B_-(N;C) = \frac{U^2}{2L} - C\frac{U}{L}.$$

Both are positive for sufficiently large N and are asymptotic to $E(N)/2$.

8.1 Fixed rational targets and the disjoint unit-sum problem

For fixed rational $r > 0$, let $k_r(N)$ be the maximum number of disjoint subsets of $[2, N]$ with reciprocal sum r. Define

$$d_r = \max\{2, \lceil 1/r \rceil\}, \qquad W_r(N) = H_N - H_{d_r - 1}.$$

Only denominators at least d_r can be used, because every unit fraction is positive.

Corollary 8 (Fixed rational packing)**.** *For every fixed rational $r > 0$ there is $C_r > 0$ such that for all large N,*

$$\left\lfloor \frac{W_r(N) - B_+(N; C_r)}{r} \right\rfloor \le k_r(N) \le \left\lfloor \frac{W_r(N) - B_-(N; C_r)}{r} \right\rfloor. \tag{23}$$

Moreover the lower bound is attained by sets having cardinalities at most

$$\frac{rN}{\log N} + O_r\big(N/(\log N)^2\big). \tag{24}$$

If $N_j(r) = \min\{N : k_r(N) \ge j\}$, then

$$\log N_j(r) = rj + H_{d_r - 1} - \gamma + \frac{(\log j)^2}{2rj} + O_r\left(\frac{\log j \log\log j}{j}\right). \tag{25}$$

Proof. The reduced denominator of a fixed rational divides D_N for all sufficiently large N. The obstruction of Section 3 applies directly, giving the upper bound. For the lower bound put $m = \lfloor (W_r(N) - B_+(N; C_r))/r \rfloor$. If $1/d_r < r$, apply Theorem 1 with $a = A = r$, $d = d_r$, and m copies of r. If $1/d_r = r$, first take the singleton $\{d_r\}$ and apply the theorem with $d = d_r + 1$ to $m - 1$ copies of r. The available mass is reduced by exactly r, so the sufficient inequality is preserved. Since $mr = L + O_r(1)$ in both cases (and likewise $(m-1)r = L + O_r(1)$), (3) gives (24).

For (25), apply the upper bound at $N = N_j(r)$ and the lower bound at $N - 1$. By minimality, $k_r(N - 1) < j$. The two inequalities imply

$$B_-(N; C_r) \le W_r(N) - rj < B_+(N - 1; C_r) + 1/N.$$

In particular $\log N = rj + O_r(1)$ by the harmonic asymptotic $H_N = \log N + \gamma + O(1/N)$. Replacing $\log N$ by $rj + O_r(1)$ in the small corrections yields (25), including the stated error. The use of $N - 1$ is essential to avoid an unjustified removal of the floor. □

Corollary 9 (Erdős Problem 296)**.** *Let $k(N)$ count the maximum number of disjoint unit-sum subsets of $[1, N]$, and $N_j = \min\{N : k(N) \ge j\}$. For some absolute constant $C > 0$ and all large N,*

$$\lfloor H_N - B_+(N; C) \rfloor \le k(N) \le \lfloor H_N - B_-(N; C) \rfloor. \tag{26}$$

In particular, $k(N) \in \{\lfloor H_N \rfloor - 1, \lfloor H_N \rfloor\}$, and

$$\log N_j = j - \gamma + \frac{(\log j)^2}{2j} + O\left(\frac{\log j \log\log j}{j}\right). \tag{27}$$

Consequently the normalized threshold correction $j(\log N_j - j + \gamma)/(\log j)^2$ tends to $1/2$.

Proof. Any unit-sum subset containing 1 is the singleton $\{1\}$, so $k(N) = 1 + k_1(N)$. Apply Corollary 8 with $d_1 = 2$; for the threshold formula replace its index by $j - 1$. The two alternatives for the integer value follow since $0 < B_-(N; C) \le B_+(N; C) < 1$ eventually. No claim is made that $H_N - k(N) \sim E(N)/2$ for arbitrary N, since the integer part can contribute a much larger error. □

Corollary 10 (Integer reciprocal sums; Problems 308 and 309)**.** *Let $h(N)$ be the maximum integer representable as the reciprocal sum of a subset of $[1, N]$, and T_j the least height permitting a representation of the integer j. For large N the positive representable integers form an initial segment, whose length $h(N)$ satisfies the same two bounds as* (26)*. Furthermore*

$$\log T_j = j - \gamma + \frac{(\log j)^2}{2j} + O\Big(\frac{\log j \log\log j}{j}\Big). \tag{28}$$

Proof. Unions of a disjoint unit-sum packing give every integer from 1 to $k(N)$. The obstruction of Section 3 applies to all integer reciprocal sums, so $h(N) \le \lfloor H_N - B_-(N; C)\rfloor$. For large N the two floors in (26) differ by at most one; hence the positive representable integers form an initial segment. The smallest missing positive integer is $h(N) + 1$, not $h(N)$. For the height formula, $T_j \le N_j$ gives the upper estimate, while the obstruction at $N = T_j$ gives $H_{T_j} - j \ge B_-(T_j; C)$. From $T_j \le N_j$ and $j \le H_{T_j}$ it follows that $\log T_j = j + O(1)$; substituting gives the lower estimate and proves (28). □

The sharp leading coefficient in Corollary 10 is the one proposed by Croot [3]. The initial-segment property itself was already implied by Croot's earlier bounds. The present conclusion concerns the sharp threshold and disjoint packings, not a new proof of the original first-order questions.

8.2 Length-restricted Egyptian-fraction hypergraphs

Let $\nu_K(r; N)$ be the maximum number of pairwise disjoint subsets of $[2, N]$ of cardinality at most K and reciprocal sum r, and $\nu_K(r)$ its analogue without a bound on the denominators. These are the matching numbers of the corresponding hypergraphs.

Corollary 11 (Sharp capacity with a cardinality restriction)**.** *For every fixed rational $r > 0$ there is a constant $C_r > 0$ such that if N is sufficiently large and*

$$K \ge \frac{rN}{\log N} + C_r \frac{N}{(\log N)^2}, \tag{29}$$

then the same two bounds as (23) *hold for $\nu_K(r; N)$. Moreover, uniformly for sufficiently large K, N,*

$$\nu_K(r; N) = \frac{\log \min\{N, K\log K\}}{r} + O_r(1), \qquad \nu_K(r) = \frac{\log K + \log\log K}{r} + O_r(1). \tag{30}$$

Proof. The lower bound under (29) is exactly the construction in Corollary 8, with its cardinality estimate (24); the upper bound follows from $\nu_K(r; N) \le k_r(N)$. For the uniform formula, m disjoint edges with at most K vertices each have total reciprocal mass $mr \le H_{mK+1} - 1$, whence $m = O_r(\log K)$ and

$$mr \le \log K + \log\log K + O_r(1).$$

For example, the first estimate follows by using $\log m \le (r/2)m + O_r(1)$ in $mr \le \log K + \log m + O(1)$; then substitute it back. The height bound also gives $mr \le H_N - 1$. These prove the upper bounds.

For the lower bound choose a sufficiently small fixed $c_r > 0$ and set

$$X = \lfloor \min\{N, c_r K \log K\}\rfloor.$$

For all large K, N, X is large, and $rX/\log X + C_r X/(\log X)^2 \le K$. This follows because the two functions $x/\log x$ and $x/(\log x)^2$ are eventually increasing and the displayed quantity at $c_r K \log K$ is $(rc_r + o(1))K < K$. Apply the first assertion at height X. It gives at least $\log X/r + O_r(1)$ edges with the desired length, yielding the bounded-height formula. Taking N sufficiently large in terms of K proves the unrestricted formula. □

In particular any fixed $c > 0$ permits the unrestricted two-term asymptotic with height $\lfloor cK \log K \rfloor$, with error $O_{r,c}(1)$. For sharp pre-rounding capacity at a given height, (29) is a sufficient length condition, not a characterization of the optimal length threshold in all parameter ranges.

9 The next term in the packing threshold

For sufficiently large arguments let $\log_1 x = \log x$ and $\log_{j+1} x = \log(\log_j x)$. Put

$$B_*(N) = \frac{U^2}{2L} + \frac{U \log U}{L},$$
$$B_{\mathrm{low}}(N; C) = B_*(N) - \frac{U}{L}(2 \log_3 U + C),$$
$$B_{\mathrm{high}}(N; C) = B_*(N) + \frac{U}{L}(4 \log_2 U + C).$$

Theorem 12 (The coefficient of the next capacity term)**.** *Under the hypotheses of Theorem 1, for a constant $C = C(a, A, d)$ and all sufficiently large N, the condition $R \le W - B_{\mathrm{high}}(N; C)$ suffices for a packing with the cardinality bound* (3)*. The existence of any packing implies $R \le W - B_{\mathrm{low}}(N; C)$. The latter implication also holds whenever all target denominators have no prime factor exceeding $\sqrt{N}$, without the other factor restrictions on their denominators.*

We prove the upper and lower bounds for the required loss separately. Only an average modular covering result is needed for the construction, since the full mass of the exceptional prime blocks can be discarded.

9.1 Modular covers outside a sparse set of primes

For an odd prime p let $f(p)$ be the least $q < p$ for which the sums $\sum_{t \in S} t^{-1}$, $S \subseteq [1, q]$, cover $\mathbb{F}_p$, where the inverses are taken in $\mathbb{F}_p$. Set $f(p) = +\infty$ if there is no such q.

Lemma 13 (Almost-all-prime cover)**.** *For an absolute constant C and $P \to \infty$, all but $P^{1/2+o(1)}$ primes $p \in [P, 2P]$ satisfy*

$$f(p) \le C(\log P)(\log_2 P)(\log_3 P)^4.$$

Proof. First, $f(p) \ll (\log p)^4$ for every sufficiently large prime. Write $z = \log p$, $T = \lceil z^4 \rceil$ and take primes $t \in (T/2, T]$. For a nonzero frequency represented by $0 < h < p$, let r_t be the least absolute residue of h/t modulo p. If $|r_t| \le p/(Kz)$, then $tr_t = h + jp$ with $|j| \le 1 + T/(Kz)$. Each nonzero $h + jp$ has at most $\log((T+2)p)/\log(T/2) \ll z/\log T$ prime divisors in the indicated interval. Thus the number of these t is $O(T/(K \log T) + z/\log T)$. For a fixed sufficiently large K, the prime number theorem leaves $\gg T/\log T$ primes t with $\|ht^{-1}/p\| > 1/(Kz)$, using any integer representative of t^{-1} modulo p and writing $\|\cdot\|$ for distance to the nearest integer. The normalized Fourier transform of a random subset sum, with each t included independently with probability $1/2$, has modulus $\prod_t |\cos(\pi h t^{-1}/p)|$. The inequality $|\cos(\pi u)| \le \exp(-c\|u\|^2)$ bounds it by $\exp(-c_K T/(z^2 \log T)) = o(1/p)$ at every nonzero frequency. Fourier inversion gives strictly positive probability at every residue.

We next describe a lifting step. Suppose that, outside at most $X^{1/2+o(1)}$ exceptional primes in each $[X, 2X]$, one has $f(\ell) \le C_1 F(X)$ for $\ell \in [X, 2X]$. Here F is one of $(\log X)^4$ and $(\log X)(\log_2 X)^4$. Put $x = \log P$ and $T = \lceil C_2 F(x) \rceil$, with C_2 large enough to cover the dyadic intervals $[x, 2x], [2x, 4x], [4x, 8x]$. In each of

$$I_1 = (x, 5x/2], \qquad I_2 = (3x, 9x/2]$$

keep the primes ℓ with $f(\ell) \le T$. The excluded primes contribute at most $x^{1/2+o(1)} \log(5x) = o(x)$ to the logarithm of the product in either interval. Consequently the two retained products M_i satisfy $\log M_i = (3/2 + o(1))x$ by the prime number theorem. Here $T = o(x)$. Put $D_i = M_i \operatorname{lcm}(1, \dots, T)$ and let $\mathcal{B}_i$ consist of all denominators ℓt with retained $\ell \in I_i$ and $1 \le t \le T$. Each denominator has a unique prime factor from $I_1 \cup I_2$, so the two pools are disjoint and all their elements are distinct within each pool.

Let A_i be the set of integers $D_i \mathcal{H}(S)$ for $S \subseteq \mathcal{B}_i$. Modulo a retained ℓ, only the terms from its own block contribute, and their possible contributions are $(D_i/\ell) \sum_{t \in J} t^{-1}$. These cover $\mathbb{F}_\ell$ since $\ell \nmid D_i/\ell$. The choices of J for distinct ℓ are independent. By the Chinese remainder theorem A_i maps onto $\mathbb{Z}/M_i\mathbb{Z}$. Select one representative for each residue and denote the resulting set by A_i'. Then

$$|A_i'| = M_i = P^{3/2+o(1)}, \qquad A_i' \subseteq [0, D_i H_{\lceil 5xT \rceil}] \subseteq [0, P^{3/2+o(1)}],$$

using $\log \operatorname{lcm}(1, \dots, T) = O(T) = o(x)$.

For completeness, the required form of the additive large sieve (cf. [9]) is

$$\sum_{\alpha} \left| \sum_{0 \le n \le H} a_n e^{2\pi i n \alpha} \right|^2 \ll (\delta^{-1} + H) \sum_n |a_n|^2$$

for points α separated by δ modulo one. Indeed, on the length-δ arc centered at each point, the central value of $|S|^2$ for the trigonometric polynomial S is bounded by its mean plus the integral of $|(|S|^2)'|$. The arcs are disjoint up to endpoints. Cauchy–Schwarz and Parseval give $\int |(|S|^2)'| \le 2\|S\|_2 \|S'\|_2 \le 4\pi H \sum |a_n|^2$.

The fractions b/p, $1 \le b < p$, $P \le p \le 2P$ prime, are $(4P^2)^{-1}$-separated modulo one. If A_i' occupies at most $p/2$ residues and m_v denotes the number of its elements in residue v, orthogonality and Cauchy–Schwarz give

$$\sum_{b=1}^{p-1} \left| \sum_{n \in A_i'} e^{2\pi i b n/p} \right|^2 = p \sum_v m_v^2 - M_i^2 \ \ge M_i^2.$$

The large sieve bounds the sum over all such primes, including all their nonzero frequencies, by $O(P^2 M_i)$ since $H = P^{3/2+o(1)} < P^2$. There are at most $O(P^2/M_i) = P^{1/2+o(1)}$ exceptional primes per pool. For every other p, both D_i are units modulo p, and each pool's subset sums occupy more than half the field. Their sum fills $\mathbb{F}_p$: for any v the image of the first pool intersects v minus the image of the second. Disjointness of the pools turns this into a subset sum with denominators at most $5xT$. The lifting step therefore replaces $F(P)$ by $O(\log P \, F(\log P))$ with the same exceptional-set exponent. Applying it twice to the Fourier bound proves the lemma. □

9.2 Usable prime blocks and the obstruction from distinct reciprocal sums

Proof of Theorem 12. For sufficiency take $Q = \lceil C_3 L U (\log U)^4 \rceil$, $Z = N/Q$. The constant covers Lemma 13 at every dyadic scale $P \in [\sqrt{N}/2, N]$. Delete the multiples of every prime $p > Z$, as in Section 4. For $\sqrt{N} < p \le Z$ also delete the full blocks of primes for which $f(p) > Q$. In a dyadic interval $[P, 2P]$ such blocks have total mass at most $P^{1/2+o(1)} H_{\lfloor N/P \rfloor}/P$. Summing over the $O(\log N)$ intervals costs $N^{-1/4+o(1)}$ in all. Each remaining non-marker prime in this range has a cover in $[1, Q]$; the dilation in Lemma 5 and the proof of Lemma 7 apply unchanged since $pQ \le N$ and $Q < 2Q_0$. A dilated candidate has only p as a prime factor exceeding y, so none lies in a deleted bad block. All these block repairs together cost $O(U/L)$ by (14). In the underlying base pool, the deletion costs from high prime powers, markers and

reserved repair pools are $o(U/L)$ by the estimates preceding (11). For the primes $p > Z$ use Lemma 4 with

$$w = \log Q = U + \log U + 4\log_2 U + O(1) = U + O(\log U).$$

Their total mass is $w^2/(2L) + O(U/L) = B_{\text{high}}(N;0) + O(U/L)$, because $(\log U)^2 = o(U)$. The available blocks therefore have mass at least $W - B_{\text{high}}(N;C)$ for sufficiently large C. Their effective-denominator ordering and arithmetic properties are those of Section 6; all additional bad blocks were supported above $\sqrt{N}$. The normalized allocation and the closures in Section 7 give (3). When $R < L/2$ the auxiliary targets used for (8) still leave a fixed proportion of the available capacity unused.

For necessity let $\mathcal{V}(q)$ be the set of distinct rational subset sums of $1, 1/2, \dots, 1/q$. Bleicher and Erdős [10, Theorem 3, $r = 3$] prove, for sufficiently large q,

$$\log|\mathcal{V}(q)| \ll \frac{q}{\log q}(\log_3 q)^2.$$

The full iterated-logarithm inequality is also displayed in [11, (2)]. Take $Q_* = \lfloor cLU/(\log_2 U)^2 \rfloor$ for a sufficiently small absolute $c > 0$, so $|\mathcal{V}(Q_*)| \le N^{1/6}$ since $\log Q_* \sim U$ and $\log_3 Q_* \sim \log_2 U$. Put $D_* = \operatorname{lcm}(1, \dots, Q_*)$. For distinct $v, w \in \mathcal{V}(Q_*)$, the integer $D_*(v - w)$ is nonzero and its absolute value is at most $D_* H_{Q_*}$. Its logarithmic absolute value is $O(Q_*)$, so it has $O(1 + Q_*/L)$ distinct prime factors exceeding N/Q_*. Taking the union over all pairs gives at most $N^{1/3}O(1 + Q_*/L) \le \sqrt{N}$ exceptional primes for large N. For $p > N/Q_*$ outside this union, reduction modulo p is injective on $\mathcal{V}(Q_*)$, since $p > Q_*$ and $p \nmid D_*$. The reciprocal sum of a nonempty subset of $[1, \lfloor N/p \rfloor]$ is positive in $\mathbb{Q}$ and hence has a residue different from that of the empty sum. Multiplying an exact representation of a target with denominator prime to p by p now excludes every multiple of p from that representation. Each exceptional block has mass at most $Q_* H_{Q_*}/N$, so their total mass is $O(Q_* H_{Q_*}/\sqrt{N}) = o(U/L)$. All these blocks are disjoint since $N/Q_* > \sqrt{N}$. As

$$\log Q_* = U + \log U - 2\log_3 U + O(1) = U + O(\log U),$$

Lemma 4 shows that the unavailable mass is at least $B_{\text{low}}(N;C)$. The obstructing denominators exceed any fixed d for sufficiently large N, as required. □

Corollary 14 (Inverse heights)**.** *For both $X_j = N_j$ and $X_j = T_j$ one has*

$$\log X_j = j - \gamma + \frac{(\log j)^2}{2j} + \frac{\log j \, \log_2 j}{j} + O\Big(\frac{\log j \, \log_3 j}{j}\Big).$$

In particular the coefficient of $(\log j)\log_2 j/j$ is one.

Proof. As in Corollary 9, the singleton $\{1\}$ and unit targets in $[2, N]$ give $\lfloor H_N - B_{\text{high}}(N;C) \rfloor \le k(N)$. The necessary obstruction applies directly to the union of the unit groups and to any single integer sum, whence $k(N) \le h(N) \le \lfloor H_N - B_{\text{low}}(N;C) \rfloor$. At $N = N_j$ use necessity there and sufficiency at $N_j - 1$ to obtain

$$B_{\text{low}}(N_j;C) \le H_{N_j} - j < B_{\text{high}}(N_j - 1;C) + 1/N_j.$$

For a single integer the obstruction at T_j gives the same lower estimate, and $T_j \le N_j$ gives the upper estimate for $\log T_j$. Both inverse heights satisfy $\log X_j = j + O(1)$ by the harmonic asymptotic and the bounds just obtained. Replacing L by $j + O(1)$ in the displayed losses, and using $H_N = \log N + \gamma + O(1/N)$, gives the asserted expansion. □

Corollary 15 (Local deletion cost)**.** *Let d_p be the minimum used to define $\mathfrak{C}_N$ in the following section. Then*

$$H_N - \mathfrak{C}_N = \frac{U^2}{2L} + \frac{U \log U}{L} + O\Big(\frac{U \log_2 U}{L}\Big).$$

Proof. For the lower bound, every nonexceptional $p > N/Q_*$ in the necessity proof has $d_p = H_{\lfloor N/p \rfloor}$, with the same negligible exceptional mass. For the upper bound work directly on the full blocks. For $p > Z$ and for bad primes discard the whole block. For a remaining good prime in $(\sqrt{N}, Z]$, take $t = \lfloor N/(pQ) \rfloor$; then $N/(2pQ) \le t < p$ and all ptu, $1 \le u \le Q$, are available. A modular cover dilated by t cancels the required residue at cost at most $2QH_Q/N$ per block. The sum is $O(U/L)$. These are upper bounds on the respective minima before any reservation of denominators; no rounding of an integer packing count enters the argument. □

10 Disjoint unit sums at the capacity of the large-prime blocks

For a prime $\sqrt{N} < p \le N$ put $q_p = \lfloor N/p \rfloor$. Since $q_p < p$, all reciprocals below can be reduced modulo p. Define

$$d_p = \min\left\{ \mathcal{H}(D) : D \subseteq [1, q_p], \sum_{b \in [1,q_p] \setminus D} b^{-1} = 0 \text{ in } \mathbb{F}_p \right\}, \qquad \mathfrak{C}_N = H_N - \sum_{\sqrt{N} < p \le N} \frac{d_p}{p}.$$

Empty retained sets are allowed. This minimum is taken over the original block of multiples of p, before any auxiliary denominators are reserved.

Theorem 16 (Attaining the local capacity for unit sums with length control)**.** *There are absolute constants C_1, C_2 such that for all sufficiently large N,*

$$\left\lfloor \mathfrak{C}_N - C_1 N^{-1/12} \right\rfloor \;\le\; k(N) \;\le\; h(N) \;\le\; \lfloor \mathfrak{C}_N \rfloor .$$

The lower bound is attained by pairwise disjoint unit-sum sets, each with at most $N/\log N + C_2 N/(\log N)^2$ denominators.

Proof. If a subset of $[1, N]$ has integral reciprocal sum, multiplication by p followed by reduction modulo p shows that its selected cofactors in $[1, q_p]$ have reciprocal sum zero in $\mathbb{F}_p$. The blocks for $p > \sqrt{N}$ are disjoint. The total mass available to such a subset is therefore at most $\mathfrak{C}_N$. Taking unions of disjoint unit-sum sets proves both upper bounds.

Put $Q_0 = \lceil L^{12} \rceil$ and $Q = 2Q_0$. For every $\sqrt{N} < p \le N$ the first step of Lemma 13 gives $f(p) \ll (\log p)^4$ and hence $f(p) \le Q$ for all sufficiently large N (enlarging N absorbs the implied constant). Thus every residue has a cover in $[1, Q]$, even at an exceptional prime of the stronger average estimate. Corollary 15 also gives $\mathfrak{C}_N = H_N + O(U^2/L)$. Choose an absolute C_1 large enough for the deletion estimate below, reserve the singleton $\{1\}$, and put $m = \left\lfloor \mathfrak{C}_N - C_1 N^{-1/12} \right\rfloor - 1 = L + O(1)$. There will be m further groups.

Reserving the denominators for exact repair. Choose distinct prime markers $N^{1/6} < P_i \le 2N^{1/6}$, $1 \le i \le m$, and put

$$X_i = \lfloor N/(C_0 P_i) \rfloor, \qquad z = (4C_0)^{-1/5} N^{1/6}, \qquad \delta = N^{-1/8}, \qquad \eta = N^{-1/4},$$

where $C_0 > 1$ is the constant of Lemma 3. For all large N,

$$z \le X_i^{1/5} < N^{1/6} < \min_j P_j, \qquad X_i \asymp N^{5/6}. \tag{31}$$

Keep $y = N^{1/20}$ in Lemmas 5 and 6; their cofactors are y-powersmooth and hence z-powersmooth. Choose the m disjoint prime sets supplied by Lemma 6. For every nonmarker

prime $z < p \le \sqrt{N}$ reserve the medium repair pools $\mathcal{C}_{i,p}$ of (9), using the dilation with $A = 2Q_0$. These sets have altogether $O(\sqrt{N}Q_0)$ elements and mass $O(Q_0/\sqrt{N})$, as in (10). Their elements have exactly one prime factor exceeding z, namely p to the first power. Independently reserve a set $\mathcal{R}$ of $\lceil 8LN\delta \rceil$ y-powersmooth integers in $[N/2, N]$; the positive-density estimate preceding Lemma 3 supplies this many. Its mass is at least $8L\delta$ and at most $O(L\delta)$.

Delete from $[2, N]$ the two reserved pools, all multiples of the markers, and all multiples of a prime power $\ell^b > z$ with $b \ge 2$. The mass deleted is

$$O\left(\frac{L}{\sqrt{z}\log z} + \frac{mL}{N^{1/6}} + \frac{Q_0}{\sqrt{N}} + L\delta\right) = O(N^{-1/12}). \tag{32}$$

Indeed $\sum_{\ell^b > z,\, b \ge 2} \ell^{-b} \ll 1/(\sqrt{z}\log z)$. For primes $\ell \le \sqrt{z}$ the geometric tail for each ℓ is at most $2/z$, and $\pi(\sqrt{z}) \ll \sqrt{z}/\log z$. For $\ell > \sqrt{z}$ the bound follows by summing $2/\ell^2$ in dyadic intervals with the same prime-counting estimate. Multiplying by H_N accounts for all multiples. The other terms in (32) bound the two reserved pools and the marker multiples, whether or not any deletions overlap.

Forming indivisible blocks. For $p > N/z$ the entire original block $\{pb : b \le q_p\}$ is still present: $b \le z$, $p^2 > N$, and p is its only prime factor exceeding z. It meets neither a marker multiple nor either reserved pool, since the unique prime exceeding z in a medium-pool denominator is at most $\sqrt{N}$. Delete an actual minimizing cofactor set in this block, at mass cost exactly d_p/p.

For $\sqrt{N} < p \le N/z$ work on the remainder of the block *after* the previous reservations. Here $pQ \le N$ for large N. Lemma 5 with $A = Q$ gives a multiplier t_p with $pt_p \ge N/(B_0Q)$ and all t_pb, $b \le Q$, y-powersmooth. Consequently the entire auxiliary set $\{pt_pb : 1 \le b \le Q\}$ is still available: its unique prime exceeding z is $p > \sqrt{N}$, it contains no marker, and all higher prime powers in its denominators are at most z. If σ_p is the remaining block mass, use the modular coverage proved above to select a subset of $[1, Q]$ whose inverse sum is $t_p p\sigma_p \bmod p$. Deleting the corresponding denominators cancels p at cost at most B_0QH_Q/N per prime. The total cost in this range is thus $O(QH_Q/z) = o(N^{-1/12})$, even counting primes by integers.

Keep every nonempty corrected block indivisible; keep the other surviving denominators as singletons. Every block weight has reduced denominator dividing

$$D' = \operatorname{lcm}(1, \dots, \lfloor z \rfloor) \prod_{\substack{z < p \le \sqrt{N} \\ p \notin \{P_1, \dots, P_m\}}} p.$$

For a large-prime block this follows from cancellation of its unique prime above $\sqrt{N}$ and the deletion of excessive higher powers. Its weight is at most $H_N/\sqrt{N}$. The total surviving weight is at least

$$H_N - 1 - O(N^{-1/12}) - \sum_{p > N/z,\ p \le N} d_p/p \ \ge\ \mathfrak{C}_N - 1 - O(N^{-1/12}) \ \ge\ m.$$

Only primes are summed here and below. This comparison charges a proxy cancellation in the lower range by its full cost; it never assumes that an optimizer for a damaged block remains feasible.

Allocation and cardinalities. Order the blocks by nondecreasing effective denominator $x(B) = |B|/\mathcal{H}(B)$ and give each allocated block to a group of least current mass. All blocks with $x(B) \le N^{1/4}$ are singletons, since a large-prime block has minimum denominator exceeding $\sqrt{N}$. They have aggregate mass at most $L/4 + O(1)$. During this initial stage the least load is at most $1/4 + o(1)$, so every load stays below $3/4 + o(1)$; groups receiving no denominator at most 100 stay below $1/4 + 1/100 + o(1)$. Every later block has mass at most η, by the singleton cutoff or the bound $H_N/\sqrt{N}$. Continue until every load reaches $1 - \delta - \eta$. No load exceeds

$1-\delta$, and exhaustion before stopping contradicts the lower bound m for available mass. Each residual then lies in $[\delta, \delta+\eta]$ and has reduced denominator dividing D'.

For clarity we verify the cardinality bound for this new cutoff. Let t_0 be the last index with $x(B) \le N^{1/3}$ and T the stopping index. The prefix through t_0 consists of singletons and has total mass $L/3 + O(1)$, by (32); none of the large-prime block cancellations changes this prefix. If $T \le t_0$, every group already has $O(N^{1/3})$ terms. Otherwise the minimum load at t_0 exceeds $3/10$: if not, every group receiving no denominator at most 100 has mass at most $3/10+\eta$, by its last addition after the initial stage or its initial bound. The at most 100 exceptional groups have total mass $O(1)$, contradicting the prefix mass. All nonexceptional groups have therefore received an addition after the initial stage by t_0. Write $v(t)$ for the minimum load at time t. From t_0 through T every nonexceptional load lies in $[v(t), v(t)+\eta]$. If W_t is the mass allocated after t through T, and $W_{i,t}$ its share in group i, then

$$W_t \ge (m - O(1))(v(T) - v(t) - \eta), \qquad W_{i,t} \le v(T) - v(t) + \eta.$$

In the first inequality the multiplier denotes the actual number m_0 of nonexceptional groups. The second holds even for an exceptional group by considering its last addition after t, or its zero increment. Hence $W_{i,t} \le W_t/m_0 + 2\eta$. Multiplying this tail inequality by the nonnegative successive increments in $x(B)$ and summing, as in (19)–(20), gives at most $N/m_0 + 2N\eta$ terms after t_0 per group; the prefix supplies at most $N^{1/3}$. Since $m = L + O(1)$, each base set thus has size at most $N/L + O(N/L^2)$.

Exact closure. For each group apply the modular cancellation of Section 7 to every nonmarker prime $z < p \le \sqrt{N}$ in its residual denominator, using its own $\mathcal{C}_{i,p}$. Each such prime appears to exponent at most one. The cofactors of the added denominators are y-powersmooth, so they introduce no other prime power exceeding z. The total mass spent, even over all groups, is $O(Q_0/\sqrt{N}) = o(\delta)$. Every residual is now between $\delta/2$ and 2δ, with z-powersmooth reduced denominator.

Use distinct remaining elements of $\mathcal{R}$ to reduce group i's residual to $[1/(2P_i),\ 1/(2P_i)+2/N]$, adding a term only while above the upper endpoint. Each starts above that interval since $N^{-1/6} = o(\delta)$. No step crosses its lower endpoint. The full pool has mass at least $8L\delta$ while all groups together require at most $2m\delta$, so all reach their intervals. For the final residual θ_i, we have $1/2 \le P_i\theta_i \le 1/2 + 2P_i/N \le 3/4$. Its reduced denominator is z-powersmooth. By (31) and Lemma 3, $P_i\theta_i$ is represented by $X_i^{1/5}$-powersmooth denominators in $[X_i, C_0X_i]$. Multiplying these denominators by P_i closes group i exactly, within $[2, N]$. Each such denominator contains its own marker and no other, and all previously used denominators avoid all markers. Hence disjointness holds throughout. A group gains at most $O(\sqrt{N}Q_0 + N\delta + N^{5/6}) = o(N/L^2)$ terms in the repairs. Restoring $\{1\}$ completes the proof. □

Corollary 17 (Integer values and inverse heights)**.** *Outside a set of at most $O(X^{11/12})$ integers $N \le X$ one has $k(N) = h(N) = \lfloor \mathfrak{C}_N \rfloor$. If $M_j = \min\{N : \mathfrak{C}_N \ge j\}$, then*

$$M_j \le T_j \le N_j \le M_j + O(M_j^{11/12}).$$

In particular $0 \le \log N_j - \log T_j \ll e^{-j/12}$ as $j \to \infty$.

Proof. The quantity $\mathfrak{C}_N$ is the greatest mass of a subset of $[1, N]$ satisfying the separate zero-residue condition at every $p > \sqrt{N}$. These conditions decouple because the prime blocks are disjoint. A feasible subset stays feasible when N is increased: old prime constraints may disappear, and primes newly entering the range divide no old element. Thus $\mathfrak{C}_n$ is nondecreasing. Moreover, if every prime factor of n is at most $\sqrt{n}$, adding n to any previous maximizer shows that $\mathfrak{C}_n - \mathfrak{C}_{n-1} \ge 1/n$.

There is an absolute $c > 0$ such that any interval $(u, v] \subseteq [X, 2X]$ of length $Y \ge 4X^{2/3}$ contains at least cY such integers, for large X. Indeed take distinct primes $X^{3/10} < p < q \le$

$X^{1/3}$. Since $pq \le X^{2/3}$, at least $Y/(2pq)$ multiples of pq lie in the interval. Their cofactor is at most $2X^{2/5} < \sqrt{X}$, and $p, q < \sqrt{X}$. Each integer under consideration has at most three prime divisors in this range and hence is counted at most three times. By partial summation of the PNT the sum of the reciprocals of these primes tends to $\log(10/9) > 0$, while the sum of their reciprocal squares tends to zero. Summing over pairs gives the assertion. Consequently

$$\mathfrak{C}_{\lfloor v \rfloor} - \mathfrak{C}_{\lfloor u \rfloor} \gg Y/X. \tag{33}$$

The theorem leaves an integer undetermined only when $\mathfrak{C}_N \in [j, j + C_1 X^{-1/12}]$ for some integer j and $X \le N \le 2X$. Since $\mathfrak{C}_N = H_N + O(U^2/L)$, only $O(1)$ such bands intersect the range in one dyadic interval. For each band the indices form an interval by monotonicity. By (33) its length is $O(X^{11/12} + X^{2/3}) = O(X^{11/12})$. A dyadic sum gives the asserted exception count.

The first two height inequalities follow from the upper bound and from taking unions. Put $X = M_j$, which exists since $\mathfrak{C}_N = H_N + o(1)$. In (33) take $u = X$ and $v = X + \left\lceil AX^{11/12} \right\rceil$ with a sufficiently large absolute A. The capacity at $\lfloor v \rfloor$ is then at least $j + C_1 X^{-1/12}$, so the theorem gives $k(\lfloor v \rfloor) \ge j$. Finally, $\mathfrak{C}_{M_j - 1} < j \le \mathfrak{C}_{M_j}$ and $\mathfrak{C}_N = H_N + o(1)$ give $\log M_j = j + O(1)$. The logarithmic bound follows. □

Scope. The construction assumes the strict inequality $1/d < a$ and the common denominator condition (1) for varying targets. For a fixed unit-fraction target, Corollary 8 handles the endpoint by reserving a singleton. Theorem 16 leaves an integer undecided when $\mathfrak{C}_N$ lies within $C_1 N^{-1/12}$ above it. Corollary 17 bounds the number of such heights.

Formal verification

The accompanying Lean 4 package verifies the stated packing thresholds, the group-size bounds, the inverse-height estimates and the exceptional-height bound. The source archive lists the verified statements and their axiom dependencies.

For Theorem 1 the formal proof obtains sufficiency from the stronger threshold of Theorem 12. It proves the short modular cover of Lemma 2 directly and establishes the lower bounds for powersmooth sets needed in Lemma 3. For the necessary side of Theorem 12, the package proves its own single-colour support bound and uses injectivity of reduction to control the mass of exceptional prime blocks.

References

[1] T. F. Bloom, *On a density conjecture about unit fractions* (with an appendix by T. F. Bloom and B. Mehta), J. Eur. Math. Soc. **27** (2025), 4563–4589. doi:10.4171/JEMS/1456.

[2] T. F. Bloom, *Erdős Problems*, entries 296, 308, and 309; consulted 25 November 2026 (including indexed copies).

[3] E. S. Croot III, *On some questions of Erdős and Graham about Egyptian fractions*, Mathematika **46** (1999), 359–372. doi:10.1112/S0025579300007828.

[4] E. S. Croot III, *On unit fractions with denominators in short intervals*, Acta Arith. **99** (2001), 99–114. doi:10.4064/aa99-2-1.

[5] P. Erdős and R. L. Graham, *Old and new problems and results in combinatorial number theory*, Monographies de L'Enseignement Mathématique, vol. 28, Université de Genève, 1980.

[6] Y. P. Liu and M. Sawhney, *On further questions regarding unit fractions*, Int. Math. Res. Not. IMRN (2026), no. 2, rnaf382. doi:10.1093/imrn/rnaf382; arXiv:2404.07113.

[7] H. Yokota, *On a problem of Bleicher and Erdős*, J. Number Theory **30** (1988), 198–207. doi:10.1016/0022-314X(88)90017-0.

[8] H. Yokota, *On the number of integers representable as sums of unit fractions, III*, J. Number Theory **96** (2002), 351–372. doi:10.1006/jnth.2002.2797.

[9] H. L. Montgomery and R. C. Vaughan, The large sieve, *Mathematika* **20** (1973), 119–134, doi:10.1112/S0025579300004708.

[10] M. N. Bleicher and P. Erdős, Denominators of Egyptian fractions. II, *Illinois J. Math.* **20** (1976), 598–613, Theorem 3.

[11] S. Bettin, L. Grenié, G. Molteni and C. Sanna, A lower bound for the number of Egyptian fractions, *Math. Comp.*, doi:10.1090/mcom/4190, arXiv:2509.10030, equation (2).