hexagon:2611.01290v1

Disjoint Equal Unit-Fraction Sums: Sharp Asymptotics and Extremal Structure

Pengcheng Chen

Abstract

Let $g(N)$ be the largest number of pairwise disjoint nonempty subsets of $\{1,\dots,N\}$ with equal reciprocal sums, allowing the common value to depend on N. We prove $g(N)\sim N/(2\log N)$ for this free-target problem of Erdős and Graham, with a common sum $1/b_N$ where $b_N \asymp N\sqrt{\log\log N}/\log N$. The proof combines a prime-power resource inequality, synchronous Chinese remainder matching, bounded-length inverse sums, and positive splitting with shared moduli. We also establish structural restrictions on near-extremizers and a simultaneous exact absorption theorem for low-load rational targets. An appendix gives the order-of-magnitude lower bound without the Bourgain–Garaev estimate.

Keywords: Egyptian fractions, disjoint equal sums, inverse sums, Chinese remainder theorem, smooth numbers.

1 Introduction and statement

For a finite set $S\subseteq\mathbb{N}$ write $\sigma(S)=\sum_{n\in S}1/n$. Define

$$g(N)=\max\left\{m:\ \begin{array}{l}\exists\, S_1,\dots,S_m\subseteq[1,N],\\ S_i\neq\varnothing,\quad S_i\cap S_j=\varnothing\ (i\neq j),\\ \sigma(S_1)=\cdots=\sigma(S_m)\end{array}\right\}.$$

Here and below an interval used as a set of integers is intersected with $\mathbb{N}$, and $[K]=\{1,\dots,\lfloor K\rfloor\}$. Erdős and Graham [1, p. 36] asked for the order of magnitude of this function and recorded a lower bound of the form $N\exp(-O(\sqrt{\log N}))$ using strong Δ-systems. The free-target question is also recorded in the commentary to Erdős Problem 296 [2]. It is distinct from the fixed-target question, where all the reciprocal sums must equal 1. For the latter, the maximum number of groups is $(1-o(1))\log N$ by the unit-fraction density theorem of Bloom [3] and the greedy observation attributed to Hunter and Sawhney in [2].

Theorem 1. *The free-target function satisfies*

$$\lim_{N\to\infty}\frac{g(N)\log N}{N}=\frac{1}{2}.$$

More precisely, there are positive integers b_N such that the maximum number of disjoint groups with common sum $1/b_N$ is $(1/2+o(1))N/\log N$, and

$$b_N\asymp\frac{N\sqrt{\log\log N}}{\log N}.$$

The bounded-length inverse-sum reservoir, shared-modulus splitting, and simultaneous allocation proved here address many disjoint targets rather than a single target on a rational lattice.

Theorem 1 is the principal result of this paper. It is proved by a resource upper bound and a matching lower construction with four ingredients: synchronous matching modulo a product of primes, bounded-length inverse-sum corrections for large primes, positive splitting of residual deficits, and sparse exact absorption in reserved local reservoirs.

The structural results are independent of the existence of a sharp lower construction. For a sequence of families with $m \sim N/(2\log N)$, Theorems 12 and 13 show that cancellation between distinct rough-part classes is forced, and that many such classes retain substantial external denominator factors. Theorem 14 proves that single-class constructions cannot reach the upper constant: their limsup constant is at most $1/2 - e^{-\gamma}/4$. Theorems 15, 17 and 18 separate modular feasibility from exact rational feasibility and quantify the limits of dilation and local cancellation.

The general absorption theorem (Theorem 19) gives simultaneous exact representations of prescribed, possibly different rational targets under explicit prime-power load and mass bounds. Appendix A proves $g(N) \asymp N/\log N$ without using the Bourgain–Garaev input; Appendix B records a weaker hierarchy from complete modular systems that does not require smooth-number asymptotics. The first appendix shares the elementary resource bound and binary-grid lemma with the main proof; neither appendix depends on the sharp lower bound.

Notation and analytic inputs

Write $\mathcal{L}(x) = \operatorname{lcm}(1, \dots, \lfloor x \rfloor)$ for $x \geq 1$. For a rational r, $\operatorname{den}(r)$ is its positive reduced denominator, including $\operatorname{den}(0) = 1$. Write $\operatorname{supp}(n)$ for the set of prime divisors of n, and $\omega(n)$ for their number. For an integer $n \geq 2$, $P^+(n)$ and $P^-(n)$ denote its largest and smallest prime factors; take $P^+(1) = 1$ and $P^-(1) = +\infty$. We use v_p for the usual valuation of a rational (with $v_p(0) = +\infty$). A rational with denominator coprime to p is reduced modulo p by inverting its denominator. All logarithms are natural unless the base is explicitly indicated; $\log_{\text{base }2}$ denotes base two. In Appendix B, $\log_j$ denotes the j-fold iterated natural logarithm. Write $H_N = \sum_{n \leq N} 1/n$ and $\vartheta(x) = \sum_{p \leq x} \log p$. Products and sums indexed by p or q in prime ranges are over primes unless otherwise stated. Variables called prime powers are p^j with $j \geq 1$. Rounding real interval endpoints to integers will be implicit.

We use the prime number theorem, its consequences

$$\log \mathcal{L}(x) = \psi(x) \sim x, \quad \sum_{p \leq x} \frac{1}{p} = \log\log x + B_1 + o(1), \quad \prod_{p \leq x} (1 - 1/p)^{-1} \sim e^{\gamma} \log x,$$

and $\sum_{p \leq x} \log p/(p-1) \sim \log x$. We also use Cauchy–Davenport, in the form $|A + B| \geq \min(p, |A| + |B| - 1)$ in $\mathbb{F}_p$. For smooth integers we use the fixed-parameter consequence of the Dickman–de Bruijn estimate (see Hildebrand–Tenenbaum [4, Theorem 1.1]): if $X, z \to \infty$ and $\log X/\log z$ ranges in a fixed compact subinterval of $(1, \infty)$, then

$$\Psi(X, z) := \#\{n \leq X : P^+(n) \leq z\} = X\big(\rho(\log X/\log z) + o(1)\big), \tag{1}$$

uniformly, where the continuous Dickman function ρ is positive at every fixed positive argument. The deeper input for the sharp lower bound is Theorem 8 of Bourgain–Garaev [5], stated in Section 4. The modular-pairing result in the structural part separately uses the additive large sieve. The internal representation and allocation lemmas are proved below.

2 Prime-power resources and the sharp upper bound

Lemma 2 (Prime-power consumption)**.** *For any disjoint nonempty sets $S_1, \dots, S_m \subseteq [1, N]$ with common sum $s = a/b > 0$ in lowest terms, put $x = N/m$ and $Y = \log N$. Then $b \mid \mathcal{L}(x)$ and $m/b \le H_N$. If $q = p^j \mid b$ with $j = v_p(b)$, let $k_i(q)$ count multiples of q in S_i. Then $k_i(q) \ge 1$ and*

$$\#\{i : k_i(q) = 1\} \le N/(pq) + 1, \qquad 2m \le N/q + N/(pq) + 1.$$

In particular $m \le (1/2 + o(1))N/Y$.

Proof. Each set requires a multiple of every prime power in b; their number is at most N/q. This gives the divisibility. The sum of all reciprocal masses is at most H_N, proving $m/b \le H_N$. In a set with exactly one multiple $n = qu$ of q, that term must have $v_p(n) = j$: if it had larger valuation, it would be the unique term of lowest reciprocal valuation. Multiplying the equality by q and reducing modulo p (all remaining terms vanish) gives $u^{-1} \equiv qs \pmod p$, a nonzero residue. Hence n lies in a single residue class modulo pq. This proves the two-multiple bound. To establish the asymptotic claim, it is enough to work along any sequence with $m \ge cN/Y$ for some fixed $c > 0$. Then $x = O_c(Y)$ and $b \ge m/H_N \gg_c N/Y^2$, so $\log b \ge Y - O_c(\log Y)$. The divisibility $b \mid \mathcal{L}(x)$ implies $x \to \infty$. For $p > \sqrt{x}$ dividing b, the divisibility $b \mid \mathcal{L}(x)$ implies $v_p(b) = 1$. We may therefore apply the stated two-multiple bound with $q = p$; rearrangement gives $p \le x(1 + 1/p)/(2 - 1/m) = (1/2 + o(1))x$. Smaller primes contribute at most $\pi(\sqrt{x}) \log x = o(x)$ to $\log b$. The larger ones have exponent one. The PNT gives $\log b \le (1/2 + o(1))x$. If $m = o(N/Y)$ no estimate is needed. This establishes the upper bound along every sequence. □

Corollary 3. *One has $g(N) \le (1/2 + o(1))N/\log N$.*

Proof. Apply Lemma 2 to a maximizing family. □

3 Two probabilistic and algebraic tools

The following elementary versions of concentration and splitting will make the order of the random choices in the construction explicit.

Lemma 4 (Bounded increments)**.** *Suppose Z_j is adapted to a finite filtration $\mathcal{F}_j$, $0 \le Z_j \le L$, and $\mathbb{P}(Z_j > 0 \mid \mathcal{F}_{j-1}) \le b_j$ for deterministic $b_j \ge 0$. If $\sum_j b_j \le B$, then*

$$\mathbb{P}\left(\sum_j Z_j > 4LB\right) \le \exp(-(5 - e)B).$$

Also, for independent Bernoulli(f) variables I_j and weights $0 \le a_j \le A$ with $\sum_j a_j \le M$, one has

$$\mathbb{P}\left(\sum_j a_j I_j > 4fM\right) \le \exp(-(5 - e)fM/A).$$

Proof. The conditional expectation of $e^{Z_j/L}$ is at most $1 + (e-1)b_j \le e^{(e-1)b_j}$. Iteration followed by Markov's inequality proves the first claim. For the second, convexity gives $e^{a_j/A} - 1 \le (e-1)a_j/A$; independence and Markov's inequality give the assertion. The same argument applies to independent indicators with unequal probabilities p_j, replacing fM by an upper bound for $\sum_j a_j p_j$. □

We shall also use the elementary lower-tail Chernoff inequality for a sum of independent Bernoulli variables: the probability of falling below half the mean μ is at most $e^{-\mu/8}$. For a hypergeometric random variable X with mean μ, the bound

$$\mathbb{P}(X \geq k) \leq (e\mu/k)^k \quad (k \geq 1) \tag{2}$$

follows from $\mathbb{E}\binom{X}{k} \leq \mu^k/k!$ and $k! \geq (k/e)^k$. The factorial-moment inequality holds because in a sample without replacement the probability that any k specified marked objects all occur is no larger than the kth power of the individual inclusion probability.

Lemma 5 (Positive splitting with a shared divisor)**.** *Let $D_1, \dots, D_k$ be positive integers all divisible by M, and let $\Delta > 0$ be rational with $\operatorname{den}(\Delta) \mid \operatorname{lcm}(D_1, \dots, D_k)$. If $M\Delta \geq 2k$, there are positive rationals r_j such that*

$$\sum_{j=1}^{k} r_j = \Delta, \qquad \operatorname{den}(r_j) \mid D_j, \qquad |r_j - \Delta/k| \leq 1/M.$$

No coprimality between the D_j is assumed.

Proof. Put $d_j = D_j/M$ and $d = \operatorname{lcm}_j d_j$. Since $\gcd_j(d/d_j) = 1$ and $dM\Delta$ is an integer, Bézout's identity gives integers a_j with $\sum_j a_j/d_j = M\Delta$. Let β_j be the fractional part of a_j/d_j. Then $\sum_j (M\Delta/k - \beta_j)$ is an integer. Round these k real numbers up or down to integers n_j preserving their sum; this is possible by rounding down first and then rounding up the required number of nonintegral entries. Set $r_j = (n_j + \beta_j)/M$. The rounding error is at most $1/M$, and $r_j \geq \Delta/(2k) > 0$. The denominator condition and equality of the total follow directly. □

4 A robust inverse-sum reservoir

We first prove a lemma with an arbitrary fixed range of polynomial scales. Its probabilistic conclusion is needed later to control the load of every individual medium-prime core, not just the total number of added denominators.

Lemma 6 (Robust bounded-length inverse sums)**.** *Fix $0 < \eta < D$ and $0 < \tau < 1$. There are constants $c, \kappa, \omega > 0$ and an integer $L \geq 3$, depending only on these parameters, with the following property. For every sufficiently large prime q and integer X with $q^\eta \leq X \leq q^D$, there is a set $W \subseteq [X/4, X]$ such that*

$$|W| \geq cX, \qquad P^+(w) \leq q^\tau \ (w \in W).$$

For every forbidden set $F_0 \subseteq W$ with $|F_0| \leq \kappa X$ and every $a \in \mathbb{F}_q$, the number of ordered L-tuples of distinct elements of $W \setminus F_0$ with sum of inverses equal to a modulo q is $\gg |W|^L/q$. Under the uniform distribution on these tuples, for every $F \subseteq W$,

$$\mathbb{P}(\textit{some selected coordinate belongs to } F) \ll |F|/X. \tag{3}$$

All constants and estimates are uniform over q, X, a, F_0, F in the stated ranges.

Proof. Write $e_q(t) = \exp(2\pi it/q)$. We use Bourgain–Garaev [5, Theorem 8]: for $1 \le M_1, M_2 < q$, coefficients bounded in modulus by 1, and fixed positive integers k_1, k_2, the absolute value of

$$\sum_{u \le M_1} \sum_{v \le M_2} a_u b_v e_q(hu^{-1}v^{-1}) \qquad (h \neq 0)$$

is at most

$$\ll_{k_1,k_2} M_1 M_2 (\log q)^{C_{k_1,k_2}} \prod_{j=1}^{2} \left(\frac{M_j^{k_j-1}}{\sqrt{q}} + \frac{\sqrt{q}}{M_j^{k_j}} \right)^{1/(2k_1k_2)}. \tag{4}$$

Here C_{k_1,k_2} is a constant. Choose fixed exponents $0 < \beta_1 < \beta_2 < \min(\eta/2, \tau, 1/2)$ such that the closed interval $[\beta_1, \beta_2]$ lies strictly inside $(1/(2k), 1/(2(k-1)))$ for some integer $k \ge 2$. For each power of two A with $q^{\beta_1} \le A$ and $2A \le q^{\beta_2}$ put

$$W_A = \{ab : A \le a < 2A,\ a \text{ prime},\ X/(4A) \le b \le X/(2A),\ P^+(b) \le q^{\beta_1}/2\}.$$

Every element has a unique such factorization since a is its unique largest prime factor. The sets for different A are disjoint. They lie in $[X/4, X]$ and are q^τ-smooth. By the prime number theorem and (1), each has $\gg X/\log q$ elements: the parameter for b is in a fixed compact range greater than 1. There are $\asymp \log q$ admissible A. Hence their union W has size $M \gg X$.

For fixed A, take $M_1 = 2A - 1$, $M_2 = \lfloor X/(2A) \rfloor$ in (4), with indicator coefficients for the two defining factors. To make the uniform saving explicit, set

$$\delta = \min\{\tfrac{1}{2} - (k-1)\beta_2,\ k\beta_1 - \tfrac{1}{2}\} > 0.$$

Since $q^{\beta_1} \le M_1 \le q^{\beta_2}$, the first parenthesis is at most $2q^{-\delta}$ when $k_1 = k$. If $M_2 < q$, choose k_2 with $M_2^{k_2-1} < \sqrt{q} \le M_2^{k_2}$, so the second parenthesis is at most 2. For sufficiently large q, $M_2 \ge q^{(\eta-\beta_2)/2}$, so $k_2 \le 1 + 1/(\eta - \beta_2)$. In particular all implied constants are uniform over a fixed finite set of k_2. If $M_2 \ge q$, group the coefficients by their nonzero residues modulo q and divide by $\lfloor M_2/q \rfloor + 1$. All coefficients then have modulus at most 1 and the interval for the second variable can be taken to be $[1, q-1]$, with $k_2 = 1$. No multiples of q occur in the original support. Restoring the normalizing factor costs $O(M_2)$ rather than $q - 1$. Thus in both cases the bound on the contribution of W_A is $O(Xq^{-\omega_0})$ for some fixed $\omega_0 > 0$, after absorbing the logarithmic factors. Summing over A and decreasing the exponent if necessary gives

$$S(h) := \sum_{w \in W} e_q(h/w), \qquad |S(h)| \le Mq^{-\omega} \quad (h \neq 0) \tag{5}$$

for a fixed $\omega > 0$. The notation h/w in e_q denotes finite-field division.

Let $R_j(a)$ count ordered j-tuples from W, with repetition allowed, with inverse sum a. Orthogonality gives

$$R_j(a) = M^j/q + O(M^j q^{-j\omega}).$$

Choose a fixed integer $L \ge 3$ with $(L-2)\omega > 1$. For $j = L, L-1, L-2$ the error is $o(M^j/q)$ uniformly. The number of L-tuples containing a member of F_0 is at most $2L|F_0|M^{L-1}/q$. Those with a repeated coordinate number at most $2\binom{L}{2}M^{L-1}/q$, by fixing the repeated value and using R_{L-2}. Taking $\kappa \le c/(8L)$ ensures that the forbidden-coordinate count is at most $M^L/(4q)$, since $M \ge cX$. The repeated-coordinate count is $o(M^L/q)$ because $M \to \infty$. The number of valid tuples is therefore $\gg M^L/q$ for all sufficiently large q. The same bound on tuples meeting a given F, divided by this lower bound, proves (3). This argument does not presume that deletion preserves the Fourier estimate itself. □

5 An exact binary grid at a small modulus

For a parameter $z \to \infty$ (local to this section), choose an integer ℓ with $2^\ell \asymp (\log z)^2$, set $J = \ell + \lceil \log_{\text{base }2} z \rceil$, and put

$$Q = 2^J \prod_{3 \le p \le z} p^{a_p}, \qquad a_p = \lfloor \log z / \log p \rfloor,$$

$$D = \{2, 4, \dots, 2^J\} \cup \bigcup_{3 \le p \le z} \bigcup_{j=1}^{a_p} \{p^j 2^{\ell+h} : 0 \le h < \lceil \log_{\text{base }2} p \rceil\}.$$

The elements are distinct divisors of Q, and $\mathcal{L}(z) \mid Q$, $\log Q \sim z$, $\max D = O(z^2 (\log z)^2) \le z^3$. Also $\sigma(D) \le 1 + 2^{1-\ell} \sum_{3 \le p \le z} 1/(p-1) = 1 + o(1)$.

Lemma 7 (Small exact grid)**.** *For the Q, D just constructed, every rational in $[1/8, 3/4] \cap Q^{-1}\mathbb{Z}$ is $\sigma(A)$ for some $A \subseteq D$.*

Proof. The subset sums of Q/d cover every residue modulo Q. For an odd p, start at exponent a_p and go down. At level j the block weights, divided by $p^{a_p - j}$ and reduced modulo p, are a nonzero scalar times reversed binary powers $1, 2, \dots, 2^{\lceil \log_{\text{base }2} p \rceil - 1}$. Their subset sums cover every residue modulo p; blocks at lower levels do not change earlier digits, and every other odd prime-power modulus is unaffected. Finally the pure powers of two fill the 2^J residue by binary representation without disturbing the odd moduli. This proves completeness. All integer subset sums are nonnegative and less than $9Q/8$ for large z. For $Q/8 \le w \le 3Q/4$ the only such integer congruent to w modulo Q is w itself, proving the lemma. □

6 Local exact representations with sparse denominators

This section proves the small-scale representation lemma that will be applied at every reserved anchor. For the sharp construction and the general absorption theorem put

$$Y = \log N, \qquad \Lambda = \log Y, \qquad z = Y^{1/4}, \qquad y = Y^4, \qquad M_0 = \lceil z^3 \rceil, \tag{6}$$

$$K_- = M_0 y^{9/4}, \qquad K_+ = M_0 y^{47/20}. \tag{7}$$

Let P be a square-free product of primes in $[Y/\Lambda^2, 2Y]$. Later it will be a particular product chosen by synchronous matching.

Lemma 8 (Local sparse-denominator absorption)**.** *There are constants $H > 0$ and $\epsilon_0 > 0$ such that the following holds for all sufficiently large N, uniformly for integers $K' \in [K_-, K_+]$ and the products P just specified. Let $x_1, \dots, x_h$ be positive rationals with $b_i = \operatorname{den}(x_i)$ satisfying*

$$b_i \mid \mathcal{L}(y), \quad (b_i, P) = 1, \quad \log b_i \le 6Y, \tag{8}$$

$$\frac{HY}{K'\Lambda} \le x_i \le \frac{Y}{K'}, \qquad \sum_i x_i \le \epsilon_0, \tag{9}$$

$$\#\{i : a \mid b_i\} \le \epsilon_0 K'/a \quad (z < a \le y,\ a \text{ a prime power}). \tag{10}$$

Then there are pairwise disjoint nonempty sets $A_i \subseteq [1, K']$ with $\sigma(A_i) = x_i$, all of whose elements divide $\mathcal{L}(y)$ and are coprime to P.

Proof. We divide the proof into three steps.

Step 1: a complete small-modulus reservoir. Use the construction of Section 5 at $z = Y^{1/4}$ and denote its Q, D by Q_0, D_0. In particular $\mathcal{L}(z) \mid Q_0$, $Q_0 \mid \mathcal{L}(y)$ (the additional power of two is $O(z(\log z)^2) < y$), $\log Q_0 \sim z$, and $\max D_0 \le M_0$. By Lemma 7,

$$[1/8, 3/4] \cap Q_0^{-1}\mathbb{Z} \subseteq \{\sigma(A) : A \subseteq D_0\}. \tag{11}$$

Step 2: correct only the active prime powers. For a target denominator b_i, we must correct the levels $a = p^j$ with $v_p(Q_0) < j \le v_p(b_i)$. Such levels all satisfy $a > z$. The number of levels per target is $O(Y/\Lambda)$: for $p \le \sqrt{z}$ the number is $O(\sqrt{z}\log y) = o(Y/\Lambda)$, while for $p > \sqrt{z}$ the estimate follows from $\sum_p v_p(b_i)\log p \le 6Y$ and $\log p > \Lambda/8$.

For each active level $a = p^j$ use the pool

$$T_a = \{av : K'/(2a) \le v \le K'/a,\ v \mid \mathcal{L}(z),\ (v, p) = 1\}.$$

Different T_a are disjoint: a is the unique complete prime-power component exceeding z. All their elements divide $\mathcal{L}(y)$ and are coprime to P, since a is active only for P-coprime targets and $z < Y/\Lambda^2$. We have uniformly $|T_a| \gg K'/a$. To see this, put $X = K'/a$. The ratio $\log X/\log z$ stays in a fixed compact subinterval of $(1, \infty)$, so (1) counts $\gg X$ smooth integers in $[X/2, X]$. Removing integers with a prime-power component exceeding z costs at most $X\pi(z)/z = o(X)$. Excluding multiples of p costs at most X/p for large p; for each of the finitely many smaller primes, (1) at X/p and $X/(2p)$ shows that the smooth multiples have proportion $1/p + o(1)$ among the smooth integers. Thus the density remains bounded below uniformly. More explicitly, the range of $\log X/\log z$ is contained in $[23 + o(1), 203/5 + o(1)]$ (the lower endpoint comes from K_-/y). On a slightly enlarged compact interval the positive continuous Dickman function has a positive minimum. For each fixed small p, the limiting removal proportion is $1/p$; for large p use X/p. This yields a single positive density bound independent of a and K'.

In any remaining pool of size at least cX for a fixed $c > 0$, there are at least $c'p$ nonzero residue classes modulo p with at least L_0 distinct representatives, where L_0 can be taken to be a sufficiently large fixed integer. Indeed each class has at most $X/p+1$ representatives, and classes with fewer than L_0 contain at most $L_0 p$ elements altogether. We have $X/p \ge K_-/y^2 \asymp Y^{7/4} \to \infty$, so the latter loss is $o(X)$ uniformly. For example, for large N, at least $(c/4)p$ classes are rich, for any fixed L_0 chosen in advance. Apply Cauchy–Davenport repeatedly to the inverses of these rich classes together with 0. Taking $L_0 > 8/c$ ensures that their L_0-fold sumset is all of $\mathbb{F}_p$. Nonzero summands can be chosen using distinct representatives even if the same residue occurs more than once; 0 means no term is needed. Thus every prescribed residue can be realized by at most L_0 distinct unused inverse terms. The constants can be chosen for a remaining density equal to half the initial lower bound.

Process each prime from its highest active exponent downward. If the current deficit e has $v_p(e) \ge -j$, choose at most L_0 unused denominators $p^j v$ from T_{p^j} with $\sum v^{-1} \equiv p^j e \pmod p$. The new deficit has p-adic valuation at least $-(j-1)$, and other prime exponents above those of Q_0 are unchanged. By (10), the total pool consumption is at most $L_0\epsilon_0 K'/a$, so choosing ϵ_0 small enough keeps at least half the initial pool. At completion the residual e_i belongs to $Q_0^{-1}\mathbb{Z}$. Each target has used $O(Y/\Lambda)$ denominators at least $K'/2$, so choosing H large in (9) ensures

$$x_i/2 \le e_i \le x_i. \tag{12}$$

Step 3: exact completion over multiple scales. Consider disjoint full dyadic intervals $[U, 2U)$ contained in

$$[K'/Y^{7/8},\ K'/M_0].$$

The ratio of the endpoints is $Y^{7/8}/M_0 \asymp Y^{1/8}$, so there are $\gg \Lambda$ full dyadic intervals, uniformly in K'. For each interval, select distinct integers $u \in [U, 2U)$ of the form $p_1p_2p_3$, with $p_1 < p_2$ in $[y^{0.7}, y^{0.72}]$ and p_3 in $[U/(p_1p_2), 2U/(p_1p_2))$. The possible exponent of p_3 relative to y lies, up to $o(1)$, between 0.77875 and 0.95, since the exponent of u lies between 2.21875 and 2.35. In particular $p_3 > p_2$ and all three primes are at most y and greater than $2Y$. The prime number theorem and Mertens' estimate give

$$\#\{u\} \gg \frac{U}{\Lambda} \sum_{y^{0.7} \le p_1 < p_2 \le y^{0.72}} \frac{1}{p_1p_2} \gg U/\Lambda, \tag{13}$$

since the sum of prime reciprocals in the indicated exponent interval tends to $\log(0.72/0.7) > 0$. Unique factorization prevents duplication. Each uD_0 lies in $[1, K']$, consists of divisors of $\mathcal{L}(y)$ coprime to P, and the copies are all disjoint. They are disjoint from the T_a because their elements have three distinct prime factors exceeding z, rather than at most one.

By (11), every copy in this dyadic interval can realize every target r/Q_0 with integer

$$A_U := \lceil Q_0/(8U) \rceil \le r \le B_U := \lfloor 3Q_0/(8U) \rfloor .$$

Indeed $ur/Q_0 \in [1/8, 3/4]$. Here $Q_0/U \to \infty$ uniformly, so $B_U \ge 2A_U$. The integer intervals $[kA_U, kB_U]$ $(k \ge 1)$ cover all integers at least A_U. Consequently a residual $e \in Q_0^{-1}\mathbb{Z}$ with $eQ_0 \ge A_U$ can be represented using $O(eU + 1)$ different copies from this dyadic interval, provided that many copies are available. For every residual in (12) and every available scale, $eU \gg Y^{1/8}/\Lambda \to \infty$. Thus (13) guarantees a safe total mass budget c_0/Λ per interval, for a sufficiently small absolute $c_0 > 0$, with all its assigned residuals completable using distinct copies. Each individual $e_i \le Y/K' = o(1/\Lambda)$. Allocate them successively to intervals with this budget, closing an interval when the next residual does not fit. Every closed interval is at least half full for large N. Since there are $\gg \Lambda$ intervals, shrinking ϵ_0 in (9) makes this allocation possible. Using the exact representations within the assigned copies completes all residuals, with no intersections among any of the pools. This proves the lemma. □

7 Synchronous initial matching

Fix $0 < \alpha < 1/10$ for the sharp lower-bound argument. All limits in that argument are first taken with α fixed. The notation $\ll_\alpha$ permits constants depending on α (including the constants in Lemma 6); all $o_\alpha(1)$ tend to zero as $N \to \infty$ with α fixed. We will prove

$$g(N) \ge (1 - C\alpha - o_\alpha(1))\frac{N}{2Y}, \tag{14}$$

with an absolute C. The local scales $y, z, M_0, K_\pm$ remain those in (6)–(7).

Choose y_0 to be the last prime in the initial segment of primes starting at Y/Λ^2 for which

$$P := \prod_{Y/\Lambda^2 \le p \le y_0} p \le N/Y^3.$$

By the prime number theorem, $y_0 \sim Y$, the next prime is $O(Y)$, and

$$N/(CY^4) \le P \le N/Y^3, \qquad R := \lfloor N/y_0 \rfloor \sim N/Y, \qquad P = o(R). \tag{15}$$

In particular P is of the form allowed by Lemma 8. Also

$$H_P := \sum_{p|P} 1/p = O(\log \Lambda/\Lambda). \tag{16}$$

For this particular bound the classical estimate $\pi(x) \ll x/\log x$ and partial summation suffice: the logarithmic length of the prime interval is $O(\log \Lambda)$ and its primes have $\log p \asymp \Lambda$. Thus no quantitative error term in the Mertens asymptotic is required.

Choose an integer $Q \mid \mathcal{L}(z)$ with $PQ \asymp R\sqrt{\Lambda}$. To justify its existence, note from (15) that $R\sqrt{\Lambda}/P$ lies between constant multiples of $Y^2\sqrt{\Lambda}$ and $Y^3\sqrt{\Lambda}$. The product of the largest powers at most z of twenty fixed small primes has order Y^5 up to fixed factors. Successively multiplying in these fixed primes gives a sequence of divisors of $\mathcal{L}(z)$ from 1 to that product with bounded successive ratios. Choose the first divisor in this sequence that is at least $R\sqrt{\Lambda}/P$; bounded successive ratios give the required size. Fix the ordering of the twenty small primes in advance. This choice of Q, like the choice of P, depends only on N, not on α. Set

$$s = \frac{1}{PQ} \asymp \frac{Y}{N\sqrt{\Lambda}}, \qquad w_* := Y/\sqrt{\Lambda}. \tag{17}$$

We use w_* solely as a mass scale; it is not an element of an inverse-sum reservoir.

For each $p \mid P$ choose $c_p \equiv ps \equiv (PQ/p)^{-1} \pmod p$ (a nonzero residue), and take $c \bmod P$ with $c \equiv c_p \pmod p$ for every such prime. We pair distinct integers $u, v \in [1, R]$ by

$$u^{-1} + v^{-1} \equiv c \pmod P. \tag{18}$$

Exclude integers u for which, for some $p \mid P$, either $p \mid u$ or $u^{-1} \equiv c_p \pmod p$. There are at most $2RH_P + 2\pi(y_0) = o(R)$ such integers. On the remaining residue classes modulo P the map $u \mapsto (c - u^{-1})^{-1}$ is a well-defined involution. Every pair of distinct residue classes interchanged by it can be matched using the smaller multiplicity, and each fixed residue class can be paired within itself. In $[1, R]$ all class multiplicities differ by at most one. Thus at most $O(P)$ additional integers remain unmatched. This gives $(1 - o(1))R/2$ pairs.

We discard every pair meeting any of the following sets of bad auxiliary integers:

- $u < \alpha R$;
- $P^+(u) \le N^{\alpha+\alpha^2}$;
- u divisible by a prime in $(N^\alpha, N^{\alpha+\alpha^2}]$ or in $(N^{1-\alpha}, N]$;
- u divisible by q^2 for a prime $q > N^\alpha$;
- u with a prime-power component $p^{v_p(u)} > y$ for some $p \le y$.

The second exceptional set has size $O(\alpha R) + o(R)$ with an absolute implied constant. Indeed for $u \ge R/Y$ in this set, the quantity $\sum_{p \le N^{\alpha+\alpha^2}} v_p(u) \log p$ equals $\log u = (1 - o(1))Y$, while its sum over all $u \le R$ is at most $R \sum_{p \le N^{\alpha+\alpha^2}} \log p/(p-1) = (\alpha + \alpha^2 + o(1))RY$. The smaller u contribute only $O(R/Y)$. The two prime-range exclusions cost $O(\alpha R) + o_\alpha(R)$ by Mertens,

since the sums of reciprocals are respectively $\log(1+\alpha)+o_\alpha(1)$ and $\log(1/(1-\alpha))+o_\alpha(1)$ or less. The square exclusion costs at most $R\sum_{q>N^\alpha} q^{-2} = o(R)$. The last exclusion costs at most $R\pi(y)/y = o(R)$: for each $p \le y$ the first forbidden power exceeds y. The first exclusion is immediate. Deleting one integer destroys at most one pair, so the surviving pairs (u_i, v_i), $1 \le i \le m$, number at least the right-hand side of (14), for some absolute C.

Define

$$C_i = \{pu_i, pv_i : p \mid P\}, \qquad r_i^{(0)} = \sigma(C_i) = H_P(1/u_i + 1/v_i).$$

Every denominator lies in $[1, N]$. Since u_i and v_i are coprime to P, its factorization identifies a unique prime $p \mid P$ with exponent one. Dividing by that prime recovers its auxiliary integer, so no denominator can occur twice, either within a group or across groups. By (18), for $p \mid P$,

$$pr_i^{(0)} \equiv u_i^{-1} + v_i^{-1} \equiv ps \pmod{p}.$$

The remaining deficit is therefore P-integral and satisfies

$$\operatorname{den}(s - r_i^{(0)}) \mid \operatorname{lcm}(Q, u_i, v_i). \tag{19}$$

Indeed the original difference has denominator dividing $\operatorname{lcm}(PQ, Pu_i, Pv_i)$, and all the factors of P cancel. Moreover $u_i, v_i \ge \alpha R$, so (16) and (17) imply $r_i^{(0)} = O_\alpha(Y \log\Lambda/(N\Lambda)) = o_\alpha(s)$.

8 Reserved core-specific anchors

We reserve the final-stage denominators *before* choosing the first-stage inverse-sum corrections. This resolves the potential dependence between a future core and the availability of its reservoir.

Write $\mathcal{Q}$ for the primes in $[N^{\alpha+\alpha^2}, N^{1-\alpha}]$. For every $q \in \mathcal{Q}$, let $X_q = \lfloor N/q \rfloor$. Apply Lemma 6, for example with $\eta = \alpha/2$, $D = 1/\alpha$, and $\tau = \alpha/2$, to fix a reservoir $W_q \subseteq [X_q/4, X_q]$. These parameters are valid for all sufficiently large N, since $X_q \gg N^\alpha$ and $q \ge N^{\alpha+\alpha^2}$. In particular

$$|W_q| \gg_\alpha N/q, \qquad P^+(w) \le N^{\alpha/2} \quad (w \in W_q). \tag{20}$$

All these sets are fixed independently of the following random reservation. Their constants, including the tuple length L, are uniform for fixed α. The actual denominators qw have exactly one prime factor exceeding N^α, namely q to exponent one. Hence qW_q are disjoint as q varies.

Call an integer E a potential core if all its prime factors lie in $(y, N^\alpha]$ and $E \le N^{1-\alpha-\alpha^2}$. The integer 1 is included. For every such E put $K_E = N/E$, so $K_E \ge N^{\alpha+\alpha^2}$. Its eligible anchors are primes t with

$$\frac{2K_E}{K_+} \le t \le \frac{K_E}{2K_-}, \qquad K_{E,t} := \lfloor K_E/t \rfloor. \tag{21}$$

For large N we have $K_- \le K_{E,t} \le K_+$ and $t > N^\alpha$. Define a full reservoir at this anchor by

$$\mathcal{R}_{E,t} = \{Etd : 1 \le d \le K_{E,t},\ d \mid \mathcal{L}(y),\ (d, P) = 1\}.$$

All its elements lie in $[1, N]$. Different (E, t) give disjoint sets: factorization uniquely recovers d from primes at most y, E from primes in $(y, N^\alpha]$, and t from the remaining prime factors. All sets are coprime to P, unlike the initial C_i.

Lemma 9 (Uniform reservation)**.** *Put $f = \Lambda^{-1/4}$. There is a choice of reserved anchors, for all potential cores at once, such that*

$$\sum_{t \text{ reserved for } E} \frac{1}{t} \gg \frac{f\Lambda}{Y} \quad \text{for every potential core } E, \tag{22}$$

and the union $\mathcal{R}^*$ of their reservoirs satisfies

$$|qW_q \cap \mathcal{R}^*| \le 4f|W_q| \quad (q \in \mathcal{Q}). \tag{23}$$

Proof. Independently reserve every eligible (E,t) with probability f. For each fixed E the interval in (21) contains $\gg \Lambda$ full dyadic intervals $[T,2T)$, since its endpoint ratio is $y^{1/10}/4$. Each has $\gg T/Y$ primes, by the prime number theorem, uniformly because $T \gg N^{\alpha+\alpha^2}/Y^{O(1)}$ and $T \le N$. The expected number reserved in it is $\gg fT/Y$. By the Bernoulli lower-tail bound, the probability of reserving fewer than half this expectation is $\exp(- \gg fT/Y)$. There are at most N potential cores, and $O(\Lambda)$ intervals for each. The union bound shows that with probability $1 - o_\alpha(1)$ every such interval supplies $\gg f/Y$ to the reciprocal sum of reserved anchors. Summing yields (22).

For fixed $q \in \mathcal{Q}$, qW_q can meet $\mathcal{R}_{E,t}$ only when $t = q$, by the unique large prime. Let $a_E = |qW_q \cap \mathcal{R}_{E,q}|$, set to zero when q is not an eligible anchor. The number of forbidden elements of qW_q is $\sum_E a_E I_E$, with independent Bernoulli(f) variables, $a_E \le K_+$ and $\sum_E a_E \le |W_q|$. Lemma 4 gives a failure probability for (23) of at most $\exp(-cf|W_q|/K_+)$. By (20) the exponent in magnitude is $\gg_\alpha fN^\alpha/K_+ \gg Y$, and there are at most N primes to consider. Another union bound makes the total failure probability $o_\alpha(1)$. Thus both desired properties hold simultaneously for some realization. Fix that realization before proceeding. □

9 Large-prime correction and conditional core loads

The correction of large primes will now use only denominators outside $\mathcal{R}^*$. We shall exploit the fact that only the two initial source integers u_i, v_i carry large primes, even though the set C_i itself contains many denominators.

9.1 Correction procedure and the residual deficit

For each $q \in \mathcal{Q}$ delete from W_q any w for which $qw \in \mathcal{R}^*$, $(w,P) > 1$, or some $p \le y$ has $p^{v_p(w)} > y$. The first exclusion costs at most $4f|W_q|$; the other two together cost at most $X_q(H_P + \pi(y)/y) = o(X_q)$, using the number of multiples of the corresponding integers up to X_q. Thus the forbidden set has size $o_\alpha(|W_q|)$ uniformly.

By (19) and the initial square exclusion, every prime $q > N^\alpha$ in the deficit has exponent one and belongs to $\mathcal{Q}$. A group may need correction at q only if $q \mid u_i$ or $q \mid v_i$. Since the auxiliary integers are pairwise distinct, the number of such groups is at most $\lfloor R/q \rfloor$. Process these candidate group positions in a deterministic order, treating a position as an empty operation if the current deficit is already q-integral. At each nonempty operation with current deficit e, choose a uniformly random ordered tuple of L distinct, currently available $w \in W_q$ with

$$\sum w^{-1} \equiv qe \pmod{q}.$$

Add the corresponding denominators qw to that group's partial set. Here qe is q-integral and the congruence makes $q(e - \sum_w 1/(qw))$ vanish modulo q, so the new deficit is q-integral. This

operation introduces no other prime factor exceeding N^α, and uses no prime from P. Also the total previous consumption from any qW_q is at most LR/q, which is $O_\alpha(|W_q|/Y)$. By Lemma 6, every required tuple exists, and its conditional selection probability obeys (3), since at every step only an $o_\alpha(1)$ fraction of the original reservoir is forbidden. No correction changes the already fixed residues at the primes of P.

Each group involves at most $2/\alpha$ large primes from u_i, v_i, so at most $\lceil 2/\alpha \rceil L$ correction denominators are added. These are at least $qX_q/4 \gg N$, uniformly. Let C_i' be the resulting partial set, and write

$$\Delta_i = s - \sigma(C_i') = (1 - o_\alpha(1))s \asymp w_*/N. \tag{24}$$

The estimate follows from the initial mass bound and the $O_\alpha(1/N)$ added mass. All deficits are positive for large N. They have no denominator prime above N^α, and their denominators are coprime to P.

Define the source list of group i to consist of u_i, v_i and each of the new correction denominators qw. (It does not contain the individual pu_i, pv_i.) Each list has length k_i with

$$2 \le k_i \le k_\alpha := 2 + \lceil 2/\alpha \rceil L. \tag{25}$$

Write its entries uniquely as $n_{ij} = E_{ij}d_{ij}v_{ij}$, where d_{ij} contains exactly the prime-power components with primes at most y, E_{ij} those with primes in $(y, N^\alpha]$, and v_{ij} those with primes above N^α. All source integers contain a prime at least $N^{\alpha+\alpha^2}$, so E_{ij} is a potential core. By the power exclusions, $d_{ij} \mid \mathcal{L}(y)$, and $(d_{ij}, P) = 1$. Removing all large-prime factors from the denominator bound gives

$$\mathrm{den}(\Delta_i) \mid \mathrm{lcm}_{1 \le j \le k_i}(Q, E_{ij}d_{ij}). \tag{26}$$

Here $\mathrm{lcm}_j(Q, E_{ij}d_{ij})$ denotes the least common multiple of Q and all $E_{ij}d_{ij}$. To verify (26), before large-prime cancellation the denominator divides the least common multiple of Q, the two original auxiliary integers and all correction denominators, by (19). After cancellation, restrict this least common multiple to primes at most N^α.

9.2 Uniform conditional load bounds

The key point is to control the number of source entries with each *exact* medium-prime core. For a potential core E and a prime power $a \le y$, or $a = 1$, define

$$D_{E,a}^{(0)} = \#\{(i,j) : E_{ij} = E,\ a \mid d_{ij}\}, \qquad D_E = D_{E,1}^{(0)}.$$

All counts include multiplicity over the source lists.

Lemma 10 (Core loads after correction)**.** *There is a realization of the first-stage correction choices for which, simultaneously for every potential core E and every a as above,*

$$D_{E,a}^{(0)} \ll_\alpha \frac{K_E \Lambda}{Y^2 a}. \tag{27}$$

Proof. We first count the original source entries. For all $X \le N$,

$$\#\{1 < v \le X : P^-(v) > N^\alpha\} \ll_\alpha X/Y. \tag{28}$$

To see this, extract a largest prime factor q from $v = qb$. Only b with $X/b > N^\alpha$ need be counted; for each, $\pi(X/b) \ll_\alpha X/(bY)$. The sum of $1/b$ over integers supported on primes in

$(N^\alpha, N]$ is bounded by the Euler product $\prod_{N^\alpha<p\le N}(1-1/p)^{-1} \ll_\alpha 1$ by Mertens. This proves (28), including the case of a vacuous range.

For an original auxiliary integer with core E we have $u = Edv \le R$, where $v > 1$ is as in (28) and $P^+(d) \le y$. Since all original source entries are distinct, their count with $a \mid d$ is at most

$$\ll_\alpha \frac{K_E}{Y^2} \sum_{\substack{P^+(d)\le y\\ a\mid d}} \frac{1}{d} \ll_\alpha \frac{K_E\Lambda}{Y^2 a}.$$

Here the reciprocal sum is bounded by $a^{-1}\prod_{p\le y}(1-1/p)^{-1} \ll \Lambda/a$, also for $a = 1$.

For the new source entries consider a candidate correction at $q \in \mathcal{Q}$. To have core E and $a \mid d$, a chosen w must be of the form $w = Ed$, with $d \le K_E/q$, $P^+(d) \le y$, and $a \mid d$. By (3), the conditional probability of selecting at least one such coordinate is at most

$$b_{q,E,a} := C_\alpha \frac{q}{N} \sum_{\substack{d\le K_E/q\\ P^+(d)\le y,\ a\mid d}} 1.$$

This holds also for an empty operation. There are at most R/q preassigned candidate positions at q. Consequently the sum of these deterministic conditional probability bounds over all positions is at most

$$\begin{aligned}\sum_{q\in\mathcal{Q}} \frac{R}{q} b_{q,E,a} &\ll_\alpha \frac{R}{N} \sum_{\substack{P^+(d)\le y\\ a\mid d}} \#\{q\in\mathcal{Q} : q \le K_E/d\}\\ &\ll_\alpha \frac{K_E}{Y^2} \sum_{\substack{P^+(d)\le y\\ a\mid d}} \frac{1}{d} \ll_\alpha \frac{K_E\Lambda}{Y^2 a}.\end{aligned}$$

The prime count in the second line is zero if $K_E/d < N^{\alpha+\alpha^2}$ and is $O_\alpha(K_E/(dY))$ otherwise. This step is uniform in both the core and the power.

At any one operation the number of new entries of the prescribed type is between 0 and L. Apply the adapted-increment part of Lemma 4, with the deterministic upper bound just obtained for B. For a sufficiently large constant in the claimed bound, the failure probability is at most

$$\exp\left(-c_\alpha K_E\Lambda/(Y^2 a)\right).$$

The bound is valid without independence of successive correction choices: the candidate positions are fixed by the initial auxiliary pairs, and at every position the probability estimate holds conditional on the full previous history, including the current deficit and the used denominators. For all potential cores and $a \le y$, the exponent in magnitude is at least $c_\alpha N^{\alpha+\alpha^2}\Lambda/Y^6 \gg Y$. There are at most $N(1+y)$ pairs (E, a), so the union of the failure events has probability $o_\alpha(1)$. Adding the deterministic original-entry count proves the lemma. Fix one successful realization. □

10 Shared moduli and positive packet splitting

We next split each Δ_i into positive packets associated with its source entries. Sharing an auxiliary modulus removes the positivity obstruction that would arise from independently prescribing the

residues of different coprime blocks. The size and the distribution of these auxiliary moduli are both essential.

For each group i, independently of other groups, construct a square-free modulus M_i as follows. Select each prime $p \in (Y^2, Y^3]$ independently with probability $3Y/(p\Lambda)$, form their product, and condition on the event

$$N^2 \le M_i \le N^4. \tag{29}$$

This sampling is well defined for large N. Indeed the unconditioned logarithm has mean

$$\frac{3Y}{\Lambda} \sum_{Y^2 < p \le Y^3} \frac{\log p}{p} = (3 + o(1))Y$$

and variance $O(Y\Lambda)$, using partial summation and $\log p \le 3\Lambda$. Chebyshev gives probability $1 - o(1)$ for (29). In particular after conditioning

$$\mathbb{P}(p \mid M_i) \ll Y/(p\Lambda) \qquad (Y^2 < p \le Y^3), \tag{30}$$

and choices for distinct groups remain independent. We have $M_i \mid \mathcal{L}(y)$ and $(M_i, P) = 1$.

For each source entry put

$$D_{ij} = E_{ij} \operatorname{lcm}(Q, M_i, d_{ij}).$$

All D_{ij} are divisible by M_i, and (26) ensures $\operatorname{den}(\Delta_i) \mid \operatorname{lcm}_j D_{ij}$. Since $M_i\Delta_i \to \infty$ uniformly, and $k_i \le k_\alpha$ is bounded, Lemma 5 supplies positive packets r_{ij} with

$$\sum_{j=1}^{k_i} r_{ij} = \Delta_i, \qquad r_{ij} = \Delta_i/k_i + O(1/M_i) \asymp_\alpha w_*/N, \qquad \operatorname{den}(E_{ij}r_{ij}) \mid \operatorname{lcm}(Q, M_i, d_{ij}). \tag{31}$$

In particular,

$$\operatorname{den}(E_{ij}r_{ij}) \mid \mathcal{L}(y), \qquad (\operatorname{den}(E_{ij}r_{ij}), P) = 1, \qquad \log \operatorname{den}(E_{ij}r_{ij}) \le 6Y \tag{32}$$

for large N: $Q = Y^{O(1)}$, $M_i \le N^4$, and $d_{ij} \le N$. No coprimality of the source cores or their small parts is needed in this splitting.

For a prime power $a \in (z, y]$ define the packet load

$$D_{E,a} = \#\{(i,j) : E_{ij} = E,\ a \mid \operatorname{den}(E_{ij}r_{ij})\}.$$

We use a load bound that is independent of the particular choice of splitting in Lemma 5.

Lemma 11 (Loads including the shared moduli)**.** *There is a realization of the M_i for which, for every potential core E and every prime power $z < a \le y$,*

$$D_{E,a} \ll_\alpha K_E/(Ya). \tag{33}$$

Proof. From the divisor in (31), a must divide at least one of Q, M_i, d_{ij}. The first is impossible because $Q \mid \mathcal{L}(z)$. If a is a proper prime power then the second is impossible because M_i is square-free. In those cases the bound follows from (27). For a prime $a \in (Y^2, Y^3]$, the additional contribution is at most

$$\sum_i h_{i,E}\, \mathbf{1}_{a \mid M_i}, \qquad h_{i,E} = \#\{j : E_{ij} = E\}.$$

The weights satisfy $h_{i,E} \le k_\alpha$ and $\sum_i h_{i,E} = D_E \ll_\alpha K_E\Lambda/Y^2$. The indicators for different i are independent and, by (30), the expectation is $O_\alpha(K_E/(Ya))$. The unequal-probability weighted version of Lemma 4 gives a failure probability $\exp(-c_\alpha K_E/(Ya))$ for a sufficiently large constant in the bound. Since $K_E \ge N^{\alpha+\alpha^2}$ and $a \le Y^4$, a union bound over at most Ny pairs (E,a) establishes (33) simultaneously. Fix one successful realization and then choose the positive splittings. For later use, note explicitly that the entire pre-splitting source contribution is $O_\alpha(K_E\Lambda/(Y^2a)) = o_\alpha(K_E/(Ya))$, while the shared-modulus contribution has the larger bound in (33). The bound holds for every permitted splitting, so there is no dependence on the splitting choices in the concentration argument. □

11 Allocation to anchors and exact completion

At this point all data from earlier random choices, including the packets, are fixed. The final random choice assigns packets to reserved anchors. We check both the total mass constraint and every local prime-power constraint of Lemma 8.

For a potential core E, put $K = K_E$, and for each of its reserved anchors t write $K_t = K_{E,t}$. A packet r_{ij} with $E_{ij} = E$ assigned to t will have local target

$$x_{ij,t} = Et\, r_{ij}. \tag{34}$$

From (31) and $EtK_t = (1+o(1))N$, uniformly,

$$x_{ij,t} \asymp_\alpha w_*/K_t = \frac{Y}{K_t\sqrt{\Lambda}} \ \text{ up to fixed factors.} \tag{35}$$

Multiplication by t can only decrease the reduced denominator of Er_{ij}. In particular (32) implies the local denominator conditions (8). This observation also means that the packet load $D_{E,a}$ is an upper bound for the number of targets needing a after scaling by any anchor.

Let $\eta_0 > 0$ be a sufficiently small constant, depending on α and the constant ϵ_0 in Lemma 8. Create

$$h_t = \lfloor \eta_0 K_t/w_* \rfloor$$

slots at each reserved anchor. The mass of any full set of h_t assigned targets is at most $C_\alpha\eta_0 \le \epsilon_0$ by (35). Also $K_t/w_* \to \infty$ uniformly, so the floor loses at most a vanishing relative amount. The total slots for core E satisfy, by (22),

$$\mathcal{H}_E := \sum_{t \text{ reserved for } E} h_t \gg_\alpha \frac{K}{w_*} \sum_{t \text{ reserved for } E} \frac{1}{t} \gg_\alpha \frac{fK\Lambda^{3/2}}{Y^2}. \tag{36}$$

But Lemma 10 bounds $D_E \ll_\alpha K\Lambda/Y^2$. For each E with a nonempty packet list, the ratio $\mathcal{H}_E/D_E$ is at least a constant times $f\sqrt{\Lambda} = \Lambda^{1/4}$ and thus tends uniformly to infinity. In particular injective assignments to slots exist for every core.

Independently for each core choose a uniform random injection of its D_E labeled packets into its $\mathcal{H}_E$ slots. For each reserved anchor t and each prime power $a \in (z,y]$, let $X_{E,t,a}$ count the assigned packets for which $a \mid \mathrm{den}(Er_{ij})$. It is hypergeometric, with mean

$$\begin{aligned}\mu_{E,t,a} &= h_t D_{E,a}/\mathcal{H}_E \\ &\ll_\alpha \frac{(K_t/w_*)\,K/(Ya)}{fK\Lambda/(w_*Y)} \ll_\alpha \frac{K_t}{af\Lambda} = o_\alpha(K_t/a),\end{aligned}$$

by (33), (36) and $f\Lambda = \Lambda^{3/4} \to \infty$. If $D_{E,a} = 0$ the assertion is immediate. Otherwise (2) gives, uniformly for large N,

$$\mathbb{P}(X_{E,t,a} > \epsilon_0 K_t/a) \le \exp(-cK_t/a)$$

for a fixed $c > 0$. Since $K_t/a \ge K_-/y \asymp Y^{23/4}$, the union bound over at most $N^2 y$ triples (E, t, a) is $o(1)$. Thus there is a fixed injection satisfying

$$X_{E,t,a} \le \epsilon_0 K_t/a \tag{37}$$

for every triple. This does not require independence between the loads for different powers at the same anchor. The random choices are made in the following order: reserve anchors, fix a successful large-prime correction history, fix successful shared moduli, choose positive splittings, and finally fix a successful injection. Each existence argument is conditional on the already fixed preceding data.

For reference the four relevant margins, with α fixed, are

Constraint	Asymptotic margin
Total slots / packets	$\gg_\alpha \Lambda^{1/4}$
Mean marked load / K_t/a	$O_\alpha(\Lambda^{-3/4})$
Individual mass / $(HY/(K_t\Lambda))$	$\gg_\alpha \sqrt{\Lambda}/H$
Individual mass / (Y/K_t)	$O_\alpha(\Lambda^{-1/2})$

All targets assigned to an anchor t of core E now meet the hypotheses of Lemma 8, with $K' = K_t$. Indeed (8) follows from (32); the sum condition follows from the slot bound, and the individual bounds from (35) since $\sqrt{\Lambda} \to \infty$. The prime-power bound follows from (37) because $\mathrm{den}(x_{ij,t}) \mid \mathrm{den}(Er_{ij})$. The lemma yields disjoint local representing sets $A_{ij} \subseteq [1, K_t]$ for the assigned packets, with all elements dividing $\mathcal{L}(y)$ and coprime to P. Define actual absorption sets

$$B_{ij} = \{E_{ij}td : d \in A_{ij}\},$$

using the anchor assigned to that packet. Then $B_{ij} \subseteq \mathcal{R}_{E_{ij},t} \subseteq [1, N]$ and

$$\sigma(B_{ij}) = \frac{\sigma(A_{ij})}{E_{ij}t} = r_{ij}. \tag{38}$$

All actual denominator collisions are excluded for the following reasons. Initially every element of C_i has a prime factor in P, and all first-stage and final-stage additions are coprime to P. Distinct first-stage reservoirs qW_q are disjoint, and repeated choices from any one are forbidden. First-stage additions avoid $\mathcal{R}^*$ by construction. Distinct final-stage reservoirs are disjoint by their unique core and anchor, and within one reservoir the local lemma gives disjoint sets. Hence all initial and added denominators occur in at most one final set and at most once in that set.

Finally put

$$S_i = C_i' \cup \bigcup_{j=1}^{k_i} B_{ij}.$$

These are nonempty pairwise disjoint subsets of $[1, N]$. Equations (24), (31), and (38) give

$$\sigma(S_i) = \sigma(C_i') + \sum_j r_{ij} = \sigma(C_i') + \Delta_i = s.$$

No groups were removed after the initial matching and exclusions. This proves (14).

12 Completion of the sharp asymptotic

For each sufficiently small fixed $\alpha > 0$, the construction gives at least $(1-C\alpha-o_\alpha(1))N/(2\log N)$ groups with the common unit-fraction sum $s = 1/(PQ)$. The absolute constant C comes solely from the initial exclusions in Section 7; all subsequent constants may depend on α without changing that loss. First taking $N \to \infty$ and then $\alpha \to 0$ yields

$$\liminf_{N\to\infty} \frac{g(N)\log N}{N} \geq \frac{1}{2}.$$

Corollary 3 gives the opposite limsup inequality. Moreover, P, Q in (17) were selected independently of α, so the same unit-fraction target has denominator $PQ \asymp N\sqrt{\log\log N}/\log N$ and achieves the asymptotic lower bound. More explicitly, let m_N be the maximum number of disjoint groups at this particular target $s_N = 1/(PQ)$, with P, Q chosen before α. For every fixed sufficiently small α, $\liminf m_N \log N/N \geq (1 - C\alpha)/2$. Taking the supremum over these fixed α gives $\liminf m_N \log N/N \geq 1/2$, and Corollary 3 bounds the limsup by $1/2$. This proves both assertions of Theorem 1.

Extremal structure and obstructions

The notation in the following structural sections is local to those sections: in particular y is a smoothness cutoff comparable to $\log N$ where indicated, not the Y^4 of the sharp lower-bound construction. The results do not use the sharp lower bound.

13 Saturation and forced nonlocal cancellation

In this section let $S_1, \dots, S_m \subseteq [1, N]$ be disjoint nonempty sets of common mass $s = a/b > 0$ in lowest terms, put $Y = \log N$, and assume $m \sim N/(2Y)$ along a sequence $N \to \infty$. Lemma 2 implies that the largest prime $p_0 \mid b$ satisfies $p_0 \sim Y$, and that for any cutoff $y \sim Y$ containing the primes of b,

$$\sum_{p\leq y,\ p\nmid b} \log p = o(Y). \tag{39}$$

Indeed $\log b \geq (1 - o(1))Y$, each prime factor is at most $(1 + o(1))Y$, and the higher powers contribute $o(Y)$. Choose $y = \max(Y, p_0)$. For fixed $0 < \delta < 1/4$ let $\mathcal{Q}_\delta = \{p \mid b : (1 - \delta)Y \leq p \leq Y\}$ and $H = |\mathcal{Q}_\delta| \sim \delta Y/\log Y$. Every such prime has exponent one in b. Put $k_i(p) = |\{n \in S_i : p \mid n\}|$ and $\mathcal{E}_p = \sum_i (k_i(p) - 2)_+$, $\mathcal{U}_p = \lfloor N/p \rfloor - \sum_i k_i(p)$. The exact accounting identity is

$$\mathcal{E}_p + \mathcal{U}_p = \lfloor N/p \rfloor - 2m + \#\{i : k_i(p) = 1\}. \tag{40}$$

Thus uniformly on $\mathcal{Q}_\delta$ the proportion of groups with $k_i(p) \neq 2$ is at most $2\delta/(1 - \delta) + o(1)$, and at p_0 the right-hand side of (40) is $o(m)$. In particular all but $o(N/p_0)$ multiples of p_0 occur in groups using exactly two such multiples. For the last assertion, the number used in groups with at least three equals at most $3\mathcal{E}_{p_0}$, and singleton groups and unused multiples have size $o(m)$.

Write $n = d_y(n)t_y(n)$, where $d_y(n)$ contains the complete prime-power components with primes at most y, and $t_y(n)$ those with primes exceeding y. In a group S_i call its nonempty classes $C = \{n \in S_i : t_y(n) = t\}$ rough-part classes. Put $r_C = \sigma(C)$, and let $e(C)$ be the part of

$\operatorname{den}(r_C)$ supported on primes exceeding y. If $e(C) = 1$, call the class locally cancelled. Genuine cross-class cancellation means that some class has $e(C) > 1$. The need for it is not merely a formal possibility: if $y < q, r < 2y$ are distinct primes, then

$$\frac{1}{qr} + \frac{1}{qr(r-1)} + \frac{1}{q(q-1)(r-1)} = \frac{1}{(q-1)(r-1)}.$$

The first two denominators have rough part qr, the last has rough part q, and the two class masses separately retain q in their denominators. Here $q-1, r-1$ are even and have no prime factor exceeding y.

Theorem 12 (Nonlocality at saturation)**.** *For every fixed $1 < A < 2$, all but $o(m)$ groups have no locally cancelled class with rough part $t \geq N/y^A$. Moreover at least $(e^{-\gamma} - o(1))m$ groups exhibit genuine cross-class cancellation, where γ is Euler's constant.*

Proof. Set $T = N/y^A$ and $L_A = \prod_{p \leq y} p^{\lfloor A \log y / \log p \rfloor}$. If $t \geq T$, every local smooth factor is at most y^A and divides L_A. The higher-power contribution to $\log L_A$ is $O(y^{A/2} \log y) = o(Y)$, so $\log L_A = (1 + o(1))Y$. For a locally cancelled class, $r_C \in L_A^{-1}\mathbb{Z}$ because $tr_C \in L_A^{-1}\mathbb{Z}$ and $(t, L_A) = 1$. Also $0 < r_C \leq s \leq H_N/m = O(Y^2/N)$. Hence there are at most $N^{o(1)}$ possible values of r_C, and $\log \operatorname{den}(r_C) \geq Y - O(\log Y)$. Uniformly, the denominator of each such class lacks only $o(H)$ primes in $\mathcal{Q}_\delta$, by comparing its logarithm with that of L_A.

If a class with this rough part contains just one multiple of some $p \in \mathcal{Q}_\delta$, it is tpv, with $v \leq Z = O(y^{A-1})$, $p \nmid v$. Reduction of ptr_C modulo p gives $ptr_C \equiv v^{-1}$, so for a fixed r_C at most Z residues of t modulo p are allowed. In particular $p \mid \operatorname{den}(r_C)$. If a class has just one multiple for at least $J = \lceil H/4 \rceil$ such primes, fix J of them. Their product P_1 is at most $N^{\delta + o(1)}$, and CRT bounds the number of allowed $t \leq N$ by $(N/P_1 + 1)Z^J \leq N^{1 - \delta(2-A)/4 + o(1)}$. Multiplying by $N^{o(1)}$ possible values of r_C and by $2^H = N^{o(1)}$ choices of primes still gives a bound N^{1-c} for some $c > 0$. Each such t has at most y^A corresponding denominators. By disjointness only $o(m)$ groups can contain one. Outside these groups, a locally cancelled large class must consume at least two multiples of each of at least $(3/4 - o(1))H$ primes in $\mathcal{Q}_\delta$.

For a fixed $p \in \mathcal{Q}_\delta$, groups with exactly two multiples of p, both from the same $t \geq T$, are rare. Such multiples are tpv_1, tpv_2 and satisfy $v_1^{-1} + v_2^{-1} \equiv t(ps) \pmod{p}$, with $ps \not\equiv 0$ modulo p. For each pair $v_1, v_2 \leq Z$ the residue of t is fixed, and $t \leq N/(p \max(v_1, v_2))$. The count of pairs and t is at most

$$\sum_{v_1, v_2 \leq Z} \left(\frac{N}{p^2 \max(v_1, v_2)} + 1 \right) \ll NZ/p^2 + Z^2 = o(m).$$

Every group with such a locally cancelled class outside the exceptional groups contributes at least $(3/4 - o(1))H$ incidences to either $k_i(p) \geq 3$ or this rare two-multiple pattern. By (40), their number is at most $(8\delta/(3(1-\delta)) + o(1))m$. Sending $\delta \downarrow 0$ after $N \to \infty$ proves the first assertion.

For the second, let $R = N/p_0$, $Z_A = y^A/p_0$. The multiples $p_0 vt$ with $v \leq Z_A$, $T \leq t \leq R/v$ and t rough comprise a set F_A of size

$$|F_A| = D_y R \sum_{v \leq Z_A} (1/v - 1/Z_A) + o(R) = ((A-1)e^{-\gamma} + o(1))R, \quad D_y = \prod_{p \leq y} (1 - 1/p).$$

Indeed $v < y$ for large N, so the decomposition is unique. Inclusion–exclusion for rough integers up to X gives $D_y X + O(2^{\pi(y)})$, and summing the endpoint errors over v gives $N^{o(1)} = o(R)$.

By saturation at p_0, all but $o(R)$ of these multiples belong to groups having exactly two p_0 multiples. The first assertion implies that, except for $o(R)$ such multiples, their classes are not locally cancelled. Each group can contribute at most two, so at least $((A-1)e^{-\gamma} - o(1))m$ groups are nonlocal. Let $A \uparrow 2$ after taking the limit. □

Theorem 13 (Residual external denominator)**.** *Fix $1 < A < 2$ and a sufficiently small fixed $\beta > 0$. Under the saturation hypothesis, the proportion of groups containing a class with $t \geq N/y^A$ and $e(C) \leq N^\beta$ has limsup at most $64\beta/(2-A)$. Consequently at least*

$$\left((A-1)e^{-\gamma} - \frac{64\beta}{2-A} - o(1)\right) m$$

groups contain a class of that rough-part size whose external denominator exceeds N^β.

Proof. Repeat the proof of Theorem 12, but allow $e(C) \leq N^\beta$. Each target mass now belongs to $(e(C)L_A)^{-1}\mathbb{Z}$; the total number of possible positive targets is at most $\sum_{e \leq N^\beta} seL_A \leq N^{2\beta+o(1)}$. Its small-prime denominator part is at least $1/(sN^\beta)$, hence it lacks at most $(\beta/\delta + o(1))H$ primes of $\mathcal{Q}_\delta$. The singleton constraint for fixed target is unchanged: $e(C)$ is coprime to p. The same CRT count, now times $N^{2\beta+o(1)}$, is sublinear when $\beta \leq \delta(2-A)/16$. Apart from $o(m)$ exceptional groups, each such class uses at least two multiples for $(3/4 - \beta/\delta - o(1))H$ primes. The rare-pair estimate and (40) bound its group proportion by $[2\delta/(1-\delta)]/(3/4 - \beta/\delta)$. Take $\delta = 16\beta/(2-A) < 1/4$, giving at most $64\beta/(2-A)$. The final assertion follows from the F_A count and saturation. □

Cancellation components. Join two distinct classes of a group if their reduced denominators share a prime exceeding y. The sum over each connected component is locally cancelled in the sense that its reduced denominator has no prime exceeding y: every external prime appearing in that component occurs in no other component, and the full sum is integral at that prime. Moreover for any class C,

$$e(C) \mid \operatorname{lcm}_{C' \neq C} \gcd(t(C), t(C')).$$

Indeed for every $q^j \mid e(C)$ there is another class with denominator valuation at least j at q, since the total is q-integral. Its rough part is a multiple of that power. This identity explains the distinction between external-denominator cancellation and merely sharing a prime. As another consequence of the proof above, apart from $o(m)$ groups, at most two components can consist entirely of classes with $t \geq N/y^A$. Each such component has positive mass at most s and reduced denominator dividing L_A, since it is y-smooth and the y-parts of every term divide L_A. Its denominator contains all but $o(H)$ primes of $\mathcal{Q}_\delta$. Three components would force at least three multiples for $(1-o(1))H$ primes, contradicting (40) in all but $O(\delta)m$ groups; then send $\delta \downarrow 0$.

14 A strict constant loss for single-class constructions

Theorem 14. *Let $S_1, \ldots, S_m \subseteq [1, N]$ be disjoint nonempty sets of common mass $s = a/b > 0$ in lowest terms. Suppose $y \asymp \log N$ contains all prime factors of b and each S_i is contained in one y-rough-part class. Then*

$$m \leq \left(\frac{1}{2} - \frac{e^{-\gamma}}{4} + o(1)\right) \frac{N}{\log N}.$$

Proof. It is enough to consider $m \gg N/Y$. As in Lemma 2, $x = N/m = O(Y)$, $\log b \geq (1-o(1))Y$, and its largest prime p satisfies $p \geq (1-o(1))Y$ and $p \leq y$. In particular $p \parallel b$. Put $R = N/p$. For fixed $k \geq 2$ and $0 < \varepsilon < 1$, set $Z_k = \lfloor y^{(1-\varepsilon)/(k-1)} \rfloor$. Let F_k consist of multiples $pvt \leq N$ with $1 \leq v \leq Z_k$, $R/Z_k \leq t \leq R/v$, and t y-rough. As before inclusion–exclusion and Mertens give

$$|F_k| = \left(\frac{(1-\varepsilon)e^{-\gamma}}{k-1} + o(1) \right) R.$$

The F_k are nested as k increases. A group containing such a multiple has all its p multiples of the form ptv_j, with $v_j \leq Z_k < p$ for large N (here $Z_2 = o(Y)$), and all share t. If it has $j \leq k$ multiples, the equality of the mass forces $\sum_{\ell=1}^{j} v_\ell^{-1} \equiv t(ps) \pmod{p}$. For each tuple the residue of t is determined and $t \leq R/\max v_\ell$. The number of such possibilities for $1 \leq j \leq k$ is at most

$$O_k(R(1 + \log Z_k + Z_k^{k-1})/p + Z_k^k) = o(R).$$

Thus all but $o(R)$ members of F_k, if used, belong to groups with at least $k+1$ multiples. Each used p multiple can be assigned credit $1/k_i(p)$ so the total credit is m; unused multiples get zero. Singleton groups contribute only $o(R)$ exceptional vertices. For a fixed J, outside the $o(R)$ exceptional vertices just counted, credit is at most $1/2$, with further decreases of $1/k - 1/(k+1)$ on F_k for $2 \leq k \leq J$. Hence

$$\frac{m}{R} \leq \frac{1}{2} - (1-\varepsilon)e^{-\gamma} \sum_{k=2}^{J} \frac{1}{(k-1)k(k+1)} + o(1).$$

Let $J \to \infty$ and $\varepsilon \downarrow 0$ after $N \to \infty$. The sum telescopes to $1/4$. Since $R \leq (1+o(1))N/Y$, the claimed bound follows. □

15 Why a shared-prime matching obstruction is insufficient

The preceding theorems concern exact rational sums. The modular matching problem on its own behaves differently. We prove an independent result explaining the failure of an upper-bound approach based solely on the number of candidate pairs with a large common factor. The analytic input here is the additive large sieve [6] for a set $V \subseteq [1, X]$:

$$\sum_{p \leq Y} p \sum_{a \bmod p} |\#\{v \in V : v \equiv a \pmod{p}\} - |V|/p|^2 \leq (X + Y^2)|V|,$$

obtained from the usual large sieve for the distinct fractions a/p $(1 \leq a < p)$ by Parseval.

Theorem 15 (Near-complete modular pairing)**.** *Fix $1 < A < 2$, $0 < \delta < 1/4$ and $\varepsilon > 0$. Put $Y = \log N$, $R = \lfloor N/Y \rfloor$ and let $\mathcal{P}$ be exactly the set of primes in $[(1-\delta)Y, Y]$. Given any nonzero residues $c_p \in \mathbb{F}_p$ for $p \in \mathcal{P}$, there exists $\beta > 0$ such that, for all but $o(|\mathcal{P}|)$ primes p, disjoint pairs (u, v) in $[1, R]$ satisfy $u^{-1} + v^{-1} = c_p$ modulo p, $\gcd(u, v) \geq N^\beta$, and cover at least $(1-\varepsilon)R$ integers. They can cover at least $1-\varepsilon$ of the integers with $t_Y(u) \geq N/Y^A$ with both endpoints large and having distinct rough parts. Furthermore, without imposing the common-factor or large-class conditions, simultaneous matching at every $p \in \mathcal{P}$ produces $(1-o(1))R/2$ disjoint partial sets in $[1, N]$, each with exactly two multiples of each p, satisfying the prescribed congruences. The partial sets need not have equal rational mass.*

Proof. Delete auxiliary integers divisible by any $p \in \mathcal{P}$; this costs $o(R)$. Partition the rest according to largest prime $q \in [N^\beta, N^{1-\beta}]$ and according to whether their rough part is at least N/Y^A. Write the entries of each block as qv, with $v \in V \subseteq [1, X_q]$, $X_q = \lfloor R/q \rfloor$. The excluded small-largest-prime integers number $O(\beta R) + o(R)$, by summing $\sum_{q \le N^\beta} v_q(u) \log q$ over $u \le R$ and comparing to $\log u$ for $u \ge R/Y$. Those with a prime above $N^{1-\beta}$ also number $O(\beta R) + o(R)$ by Mertens. For fixed β, $X_q \gg N^\beta/Y \gg Y^2$.

Let $D_p(V) = \sum_{a \bmod p} |M_p(a) - |V|/p|$ for the class sizes $M_p(a)$. By Cauchy–Schwarz and the large sieve, $\sum_{p \in \mathcal{P}} D_p(V)^2 \ll X_q^2$. Within one block the inverse condition on v is $v_1^{-1} + v_2^{-1} = qc_p$. The partner map on legal nonzero residues is an involution; pair distinct interchanged classes as fully as possible and pair the possible fixed class internally. At most $O(D_p(V) + |V|/p + 1)$ vertices are unmatched. Sum these losses over blocks and primes. Since $\sum_q X_q \ll_\beta R$, Cauchy–Schwarz gives total discrepancy $O_\beta(R\sqrt{|\mathcal{P}|}) = o(R|\mathcal{P}|)$. The other two losses sum to $o(R|\mathcal{P}|)$ as well (there are at most $2N^{1-\beta}$ blocks). Thus for all but $o(|\mathcal{P}|)$ primes the matching loss within blocks is $o(R)$. Every block pair shares q.

The large-rough-part vertices number $((A-1)e^{-\gamma} + o(1))R$ by the same rough-integer count used for F_A, now with threshold N/Y^A on u instead of pu. For a large pair with identical rough part t, write $u = tv_1$, $v = tv_2$ with $v_j \ll Y^{A-1}$. Fixing v_1, v_2 fixes t modulo p, so there are $O(RY^{A-1}/p + Y^{2A-2}) = o(R)$ possible pairs of this kind uniformly in p. Delete them. Choose β sufficiently small in terms of A, ε to absorb the $O(\beta R)$ exclusions for both the full and large vertex counts. This proves the first statements.

For simultaneous unrestricted pairing use the entire filtered interval $[1, R]$ for each p. Its class counts before filtering differ by at most one, so involution matching loses only $O(R/p + p) = o(R)$ vertices; filtering destroys at most $o(R)$ pairs uniformly. Take the minimum of these matching sizes over p. At each prime assign that many different pairs to the same number of slots and use denominators pu, pv. Since every auxiliary integer avoids every key prime, denominators of distinct primes are disjoint. Within each prime the matching is disjoint. This yields the asserted partial sets. □

16 Counting obstructions to dilation

For arbitrary disjoint equal-mass families (not necessarily near saturation), let $V(K)$ denote the number of distinct reciprocal subset sums of $[K]$, including zero. The function has been studied by Bleicher and Erdős [7] and more recently by Bettin, Grenié, Molteni and Sanna [8]; see also [9]. We give a weaker self-contained upper bound sufficient here.

Lemma 16. *For $K \to \infty$, $\log V(K) \ll K \log\log K / \log K$.*

Proof. Set $z = K/\log K$. Integers with no prime-power divisor exceeding z all divide $\mathcal{L}(z)$; their reciprocal subset sums have at most $1 + H_K\mathcal{L}(z)$ possible values. The remaining integers are contained in the union of multiples of prime powers in $(z, K]$. Prime divisors $p > z$ contribute at most $K \sum_{z < p \le K} 1/p \ll K \log\log K/\log K$ by the prime counting bound $\pi(t) \ll t/\log t$ and partial summation. Prime powers of exponent at least two contribute at most $Kz^{-1}O(\sqrt{K}\log K) = o(K/\log K)$. Their reciprocal subsets have at most $2^{O(K \log\log K/\log K)}$ possible sums. Combining the two estimates and $\log \mathcal{L}(z) = O(z)$ proves the lemma. □

Theorem 17 (Bounded-multiplicity dilation)**.** *Suppose $S_i = t_i A_i \subseteq [N]$ are disjoint nonempty sets of equal positive reciprocal sum s, where t_i are positive integers, and each t occurs at most h*

times. Then

$$m \ll \frac{hN \log\log\log N}{\log N \log\log N}.$$

In particular any such representation with $h = o(\log\log N/\log\log\log N)$ has $m = o(N/\log N)$.

Proof. Put $Y = \log N$, $L_3 = \log\log\log N$, and $K = \lfloor \gamma_0 Y \log Y/L_3 \rfloor$ for a sufficiently small absolute $\gamma_0 > 0$. Lemma 16 gives $V(K) \le N^{1/2}$. There are at most N/K different $t \le N/K$. For $t > N/K$, the corresponding $A_i \subseteq [K]$ has reciprocal sum ts; different t give different sums. Thus at most $V(K)$ additional t occur. Multiplying by h and using $N^{1/2} = o(N/K)$ proves the theorem. □

17 Cancellation across rough parts

The previous obstruction does not preclude large multiplicity for one dilation factor. The next result allows arbitrarily many groups per factor, and even several rough-part classes inside a group.

Let $S_1, \dots, S_m \subseteq [1, N]$ be disjoint nonempty sets of common mass $s = a/b > 0$ in lowest terms, and let $\mathcal{P}$ be any finite prime set containing $\mathrm{supp}(b)$. Write each n uniquely as $n = d(n)t(n)$, where $d(n)$ is supported on $\mathcal{P}$ and $t(n)$ is coprime to every prime in $\mathcal{P}$. For a group S_i and integer t, let $S_{i,t} = \{n \in S_i : t(n) = t\}$. A nonempty class is called *locally cancelled* if the reduced denominator of its reciprocal sum is supported on $\mathcal{P}$. Equivalently, if $\sum_{n \in S_{i,t}} 1/d(n) = u/v$ in lowest terms, then $t \mid u$; indeed $(t, v) = 1$.

For this section write $M(\mathcal{P}) = \sum_{q \in \mathcal{P}} \log q/(q-1)$.

Theorem 18 (Rough-part obstruction)**.** *Suppose all the nonempty classes $S_{i,t}$ are locally cancelled. Then, with an absolute implied constant,*

$$m \ll \frac{(1 + M(\mathcal{P}))N}{\log N \log\log\log N}.$$

In particular, if $M(\mathcal{P}) = o(\log\log\log N)$, then $m = o(N/\log N)$.

More generally, if $m \ge cN/\log N$ and $M(\mathcal{P}) = o(\log\log\log N)$, there is a prime $p \mid b$ with $p \asymp_c \log N$ such that only $o(m)$ of the groups can contain a multiple of p in a locally cancelled class. Thus almost all groups require cancellation between different rough-part classes involving a multiple of p.

Proof. Use Y, L_3, K as in Theorem 17, decreasing the small positive constant γ_0 if necessary. If $m < N/(YL_3)$, the first bound is immediate. Otherwise $x = N/m \le YL_3$, and Lemma 2 gives $\log b \ge (1 - o(1))Y$. If p is the largest prime divisor of b, the contribution from prime powers beyond the first is $O(\sqrt{x}\log x) = o(Y)$, so $\log b \le \vartheta(p) + o(Y)$. By the PNT,

$$Y/2 \le p \le x \le YL_3 \tag{41}$$

for large N. Each group contains a multiple of this p.

First count the denominators that can belong to a locally cancelled class with $t \ge N/K$. Such a class has $d(n) \le K$, and its unscaled reciprocal sum $u/v > 0$ is one of at most $V(K)$ possible fractions with numerator $u \le H_K \mathcal{L}(K)$. Moreover $t \mid u$. Writing $\tau(u)$ for the number of positive divisors of u, the elementary maximal-divisor estimate $\log \tau(u) \ll$

$\log u/\log\log u$ for large u implies, uniformly for $1 \le u \le H_K\mathcal{L}(K)$, that $\log\tau(u) \ll K/\log K$. For completeness, this estimate follows by separating prime factors at $\sqrt{\log u}$: the small primes contribute $O(\sqrt{\log u}\log\log u)$ to $\log\tau(u)$, while the sum of exponents of the large primes is at most $2\log u/\log\log u$ and $\log(e+1) \le e\log 2$. Thus the number of possible t is at most $V(K)\exp(O(K/\log K))$. Each such t has at most K multiples up to N. By Lemma 16 and a sufficiently small γ_0, the total number of possible denominators in these classes is at most $N^{1/2}$ for all large N.

The multiples of p below N/L_3 number at most $N/(pL_3)$. For a remaining multiple $n = pu \ge N/L_3$ in a class with $t < N/K$, since $p \in \mathcal{P}$ we have $d(n) = pd(u)$ and therefore

$$d(u) = \frac{n}{pt} > \frac{K}{pL_3}.$$

By (41) and the definition of K, $\log(K/(pL_3)) \ge L_3/2$ for large N (the ratio is at least a constant times $\log Y/L_3^3$). On the other hand,

$$\begin{aligned}\sum_{u\le N/p}\log d(u) &= \sum_{q\in\mathcal{P}}\log q\sum_{j\ge1}\left\lfloor\frac{N/p}{q^j}\right\rfloor\\ &\le (N/p)M(\mathcal{P}).\end{aligned}$$

Markov's inequality bounds the number of these u by $2(N/p)M(\mathcal{P})/L_3$. Combining the three counts, the total number of multiples of p available in locally cancelled classes is at most

$$N^{1/2} + \frac{N}{pL_3} + \frac{2NM(\mathcal{P})}{pL_3} \ll \frac{(1+M(\mathcal{P}))N}{YL_3}.$$

Disjointness and the compulsory p-multiple prove the first assertion. If $m \ge cN/Y$, then $p \le Y/c$, and the same count is $o(m)$ under the moment hypothesis, whether or not all classes are locally cancelled. For each of the other groups a class containing a multiple of p has a reduced denominator with a prime factor $q \notin \mathcal{P}$. As the full sum has no such factor, another class in that group must also have a denominator divisible by q; otherwise the negative valuation of the first class could not be cancelled. This proves the final assertion. □

General absorption beyond the equal-target problem

18 Limits of independent positive splitting

Let $\theta = T/E \in (0,1)$ be reduced and write $E = \prod_{j=1}^{k} E_j$, where the $E_j > 1$ are pairwise coprime. There exists a decomposition $\theta = \sum_j r_j$ with $r_j > 0$ and $\mathrm{den}(r_j) \mid E_j$ if and only if

$$\sum_j \frac{a_j}{E_j} = \theta, \qquad a_j \equiv T(E/E_j)^{-1} \pmod{E_j}, \quad 1 \le a_j < E_j.$$

Indeed any such r_j must be a_j/E_j since $r_j < 1$ and reduction modulo E_j of the equation after multiplication by E fixes its numerator. The right side of the displayed test can differ from the

left by a positive integer: for $1/15$, the representatives are $2/3$ and $2/5$. Lemma 5 resolves this obstruction by admitting a shared denominator.

More quantitatively, if $E = N^{d+o(1)}$ is square-free, $\omega(E) = o(\log N)$ and $\sum_{p|E} 1/p = o(1)$, take $\theta_0 = N^{-c+o(1)}$ with $0 < c < d/2$. Among the integers T in $[\theta_0 E/2, \theta_0 E]$ coprime to E, almost none give a positive splitting with nontrivial coprime denominator factors. For a fixed factorization $E = UV$ such a splitting $T/E = a/U + b/V$ has at most $\theta_0^2 E$ choices; taking the union over at most $2^{\omega(E)} = N^{o(1)}$ factorizations gives $o(\theta_0 E)$. There are $(1/2 + o(1))\theta_0 E$ admissible T, by the union bound on multiples of primes dividing E. The latter bound includes at most $\omega(E)$ rounding errors, negligible here. More than two packets would give such a splitting by grouping packets. This demonstrates why positive splitting with common moduli is a genuinely different algebraic mechanism.

19 Simultaneous exact absorption of low-load targets

Return to the scales of (6)–(7), and to the notation $Y = \log N$, $\Lambda = \log Y$, $y = Y^4$. Write P for any square-free product of primes in $[Y/\Lambda^2, 2Y]$. Let $w = w(N)$ satisfy $2HY/\Lambda \le w \le Y$ (where H is the constant in Lemma 8).

Theorem 19. *There are constants $c_1, c_2 > 0$ such that for sufficiently large N the following is true. Given positive rationals r_i $(1 \le i \le m)$ with*

$$\frac{2HY}{N\Lambda} \le r_i \le \frac{w}{N}, \quad b_i = \mathrm{den}(r_i) \mid \mathcal{L}(y), \quad (b_i, P) = 1, \quad \log b_i \le 6Y,$$

assume $m \le c_1 N/w$ and, for each prime power $a \in (z, y]$, $\#\{i : a \mid b_i\} \le c_2 N/a$. Then there are pairwise disjoint sets $S_i \subseteq [1, N]$ with $\sigma(S_i) = r_i$, all elements coprime to P.

Proof. Use all y-rough integers t in the interval $[2N/K_+, N/(2K_-)]$, that is, integers without prime factors at most y. Set $K_t = \lfloor N/t \rfloor$. Then $K_t \in [K_-, K_+]$. We need the elementary sieve estimate

$$\sum_t \frac{1}{t} = \frac{e^{-\gamma}}{10} + o(1). \tag{42}$$

For clarity, this follows by truncated inclusion–exclusion: put $D_y = \prod_{p \le y}(1 - 1/p)$ and $J = \lceil 100 \log\log y \rceil$. The two adjacent Bonferroni sums of lengths $J, J+1$ for the main term differ from D_y by at most $\sum_{j \ge J} H^j/j!$, where $H = \sum_{p \le y} 1/p = \log\log y + O(1)$. This is $o(1/\log y)$ by Stirling's bound. Rounding errors in either sum are at most $y^{J+2} = N^{o(1)}$. Thus uniformly at $T \asymp N/Y^{O(1)}$ the count of rough integers up to T is $D_y T + o(T/\log y) + O(N^{o(1)})$. Partial summation and Mertens' formula prove (42), since the endpoint ratio is $y^{1/10}/4$.

Give each t $h_t = \lfloor \eta_0 K_t/w \rfloor$ slots for small fixed $\eta_0 > 0$. Their total $H_* \gg N/w$ by (42). Choose c_1 small enough that there are at least m slots. A uniformly random injection of labeled targets into slots gives, for each t, a, a hypergeometric number of marked targets with mean $\le h_t c_2 N/(aH_*) \ll c_2 K_t/a$. Choose c_2 small enough that the mean is at most $\epsilon_0 K_t/(10a)$. The elementary bound (2) shows that the probability of exceeding $\epsilon_0 K_t/a$ is $\exp(- \gg K_t/a)$. Since $K_t/a \gg Y^{23/4}$, a union bound over at most Ny events gives a successful allocation. At each t, scale the allocated targets by t. Their denominators are unchanged, because t is coprime to $\mathcal{L}(y)$, and their individual masses are in the range of (9): use $tK_t = N(1 - o(1))$ and $w \le Y$. Their total mass is at most η_0; choose $\eta_0 \le \epsilon_0$. Lemma 8 represents these targets by disjoint sets A_i of divisors of $\mathcal{L}(y)$, coprime to P. Set $S_i = tA_i$. Different t give disjoint sets by unique factorization of the rough part. This completes the proof. □

20 Consequences of the robust inverse theorem

Lemma 6 implies the following finite-field covering property: if q is prime, $q^\eta \le X < q$ and $V \subseteq [1, X]$ omits only $o(X)$ integers, every $a \in \mathbb{F}_q$ is a sum of a bounded number of distinct inverses of elements of V. Indeed choose the reservoir W of the lemma, and forbid $W \setminus V$. The lemma is stronger for applications: it gives positive-density reservoirs supported on small primes and controls conditional probabilities for sequential sampling.

For a typical application, suppose disjoint partial representations have denominators containing q to exponent at most one, and at most $o(N/q)$ of the groups require correction at this prime. For fixed $\beta, \xi > 0$ and $N^\beta < q \le N^{1-\xi}$, let $X = \lfloor N/q \rfloor$ and apply Lemma 6 with $\eta = \xi/2$, $D = 1/\beta$, and $\tau = \beta/2$. The correction terms qw have unique largest prime q, and the other prime factors are at most $N^{\beta/2}$. As long as the forbidden portion of each reservoir is $o(X)$ uniformly, every group can be corrected at q using L terms by prescribing the negative residue of q times its current sum. Pools at different q do not intersect, and correcting one such prime does not reintroduce another. The total per-prime consumption $L\, o(N/q)$ is admissible. This is a modulus-reduction statement, not an assertion that arbitrary global rational deficits are exactly representable.

Scope of simultaneous absorption

The low-load hypothesis is essential. For example take $m \asymp N/(\log N)$ identical targets whose reduced denominator contains a prime $q \asymp (\log N)^2$. There are at most $N/q = o(m)$ multiples of q in $[1, N]$, regardless of how small the targets are. Neither smoothness of denominators nor a bound on their total mass replaces the resource hypotheses. Theorem 19 concerns a fixed polylogarithmic smoothness cutoff; its constants do not justify passing to a growing exponent. It also does not handle arbitrary prior deletions in the final representation stage. The proof of Lemma 6 uses an exact published estimate; the rest of this paper uses elementary probability, PNT, Mertens, fixed-parameter smooth-number asymptotics and Cauchy–Davenport.

21 Discussion and remaining questions

The sharp construction and the independent order-of-magnitude proof address the free-target problem, not the fixed-target packing problem. The stability statements concern arbitrary near-extremizing families and do not assume the sharp lower bound. The simultaneous absorption theorem requires explicit prime-power load bounds and does not assert robustness under arbitrary prior deletions in its final-stage pools. The limiting distribution of cancellation networks and possible further restrictions on extremal families remain open. The primary lower-bound proof uses the Bourgain–Garaev estimate as an external theorem; the two appendices provide routes independent of that estimate. The proof is asymptotic: constants in the inverse-sum lemma and the local smooth-number density estimates can be very small or very large, and no practical numerical threshold for N is claimed.

A An independent order-of-magnitude proof

Theorem 20. *For an absolute constant $c > 0$ and all sufficiently large N,*

$$cN/\log N \le g(N) \le (1/2 + o(1))N/\log N.$$

The lower bound can be obtained with a common unit-fraction target without using the Bourgain–Garaev estimate.

The upper bound is Corollary 3. For the lower bound we use only the prime number theorem, Mertens' formulas, Cauchy–Davenport, the fixed-parameter smooth-number estimate (1), and the binary-grid Lemma 7. The local parameters y, z, K in this appendix are independent of those used in the sharp lower bound.

A.1 Local multiple representations at small mass

Lemma 21. *There exist absolute constants $H, \eta > 0$ such that for all sufficiently large y, all integers $y^3 \le K \le y^{31/10}$ and all $x \in \mathcal{L}(y)^{-1}\mathbb{Z}$ with*

$$Hy/(K \log y) \le x \le 2Hy/(K \log y),$$

there are $\lfloor \eta K/y \rfloor$ disjoint nonempty representing sets $A_i \subseteq [1, K]$, all elements dividing $\mathcal{L}(y)$, with $\sigma(A_i) = x$.

Proof. Take $z = y^{1/4}$ and the Q, D from Lemma 7. Then $Q \mid B := \mathcal{L}(y)$, since its extra power of two is $O(z(\log z)^2) < y$. Put $M = \lceil z^3 \rceil$. For every level $q = p^j$ with $v_p(Q) < j \le v_p(B)$, reserve

$$\mathcal{T}_q = \{qv : K/(2q) \le v \le K/q,\ v \mid \mathcal{L}(z),\ (v, p) = 1\}.$$

These pools are disjoint since $q > z$ is the unique complete prime-power component exceeding z. All their integers divide B. With $X = K/q$, the ratio $\log X/\log z$ lies in $[8, 12.4]$. Thus smooth-number asymptotics count $\gg X$ z-smooth integers in $[X/2, X]$. Those with a forbidden higher prime power number at most $X\pi(z)/z = o(X)$. To exclude multiples of p, use the bound X/p for sufficiently large p; for the finitely many small p, the smooth-number asymptotic applied also at X/p and $X/(2p)$ shows that the excluded proportion among smooth integers tends to $1/p$. Hence $|\mathcal{T}_q| \ge c_0 X$ uniformly.

While at least $c_0 X/2$ elements remain, at least $c_1 p$ nonzero residues of v have at least L unused representatives for any fixed L: a residue has at most $X/p + 1$ elements and the non-rich residues have at most $Lp = o(X)$ since $X/p \ge y \to \infty$. Invert the rich residues, adjoin zero, and apply Cauchy–Davenport repeatedly. For sufficiently large fixed L the L-fold sumset is all of $\mathbb{F}_p$. Repeated nonzero residues can use distinct representatives, and zero requires no element. Each group can therefore prescribe any needed inverse sum at this level using at most L distinct unused v's. If $h = \lfloor \eta K/y \rfloor$ with $L\eta < c_0/2$, the total consumption per level is at most $Lh \le L\eta K/q$, so this greedy operation always works.

Starting with deficit x at every group, correct levels of each prime from high to low: when the deficit e has $v_p(e) \ge -j$, choose $\sum v^{-1} \equiv p^j e \pmod p$ and subtract the terms $1/(p^j v)$. This removes that digit of the denominator. The other prime factors of v lie in $\mathcal{L}(z)$ and cannot reintroduce already corrected levels. There are $O(y/\log y)$ levels altogether, and every correction denominator is at least $K/2$. Choosing H sufficiently large leaves a deficit $e_i \in Q^{-1}\mathbb{Z}$ in $[Hy/(2K \log y), 2Hy/(K \log y)]$ for every group.

To fill these deficits use distinct scaled copies uD, with $U = K/(2M)$ and $U \le u \le 2U$. Take $u = p_1 p_2 p_3$ where $p_1 < p_2$ are primes in $[y^{0.7}, y^{0.72}]$ and p_3 is prime in $[U/(p_1 p_2), 2U/(p_1 p_2)]$. Since U is between constant multiples of $y^{2.25}$ and $y^{2.35}$, the exponent for p_3 lies between $0.81 + o(1)$ and $0.95 + o(1)$. All three primes are distinct, exceed z and do not exceed y. PNT counts $\gg U/(p_1 p_2 \log y)$ choices for p_3. Mertens implies $\sum_{p_1 < p_2} 1/(p_1 p_2) \gg 1$, so unique

factorization yields $\gg U/\log y$ copies. Each consists of divisors of B, lies in $[1, K]$, and has three distinct primes exceeding z. Thus the copies and the earlier correction pools are pairwise disjoint.

By Lemma 7, each copy can represent every r/Q with $a = \lceil Q/(8U)\rceil \le r \le b = \lfloor 3Q/(8U)\rfloor$. For large y, $Q/U \to \infty$ and $b \ge 2a$. All integers $r \ge a$ can be written as sums of $O(r/a + 1)$ integers in $[a, b]$, since the intervals $[ka, kb]$ cover all integers above a. Every e_iQ is such an integer: $e_iU \asymp y^{1/4}/\log y \to \infty$. So each group requires $O(e_iU + 1) = O(Hy/(M \log y))$ copies. Total demand is $O(H\eta K/(M \log y))$, whereas the supply is $\gg K/(M \log y)$. Shrink η after fixing H to ensure sufficiency. All deficits can be filled exactly with disjoint copies, proving the lemma. □

A.2 Global construction and the lower bound

Take $y = 2 \log N$ and $B = \mathcal{L}(y) = N^{2+o(1)}$. First take a target on the full B^{-1} grid. Set $s = B^{-1}\lceil 3HyB/(2N \log y)\rceil \sim 3Hy/(2N \log y)$. For all rough t in $[N/y^{31/10}, N/(2y^3)]$ let $K_t = \lfloor N/t\rfloor$ and $x_t = ts$. Then $K_t \in [y^3, y^{31/10}]$ for large N, $x_t \in B^{-1}\mathbb{Z}$, and $x_tK_t \sim 3Hy/(2 \log y)$ uniformly. Lemma 21 gives $\lfloor \eta K_t/y\rfloor$ disjoint sets of mass x_t consisting of divisors of B. Multiplying each denominator by t gives sets in $[1, N]$ of mass s. Sets from different t are disjoint because the part of every denominator supported on primes above y is exactly t.

For any T in this interval, inclusion–exclusion counts the y-rough integers up to T as $D_yT + O(2^{\pi(y)})$, where $D_y = \prod_{p\le y}(1 - 1/p) \sim e^{-\gamma}/\log y$. The error is $N^{o(1)}$. Partial summation gives

$$\sum_{\substack{N/y^{31/10}\le t\le N/(2y^3)\\ t \text{ rough}}} \frac{1}{t} \longrightarrow e^{-\gamma}/10.$$

Floors are negligible uniformly since $K_t/y \to \infty$. Thus the number of groups is at least $(\eta e^{-\gamma}/10 + o(1))N/y \gg N/\log N$, proving the order lower bound.

For a unit-fraction target there is also a simpler modification avoiding a delicate choice of y. Use the same $y = 2 \log N$ and the full B, but pick $B_1 \mid B$ with $B_1 = (1+o(1))N \log y/((3/2)Hy)$, as follows. The product of primes in $[\sqrt{y}, y]$ is $N^{2+o(1)}$ and each jump is at most y. Start with the product of consecutive such primes that stops before exceeding T_0/y^2, where $T_0 = N \log y/((3/2)Hy)$. The ratio T_0 to the resulting product lies in $[y^2, y^3]$ up to harmless endpoint equality. The largest powers of, say, the first ten fixed primes bounded by y have product $\gg y^{10}$, and all avoid the selected large primes. To obtain the stated $1 + o(1)$ approximation, use the powers of ten fixed primes other than $2, 3$, each bounded by y. Increase their product by multiplying in one such prime at each step, stopping just before it would exceed the remaining ratio divided by $y^{1/4}$. The available product is $\gg y^{10}$ and the jumps are bounded by a fixed constant C, so the new remaining ratio lies between $y^{1/4}$ and $Cy^{1/4}$; use powers of 2 and 3 to approximate that remaining ratio from below. This last step follows from density of multiples of $\log 3$ modulo $\log 2$ (irrationality of $\log 3/\log 2$), using finitely many powers of 3 for any fixed precision, then powers of 2 to fill the size. A diagonal choice of the precision gives $1 + o(1)$; the individual powers of 2 and 3 are at most y for large y. Thus B_1 divides B and has the prescribed asymptotic. Take $s = 1/B_1$ in the preceding argument, which only used the asymptotic size of s and $s \in B^{-1}\mathbb{Z}$. This proves the unit-fraction assertion and completes the independent proof of Theorem 20.

B An elementary hierarchy via complete systems and coprime dilation

For an integer $B \geq 2$ and a set D of positive divisors of B, let $K(D) = \max D$ and $R(D) = \sum_{d \in D} 1/d$. Call (B, D) complete if the subset sums of the integers B/d ($d \in D$) cover all residues modulo B.

Lemma 22 (Coprime dilation)**.** *If (B, D) is complete then $g(\lfloor BR(D)K(D) \rfloor) \geq \varphi(B)$.*

Proof. For each $r \in (\mathbb{Z}/B\mathbb{Z})^\times$ choose a subset $A_r \subseteq D$ with $t_r = \sum_{d \in A_r} B/d \equiv r \pmod{B}$. Then $1 \leq t_r \leq BR(D)$ and $(t_r, B) = 1$. Set $S_r = \{t_r d : d \in A_r\}$. Its sum is $1/B$. The part of any $n \in S_r$ supported on primes not dividing B is exactly t_r. Thus a collision between S_r and $S_{r'}$ implies $t_r = t_{r'}$, hence $r = r'$. The size bound on the denominators follows from their definition. □

For this section $\log_j$ denotes the j-fold iterated natural logarithm. Define

$$F_k(X) = (\log_k X)^2 \prod_{j=1}^{k-1} \log_j X.$$

We construct complete systems with $B \asymp X$, $K(D) \ll_k F_k(X)$, bounded $R(D)$ and $\varphi(B)/B \gg_k 1$. It is useful to keep a slightly stronger invariant. For a finite prime set $\mathcal{P}$, let

$$M(\mathcal{P}) = \sum_{q \in \mathcal{P}} \frac{\log q}{q - 1}. \tag{43}$$

Proposition 23 (Recursive complete systems)**.** *For each fixed $k \geq 1$ and each fixed finite set $\mathcal{F}$ of odd primes, for all sufficiently large real X there exists a complete system (B, D) with the following properties:*

$$\begin{aligned} &X/2 < B \leq X, \qquad \mathcal{F} \cap \operatorname{supp}(B) = \varnothing, \\ &K(D) \ll_{k,\mathcal{F}} F_k(X), \qquad R(D) \ll_{k,\mathcal{F}} 1, \\ &\varphi(B)/B \gg_{k,\mathcal{F}} 1, \qquad M(\operatorname{supp}(B)) \ll_{k,\mathcal{F}} 1. \end{aligned}$$

Here $\operatorname{supp}(B)$ is the set of prime divisors of B. Moreover if $2^J \parallel B$, then D contains every 2^j for $1 \leq j \leq J$, and $J \to \infty$ as $X \to \infty$.

Base case $k = 1$. Put $y = 3 \log X$ and $J_0 = \lceil \log_{\text{base}\,2} y \rceil$, where $\log_{\text{base}\,2}$ denotes logarithm to base 2. Choose primes in $(y/2, y]$ successively until adding the next prime to 2^{J_0} times the current product would exceed X. This stopping point exists by the PNT, since the log of the full prime product is $(1/2 + o(1))y > \log X$. All these primes avoid $\mathcal{F}$ for large X. If the current product including 2^{J_0} is C, then $C \leq X$ and $X/C < y$ (or $\leq y$, which is sufficient). Multiply C by the largest power of two not exceeding X/C. The resulting integer is

$$B = 2^J \prod_{p \in P} p \in (X/2, X], \qquad 2^J \leq 2y^2,$$

where P is the chosen prime set and $J \geq J_0$.

For each $p \in P$ put $k_p = \lceil \log_{\text{base }2} p \rceil$ and take

$$D = \{2^j : 1 \le j \le J\} \ \cup \bigcup_{p \in P} \{2^j p : 0 \le j < k_p\}.$$

All elements divide B. In the p-block the weights $B/(2^j p)$, after multiplication by an invertible constant and reversal of order, are the binary weights $1, 2, \dots, 2^{k_p - 1}$ modulo p. Their subset sums cover $\mathbb{F}_p$. They vanish modulo every other prime of P. Fix all odd-prime residues first. The reserved pure powers of two then give every residue modulo 2^J and vanish modulo all odd primes. The Chinese remainder theorem proves completeness.

The maximum denominator is at most $2y^2$. The mass and totient ratio satisfy

$$R(D) \le 1 + 2 \sum_{y/2 < p \le y} 1/p = 1 + O(1/\log y), \qquad \frac{\varphi(B)}{B} = \frac{1}{2} \prod_{p \in P} (1 - 1/p) = \frac{1}{2} - o(1).$$

Finally $\sum_{y/2 < p \le y} \log p/(p-1) = O(1)$, so (43) is bounded. This proves the base case. □

Inductive step. Assume the proposition for k and all fixed forbidden sets. Fix an odd prime $\ell \notin \mathcal{F}$. Apply the induction hypothesis with forbidden set $\mathcal{F} \cup \{\ell\}$ at $X' = 2y$, where $y = 3 \log X$. Obtain a complete system (Q, D_0) with

$$y < Q \le 2y, \qquad K_0 = K(D_0) \ll F_k(2y) = o(y), \qquad R(D_0) \le R_0,$$

where R_0 is constant for this fixed induction step. Write $2^J \parallel Q$. Every prime divisor of Q must divide at least one $d \in D_0$: otherwise all Q/d would be zero modulo that prime, contradicting completeness. Thus all prime divisors of Q are at most K_0. In particular every $p \in (y/2, y]$ is coprime to Q and, for large y, to ℓ.

Let W be the set of integer subset sums of Q/d for $d \in D_0$. Completeness gives $|W| \ge Q$, while $W \subseteq [0, R_0 Q]$. For each prime $p \in (y/2, y]$, each residue class modulo p contains at most $1 + 4R_0$ elements of that interval. Therefore the set

$$T_p = \Big\{ \sum_{d \in A} d^{-1} \pmod p : A \subseteq D_0 \Big\}$$

has cardinality at least $Q/(1 + 4R_0) \ge p/(1 + 4R_0)$, because multiplication by Q^{-1} is invertible. Choose a fixed positive integer $h > 2(1 + 4R_0)$. By repeated Cauchy–Davenport, for sufficiently large p,

$$T_p + \ell^{-1} T_p + \cdots + \ell^{-(h-1)} T_p = \mathbb{F}_p.$$

The sets $\ell^i D_0$ for $0 \le i < h$ are disjoint because $\ell \nmid Q$. Put $E = \bigcup_{i=0}^{h-1} \ell^i D_0$; its inverse subset sums cover every residue modulo each such p, and $K(E) = O(K_0) = o(y)$.

Start with $C = \ell^{h-1} Q$ and multiply by primes in $(y/2, y]$ until the next prime would push the product past X. The PNT guarantees a stopping point, since $C = O(y)$ and the log of the full prime product is $(1/2 + o(1))y$. As before choose an additional power $2^v \le y$ to obtain

$$B = 2^v \ell^{h-1} Q \prod_{p \in P} p \in (X/2, X].$$

Let $j_\ell = \lceil \log_{\text{base }2} \ell \rceil$ and define

$$\begin{aligned} D = D_0 &\cup \{2^j : 1 \le j \le J + v\} \\ &\cup \{\ell^i 2^j : 1 \le i < h, \ 0 \le j < j_\ell\} \cup \bigcup_{p \in P} pE. \end{aligned}$$

Every element divides B, since $J \to \infty$ and $\ell \nmid Q$. The outer blocks pE are disjoint from each other and from the other blocks: all their cofactors are less than $y/2$, whereas their designated prime is greater than $y/2$. The ℓ-correction blocks avoid D_0 and the pure powers of two. The overlap of the two sets of pure powers is taken only once.

To show completeness, first use the outer blocks to set each residue modulo $p \in P$. Next set the residue modulo ℓ^{h-1}, correcting the digits from least significant to most significant by using the blocks $\ell^i 2^j$ in order $i = h-1, h-2, \dots, 1$. In block i, the weight $B/(\ell^i 2^j)$ has ℓ-valuation $h-1-i$; after removing this power, the remaining weights modulo ℓ are invertible multiples of reversed binary weights. Hence one can correct the next digit without changing any lower digit. These blocks do not affect the outer primes.

Now use D_0 with its pure powers of two reserved, to correct the residue modulo the odd part of Q. Completeness of (Q, D_0) implies that this is possible, since the weights of pure powers of two vanish modulo the odd part of Q. Multiplication by B/Q is invertible modulo that odd part. This step does not change the outer or ℓ-residues. Finally the pure powers of two through 2^{J+v} correct the residue modulo 2^{J+v} and leave all odd residues unchanged. CRT completes the proof of coverage.

Since $2^J \le K_0$ and $2^v \le y$, all denominators are $O(yK_0)$. The mass is bounded by

$$R(D_0) + 1 + \sum_{i=1}^{h-1} \sum_{j=0}^{j_\ell - 1} \frac{1}{\ell^i 2^j} + R(E) \sum_{p \in P} \frac{1}{p} = O(1).$$

All constants here are for fixed $k, \mathcal{F}$. The totient ratio stays bounded below, as it is multiplied by $(1-1/\ell)\prod_{p\in P}(1-1/p) \gg 1$. The new prime moment is at most the old moment plus $\log \ell/(\ell-1)$ and $\sum_{y/2<p\le y} \log p/(p-1) = O(1)$. The forbidden primes are avoided. Finally

$$yF_k(2y) \asymp F_{k+1}(X),$$

which proves the required denominator bound and the induction. □

Corollary 24. *For every fixed integer $k \ge 1$, $g(N) \gg_k N/F_k(N)$. In particular this approach alone yields $g(N) \ge N/(\log N)^C$ for every fixed $C > 1$, eventually.*

Proof. Fix k and use Proposition 23 with empty forbidden set at $X = \delta_k N/F_k(N)$, where $\delta_k > 0$ is sufficiently small. Since F_k is eventually increasing and $B \le X$, we have $BR(D)K(D) \le N$ for large N. Lemma 22 gives $g(N) \ge \varphi(B) \gg_k B \gg_k N/F_k(N)$. Taking $k = 2$ gives $g(N) \gg N/(\log N(\log_2 N)^2)$, which dominates $N/(\log N)^C$ for any fixed $C > 1$. This proof of a weaker lower bound is independent of smooth-number estimates. □

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] T. F. Bloom, *Erdős Problem #296*, `https://www.erdosproblems.com/296` (accessed 28 November 2026). The free-target equal-sums variant is in the commentary; the fixed-target result is a separate question.

[3] 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.

[4] 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.

[5] J. Bourgain and M. Z. Garaev, Sumsets of reciprocals in prime fields and multilinear Kloosterman sums, *Izv. Math.* **78** (2014), no. 4, 656–707, Theorem 8 (p. 660). doi:10.1070/IM2014v078n04ABEH002703; arXiv:1211.4184.

[6] H. L. Montgomery and R. C. Vaughan, *Multiplicative Number Theory II: Primes and Sieves*, Cambridge University Press; additive large sieve.

[7] M. N. Bleicher and P. Erdős, The number of distinct subsums of $\sum_{i=1}^{N} 1/i$, *Math. Comp.* 29 (1975), 29–42.

[8] S. Bettin, L. Grenié, G. Molteni and C. Sanna, A lower bound for the number of Egyptian fractions, *Math. Comp.*, published electronically 22 January 2026. doi:10.1090/mcom/4190.

[9] T. F. Bloom, *Erdős Problem #320*, `https://www.erdosproblems.com/320`.