hexagon:2611.01226v1

SHARP SECOND-ORDER ENTROPY ASYMPTOTICS FOR DISJOINT EGYPTIAN FRACTION REPRESENTATIONS

PENGCHENG CHEN

Abstract. Let $\mathcal{P}(N,m)$ count ordered pairwise disjoint subsets $S_1,\dots,S_m\subseteq[N]$ of reciprocal weight one each. For all feasible sequences $m\sim\log N$ we obtain

$$\log\mathcal{P}(N,m)=N\log m-\frac{3N(\log\log N)^2}{2\log N}+O\left(\frac{N\log\log N\log\log\log N}{\log N}\right).$$

The coefficient $3/2$ is the sum of losses from forced omissions and restrictions on the number of labels within large-prime multiple blocks. A robust rational-grid reservoir, tolerant of missing denominators and forbidden prime factors, allows the construction to attain both losses. We further prove that the scalar integer representation threshold and the disjoint unit-sum threshold coincide for all sufficiently large m, with the same entropy expansion for partitions of any prescribed minimal scalar representation support. All completions preserve the original free labels and are injective. The threshold equality and the entropy estimates are independent of the sharp scalar capacity asymptotics. We also give the fixed-dimensional entropy formula for arbitrary fixed positive rational targets.

1. Introduction and main results

For a finite $A\subset\mathbb{N}$ let $w(A)=\sum_{n\in A}1/n$, let $[X]=\{1,\dots,\lfloor X\rfloor\}$, and let $H_N=w([N])$. All logarithms are natural. Define

$$\begin{aligned}\mathcal{P}(N,m)&=\#\{(S_1,\dots,S_m):S_i\subseteq[N],\ S_i\cap S_j=\varnothing\ (i\neq j),\ w(S_i)=1\},\\ N_*(m)&=\min\{N:\mathcal{P}(N,m)>0\},\\ T_*(m)&=\min\{N:\exists A\subseteq[N],\ w(A)=m\}.\end{aligned}$$

Minima over empty sets are $+\infty$. Plainly $T_*(m)\le N_*(m)$ and $\mathcal{P}(N,m)\le(m+1)^N$, by assigning an unused label or one of the m used labels to each integer. For $A\subseteq[N]$ with $w(A)=m$, define

$$\mathcal{D}(A,m)=\#\{(S_1,\dots,S_m):A=S_1\sqcup\cdots\sqcup S_m,\ w(S_i)=1\}. \tag{1}$$

For large N, abbreviate

$$\ell=\log N,\quad L=\log\ell,\quad \Delta_N=L^2/\ell,\quad B_N=NL^2/\ell,\quad E_N=NL\log L/\ell. \tag{2}$$

These scales are used only for sufficiently large N. In the scalar problem, Croot [4] proved that the harmonic capacity loss has order Δ_N. For a fixed positive rational target, Conlon, Fox, He, Mubayi, Pham, Suk and Verstraëte [3] determined the precise exponential representation count. Here the number of labels grows with N and the question is to determine the first nontrivial entropy deficit, even at exact feasibility.

An earlier preprint [1, Corollaries 14 and 17] gave the same two-term expansion for $\log T_*(m)$ and $\log N_*(m)$ separately. It also showed that, for all but $O(X^{11/12})$ heights $N\le X$, the largest integer representable in $[N]$ equals the maximum number of disjoint unit sums in $[N]$. Here we

Date: November 27, 2026 (Version 0.4.1).

2020 *Mathematics Subject Classification.* 11D68, 11P70, 05A16, 94A17.

Key words and phrases. Egyptian fractions, disjoint representations, entropy, capacity, robust modular covering.

prove $N_*(m) = T_*(m)$ for every sufficiently large m and determine the second-order entropy of ordered partitions, including those of every prescribed scalar support at the threshold.

Theorem 1.1 (Sharp entropy on a prescribed critical support)**.** *Let $N = T_*(m)$, where $m \to \infty$ through the positive integers. For every $A \subseteq [N]$ with $w(A) = m$,*

$$\log \mathcal{D}(A, m) = N \log m - \tfrac{3}{2} B_N + O(E_N), \tag{3}$$

with an absolute error bound uniform in A.

Corollary 1.2 (Exact threshold equality)**.** *For every sufficiently large m, $N_*(m) = T_*(m)$, and for $N = N_*(m)$ one has*

$$\log \mathcal{P}(N, m) = N \log m - \frac{3N(\log m)^2}{2m} + O\Big(\frac{N \log m \log\log m}{m}\Big). \tag{4}$$

Theorem 1.3 (Uniform entropy in the feasible region)**.** *Along any sequence of positive integers N, m with $N \to \infty$ and $m \sim \log N$, if $\mathcal{P}(N, m) > 0$, then*

$$\log \mathcal{P}(N, m) = N \log m - \tfrac{3}{2} B_N + O(E_N). \tag{5}$$

There is an absolute C_0 such that feasibility is guaranteed, under the same growth condition, whenever $H_N - m \geq C_0 \Delta_N$.

In particular, for $m = \lfloor H_N - C\Delta_N \rfloor$ with a sufficiently large absolute C, this implies $\log \mathcal{P}(N, m) \sim N \log\log N$ and determines its finer deficit relative to $N \log m$.

The arithmetic tools are first proved in Section 2. Sections 3 and 4 give high-precision completion and local saturation; Section 5 supplies reservoirs inside a given support. The robust grid construction in Section 6 allows many local modular completions without sacrificing the bulk of a large-prime block. Sections 7–10 integrate the entropy losses and prove the main theorems. The final section gives the fixed-vector entropy formula. The proofs of the entropy theorems and the threshold equality are independent of the sharp scalar capacity asymptotics in the companion paper [2].

2. Arithmetic preliminaries

Put $K_y = \operatorname{lcm}(1, \dots, \lfloor y \rfloor)$ for $y \geq 1$. Let $P^*(n)$ be the largest prime-power divisor of n, with $P^*(1) = 1$. Then $P^*(n) \leq y$ is equivalent to $n \mid K_y$. We use the quantitative prime number theorem (PNT), $\pi(x) = \mathrm{Li}(x) + O(xe^{-c\sqrt{\log x}})$ ([5, Chapter 6]). It gives $\log K_y = (1 + o(1))y$ and $\pi(2y) - \pi(y) \asymp y/\log y$. Whenever a rational is reduced modulo a prime, its reduced denominator is understood to be coprime to that prime. All unions or sums indexed by q in the constructions below are over prime powers $q = p^a$ unless otherwise stated.

Lemma 2.1 (Fourier coverage by auxiliary primes)**.** *Let $X \to \infty$, let M be a fixed positive power of $\log X$ up to rounding, and let k and $0 < \eta < 1/2$ depend on X. Suppose*

$$(1 + M\eta)\frac{\log X}{\log M} = o(k), \qquad k\eta^2/\log X \to \infty. \tag{6}$$

For every prime $p \leq X$ and every set B of k distinct primes in $[M, 2M] \setminus \{p\}$, the reciprocal subset sums of B cover $\mathbb{F}_p$, uniformly in p, B for all sufficiently large X.

Proof. Fix $h \in \{1, \dots, p-1\}$. If $\|hb^{-1}/p\| \leq \eta$, using distance to the nearest integer, then $bv = h + jp$ for integers v, j with $|v| \leq \eta p$ and $|j| \leq 2M\eta + 1$. All integers $h + jp$ are nonzero, and their sizes are at most $(2M\eta + 2)X$. Each has $O(\log X/\log M)$ distinct prime divisors in

$[M, 2M]$. Thus at most $O((1 + M\eta) \log X/\log M) = o(k)$ of the auxiliary primes are exceptional. For the sum Z of the inverses of a uniformly random subset of B,

$$|\mathbb{E}e^{2\pi ihZ/p}| = \prod_{b\in B} |\cos(\pi hb^{-1}/p)| \le e^{-ck\eta^2}$$

for some absolute $c > 0$, since at least half the primes are nonexceptional. Fourier inversion gives $\mathbb{P}(Z = a) \ge p^{-1}(1 - (p-1)e^{-ck\eta^2}) > 0$ for every $a \in \mathbb{F}_p$, by (6). This proves the claim. The proof is an elementary variant of the auxiliary-prime argument in [4, Proposition 2]. □

Lemma 2.2 (Mixed-radix denominators)**.** *If $0 \le r < 1$ and $K_J r \in \mathbb{Z}$ for an integer $J \ge 2$, then $r = w(E)$ for some subset of*

$$R(J) = \{K_q/u : q = p^a \le J,\ 1 \le u < p\}. \tag{7}$$

Each indicated denominator is integral and has largest prime-power divisor q. Moreover $w(R(J)) = O(1)$ uniformly in J.

Proof. At $q = p^a$ the relation $K_q = pK_{q-1}$ holds. From the largest prime power downwards, choose a digit $u_q \in \{0, \dots, p-1\}$ congruent to the current numerator when the residual is written over K_q. Subtract u_q/K_q; the new denominator divides K_{q-1}. The final residual is an integer. The digit sum is nonnegative and is at most

$$\sum_{q=p^a\le J} \frac{p-1}{K_q} = \sum_{q=p^a\le J} \left(\frac{1}{K_{q-1}} - \frac{1}{K_q}\right) = 1 - \frac{1}{K_J}.$$

Hence the final residual is zero. Every nonzero digit divides K_q and is coprime to p because $1 \le u_q < p$. Thus u_q/K_q is the reciprocal of K_q/u_q and $P^*(K_q/u_q) = q$. These denominators are distinct across different q, and within a fixed q different u give different denominators. Finally $w(R(J)) = \sum_{q=p^a\le J} p(p-1)/(2K_q) \ll 1$, since $K_q \ge e^{q/2}$ for large q. □

Lemma 2.3 (Pools for intermediate prime powers)**.** *Fix a positive integer $h \ge 2$. For a sufficiently large prime power $q = p^a$, choose h distinct primes $G \subset (q/2, q)$, put $b = \prod_{v\in G} v$, and let*

$$P(q, G) = \{qbu : 1 \le u < p,\ (u, b) = 1\}. \tag{8}$$

Then $P^*(n) = q$ and $n < q^{h+2}$ for each $n \in P(q, G)$, its weight is $O_h((1 + \log q)/q^{h+1})$, and

$$\{q\, w(E) \bmod p : E \subseteq P(q, G)\} = \mathbb{F}_p. \tag{9}$$

For any fixed $h \ge 2$ one may choose $\gg_h (q/\log q)^{h-1}$ such pools that are pairwise disjoint, uniformly for large q.

Proof. None of the auxiliary primes is p: if $a = 1$, the endpoint $q = p$ is excluded, and otherwise $p \le q/2$. Also $u < p$ is coprime to p and by definition to b. Hence the largest prime-power divisor is q, and $qbu < q^{h+2}$. The weight is at most $H_{p-1}/(qb) \ll_h (1 + \log q)/q^{h+1}$.

If $a \ge 2$, all $u \in [p-1]$ are admissible, so $(bu)^{-1}$ ranges over the nonzero residues and the empty subset supplies zero. If $a = 1$, at most h of the u are excluded, since any excluded $u < q$ divisible by an auxiliary prime above $q/2$ must equal that prime. The set D of admissible inverses $(bu)^{-1}$ thus has size at least $p - 1 - h$. For any $r \in \mathbb{F}_p$, $|D \cap (r - D)| \ge 2|D| - p > 1$ for large $p = q$. Since p is then odd, at most one element has $v = r - v$. Therefore r is a sum of two distinct elements of D, proving (9).

Let s be the number of primes in $(q/2, q)$, so $s \asymp q/\log q$. Greedily choose h-element subsets with pairwise intersections of size at most $h-2$. Each chosen subset precludes at most $h(s-h)+1 = O_h(s)$ candidates, so there are $\gg_h s^{h-1}$ chosen subsets. If different chosen sets have products b, b', then $b/(b, b')$ contains at least two distinct primes above $q/2$ and exceeds q for large q. An equality $qbu = qb'u'$ would force $b/(b, b') \mid u' < q$, which is impossible. This proves disjointness. □

Lemma 2.4 (Descending elimination)**.** *Let $\mathcal{Q}$ be a finite set of prime powers containing all smaller powers of the same prime whenever it contains a power. Call a rational admissible if each prime power dividing its reduced denominator lies in $\mathcal{Q}$. Suppose that for every $q = p^a \in \mathcal{Q}$ with $q > L$ there is a pool $P(q)$ of integers, all having largest prime-power divisor q and all prime-power divisors in $\mathcal{Q}$, with the covering property* (9)*. Then any admissible rational r can be reduced, by subtracting weights of subsets of these pools in decreasing order of q, to a rational r' with $K_L r' \in \mathbb{Z}$. The total subtracted weight is between zero and $\sum_{q>L} w(P(q))$.*

Proof. At the step $q = p^a$, the remaining denominator contains no prime power larger than q by induction. Multiplying the residual by q gives a rational that is reducible modulo p. Select $E \subseteq P(q)$ by the covering property to cancel this residue. The new denominator has p-adic exponent at most $a-1$ and contains no prime power larger than q. The other prime powers are admissible because all denominators subtracted are admissible and admissibility is closed under least common multiples of denominators. Hence the induction continues. The pools for distinct q are disjoint by P^*; all subtracted weights are nonnegative and bounded by the total pool weights. At the end the remaining prime powers are at most L. □

3. High-precision completion and large-prime blocks

Throughout the growing-label argument put

$$M = \lceil \ell^{1000} \rceil, \quad Y = N/(4M), \quad V = \{n \le N : P^*(n) \le Y\}, \quad D = [N] \setminus V. \tag{10}$$

Then $Y > \sqrt{N}$ and $4M < Y$ for large N. The set D is the disjoint union of the blocks $p[z_p]$ for primes $Y < p \le N$, where $z_p = \lfloor N/p \rfloor$, and the set

$$D_2 = \{n \le N : \exists p^a \mid n,\ a \ge 2,\ p^a > Y\}.$$

Indeed two large prime powers of distinct primes cannot both divide $n \le N$, and $p > Y$ implies $p^2 > N$. The number of higher prime powers up to N is $O(\sqrt{N}\ell)$. By the union bound,

$$w(D_2) \ll \sqrt{N}\ell^2/Y = O(N^{-1/2}\ell^{1002}). \tag{11}$$

Partial summation with the PNT gives $\sum_{Y<p\le N} 1/p = O(L/\ell)$; as $H_{z_p} = O(L)$, it follows that

$$w(D) = O(\Delta_N), \qquad |D_2| \le N w(D_2) = O(\sqrt{N}\ell^{1002}). \tag{12}$$

For later use, set $u_p = \log(N/p)$. For any fixed $C > 0$, uniformly for $0 \le a \le b \le CL$,

$$\sum_{a \le u_p < b} z_p = \frac{N}{\ell}(b-a) + O_C\left(\frac{N}{\ell} + \frac{NL^2}{\ell^2}\right). \tag{13}$$

Here the sum is over primes $p \le N$. To prove this, replace each z_p by N/p, at total cost $O(\pi(N)) = O(N/\ell)$. Partial summation of N/t with the quantitative PNT gives the integral $N\int_a^b (\ell - u)^{-1}\,du$, with error $O_C(N(1+L)e^{-c'\sqrt{\ell}})$; endpoint inclusion changes the estimate by at most $O(e^{CL})$. Expanding the integrand proves (13). In particular $|D| = O(NL/\ell)$. If g is piecewise continuous, supported in $[0, CL]$ with $\|g\|_\infty$ and total variation both $O(L)$, partial summation of (13) gives

$$\sum_p z_p g(u_p) = \frac{N}{\ell}\int_0^{CL} g(u)\,du + O_C(NL/\ell). \tag{14}$$

The error term uses $(N/\ell + NL^2/\ell^2)O(L) = O(NL/\ell)$. Jumps can be included by a consistent endpoint convention.

Lemma 3.1 (High-precision disjoint reservoirs)**.** *For $b \in \{200, 201\}$ and any integer $1 \le s \le 2\ell$, there are pairwise disjoint pools $Q_1, \dots, Q_s \subseteq V \cap [N/2] \setminus \{1\}$, each of weight $O(\ell^{-b})$, such that every rational r with $K_Y r \in \mathbb{Z}$ and*

$$d_b/2 \le r \le 2d_b, \qquad d_b = 1/(8\lceil \ell^b \rceil), \tag{15}$$

can be represented by a subset of each Q_i. All bounds are uniform in s.

Proof. Let $J_1 = \lfloor \ell/10 \rfloor$, $k = \lceil \ell^{750} \rceil$. For $N^{1/207} < q = p^a \le Y$, choose for each label i a disjoint set of k primes $B_{i,q} \subseteq [M, 2M] \setminus \{p\}$. There are more than sk available primes, since $sk = O(\ell^{751})$ and $M/\log M \gg \ell^{751}$. Put $P_i(q) = \{qv : v \in B_{i,q}\}$. All these denominators have $P^* = q$ and are at most $N/2$. With $X = N$, $\eta = \ell^{-252}$, the two conditions of Lemma 2.1 hold: the exceptional bound is $O(\ell^{749}/\log \ell) = o(k)$ and $k\eta^2/\ell \asymp \ell^{245}$. Thus (9) holds for each label. Their total weight per label is $O((k/M)H_N) = O(\ell^{-249})$.

For $J_1 < q \le N^{1/207}$ use Lemma 2.3 with $h = 204$, choosing s pairwise disjoint pools at each q. The number of choices is $\gg (q/\log q)^{203} \gg s$. Denominators are below $q^{206} \le N^{206/207} < N/2$; their weight per label is $O((1 + \log J_1)J_1^{-204}) = o(\ell^{-201})$. Pools for different q are disjoint by their largest prime-power divisor. They all lie in V.

Take distinct primes $t_i \in [T, 2T]$, $T = \lceil \ell^b \rceil$, and set $R_i = t_i R(J_1)$, using the mixed-radix pool of Lemma 2.2. Each weighs $O(T^{-1})$, has largest prime-power index t_i, and has denominators at most $2TK_{J_1} = N^{1/10+o(1)} < N/2$. The only possible collision with a previous pool would be for $q = t_i$, an intermediate prime power. Its denominators divided by t_i contain an auxiliary prime above $t_i/2 > J_1$, whereas those of R_i do not. This rules out collisions for all labels. Let $Q_i = R_i \cup \bigcup_{q > J_1} P_i(q)$. Descending elimination (Lemma 2.4) subtracts weight $o(d_b)$ and leaves r' with $K_{J_1} r' \in \mathbb{Z}$ and $0 < r' \le 2d_b$. Then $0 < t_i r' \le 1/2$, so the mixed-radix lemma, applied to $t_i r'$ and scaled by t_i, completes r' inside R_i. □

Corollary 3.2 (Coarse scalar threshold)**.** *At $N = T_*(m)$, as $m \to \infty$, one has $m \sim \ell$ and $0 \le H_N - m = O(\Delta_N)$.*

Proof. Take one pool from Lemma 3.1 with $b = 200$. If $m \sim \ell$ and $H_N - m \ge C\Delta_N$ for a sufficiently large C, then $w(V \setminus Q_1) \ge m$. Remove denominators from $[N^{1/4}, N^{1/2}]$, all of which are smooth before excluding the pool, until the remaining weight first falls below $m - d_{200}$. The available weight is $(1/4 + o(1))\ell$, exceeding the required removal $o(\ell)$, and the last increment is at most $N^{-1/4}$. The residual is admissible and lies in $[d_{200}/2, 2d_{200}]$. Completing it gives the scalar representation. Now take $N = \lceil \exp(m - \gamma + C'(\log m)^2/m) \rceil$ with C' sufficiently large; the harmonic asymptotic gives the upper bound, and $H_{T_*(m)} \ge m$ gives the lower bound. Neither disjoint partitions nor any sharp scalar constant has been used. □

4. Local saturation at the scalar threshold

For primes $p > Y$ let

$$\lambda_p = \min\{w(E) : E \subseteq [z_p],\ w(E) \equiv H_{z_p} \pmod p\}, \qquad \Lambda(N) = \sum_{Y < p \le N} \lambda_p/p. \tag{16}$$

For $A \subseteq [N]$ with integral weight m, put $A_p = \{j \le z_p : pj \in A\}$. Multiplying $w(A)$ by p and reducing modulo p gives $w(A_p) \equiv 0 \pmod p$ (all other terms vanish modulo p). Thus $e_p := w([z_p] \setminus A_p) - \lambda_p \ge 0$. Since V, D_2 and the large-prime blocks partition $[N]$, the following equality is exact:

$$H_N - m - \Lambda(N) = w(V \setminus A) + \sum_{p > Y} \frac{e_p}{p} + w(D_2 \setminus A). \tag{17}$$

In particular $H_N - m \ge \Lambda(N)$. We only need $0 \le \Lambda(N) \le w(D) = O(\Delta_N)$ in this paper.

Lemma 4.1 (Local saturation)**.** *At $N = T_*(m)$ for all sufficiently large m,*

$$0 \le H_N - m - \Lambda(N) = O(\ell^{-200}). \tag{18}$$

Consequently every representation support A of weight m satisfies

$$W := w(V \setminus A) = O(\ell^{-200}), \quad [\ell^{160}] \subseteq A, \quad \sum_{p>Y} e_p/p = O(\ell^{-200}). \tag{19}$$

The constants are independent of A.

Proof. For each $p > Y$, choose a deletion attaining λ_p, and retain its complementary multiple block $F_p \subseteq p[z_p]$. Its weight has reduced denominator dividing K_Y: the factor p cancels by the congruence, and all cofactors $j \le z_p \le 4M$ are below Y. The union F of these retained blocks is disjoint from V and has weight at most $w(D)$. The exact available weight is

$$w(V \cup F) = H_N - \Lambda(N) - w(D_2).$$

To work in $[N-1]$, delete N if it is smooth, or delete the entire retained block containing N if N is a multiple of a prime above Y. If $N \in D_2$ it is already absent. The further weight lost is at most $1/N + \max_{p>Y} H_{z_p}/p \ll ML/N = o(\ell^{-200})$. Deleting an entire block preserves the property that the retained rough weight has denominator dividing K_Y.

Reserve a single pool from Lemma 3.1 with $b = 200$, contained in $V \cap [N/2]$. If $H_N - m - \Lambda(N) > C\ell^{-200}$ with C sufficiently large, the remaining weight after these deletions and the pool reservation is at least m, by (11). The excess is $O(\Delta_N)$ by Corollary 3.2. As in its proof, delete from $[N^{1/4}, N^{1/2}]$ until the residual is $d_{200} + O(N^{-1/4})$. These denominators are smooth; the interval has available weight $(1/4 + o(1))\ell$. The residual has denominator dividing K_Y and is in the required range for Lemma 3.1. Completing it gives an integral representation of m in $[N-1]$, a contradiction. This proves (18). The first and last estimates of (19) follow from nonnegativity in (17). If $n \le \ell^{160}$ were missing, then $n \in V$ and $W \ge 1/n \ge \ell^{-160}$, contrary to the first estimate. This proves the initial-interval claim. □

5. Disjoint reservoirs inside the prescribed support

The following lemma is applied to the smooth part of a representation support, or to V itself. Its reservoirs must lie entirely in the prescribed smooth set.

Lemma 5.1 (Internal reservoirs)**.** *Let $A \subseteq V$ have $w(V \setminus A) = O(\ell^{-200})$ and $[\ell^{160}] \subseteq A$, and let s be any positive integer at most 2ℓ. There exist a set $F \subseteq A \setminus \{1\}$, a set of bad prime powers satisfying* (20)*, a downward-closed set $\mathcal{Q}$ of prime powers (in the sense of Lemma 2.4), and pairwise disjoint sets $Q_1, \dots, Q_s \subseteq A \setminus (F \cup \{1\})$ with the following properties:*

1. *$F = A \cap \bigcup_{q \text{ bad}} \{n : q \mid n\}$ has weight $O(\ell^{-199}L) = o(1)$ and cardinality $O(N\ell^{-200}L)$. All bad powers exceed $2M_0$, where $M_0 = \lceil \ell^{24} \rceil$, and $\mathcal{Q}$ is exactly the set of prime powers at most Y not divisible by any bad power. Every $n \in A \setminus F$ has only such prime-power divisors;*
2. *with $T_0 = \lceil \ell^3 \rceil$ and $d_0 = 1/(8T_0)$, each Q_i has weight $2d_0 \le w(Q_i) = O(\ell^{-3})$;*
3. *any admissible rational r (relative to $\mathcal{Q}$) with $d_0/2 \le r \le 2d_0$ is $w(E)$ for some $E \subseteq Q_i$, for every i.*

All bounds are uniform in A, s subject to a uniform bound on W in (19)*. If $A = V$, the construction has $F = \varnothing$ and $\mathcal{Q}$ equal to all prime powers at most Y.*

Proof. Put

$$M_0 = \lceil \ell^{24} \rceil, \quad k_0 = \lceil \ell^{18} \rceil, \quad L_1 = \lfloor \ell/10 \rfloor, \quad L_0 = \lfloor \log \ell \rfloor.$$

Freezing. For every $2M_0 < q = p^a \le Y$, form the candidate set

$$B_q = \{b \in [M_0, 2M_0] : b \text{ prime},\ b \ne p,\ qb \in A\}.$$

Call q bad if $|B_q| < sk_0$. All candidate denominators qb, whether in A or not, are in V and at most $2M_0Y < N$, with $P^*(qb) = q$. There are $\asymp M_0/\log M_0$ primes in the auxiliary interval, while $sk_0 = O(\ell^{19}) = o(M_0/\log M_0)$. For each bad q, the missing candidates have total weight $\gg 1/(q \log M_0)$. Missing candidates for different q are distinct. Consequently

$$\sum_{q \text{ bad}} \frac{1}{q} \ll W \log M_0 = O(\ell^{-200} \log \ell). \tag{20}$$

Set

$$F = A \cap \bigcup_{q \text{ bad}} \{n \le N : q \mid n\}.$$

Equation (20) gives $w(F) \le H_N \sum_{q \text{ bad}} q^{-1} = O(\ell^{-199}L)$ and $|F| \le N \sum_{q \text{ bad}} q^{-1} = O(N\ell^{-200}L)$. The admissible set $\mathcal{Q}$ consists of the prime powers at most Y that are not divisible by any bad prime power. This set is downward closed, includes all prime powers at most $2M_0$, and contains the prime powers of every denominator in $A \setminus F$. The set F excludes 1. If $A = V$, no q is bad by the prime supply bound, and $F = \varnothing$.

Large prime powers. For each admissible $2M_0 < q = p^a \le Y$, choose disjoint k_0-element subsets $B_{i,q}$ of B_q for $1 \le i \le s$; this is possible by admissibility. Set $P_i(q) = \{qb : b \in B_{i,q}\}$. All denominators belong to $A \setminus F$, since q is admissible and each auxiliary prime $b \le 2M_0$ is distinct from p. Their largest prime-power divisor is q. Lemma 2.1 with $X = N$, $\eta_0 = \ell^{-8}$, $M = M_0$, $k = k_0$ applies: the exceptional bound is $O(\ell^{17}/\log \ell) = o(k_0)$ and $k_0\eta_0^2/\ell \asymp \ell \to \infty$. Thus (9) holds separately for each pool. For each i their total weight is $O((k_0/M_0)H_N) = O(\ell^{-5})$.

Intermediate prime powers. For $L_1 < q \le 2M_0$ use Lemma 2.3 with $h = 4$, selecting s disjoint pools $P_i(q)$. There are enough choices since $(q/\log q)^3 \gg \ell^3/(\log \ell)^3 \gg s$ uniformly in this range. All these denominators are smaller than $q^6 \le (2M_0)^6 = O(\ell^{144}) < \ell^{160}$. Hence they all belong to A by (19). Their prime powers are at most $2M_0$, so they avoid F. They have the required modular covering property. The total weight for each i is $O((1+\log L_1)L_1^{-4}) = o(\ell^{-3})$. Combining the large and intermediate pools gives total weight $o(d_0)$ per label. Different q have different largest prime-power indices, and the choices for the same q are disjoint across labels.

Small prime powers, with a separate scaling for each label. Choose distinct primes $t_1, \dots, t_s \in [T_0, 2T_0]$ by the PNT. For $L_0 < q \le L_1$ construct one unscaled pool $P^{\text{small}}(q)$ by Lemma 2.3 with $h = 4$. Put $P^{\text{small}} = \bigcup_{L_0 < q \le L_1} P^{\text{small}}(q)$ and define

$$R_i = t_i\big(P^{\text{small}} \cup R(L_0)\big), \qquad 1 \le i \le s.$$

The unscaled union has weight $O(1)$: the first part weighs $O((1 + \log L_0)L_0^{-4}) = o(1)$ and the mixed-radix part has bounded weight by Lemma 2.2. Thus $w(R_i) = O(T_0^{-1})$. Its denominators are at most $\max(2T_0L_1^6, 2T_0K_{L_0}) = O(\ell^9)$, because $K_{L_0} = \ell^{1+o(1)}$. They belong to A by (19). Their largest prime-power divisor is t_i since all unscaled prime powers are at most $L_1 < t_i$. Hence they avoid F and are disjoint across labels. Within each R_i, dividing by t_i separates the pools with different prime-power indices and the mixed-radix part. The only possible intersection with a previously constructed pool is an intermediate pool for $q = t_i$ (indeed $L_1 < t_i < 2M_0$). Such a denominator divided by t_i has an auxiliary prime factor above $t_i/2 > L_1$, whereas every prime-power divisor of the unscaled denominators of R_i is at most L_1. This excludes that intersection, for every pair of labels.

Completion and total weight. Define $Q_i = R_i \cup \bigcup_{q \in \mathcal{Q},\, q > L_1} P_i(q)$. The preceding checks prove membership in $A \setminus (F \cup \{1\})$ and disjointness. Since $2t_i \in t_iR(L_0)$, $w(Q_i) \ge 1/(2t_i) \ge 1/(4T_0) = 2d_0$, and the weight upper bound is $O(\ell^{-3})$.

Let r be admissible and in $[d_0/2, 2d_0]$. Apply Lemma 2.4 to the large and intermediate pools for label i (they have only admissible prime-power divisors). It leaves r' with $K_{L_1}r' \in \mathbb{Z}$ and $d_0/2 - o(d_0) \le r' \le 2d_0$. Now $t_i r'$ has denominator dividing K_{L_1} and

$$1/16 - o(1) \le t_i r' \le 1/2.$$

Apply Lemma 2.4 to the unscaled pools $P^{\mathrm{small}}(q)$, with all prime powers at most L_1 admissible and cutoff L_0. Their total weight is $o(1)$, so the remaining rational is in $(0,1)$ and has denominator dividing K_{L_0}. Lemma 2.2 completes it inside $R(L_0)$. Scaling every denominator in these last two steps by t_i represents r' inside R_i. All pieces are disjoint, which proves property (3) and the lemma. □

6. A robust rational-grid reservoir

We now construct small-denominator modular covering even when some candidate denominators are absent. We state the lemma with a fixed bound on the missing weight and with uniform hypotheses in the forbidden prime set, as needed for many labels simultaneously.

Theorem 6.1 (Robust rational-grid covering)**.** *Let $y \to \infty$ and $D_0 = \lceil C_1 y(\log y)^{12} \rceil$ with a sufficiently large absolute C_1. Let $\mathcal{B} \subseteq [D_0]$ satisfy*

$$w([D_0] \setminus \mathcal{B}) \le C_2(\log y)^{-28}, \tag{21}$$

where C_2 is fixed. Let $\mathcal{T}$ be a set of primes greater than $3(\log y)^{12}$, with $|\mathcal{T} \cap [y]| \log y = o(y)$. There are an integer $K' \mid K_y$ with $\log K' = (1+o(1))y$ and a fixed pool $G \subseteq \mathcal{B}$ such that each $n \in G$ divides K' and has no prime factor in $\mathcal{T}$. Every point of $(K')^{-1}\mathbb{Z} \cap [1/4, 1/2]$ is $w(E)$ for some $E \subseteq G$. In particular, for every prime $p > y$ with $K' > 4p$, the reciprocal subset sums of G cover $\mathbb{F}_p$ with weights in $[1/4, 1/2]$. The conclusions are uniform if the $o(y)$ hypothesis holds uniformly.

Proof. Let $J = \lfloor (\log\log y)/2 \rfloor$. For each prime power $q = s^a \in (J, y]$, put $M_q = \lceil (\log q)^{12} \rceil$ and $k_q = \lceil (\log q)^{10} \rceil$. Candidate denominators are qv, where v ranges over primes in $[M_q, 2M_q] \setminus \{s\}$. Uniformly for $q > J$ and large y, we have $2M_q < q$ since $J \to \infty$. All candidates have $P^* = q$, divide K_y, and are at most D_0. Call q damaged if fewer than k_q candidates belong to $\mathcal{B}$. For a damaged q, the PNT shows that at least $\gg M_q/\log M_q$ candidates are missing, so their missing weight is $\gg 1/(q \log M_q)$. The missing candidates are disjoint across q by P^*. Writing $\mathcal{J}$ for the damaged prime powers and $L_y = \log y$, we obtain

$$\sum_{q \in \mathcal{J}} \frac{1}{q} \ll L_y^{-28} \log L_y. \tag{22}$$

In particular no $q \le 3L_y^{12}$ is damaged, since its reciprocal would be at least $(3L_y^{12})^{-1}$.

Delete from the prime-power structure of K_y each power of a prime in $\mathcal{T}$, and each damaged q along with all higher powers of that prime divisible by q. The remaining powers are downward closed. Let K' be the product of the maximal remaining powers, one for each surviving prime. Each damaged q accounts for a loss of at most $\log y$ in $\log K_y$, and the number of damaged powers is at most $y \sum_{q \in \mathcal{J}} q^{-1}$. Hence

$$0 \le \log K_y - \log K' \le y \log y \sum_{q \in \mathcal{J}} q^{-1} + |\mathcal{T} \cap [y]| \log y = o(y).$$

The PNT gives $\log K_y = (1+o(1))y$, proving the claimed size. Every prime power at most J survives.

For each surviving $q = s^a > J$ choose a fixed set of k_q usable candidate primes and let $P(q)$ consist of their products with q. Each auxiliary prime is at most $3L_y^{12}$ and is distinct from s. No

such prime is in $\mathcal{T}$ or is damaged. Thus all selected denominators divide K', and their prime-power structure is admissible for descending elimination. Lemma 2.1, with $X = q$ and $\eta = (\log q)^{-4}$, applies uniformly: its exceptional bound is $O((\log q)^9/\log\log q) = o(k_q)$ and $k_q\eta^2/\log q \to \infty$. Therefore the pools satisfy (9). Their total weight is

$$\sum_{q>J} w(P(q)) \ll \sum_{n>J} \frac{1}{n(\log n)^2} = O(1/\log J) = o(1).$$

For the mixed-radix pool $R(J)$, the largest denominator is at most $K_J = (\log y)^{1/2+o(1)}$. All integers up to K_J belong to $\mathcal{B}$: otherwise the missing weight is at least $1/K_J$, contradicting (21). Every denominator of $R(J)$ divides K', and these denominators have no forbidden prime factors. Set $G = R(J) \cup \bigcup P(q)$, a disjoint union indexed by the surviving prime powers above J. It is fixed before choosing a target.

For any $r \in (K')^{-1}\mathbb{Z} \cap [1/4, 1/2]$, Lemma 2.4 reduces it to a positive rational below 1 with denominator dividing K_J; positivity follows from the $o(1)$ bound just proved. Lemma 2.2 then completes it exactly in $R(J)$. No denominator is repeated. Finally, if $p > y$ and $K' > 4p$, then $p \nmid K'$, and any residue h is $v/K' \pmod p$ for some integer $v \in [K'/4, K'/2]$, because that interval has length greater than p and one may prescribe $v \equiv hK' \pmod p$. This proves modular coverage, including the real weight range. □

Corollary 6.2 (Local deletion bound)**.** *For the parameters in* (10)*, uniformly for primes $p > Y$ and $z_p \geq 2D_0$, where $y = \lceil 2\ell \rceil$ and $D_0 = \lceil C_1 y(\log y)^{12} \rceil$, one has*

$$\lambda_p \ll \ell L^{12}/z_p. \tag{23}$$

Proof. Apply Theorem 6.1 with $\mathcal{B} = [D_0]$, $\mathcal{T} = \varnothing$. Its K' is $N^{2+o(1)} > 4p$ and $p > y$. Let $s = \lfloor z_p/D_0 \rfloor$, so $s \geq z_p/(2D_0)$ and $s < p$. Represent the residue $sH_{z_p} \pmod p$ using a subset of $[D_0]$ of weight at most $1/2$, and multiply its denominators by s. The resulting distinct denominators lie in $[z_p]$, their weight has the required residue and is at most $1/(2s) \ll D_0/z_p$. This proves the bound. □

7. The necessary entropy deficit

The upper bound for $\mathcal{P}(N, m)$ does not require a fixed support. We may count all label assignments on $[N]$, including label zero for unused denominators, and enforce only restrictions within disjoint large-prime blocks.

Proposition 7.1 (Universal upper bound)**.** *If $m \sim \ell$ and $\mathcal{P}(N, m) > 0$, then*

$$\log \mathcal{P}(N, m) \leq N \log m - \tfrac{3}{2} B_N + O(E_N). \tag{24}$$

Proof. For a prime $p > Y$, every label $i \neq 0$ has the property that the cofactors j of its denominators in $p[z_p]$ have reciprocal sum zero modulo p. If $z_p \leq \ell/2$, put $t = \lfloor \ell/2 \rfloor$. By the PNT, $K_t H_t = N^{1/2+o(1)} < Y < p$ for large N, so $K_{z_p} H_{z_p} \leq K_t H_t < p$. A nonempty zero-sum subset E would give a positive integer $K_{z_p} w(E) < p$ divisible by p, which is impossible. Thus all labels in these blocks are zero.

If a nonzero label uses $k \geq 1$ distinct cofactors $j \leq z_p$, multiplying the zero-sum congruence by their product gives a positive integer divisible by p and bounded by $k z_p^{k-1} \leq z_p^k$. Hence $k \geq \log p/\log z_p$ whenever $z_p \geq 2$. The number of nonzero labels appearing in a block is at most $R_p = \lfloor z_p \log z_p/\log p \rfloor$. The number of block assignments is at most

$$\sum_{j=0}^{\min(R_p, m)} \binom{m}{j} (j+1)^{z_p} \leq (m+1)^{R_p+1} (R_p+1)^{z_p}. \tag{25}$$

For $L \le u_p \le 2L$, we have $z_p \asymp e^{u_p}$ and $\log p = \ell - u_p$, so

$$\log(R_p+1) \le u_p - L + O(\log L), \qquad \frac{R_p+1}{z_p}\log(m+1) = O(L^2/\ell).$$

In the first bound $R_p+1 \ll e^{u_p}L/\ell$ follows from $u_p \ge L$; in the second use $z_p \ge \ell-1$. Thus relative to the trivial block count $(m+1)^{z_p}$, the entropy lost in this range is at least $z_p(2L-u_p-O(\log L))$, since $\log(m+1) = L + o(1)$. Outside the two specified ranges use the trivial block bound. The short gap of length $\log 2$ before L costs at most $O(NL/\ell)$ by (13). The $O(\log L)$ per-denominator error in the middle range sums to $O(E_N)$. Using (14) with the function

$$g(u) = \begin{cases} L & 0 \le u \le L, \\ 2L-u & L < u \le 2L, \\ 0 & u > 2L, \end{cases} \tag{26}$$

we obtain a deficit of at least $(N/\ell)\int g(u)\,du - O(E_N) = \frac{3}{2}B_N - O(E_N)$ relative to $N\log(m+1)$. Since $N\log(m+1) - N\log m = O(N/\ell)$, this proves (24). All blocks used are disjoint because their primes exceed Y for large N. □

8. Local multi-label completion in large-prime blocks

We next construct enough assignments to match the profile (26). In this section fix $m \sim \ell$. For each prime $p > Y$ suppose that a cofactor set $A_p \subseteq [z_p]$ with $w(A_p) \equiv 0 \pmod p$ has been prescribed. Define $e_p = w([z_p] \setminus A_p) - \lambda_p$. Assume

$$\sum_{p>Y} e_p/p = O(\ell^{-200}). \tag{27}$$

This holds for a minimal scalar representation by Lemma 4.1, and also if each A_p is chosen as a complement of a local optimum, in which case every $e_p = 0$. We use labels $2,\dots,m$ with label m as a common final label. Label 1 will later be assigned $\{1\}$. Let $\mathcal{Q}_{\mathrm{bad}}$ be a set of global bad prime powers, either empty or as in Lemma 5.1; in particular they exceed $2^{\lceil \ell^{24}\rceil}$ and their reciprocal sum is $O(\ell^{-200}L)$ by (20).

Lemma 8.1 (Local assignments attaining the entropy profile)**.** *There is a family of assignments of the entire rough support $\bigcup_{p>Y} pA_p$ to labels $2,\dots,m$ with the following properties.*

1. *In every block, each label has cofactor reciprocal sum zero modulo p. Every nonfinal label has no cofactor divisible by a global bad power. If $\mathcal{Q}_{\mathrm{bad}}$ is empty, this assertion applies to every label.*
2. *The assignments preserve all denominators of the prescribed support. There are fixed sets $U_p \subseteq A_p$ and fixed allowed label sets of size r_p+1, with $0 \le r_p \le m-2$, such that every assignment on all the U_p to their respective label sets admits a completion. It can be recovered from the completed assignment by restriction to pU_p.*
3. *The logarithm of the number of these free assignments, denoted by $\mathcal{L}$, satisfies*

$$\log \mathcal{L} \ge \left(\sum_{p>Y} z_p\right)\log m - \tfrac{3}{2}B_N - O(E_N). \tag{28}$$

The bounds are uniform under (27) *and the bound on the global bad powers.*

Proof. Call a prime exceptional if $e_p > \ell^{-10}$. Then the total size of its blocks is at most

$$\sum_{p \text{ exceptional}} z_p \le N\ell^{10}\sum_{p>Y} e_p/p = O(N\ell^{-190}). \tag{29}$$

Freeze their whole prescribed supports, assigning them to label m. For nonexceptional primes with $z_p \geq 2D_0$, Corollary 6.2 gives

$$w([z_p] \setminus A_p) \ll \ell L^{12}/z_p + \ell^{-10}. \tag{30}$$

In every block, assign all cofactors divisible by any global bad power to label m; their number is at most $z_p \sum_{q \in \mathcal{Q}_{\mathrm{bad}}} 1/q = O(z_p \ell^{-200} L)$. Such a fixation is made before any free assignment. It is redundant when there are no global bad powers.

Define the three dividing points

$$Z_0 = \ell L^{80}, \qquad Z_1 = \ell^2/L^{80}, \qquad Z_2 = \ell^2 L^{80}. \tag{31}$$

For nonexceptional primes outside $Z_0 \leq z_p \leq Z_1$ or $z_p \geq Z_2$, simply freeze the entire prescribed block to label m. In these frozen blocks put $U_p = \varnothing$ and $r_p = 0$. Congruence zero holds by the hypothesis on A_p.

Intermediate blocks. Suppose $Z_0 \leq z = z_p \leq Z_1$ and p is nonexceptional. Let $y = \lceil 2\ell \rceil$, $D_0 = \lceil C_1 y (\log y)^{12} \rceil$, as before. Set

$$T = \left\lfloor \frac{z}{\ell L^{40}} \right\rfloor, \qquad r = \left\lfloor \frac{T}{10 \log T} \right\rfloor. \tag{32}$$

Here $T \gg L^{40}$, $T \ll \ell/L^{120}$, and $1 \leq r \leq m-2$ for large N. Choose r distinct primes $t_i \in [T, 2T]$ and fix a bijection to r distinct nonfinal labels (for example labels $2, \ldots, r+1$). Put $\mathcal{T} = \{t_i\}$. The supply follows from the PNT. All $t_i > 3(\log y)^{12}$ and $|\mathcal{T} \cap [y]| \log y \leq r \log y = o(y)$ uniformly. Also $2TD_0 \ll z/L^{28}$, so for each i the set

$$\mathcal{B}_i = \{n \in [D_0] : t_i n \in A_p\}$$

is well defined, and by (30),

$$\begin{aligned} w([D_0] \setminus \mathcal{B}_i) &\leq t_i w([z] \setminus A_p) \\ &\ll T(\ell L^{12}/z + \ell^{-10}) = O(L^{-28}). \end{aligned} \tag{33}$$

This uses $T \leq \ell$ for the last term. Since $\log y \sim L$, the robust grid hypotheses hold uniformly. Theorem 6.1 provides a pool $G_i \subseteq \mathcal{B}_i$, none of whose denominators is divisible by any t_j, covering all residues modulo p (its grid denominator is $N^{2+o(1)} > 4p$). Set $Q_{i,p} = t_i G_i$. Scaling by t_i preserves modular coverage because $t_i < p$. These pools lie in A_p and are pairwise disjoint: an equality $t_i n = t_j n'$ with $i \neq j$ would force $t_i \mid n'$, contrary to the forbidden-prime condition. They also avoid every cofactor containing a global bad power, since their largest possible cofactor is $O(\ell^2 L^{52}) < 2\lceil \ell^{24} \rceil$. Thus no pool intersects the preassigned frozen cofactors. The pools use at most $rD_0 = O(z/L^{28})$ cofactors in total.

Large blocks. For nonexceptional primes with $z = z_p \geq Z_2$, take $T = \lceil \ell L^{40} \rceil$ and $r = m-2$. There are more than r primes in $[T, 2T]$; these primes exceed y, so $\mathcal{T} \cap [y] = \varnothing$. The maximum scaled denominator is $2TD_0 = O(\ell^2 L^{52}) = O(z/L^{28})$. The right side of (33) is now $O(L^{-28}) + O(\ell^{-9} L^{40}) = O(L^{-28})$. The same construction yields disjoint pools, avoiding the fixed cofactors, with total size $rD_0 = O(z/L^{68})$.

Completion and recovery. In either active range, define U_p by removing the preassigned cofactors and all the local pools from A_p, and put $r_p = r$. The freely available labels for U_p are the chosen r_p labels and label m. Choose their assignments arbitrarily. For each chosen nonfinal label i with initial cofactor sum $h_i \pmod p$, select a subset of its fixed pool $Q_{i,p}$ with sum $-h_i \pmod p$; this is possible by modular coverage. Give all unused pool cofactors to label m. The sum of all cofactors in A_p is zero modulo p, so label m also has zero cofactor sum. All other labels not selected have no cofactors in this block. The resulting assignment of A_p is a partition; restriction to U_p recovers its free labels exactly. All choices of correction subsets may be made

deterministic by a fixed ordering. For frozen blocks there is just one assignment. This proves the first two claims, since no nonfinal cofactor has a global bad divisor.

Entropy calculation. In active blocks the number of missing cofactors is at most $zw([z] \setminus A_p)$, since each reciprocal is at least $1/z$. In both ranges this weight is $O(L^{-68}) + O(\ell^{-10})$ by (31) and (30). Together with pool and frozen counts it follows uniformly that

$$|U_p| = z_p(1 - O(L^{-28})). \tag{34}$$

In the intermediate range (32) implies

$$\log(r_p + 1) = \log z_p - L + O(\log L).$$

Indeed $\log T = \log z_p - L - 40 \log L + O(1)$, and $\log \log T = O(\log L)$ because $T \le \ell$; the integer parts cause a bounded error as $T \gg L^{40}$. Hence the deficit of this block relative to $z_p \log m$ is $z_p(2L - u_p + O(\log L)) + O(z_p L^{-27})$. In the large range the deficit is $O(z_p L^{-27} + z_p/\ell)$, by $r_p + 1 = m - 1$. The sum of these lower-order terms over all active primes is $O(E_N)$, since $\sum_{p>Y} z_p = O(NL/\ell)$.

Exceptional blocks cost $O(N\ell^{-190}L) = o(E_N)$ by (29). The remaining frozen blocks below Z_0 contribute at most $z_p \log m$ each. On $u_p \le L$ this matches $z_p g(u_p)$ up to $O(z_p)$; the extra range to $\log Z_0 = L + 80 \log L$ has length $O(\log L)$ and total cost $O(E_N)$ by (13). Likewise the frozen transition between Z_1 and Z_2 has logarithmic width $160 \log L$ and costs $O(E_N)$. Replacing $\log z_p$ by u_p in the active intervals changes it by $O(1)$, and rounding at the cutoffs changes the logarithmic widths by at most $O(1)$. Thus the total deficit is at most $\sum_{p>Y} z_p g(u_p) + O(E_N) = \frac{3}{2} B_N + O(E_N)$, using (14) and (26). Since the free assignment count is $\mathcal{L} = \prod_{p>Y} (r_p+1)^{|U_p|}$, this proves (28). □

Remark 8.2 (Compatibility of the local and global arithmetic). For each label, a block contribution is $p^{-1} \sum_{j \in C_{i,p}} 1/j$, with $\sum_{j \in C_{i,p}} 1/j \equiv 0 \pmod p$. Its reduced denominator therefore does not contain p. Every other prime power in that denominator divides some cofactor $j \le z_p < 4M < Y$. For nonfinal labels none of these powers is divisible by a global bad power. This gives exactly the admissibility condition of Lemma 5.1, after summation across blocks. In the absence of global bad powers it gives the K_Y-denominator condition for every label. No assumption is made about the reduced denominator of an arbitrary set assigned to the final label in the saturated case; its exact sum will follow from conservation of total weight.

9. A balanced random assignment lemma

We record the probabilistic part in a form independent of the arithmetic construction.

Lemma 9.1 (Balanced high-denominator labels)**.** *Let $B \subseteq [N] \setminus \{1\}$, $J \sim \ell$, and let $c_1, \dots, c_J \ge 0$ satisfy $\max_i c_i = o(1)$ and*

$$w(B) + \sum_{i=1}^{J} c_i = J + o(Jd_0), \qquad U = B \cap (N^{3/4}, N], \quad w(U) = (1/4 + o(1))\ell, \tag{35}$$

where $d_0 = 1/(8\lceil \ell^b \rceil)$ with $b \in \{3, 201\}$. There is a fixed partition of $B \setminus U$ into sets $B_1^{\mathrm{low}}, \dots, B_J^{\mathrm{low}}$ such that at least $(1 - o(1))J^{|U|}$ functions $\sigma : U \to [J]$ have

$$\left| c_i + w(B_i^{\mathrm{low}}) + w(\sigma^{-1}(i)) - 1 \right| \le d_0/4 \quad \text{for every } i \in [J]. \tag{36}$$

Proof. Assign $B \setminus U$ in increasing order of denominators to the label whose current adjusted load $c_i + w(B_i^{\mathrm{low}})$ is smallest, breaking ties by the label index. Write $a_i = w(B_i^{\mathrm{low}})$ after this process. The final average adjusted load is $1 - w(U)/J + o(d_0) = 3/4 + o(1)$ by (35). Set $n_0 = \lceil \ell^{205} \rceil$.

In particular $n_0^{-1} = o(d_0)$ for either exponent, and $H_{n_0} = O(L)$. If a group last received a denominator at most n_0, its final adjusted load is at most

$$\frac{1}{2} + \frac{H_{n_0} + \sum_i c_i}{J} = 1/2 + o(1).$$

Here the pre-addition minimum is at most the pre-addition average, and $1/n \le 1/2$ since $1 \notin B$. A group receiving nothing has load $c_i = o(1)$. Thus a group of maximal final load must have received its last denominator above n_0. Immediately before that addition it had minimum load, which cannot exceed the final minimum because adjusted loads only increase. Its last increment is at most $1/n_0$. Consequently

$$\max_i(c_i + a_i) - \min_i(c_i + a_i) \le 1/n_0 = o(d_0). \tag{37}$$

Give the denominators in U independent uniform labels in $[J]$. If $Z_i = w(\sigma^{-1}(i))$, then $\mathbb{E}Z_i = w(U)/J$ and

$$\mathrm{Var}(Z_i) \le \frac{1}{J} \sum_{n > N^{3/4}} n^{-2} = O((JN^{3/4})^{-1}).$$

The average adjusted load plus $\mathbb{E}Z_i$ is $1 + o(d_0)$, so (37) gives $c_i + a_i + \mathbb{E}Z_i = 1 + o(d_0)$ uniformly. Chebyshev's inequality and the union bound show that the probability some $|Z_i - \mathbb{E}Z_i| > d_0/8$ is $O(N^{-3/4} d_0^{-2}) = o(1)$. For large N the deterministic error is at most $d_0/8$. Thus (36) holds with probability $1 - o(1)$, which is exactly the asserted fraction of all $J^{|U|}$ assignments. □

10. Global completion and the main theorems

We now combine the arithmetic and probabilistic constructions, first for any prescribed support satisfying the two saturation estimates, then for larger available capacity.

Proposition 10.1 (Prescribed saturated supports)**.** *Suppose $m \sim \ell$, $A \subseteq [N]$, $w(A) = m$, and*

$$w(V \setminus A) + \sum_{p > Y} e_p/p = O(\ell^{-200}). \tag{38}$$

Then $\log \mathcal{D}(A, m) = N \log m - \frac{3}{2} B_N + O(E_N)$, with a uniform error bound when the constant in (38) is fixed.

Proof. Write $A^{\mathrm{sm}} = A \cap V$. Condition (38) implies $[\ell^{160}] \subseteq A^{\mathrm{sm}}$ just as in Lemma 4.1. Apply Lemma 5.1 to A^{sm} with $s = m - 2$, indexing its pools Q_i by $2 \le i \le m - 1$. Let F_{bad} be its frozen smooth set and $\mathcal{Q}_{\mathrm{bad}}$ its bad powers. Fix $S_1 = \{1\}$. Let $Q = \bigcup_{i=2}^{m-1} Q_i$, $d = 1/(8\lceil \ell^3 \rceil)$, $J = m - 1$, and set

$$B = A^{\mathrm{sm}} \setminus (\{1\} \cup F_{\mathrm{bad}} \cup Q), \qquad U = B \cap (N^{3/4}, N].$$

The disjoint smooth pools are fixed before free assignments. By Lemma 8.1, the large-prime part of A admits $\mathcal{L}$ freely distinguishable local assignments. For each such assignment complete its local blocks by the deterministic rule in that lemma, then assign F_{bad} and $A \cap D_2$ to label m. Let F_i be the already assigned set of label i (outside $B, Q, \{1\}$), and put $f_i = w(F_i)$. Their total weight is $w(A \setminus V) + w(F_{\mathrm{bad}}) = O(\Delta_N) = o(1)$, independent of the free assignment. Thus $\max_i f_i = o(1)$. By Remark 8.2, each f_i for $i < m$ is admissible relative to the internal pools. The total mass identity is

$$w(B) + \sum_{i=2}^{m} f_i + w(Q) = J.$$

To account for unused pool denominators in the final label, choose initial target loads

$$b_i = 1 - d \quad (2 \le i \le m - 1), \qquad b_m = 1 - w(Q) + (m - 2)d.$$

Here $w(Q) \geq 2(m-2)d$, $w(Q) = O(\ell^{-2})$, and $\sum_{i=2}^{m} b_i = J - w(Q)$. Put $c_i = f_i + 1 - b_i$. All offsets are nonnegative, $\max c_i = o(1)$, and $w(B) + \sum c_i = J$ exactly. Since $w(D) = O(\Delta_N)$, $w(V \setminus A) = O(\ell^{-200})$, and the removed frozen sets and pools have weight $o(1)$, we have $w(U) = \ell/4 + O(1)$. Lemma 9.1, with exponent 3, produces at least $(1 - o(1))J^{|U|}$ high-denominator assignments, after a fixed low-denominator partition into B_i^{low} (which may depend on the local assignment through its offsets). For every successful assignment σ, the residual of label $i < m$ is

$$r_i = 1 - f_i - w(B_i^{\mathrm{low}}) - w(\sigma^{-1}(i)) \in [3d/4, 5d/4].$$

All its denominator prime powers are admissible, since those of f_i and of the elements of B are admissible. Choose $E_i \subseteq Q_i$ with $w(E_i) = r_i$ by Lemma 5.1. Form

$$S_i = F_i \cup B_i^{\mathrm{low}} \cup \sigma^{-1}(i) \cup E_i \quad (2 \leq i < m),$$
$$S_m = F_m \cup B_m^{\mathrm{low}} \cup \sigma^{-1}(m) \cup \bigcup_{2 \leq i < m} (Q_i \setminus E_i).$$

All pieces are disjoint, the first $m-1$ labels have weight one, and the union is exactly A. Conservation of weight gives $w(S_m) = m - (m-1) = 1$.

The map from local and global free assignments to these partitions is injective. Indeed, the sets pU_p and U are fixed independently of both assignments, are pairwise disjoint, and are outside every correction pool used at their respective stages or any later stage. For every label i the final intersections $S_i \cap pU_p$ recover the local free labels, and $S_i \cap U = \sigma^{-1}(i)$ recovers the global ones. Deterministic choices of valid subsets can again be made using fixed total orders. The balancing probability tends to one uniformly over the local assignments, since its estimates use only uniform offset and weight bounds. Thus

$$\mathcal{D}(A, m) \geq (1 - o(1))\mathcal{L}(m-1)^{|U|}. \tag{39}$$

Finally, the loss of smooth denominators outside U is at most

$$\begin{aligned} |V| - |U| &\leq N w(V \setminus A) + |F_{\mathrm{bad}}| + N w(Q) + N^{3/4} + 2 \\ &= O(N\ell^{-2}) + O(N\ell^{-200}L) + O(N^{3/4}). \end{aligned}$$

This multiplied by $\log m$ is $o(E_N)$. Since $|D_2| \log m = o(E_N)$ by (12), and V, D_2, and $p[z_p]$ partition $[N]$, combining (39) and (28) proves the lower bound $N \log m - \frac{3}{2}B_N - O(E_N)$. The upper bound follows from Proposition 7.1 because every partition counted by $\mathcal{D}$ is counted by $\mathcal{P}$. □

Proof of Theorem 1.1 and Corollary 1.2. Corollary 3.2 and Lemma 4.1 verify the hypotheses of Proposition 10.1 for every minimal scalar representation A. This proves (3). In particular the number of partitions is positive for all sufficiently large m; hence $N_*(m) \leq T_*(m)$, and the reverse inequality is immediate by union of the m sets. At their common value N, Proposition 7.1 and the lower bound for any such A give $\log \mathcal{P}(N, m) = N \log m - \frac{3}{2}B_N + O(E_N)$. As $\ell = m - \gamma + O(\Delta_N)$ at $T_*(m)$, replacing ℓ by m and L by $\log m$ in the correction terms gives (4), with the resulting errors absorbed by its remainder. No two-term scalar threshold result is required for this step. □

Proposition 10.2 (Available slack)**.** *There is an absolute $C_* > 0$ such that for $m \sim \ell$ and*

$$H_N - m - \Lambda(N) \geq C_* \ell^{-200}, \tag{40}$$

one has $\mathcal{P}(N, m) > 0$ and $\log \mathcal{P}(N, m) = N \log m - \frac{3}{2}B_N + O(E_N)$.

Proof. For every $p > Y$, choose A_p as the complement of a deletion attaining λ_p, so $w(A_p) \equiv 0 \pmod p$ and $e_p = 0$. Keep no elements of D_2. Let $F^* = \bigcup_{p>Y} pA_p$. Then $w(V) + w(F^*) = H_N - \Lambda(N) - w(D_2)$. Reserve label 1 for $\{1\}$ and use Lemma 3.1 with $b = 201$ and $s = m - 1$ to obtain disjoint pools $Q_i \subseteq V$ for every label $2 \leq i \leq m$. Put $Q = \bigcup_i Q_i$ and $d = d_{201}$. The

total pool weight is $O(\ell^{-200})$. Choose C_* greater than the absolute constants in that bound and (11). Then the smooth set $V \setminus (\{1\} \cup Q)$, together with F^*, has weight at least $m-1$. Trim this smooth set by removing denominators in $[N^{1/4}, N^{1/2}]$ until the remaining smooth set B satisfies

$$w(B) + w(F^*) = (m-1)(1-d) + O(N^{-1/4}).$$

The required removal is $o(\ell)$ since $m \sim \ell$; the available weight is $(1/4 + o(1))\ell$. All these choices are independent of future free labels. In particular $U = B \cap (N^{3/4}, N]$ has $w(U) = \ell/4 + O(1)$ and $|V| - |U| \le Nw(Q) + N^{3/4} + N^{1/2} + 2$.

Use Lemma 8.1 on the optimal blocks with the global bad set empty. Each of its $\mathcal{L}$ free assignments completes to a partition of F^*, with all label weights f_i having denominator dividing K_Y by Remark 8.2. Also $0 \le f_i \le w(D) = o(1)$. Apply Lemma 9.1 with offsets $c_i = f_i + d$, now using exponent 201. Its mass assumption holds because $w(B) + \sum c_i = m - 1 + O(N^{-1/4})$; the error is $o((m-1)d)$. The increased greedy cutoff $\lceil \ell^{205} \rceil$ gives a load error $O(\ell^{-205}) = o(d)$ and the random failure probability is $O(N^{-3/4}\ell^{402}) = o(1)$. Thus the residuals of all $m-1$ labels are in $[d/2, 2d]$ and have denominator dividing K_Y. Complete each from its own Q_i, leaving unused correction denominators unused. All m labels now have weight one and are disjoint.

As before, the final intersections with pU_p and U recover the free assignments, so the lower bound is $(1 - o(1))\mathcal{L}(m-1)^{|U|}$. The smooth denominator loss is $o(E_N / \log m)$ from the bound on $|V| - |U|$; D_2 is negligible as well. Lemma 8.1 and Proposition 7.1 give the claimed estimate. □

Proof of Theorem 1.3. For a feasible (N, m) with $m \sim \ell$, choose a representation support A as the union of any existing ordered partition. If $H_N - m - \Lambda(N) < C_*\ell^{-200}$, the identity (17) implies (38); Proposition 10.1 gives the lower bound, and Proposition 7.1 gives the upper bound. Otherwise Proposition 10.2 applies. The constants in the first case are uniform because C_* is fixed, and in the second case the construction uses fixed parameters. This covers arbitrary alternating sequences of the two cases. For the final sufficiency statement, $\Lambda(N) \le w(D) = O(\Delta_N)$ implies (40) whenever $H_N - m \ge C_0\Delta_N$ for C_0 sufficiently large. □

Remark 10.3 (Combining with the scalar capacity expansion). The proof of the threshold equality uses the coarse scalar bound of Corollary 3.2. Combining Corollary 1.2 with the two-term scalar expansion proved in the companion manuscript [2] gives

$$\log N_*(m) = m - \gamma + \frac{(\log m)^2}{2m} + \frac{\log m \log\log m}{m} + O\Big(\frac{\log m \log\log\log m}{m}\Big).$$

This is the only statement here that invokes the sharper capacity asymptotic of that manuscript.

11. Fixed vectors of rational targets

For a fixed positive integer r and a fixed vector $\mathbf{x} = (x_1, \dots, x_r) \in \mathbb{Q}^r_{>0}$, let

$$\mathcal{P}_{\mathbf{x}}(N) = \#\{(S_1, \dots, S_r) : S_i \subseteq [N],\ S_i \cap S_j = \varnothing\ (i \ne j),\ w(S_i) = x_i\}.$$

This is a fixed-dimensional analogue, not a substitution of a growing dimension into an uncontrolled error term.

Theorem 11.1 (Exact fixed-target entropy)**.** *There is a unique vector $\boldsymbol{\lambda} \in (0, \infty)^r$ satisfying*

$$x_i = \int_0^1 \frac{e^{-\lambda_i/t}}{t(1 + \sum_{j=1}^r e^{-\lambda_j/t})}\, dt, \qquad 1 \le i \le r. \tag{41}$$

For $t \in (0,1]$ put $Z(t) = 1 + \sum_{j=1}^{r} e^{-\lambda_j/t}$, $p_0(t) = 1/Z(t)$ and $p_i(t) = e^{-\lambda_i/t}/Z(t)$ for $i \geq 1$. Then

$$\log \mathcal{P}_{\mathbf{x}}(N) = N\mathcal{H}(\mathbf{x}) + o(N), \tag{42}$$

$$\begin{aligned}\mathcal{H}(\mathbf{x}) &= -\int_0^1 \sum_{i=0}^{r} p_i(t) \log p_i(t)\, dt \\ &= \min_{\boldsymbol{\lambda} \geq 0} \left(\sum_{i=1}^{r} \lambda_i x_i + \int_0^1 \log \left(1 + \sum_{j=1}^{r} e^{-\lambda_j/t} \right) dt \right). \end{aligned} \tag{43}$$

Proof. Let Φ denote the objective in (43). It is continuous on the closed nonnegative orthant by dominated convergence and coercive because each $x_i > 0$. It is strictly convex: for each $t > 0$, the Hessian of $\log Z(t)$ is t^{-2} times the covariance matrix of the first r indicators for a random label in $\{0, 1, \ldots, r\}$ with all probabilities positive. Any nonzero vector has strictly positive variance since label zero and each other label have positive probability. Thus the integrated objective is strictly convex (one can use pointwise strict convexity even on the boundary), and the minimizer is unique.

No minimizer lies on the boundary. If $\lambda_i = 0$, increasing it by $h \in (0,1)$ decreases the logarithm term by at least $(1 - e^{-h/t})/(r+1)$ for each t (use $\log(1-u) \leq -u$ and $Z(t) \leq r+1$). On $t \in [h, 1]$, $1 - e^{-h/t} \geq h/(2t)$. Hence the change in Φ is at most $hx_i - h\log(1/h)/(2(r+1)) < 0$ for small h. At the interior minimizer differentiation under the integral is justified by the exponential decay near $t = 0$ of $t^{-1}e^{-\lambda_i/t}$, uniformly in a small neighborhood. The first derivative equations are (41). The identity $-\sum p_i \log p_i = \log Z + \sum_{i \geq 1} \lambda_i p_i / t$ proves (43) for this vector.

For the upper bound, the product $\prod_{n \leq N} (1 + \sum_{i=1}^{r} e^{-N\lambda_i/n})$ expands over every assignment including unused denominators. Each exact representation has weight $e^{-N\sum_i \lambda_i x_i}$, so division and a Riemann sum give $\limsup N^{-1} \log \mathcal{P}_{\mathbf{x}}(N) \leq \Phi(\boldsymbol{\lambda}) = \mathcal{H}(\mathbf{x})$.

For the matching lower bound, take $A = V$, apply Lemma 5.1 with $s = r$ (all prime powers up to Y admissible), and reserve its disjoint pools Q_i. Put $U = (V \setminus \bigcup_i Q_i) \cap [N^{1/2}, N]$; then $|U| = N - o(N)$ by (12) and the pool weight bound. Set

$$\mu_i = \sum_{n \in U} \frac{p_i(n/N)}{n} = x_i + o(1), \qquad \widetilde{p}_i(n) = \frac{x_i - d_0}{\mu_i} p_i(n/N) \quad (i \geq 1), \qquad \widetilde{p}_0(n) = 1 - \sum_{i=1}^{r} \widetilde{p}_i(n).$$

The equality for μ_i follows from the Riemann sum for the bounded, continuously extended function $p_i(t)/t$, with only $o(N)$ grid points removed. These are valid probabilities for large N: the factors for $i \geq 1$ tend to one, and $p_0(t) \geq 1/(r+1)$. Label independently with these probabilities. Each random weight $X_i = w(\sigma^{-1}(i))$ has mean $x_i - d_0$ and variance at most $\sum_{n \geq N^{1/2}} n^{-2} = O(N^{-1/2})$. Chebyshev's inequality gives, with probability $1 - o(1)$, that every residual $x_i - X_i$ is in $[d_0/2, 2d_0]$. Since each x_i is a fixed rational, its denominator divides K_Y for large N; the residuals are admissible. Complete them using the separate pools Q_i. Unused labels and unused correction denominators remain unused.

To count successful assignments under this nonuniform distribution, let $I = -\log \mathbb{P}(\sigma)$ be their self-information, a sum of independent random variables. Uniform convergence of the adjusted probabilities to their original values on the grid, continuity of categorical entropy on the closed simplex (with $0 \log 0 = 0$), and $|[N] \setminus U| = o(N)$ imply $\mathbb{E}I = N\mathcal{H}(\mathbf{x}) + o(N)$. Also $\operatorname{Var}(I) = O(N)$, since $u(\log u)^2$ is bounded on $[0,1]$ and the number of labels is fixed; terms of probability zero do not occur. With probability $1 - o(1)$ we have $|I - \mathbb{E}I| \leq N^{3/4}$. Intersect this event with successful completion. Every assignment in the intersection has probability at most $e^{-N\mathcal{H}(\mathbf{x}) + o(N)}$ with a deterministic error uniform on the event, so its cardinality is at least $e^{N\mathcal{H}(\mathbf{x}) - o(N)}$. Intersection

of each final S_i with U recovers all its used labels and hence the unused label as well. The completion is injective and proves the lower bound. This completes (42). □

Remark 11.2 (Comparison with the one-target result). For $r = 1$, (41) reduces to $x = \int_0^1 (t(1 + e^{\lambda/t}))^{-1} dt$. The entropy in bits is $\mathcal{H}(x)/\log 2$, exactly the constant c_x in [3]; their example is $c_1 \approx 0.91117$. For several targets, the shared unused label in $Z(t)$ is essential. In particular $\mathcal{H}(\mathbf{x}) < \log(r + 1)$ since for every $t > 0$ the optimal positive parameters give a nonuniform categorical distribution. When $r \geq 2$, one has the strict competition inequality

$$\mathcal{H}(\mathbf{x}) < \sum_{i=1}^{r} \mathcal{H}(x_i).$$

Indeed, at every $t > 0$ the label is equivalent to its vector of r mutually exclusive indicators. Their joint entropy is strictly smaller than the sum of the marginal binary entropies, since they are not independent (all the positive labels have nonzero probabilities). Each marginal has integral constraint x_i, so its integrated binary entropy is at most $\mathcal{H}(x_i)$ by the one-dimensional variational formula and the relative-entropy inequality.

When $x_i = x$ for every i, uniqueness gives $\lambda_i = \lambda$ and $x = \int_0^1 [t(r + e^{\lambda/t})]^{-1} dt$. One also has

$$(1 - e^{-(r+1)x}) \log(r + 1) \leq \mathcal{H}(x, \ldots, x) < \log(r + 1).$$

For the lower bound consider the categorical distribution uniform on $r + 1$ labels for $e^{-(r+1)x} \leq t \leq 1$, and concentrated on label zero otherwise. It has the required integrals x and the displayed entropy. Its entropy is at most the minimum in (43): pointwise the inequality $-\sum_{i=0}^{r} a_i \log a_i \leq \log Z + \sum_{i \geq 1} \lambda_i a_i / t$ follows from nonnegativity of relative entropy. Thus the fixed-dimensional rate tends to $\log r$ as $r \to \infty$ for each fixed $x > 0$, but this observation alone does not control a growing number of groups at finite N.

12. Scope and proof dependencies

The entropy constant 3/2 follows from the integral of the necessary profile (26), matched by the two-scale assignment construction. The local robust covering theorem is used only at $y = \lceil 2 \log N \rceil$ and needs no adaptive covering improvement from the companion scalar paper. The saturation argument is applied to the exact scalar threshold, not presumed for arbitrary representation supports away from it. In the larger-slack regime, local optimal supports are chosen anew and smooth trimming takes place only below $N^{1/2}$. The exponent 205 in Lemma 9.1 ensures that the stronger correction scale d_{201} is handled without assuming the older $O(\ell^{-5})$ greedy error is sufficient.

All error terms describe asymptotic regimes. The large powers in the constructions serve to separate scales. The accompanying Lean 4 formalization provides proofs of the entropy theorems and the threshold equality through the module `Egyptian.PaperII`. The module `Egyptian.Combined` combines the two papers to prove the scalar expansion at the common threshold.

References

[1] P. Chen, *Local Capacity and Thresholds for Simultaneous Egyptian Fraction Packings*, Hexagon preprint, 26 November 2026. hexagon:2611.01130v1.

[2] P. Chen, *Sharp Two-Term Capacity Asymptotics for Egyptian Fractions via Rational Grid Representations*, companion preprint, 2026.

[3] D. Conlon, J. Fox, X. He, D. Mubayi, H. T. Pham, A. Suk, and J. Verstraëte, A question of Erdős and Graham on Egyptian fractions, *Discrete Analysis* 2025:28, 13 pp. doi:10.19086/da.154329.

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

[5] H. L. Montgomery and R. C. Vaughan, *Multiplicative Number Theory I: Classical Theory*, Cambridge Studies in Advanced Mathematics 97, Cambridge University Press, 2007, Chapter 6.