hexagon:2611.01225v1

# SHARP TWO-TERM CAPACITY ASYMPTOTICS FOR EGYPTIAN FRACTIONS VIA RATIONAL GRID REPRESENTATIONS

PENGCHENG CHEN

Abstract. Let $T_*(m)$ be the least $N$ for which the integer $m$ is a sum of distinct reciprocals of positive integers at most $N$. We prove the two-term expansion

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

The leading correction has the constant conjectured by Croot. The construction uses a fixed rational-grid reservoir and an adaptive scaling argument. For a prime $p$, if $f(p)$ is the least length of an initial interval whose reciprocal subset sums cover $\mathbb{F}_p$, we also obtain $f(p) = \log p(\log\log p)^{1+o(1)}$. The two capacity terms result from an upper bound for the local deletion cost and an exceptional-prime count based on the number of distinct rational subset sums. All auxiliary covering and exact-completion steps are included.

## 1. Introduction and main theorems

For a finite set $A \subset \mathbb{N}$, put $w(A) = \sum_{n\in A} 1/n$. Write $[X] = \{1, \dots, \lfloor X \rfloor\}$, $H_N = w([N])$, and

$$T_*(m) = \min\{N \in \mathbb{N} : \exists A \subseteq [N],\ w(A) = m\}.$$

A minimum of an empty set is interpreted as $+\infty$. All logarithms are natural. For large $N$ define

$$\ell = \log N, \quad L = \log \ell, \quad \Delta_N = L^2/\ell, \quad a_N = \frac{L^2}{2\ell} + \frac{L\log L}{\ell}, \quad b_N = \frac{L\log\log L}{\ell}. \tag{1}$$

All asymptotic statements involving iterated logarithms concern sufficiently large arguments.

Croot [3] proved capacity-loss bounds with constants $1/2$ and $9/2$ on the scale $\Delta_N$, and conjectured that the constant $1/2$ is sharp. Yokota [6] improved the lower estimate for the number of representable integers, obtaining the constant $\pi^2/3$. The two terms in the capacity loss and the corresponding expansion of $\log T_*(m)$ already appear in the author's work on simultaneous packings [2]. Here we give an independent scalar proof using rational grids and their projections modulo primes. We also obtain quantitative bounds for the modular covering length $f(p)$ and the rational-grid height $G(y)$, defined below.

**Theorem 1.1** (Two-term capacity bound)**.** *There is an absolute constant $C > 0$ such that for all sufficiently large $N$ and all positive integers $m$:*

1. *If $H_N - m \ge a_N + Cb_N$, then $m = w(A)$ for some $A \subseteq [N]$.*
2. *If $H_N - m < a_N - Cb_N$, then there is no such set.*

**Corollary 1.2** (Inverse threshold)**.** *As $m \to \infty$,*

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

*In particular, the first two corrections are unchanged upon writing $T_*(m)e^{\gamma-m} - 1$ in place of $\log T_*(m) - (m - \gamma)$.*

*Date*: Version 0.4.1 — November 27, 2026.

2020 *Mathematics Subject Classification.* 11D68, 11P70, 11B75.

*Key words and phrases.* Egyptian fractions, reciprocal subset sums, prime powers, capacity thresholds.

For a prime $p$, let $f(p)$ be the least integer $z < p$ such that all residues in $\mathbb{F}_p$ are reciprocal subset sums of $[z]$. Existence follows, for example, by using singletons in $[p-1]$ and the empty set. Eppstein [4] suggested that $f(p)$ might have order $\log p$. The distinction between the number of subsets and the number of distinct rational sums is essential here; the latter was studied by Bleicher and Erdős [1].

**Theorem 1.3** (Short modular coverage)**.** *For sufficiently large primes $p$, every residue in $\mathbb{F}_p$ has a representation by distinct reciprocals of denominators at most*

$$C \log p \, \log\log p \, (\log\log\log p)^{12}, \tag{3}$$

*with real reciprocal sum in $[1/4, 1/2]$. Furthermore,*

$$f(p) \gg \frac{\log p \, \log\log p}{\log\log\log p}. \tag{4}$$

*Consequently $f(p) = \log p (\log\log p)^{1+o(1)}$.*

Let $K_y = \operatorname{lcm}(1, \dots, \lfloor y \rfloor)$. Define $G(y)$ to be the smallest $B$ such that every point in $K_y^{-1}\mathbb{Z} \cap [1/4, 1/2]$ is a sum of distinct reciprocals with denominators at most $B$. Our grid estimates give a corresponding result.

**Theorem 1.4** (Rational-grid size)**.** *As $y \to \infty$,*

$$\left(\frac{1}{\log 2} - o(1)\right) y \log y \le G(y) \ll y \log y (\log\log y)^{12}.$$

*In particular $G(y) = y(\log y)^{1+o(1)}$.*

We distinguish two constructions: a fixed bounded-weight reservoir, used in the exact-completion step, and an adaptively scaled representation, used to improve the maximum denominator. The proof of Theorem 1.1 is entirely scalar. The companion manuscript [7] proves equality of the scalar and disjoint-partition thresholds, and a sharper entropy estimate; none of its results is needed here.

## 2. An auxiliary modular subset-sum lemma

For a positive integer $n$, write $P^*(n)$ for its largest prime-power divisor, with $P^*(1) = 1$. Thus $P^*(n) \le y$ if and only if $n \mid K_y$, where

$$K_y = \operatorname{lcm}(1, \dots, \lfloor y \rfloor).$$

We use the classical quantitative prime number theorem (PNT), $\pi(x) = \operatorname{Li}(x) + O(xe^{-c\sqrt{\log x}})$ for an absolute $c > 0$ (see [5, Chapter 6]). It gives $\log K_y = \psi(y) = (1+o(1))y$ and $\pi(2y) - \pi(y) \asymp y/\log y$. For a prime $s$, reduction of a rational modulo $s$ is understood only when its reduced denominator is coprime to $s$.

**Lemma 2.1** (Auxiliary primes)**.** *Let $x$ be sufficiently large, and put*

$$M = \lceil (\log x)^{12} \rceil, \qquad k = \lceil (\log x)^{10} \rceil.$$

*For every prime $s \le x$, every set $B$ of $k$ distinct primes in $[M, 2M]$ with $s \notin B$, and every $h \in \mathbb{F}_s$, there is a subset $E \subseteq B$ such that $\sum_{b \in E} b^{-1} = h$ in $\mathbb{F}_s$.*

*Proof.* Set $\eta = (\log x)^{-4}$ and fix $h \in \{1, \dots, s-1\}$. If $\|hb^{-1}/s\| \le \eta$, where the norm is distance to the nearest integer and any integer representative of $hb^{-1}$ is used, there is an integer $v$ with $|v| \le \eta s$ and

$$bv = h + js, \qquad |j| \le 2M\eta + 1.$$

Every integer $h + js$ is nonzero and has absolute value at most $(2M\eta + 2)x$. It has at most $O(\log x/\log M)$ distinct prime divisors in $[M, 2M]$; here $M$ is a fixed power of $\log x$ up to rounding. Hence the number of exceptional $b$ is at most

$$O\left((1 + M\eta)\frac{\log x}{\log M}\right) = O\left(\frac{(\log x)^9}{\log\log x}\right) = o(k),$$

uniformly in $h, s, B$.

Let $Z = \sum_{b\in B} \xi_b b^{-1}$ in $\mathbb{F}_s$, with independent Bernoulli(1/2) variables $\xi_b$. The elementary bound $|\cos(\pi u)| \le \exp(-c\,\|u\|^2)$ for an absolute $c > 0$ gives

$$|\mathbb{E}e^{2\pi i h Z/s}| = \prod_{b\in B} |\cos(\pi h b^{-1}/s)| \le e^{-ck\eta^2/2}.$$

For large $x$ this holds at every nonzero frequency, since at least $k/2$ factors are nonexceptional. Fourier inversion now yields, for each $a \in \mathbb{F}_s$,

$$\mathbb{P}(Z = a) \ge \frac{1 - (s-1)e^{-ck\eta^2/2}}{s} > 0,$$

because $s \le x$ and $k\eta^2 \ge (\log x)^2$. A positive probability gives a subset with the required sum. □

This elementary argument is a deliberately nonoptimal version of the auxiliary-prime method in [3, Proposition 2]. Its important feature here is uniformity in the smaller prime modulus $s$.

## 3. Rational grid representations

**Lemma 3.1** (Mixed-radix expansion)**.** *For every integer $J \ge 2$ and every $0 \le r < 1$ with $K_J r \in \mathbb{Z}$, there are digits $u_q$ indexed by prime powers $q = s^a \le J$, with $0 \le u_q \le s - 1$, such that*

$$r = \sum_{q=s^a\le J} \frac{u_q}{K_q}. \tag{5}$$

*For each nonzero digit, $K_q/u_q$ is an integer with $P^*(K_q/u_q) = q$. In particular the nonzero terms in (5) are distinct unit fractions with denominators dividing $K_J$.*

*Proof.* At each prime power $q = s^a$, $K_q = sK_{q-1}$, and between prime powers the least common multiple is constant. Starting with the largest prime power, choose the digit to be the residue of the current numerator modulo $s$ when the current rational is written with denominator $K_q$. Subtracting $u_q/K_q$ leaves a rational with denominator dividing $K_{q-1}$. Repetition leaves an integer after the last step. The digit sum is between 0 and

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

Since the input was in $[0, 1)$, the remaining integer is zero. If $u_q \ne 0$, then $u_q < s \le q$ gives $u_q \mid K_q$ and $s \nmid u_q$. Thus the prime power $q$ remains a divisor of $K_q/u_q$ and no larger prime power can divide it. This also proves distinctness across different $q$. □

**Theorem 3.2** (A fixed rational-grid reservoir)**.** *For all sufficiently large real $y$, there exists a set $G_y$ of positive integers such that*

$$n \mid K_y \ (n \in G_y), \qquad \max G_y \ll y(\log y)^{12}, \qquad w(G_y) \ll 1. \tag{6}$$

*Every rational number $r \in [1/4, 1/2]$ with $K_y r \in \mathbb{Z}$ is $w(E)$ for some $E \subseteq G_y$. The implied constants are absolute, and $G_y$ is independent of $r$.*

*Proof.* Let $J = \lfloor (\log y)/2 \rfloor$. For each prime power $q = s^a$ in $(J, y]$, put

$$M_q = \lceil (\log q)^{12} \rceil, \qquad k_q = \lceil (\log q)^{10} \rceil.$$

Choose $k_q$ distinct primes $B_q \subseteq [M_q, 2M_q] \setminus \{s\}$, using the PNT. For all sufficiently large $y$, $2M_q < q$ uniformly for $q > J$: indeed $(\log q)^{12}/q \to 0$ as $q \to \infty$ and $J \to \infty$. Define

$$P_q = \{qb : b \in B_q\}, \qquad P = \bigcup_{J < q \le y} P_q,$$

where all unions indexed by $q$ are over prime powers. Each element of $P_q$ has largest prime-power divisor $q$, since $b < q$ is prime and $b \neq s$. It divides $K_y$, and it is at most $2qM_q \ll y(\log y)^{12}$. The pools $P_q$ are pairwise disjoint. By Lemma 2.1 with $x = q$,

$$\{q\, w(E) \bmod s : E \subseteq P_q\} = \mathbb{F}_s. \tag{7}$$

Moreover,

$$w(P) \le \sum_{J < q \le y} \frac{k_q}{qM_q} \ll \sum_{n > J} \frac{1}{n(\log n)^2} \ll \frac{1}{\log J} = o(1). \tag{8}$$

Given $r$ as in the theorem, process the prime powers in $(J, y]$ in decreasing order. At the step $q = s^a$, let $\rho$ be the current residual. Its denominator divides $K_q$ by induction (for the first step, it divides $K_y = K_q$). In particular, the reduced denominator of $q\rho$ is coprime to $s$. Choose $E_q \subseteq P_q$ so that $q\, w(E_q) \equiv q\rho \pmod s$, by (7). Then the reduced denominator of $\rho - w(E_q)$ divides $K_{q-1}$: it divides $K_q$, and its $s$-adic exponent is at most $a - 1$. No larger prime-power factor is introduced. This proves the induction, since $K_j$ is constant between consecutive prime powers. After all steps the residual $\rho'$ satisfies

$$K_J \rho' \in \mathbb{Z}, \qquad 1/4 - o(1) \le \rho' \le 1/2,$$

by (8). In particular, $0 < \rho' < 1$ for large $y$.

Apply Lemma 3.1 to $\rho'$. All its possible denominators belong to the fixed set

$$R_J = \{K_q/u : q = s^a \le J,\ 1 \le u < s\}.$$

This set is disjoint from $P$ by the largest prime-power divisor, and $\max R_J \le K_J = y^{1/2 + o(1)}$. Also

$$w(R_J) = \sum_{q = s^a \le J} \frac{s(s-1)}{2K_q} \ll 1,$$

since $K_q \ge e^{q/2}$ for all sufficiently large $q$ by the PNT. Defining $G_y = P \cup R_J$ proves (6). The selected $E_q$, together with the mixed-radix denominators, form a subset of $G_y$ of weight exactly $r$. No denominator is repeated. □

## 4. Adaptive scaling and modular covering

**Lemma 4.1** (Projection of a grid)**.** *If $y < p$ and $K_y > 4p$, then for each $h \in \mathbb{F}_p$ there is a rational $r \in [1/4, 1/2]$ with $K_y r \in \mathbb{Z}$ and $r \equiv h \pmod p$. Thus any representation theorem for this grid using denominators dividing $K_y$ gives a corresponding modular covering theorem.*

*Proof.* The integer interval $[K_y/4, K_y/2]$ has length greater than $p$. Choose an integer $v$ in it congruent to $hK_y$ modulo $p$. Then $r = v/K_y$ has the claimed properties because $p \nmid K_y$. □

Applying Lemma 4.1 with $y = \lceil 2 \log p \rceil$ to Theorem 3.2 gives, uniformly for large primes $p$, a covering of every residue with real weight in $[1/4, 1/2]$ using denominators at most $C \log p (\log\log p)^{12}$, all coprime to $p$. Indeed $\log K_y = (1 + o(1))y$ gives $K_y = p^{2 + o(1)}$.

**Lemma 4.2** (Adaptive grid bound)**.** *Let $F : [2,\infty) \to [1,\infty)$ be nondecreasing with $F(x) = (\log x)^{O(1)}$. Suppose that for every prime $s$, every residue in $\mathbb{F}_s$ is $w(B) \pmod s$ for some set of distinct positive integers $B \subseteq [F(s)]$ coprime to $s$, with $w(B) \le 1$. Then, for all sufficiently large $y$, every $r \in K_y^{-1}\mathbb{Z} \cap [1/4, 1/2]$ is a sum of distinct reciprocals of divisors of $K_y$ at most $O(yF(y))$.*

*Proof.* Set $J = \lfloor (\log y)/2 \rfloor$. For every prime power $q = s^a \in (J, y]$, choose an integer $t_q$ coprime to $s$ with

$$\min\{\sqrt{q}, y/q\} \le t_q \le \min\{\sqrt{q}, y/q\} + 2.$$

Such an integer exists since at most one of any two consecutive integers is divisible by $s$. At the stage $q$ of descending elimination, let $h$ be the residue of $q$ times the current residual modulo $s$. The assumption supplies $B_q \subseteq [F(s)]$ with $w(B_q) \le 1$ and $w(B_q) \equiv t_q h \pmod s$, each $b \in B_q$ coprime to $s$. Subtract the sum of the reciprocals with denominators $qt_q b$ for $b \in B_q$. Multiplying that sum by $q$ gives $w(B_q)/t_q \equiv h \pmod s$.

Uniformly for $q > J$, one has

$$t_q F(s) \le (\sqrt{q} + 2)F(q) < q$$

for large $y$. Thus every chosen denominator has largest prime-power divisor precisely $q$, divides $K_y$, and is at most $3yF(y)$ because $qt_q \le y + 2q \le 3y$. The same descending induction as in Theorem 3.2 is therefore valid; disjointness follows from the largest prime-power divisor. The total actually subtracted weight, independently of the required residues, is bounded by

$$\begin{aligned}\sum_{J<q\le y} \frac{1}{qt_q} &\le \sum_{n>J} n^{-3/2} + \frac{1}{y}\#\{q \le y : q \text{ a prime power}\} \\ &= O(J^{-1/2}) + O(1/\log y) = o(1).\end{aligned}$$

The prime-power count follows from the PNT and $O(\sqrt{y}\log y)$ higher powers. The residual is positive, below 1, and has denominator dividing $K_J$. Lemma 3.1 completes it with denominators at most $K_J = y^{1/2+o(1)}$. All final denominators are distinct. This argument bounds the selected weights, not the weight of a fixed reservoir containing every possible adaptive selection. □

For the finitely many small primes not covered by the first projection, choose $s-1$ distinct integers congruent to $1 \pmod s$ and large enough that their total reciprocal weight is at most 1. Subsets of these supply all residues. Enlarging a constant therefore yields a nondecreasing $F(x) \ll \log(2x)(\log\log(4x))^{12}$ satisfying Lemma 4.2. It follows that

$$G(y) \ll y \log y (\log\log y)^{12}. \tag{9}$$

A second projection with $y = \lceil 2\log p \rceil$ proves (3), including the real weight constraint.

## 5. The distinct-sum obstruction

Define $\mathcal{S}(z) = \{w(E) : E \subseteq [z]\}$, a set of rational values, not a multiset of representations.

**Lemma 5.1** (Distinct reciprocal subset sums)**.** *For large real $z$,*

$$\log|\mathcal{S}(z)| \ll \frac{z\log\log z}{\log z}. \tag{10}$$

*Proof.* This elementary bound is of the type obtained by Bleicher and Erdős [1]. Set $Y_z = z/\log z$ and split $[z]$ into $V_z = \{n : P^*(n) \le Y_z\}$ and $D_z = [z] \setminus V_z$. The values of $w(E)$ with $E \subseteq V_z$ lie

in $K_{Y_z}^{-1}\mathbb{Z} \cap [0, H_{\lfloor z \rfloor}]$, so their number is at most $1 + K_{Y_z} H_{\lfloor z \rfloor}$. The PNT gives the logarithm of this bound as $O(z/\log z)$. For the other set, use

$$|D_z| \le \sum_{Y_z < q \le z} \left\lfloor \frac{z}{q} \right\rfloor \ll \frac{z \log\log z}{\log z}.$$

Indeed partial summation over primes gives $\sum_{Y_z<p\le z} p^{-1} = O(\log\log z/\log z)$, and the higher prime powers contribute $O(\sqrt{z}(\log z)^2)$, which is negligible. Splitting each subset across the two sets proves $|\mathcal{S}(z)| \le (1 + K_{Y_z} H_{\lfloor z \rfloor}) 2^{|D_z|}$ and hence (10). □

*Completion of the proof of Theorem 1.3.* For $z = f(p) < p$, reduction modulo $p$ cannot yield more distinct residues than there are distinct rational subset sums. Thus $p \le |\mathcal{S}(z)|$. The number of subsets gives $z \ge \log p/\log 2$, and the bound (3) supplies the upper estimate. In particular $\log z \sim \log\log p$ and $\log\log z \sim \log\log\log p$. Applying (10) gives (4). Together the bounds prove the asserted logarithmic scale. In particular the proposed order $\Theta(\log p)$ fails already by the distinct-sum obstruction. □

*Proof of Theorem 1.4.* The upper bound is (9). Take a prime $q \le y$ with $q \sim y$, which exists by the PNT, and write $B = G(y)$. If $B \ge q^2$ there is nothing to prove. Otherwise $v_q(K_y) = 1$, and every denominator at most $B$ has $q$-adic exponent at most one. The residues of $qr$ as $r = v/K_y$ ranges over the grid fill $\mathbb{F}_q$, since $K_y/4 > q$. In any representation of $r$, only denominators divisible by $q$ contribute to $qr$ modulo $q$. There are at most $\lfloor B/q \rfloor$ such denominators, so $q \le 2^{\lfloor B/q \rfloor}$, giving $B \ge q \log q/\log 2$. This proves the stated lower bound. One may also use Theorem 1.3 directly here: these multiples force $\lfloor B/q \rfloor \ge f(q)$ if $B < q^2$. □

## 6. Large prime blocks and the local loss

Fix the following parameters throughout the capacity argument:

$$M_N = \lceil \ell^{1000} \rceil, \qquad Y = N/(4M_N), \qquad V = \{n \in [N] : P^*(n) \le Y\}. \tag{11}$$

For all large $N$, $Y > \sqrt{N}$ and $4M_N < Y$. For primes $p > Y$ put

$$z_p = \lfloor N/p \rfloor, \quad \lambda_p(z_p) = \min_{\substack{E \subseteq [z_p] \\ w(E) \equiv H_{z_p} \pmod p}} w(E), \quad \Lambda(N) = \sum_{Y < p \le N} \frac{\lambda_p(z_p)}{p}. \tag{12}$$

The minima exist since the whole interval is an admissible deletion. Distinct blocks $p[z_p]$ are disjoint and $z_p < p$.

**Lemma 6.1** (Weighted prime sums)**.** *Uniformly for $2 \le Z$ with $\log Z = O(L)$,*

$$\sum_{N/Z < p \le N} \frac{1}{p} \ll \frac{1 + \log Z}{\ell}, \tag{13}$$

$$\sum_{N/Z < p \le N} \frac{H_{\lfloor N/p \rfloor}}{p} = \frac{(\log Z)^2}{2\ell} + O\Big(\frac{1 + \log Z}{\ell}\Big). \tag{14}$$

*The constants may depend on the fixed bound in $\log Z = O(L)$.*

*Proof.* Partial summation with the quantitative PNT in Section 2 gives (13). More explicitly the integral is $\int_{N/Z}^{N} (t \log t)^{-1}\, dt$ and the error from the prime-counting remainder is $O((1+\log Z) e^{-c'\sqrt{\ell}})$ for some $c' > 0$. For (14), replace $H_{\lfloor N/p \rfloor}$ by $\log(N/p) + O(1)$ at the cost in (13). The remaining weighted sum has main integral

$$\int_{N/Z}^{N} \frac{\log(N/t)}{t \log t}\, dt = \int_0^{\log Z} \frac{u}{\ell - u}\, du = \frac{(\log Z)^2}{2\ell} + O\left(\frac{(\log Z)^3}{\ell^2}\right).$$

The prime-counting remainder contributes $O((1+\log Z)^2 e^{-c'\sqrt{\ell}})$ by the boundary term and the integral of the absolute derivative. Both errors are absorbed in the displayed bound. □

**Proposition 6.2** (Two-term local loss)**.** *With the definitions in* (11)–(12)*,*

$$\Lambda(N) = \frac{L^2}{2\ell} + \frac{L\log L}{\ell} + O\Big(\frac{L\log\log L}{\ell}\Big) = a_N + O(b_N). \tag{15}$$

*For every fixed $\varepsilon > 0$, one also has*

$$\sum_{Y<p\le N/\ell^{1+\varepsilon}} \frac{\lambda_p(z_p)}{p} = O(L/\ell) = o(\Delta_N). \tag{16}$$

*If $w(A)$ is an integer for $A \subseteq [N]$, then $H_N - w(A) \ge \Lambda(N)$.*

*Proof.* By Theorem 1.3, the residues $H_{z_p}$ have representations of weight at most $1/2$ within $[z_p]$ whenever $z_p \ge Z_+ - 1$, where

$$Z_+ = C_+ \ell L (\log L)^{12}$$

with a sufficiently large constant $C_+$. The floor is harmless with this choice. Hence $\lambda_p(z_p) \le 1/2$ for $Y < p \le N/Z_+$. This proves (16) by (13), since $Z_+ = o(\ell^{1+\varepsilon})$. Elsewhere use $\lambda_p(z_p) \le H_{z_p}$. The resulting upper bound is

$$\Lambda(N) \le \frac{(\log Z_+)^2}{2\ell} + O(L/\ell).$$

Here and below $\log Z_+ = L + \log L + 12\log\log L + O(1)$. Its square gives the upper half of (15).

For the lower half put $Z_- = \ell L/(\log L)^2$ and $K = K_{Z_-}$. By Lemma 5.1,

$$\log|\mathcal{S}(Z_-)| \ll \frac{Z_-\log\log Z_-}{\log Z_-} \ll \frac{\ell}{\log L} = o(\ell).$$

For each nonzero $r \in \mathcal{S}(Z_-)$, the integer $Kr$ is positive, with $\log(Kr) \le \log(KH_{\lfloor Z_-\rfloor}) = O(Z_-)$. Thus it has at most $O(Z_-/\ell)$ distinct prime divisors above $N/Z_-$ (since those primes have logarithm $(1-o(1))\ell$). The number of primes above $N/Z_-$ dividing any such integer is at most $N^{o(1)}$. Outside this exceptional set, every nonempty subset of $[z_p] \subseteq [Z_-]$ has nonzero weight modulo $p$: otherwise $p \mid Kr$, since $p > Z_-$ for large $N$. Consequently $\lambda_p(z_p) = H_{z_p}$ for these primes. Their exceptional counterparts have total possible contribution at most $N^{o(1)} Z_- H_{\lfloor Z_-\rfloor}/N = N^{-1+o(1)}$. We obtain

$$\Lambda(N) \ge \frac{(\log Z_-)^2}{2\ell} - O(L/\ell) - N^{-1+o(1)}.$$

Since $\log Z_- = L + \log L - 2\log\log L$, this proves (15) by expanding both squares; terms of size $(\log L)^2/\ell$ are $O(b_N)$.

For the last claim multiply $w(A)$ by $p$ and reduce modulo $p$. As $p^2 > N$, this yields $w(\{j \le z_p : pj \in A\}) \equiv 0 \pmod p$. The complementary set has weight at least $\lambda_p(z_p)$. The prime blocks are disjoint, so summing their deletion costs gives the necessary inequality. □

## 7. From local loss to the sharp integral threshold

For each prime $p > Y$, choose $E_p \subseteq [z_p]$ attaining the minimum in (12), and let

$$F_p = \{pj : j \in [z_p] \setminus E_p\}, \qquad F = \bigcup_{Y<p\le N} F_p.$$

Then $w(F_p) = p^{-1}w([z_p] \setminus E_p)$ has reduced denominator dividing $K_Y$: the factor $p$ cancels by the defining congruence, and all other denominator prime powers are at most $z_p < Y$. Thus $K_Y w(F) \in \mathbb{Z}$. We also have

$$w(F) \le \sum_{Y<p\le N} \frac{H_{z_p}}{p} = O(\Delta_N) = o(1) \tag{17}$$

by Lemma 6.1.

Define the set of omitted large higher prime-power multiples by

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

The number of prime powers $s^a \le N$ with $a \ge 2$ is $O(\sqrt{N}\log N)$. Consequently

$$w(D_2) \le \sum_{\substack{Y<s^a\le N\\ a\ge 2}} \frac{H_{\lfloor N/s^a\rfloor}}{s^a} \ll \frac{\sqrt{N}\ell^2}{Y} = o(\Delta_N). \tag{18}$$

Here $w(D_2)$ denotes the actual set weight, not the possibly repeated sum on the right. The set $D_2$ is disjoint from all the prime blocks above $Y$: for different primes their large powers would have product greater than $N$, and a prime $p > Y$ cannot have $p^2 \le N$. These observations give the exact identity

$$w(V) + w(F) = H_N - \Lambda(N) - w(D_2). \tag{19}$$

**Lemma 7.1** (Scaled grid completion)**.** *There is a fixed set $Q \subseteq [N]$, disjoint from $F$, with $w(Q) = O(\ell^{-3})$, that represents every rational $r$ satisfying $K_Y r \in \mathbb{Z}$ and $1/(4t) \le r \le 1/(2t)$, where $t = \lceil \ell^3 \rceil$.*

*Proof.* Apply Theorem 3.2 at $y = Y$ and take $Q = tG_Y = \{tn : n \in G_Y\}$. It is a set of distinct integers of weight $w(G_Y)/t = O(\ell^{-3})$. Its largest element is $O(tY(\log Y)^{12}) < N$ for large $N$ by (11). Every prime factor of any $tn$ is at most $Y$, since $n \mid K_Y$ and $t < Y$. Thus $Q \cap F = \varnothing$. Some elements of $Q$ may have a prime-power divisor above $Y$; we do not require $Q \subseteq V$. For a target $r$ in the stated interval, $tr \in [1/4, 1/2]$ and $K_Y tr \in \mathbb{Z}$. The grid theorem represents $tr$ inside $G_Y$. Scaling the selected denominators by $t$ gives $r$ inside $Q$. □

*Proof of Theorem 1.1.* The necessary condition follows from Proposition 6.2 by taking $C$ larger than the absolute error constant in (15). For sufficiency take $C$ larger still and suppose $H_N - m \ge a_N + Cb_N$. Since $w(D_2) = O(N^{-1/2}\ell^{1002})$ and $w(Q) = O(\ell^{-3})$, both $o(b_N)$, (19) and Lemma 7.1 give

$$w(V \setminus Q) + w(F) \ge m$$

for all sufficiently large $N$, uniformly in $m$. Set $d = 3/(8t)$ and $b = m - d - w(F) > 0$, using (17) and $m \ge 1$. We seek $B \subseteq V \setminus Q$ with $w(B) = b + O(N^{-1/4})$.

For clarity, here is a uniform construction even for small $m$. Let $L_0 = (V \setminus Q) \cap [N^{1/4}]$ and $L_1 = (V \setminus Q) \setminus L_0$. We have $w(L_0) \le \ell/4 + O(1)$ and $w(L_1) = 3\ell/4 + O(1)$, since the total loss $w([N] \setminus V) + w(Q)$ is $O(\Delta_N)$ by Lemma 6.1 and (18). If $b \ge w(L_0)$ start with $L_0$; otherwise start with the empty set. In either case the starting weight is at most $b$, and adding all of $L_1$ would reach at least $b$. Add elements of $L_1$ in any fixed order, stopping before the first addition that would exceed $b$ (or when equality holds). Each such reciprocal is at most $N^{-1/4}$. Thus we obtain a subset $B$ with $0 \le b - w(B) \le N^{-1/4}$.

The residual

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

has denominator dividing $K_Y$ and lies in $[1/(4t), 1/(2t)]$ for large $N$. Lemma 7.1 gives $E \subseteq Q$ with $w(E) = r$. The three sets $F, B, E$ are pairwise disjoint and lie in $[N]$, so their union has

weight $m$. Notice that only $Q \cap V$ was withheld from $V$; it is immaterial whether the final selected $E$ contains an element of $Q \setminus V$. This proves the sufficient condition. □

*Proof of Corollary 1.2.* Write

$$A_m = \frac{(\log m)^2}{2m} + \frac{\log m \log\log m}{m}, \qquad B_m = \frac{\log m \log\log\log m}{m}.$$

For a fixed sufficiently large $C_1$, take

$$N_\pm = \lfloor \exp(m - \gamma + A_m \pm C_1 B_m) \rfloor .$$

The harmonic estimate $H_N = \log N + \gamma + O(N^{-1})$ and the fact that $\log N_\pm = m + O(1)$ give $a_{N_\pm} = A_m + o(B_m)$, $b_{N_\pm} \sim B_m$ and $H_{N_\pm} - m = A_m \pm C_1 B_m + O(N_\pm^{-1})$. Thus Theorem 1.1, with $C_1$ larger than its constant, shows that $N_- < T_*(m) \le N_+$ by monotonicity. Taking logarithms proves (2). The squared exponential correction $O((\log m)^4/m^2)$ is $o(B_m)$, which justifies the corresponding relative expansion of $T_*(m)$. □

## 8. Scope and dependencies

The principal loss from large prime blocks is evaluated in Proposition 6.2. Its second term uses an exceptional-prime bound rather than asserting that every block of length $Z_-$ is forbidden. The first grid construction in Theorem 3.2 is fixed and has bounded total weight; the adaptive construction in Lemma 4.2 need not have this stronger reservoir property. The exact scalar completion uses only the fixed construction. This distinction avoids inferring a bound on the union of adaptive selections from a bound on a single selection.

The sharp constant $1/2$ follows in particular from Theorem 1.1; the additional term has coefficient 1. The logarithmic scale of modular covering is matched but the remaining iterated-log factors in the upper and lower bounds are not resolved. All bounds are asymptotic and the constructions do not furnish practical numerical thresholds for moderate $N$.

*Lean formalization.* The accompanying Lean 4 project formalizes the results of this paper in the module `Egyptian.PaperI`. For Lemma 6.1, it proves both estimates uniformly over the stated range of $Z$, including $Z = 2$.

## References


[1] M. N. Bleicher and P. Erdős, Denominators of Egyptian fractions II, *Illinois J. Math.* **20** (1976), 598–613. doi:10.1215/ijm/1256049650.

[2] P. Chen, *Local Capacity and Thresholds for Simultaneous Egyptian Fraction Packings*, Hexagon preprint 2611.01130v1 (2026).

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

[4] D. Eppstein, Modular Egyptian fractions, online note, 12 December 2006. Permalink.

[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.

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

[7] P. Chen, *Sharp Second-Order Entropy Asymptotics for Disjoint Egyptian Fraction Representations*, companion preprint, 2026.