hexagon:2611.01295v1

Robust Multivariate Local Limit Theorems for Reciprocal Subset Sums

Pengcheng Chen Meihong Ai

28 November 2026

Abstract

We prove a protected-reservoir local limit theorem for mutually exclusive random labels on sets of integer denominators. The number of labels may grow as a fixed power of the logarithm of the height, arbitrary probabilities are allowed on the main set starting at N^{κ} for any fixed $\kappa > 1/2$, and adversarial deletions from the upper-half reservoir are charged through a uniform linear-cost pruning estimate. A single Fourier argument treats reciprocal sums and any selected subset of the cardinality coordinates, with a Gaussian prefactor on the actual rational lattice. Synchronized divisor fibers control the reciprocal frequencies, while surviving pairs of denominators in the ratio $2:3$ control the length frequencies. Applications include rational interval filling, an exact fractional allocation criterion, a robust blocking threshold for several disjoint representations of prescribed lengths, and quantitative entropy bounds for disjoint representations of fixed vectors of targets, with or without prescribed lengths. We characterize the interior fractional feasibility region for the latter constraints and obtain a logarithmic deficit of order $N \log\log N/\log N$ for $O(N \log\log N/\log N)$ adversarial deletions. A second-scale modular cancellation argument also yields robust simultaneous representations of a fixed rational target near the harmonic capacity, and explicit exceptional-height asymptotics in the unit-target case. The results distinguish quantitative aperiodicity from abstract generation of the underlying lattice.

Keywords: Egyptian fractions; local limit theorem; categorical sampling; exact cardinality; adversarial deletion; entropy.
2020 MSC: 11D68, 11B75, 60F05, 05C65.

Contents

1 Introduction and main results **2**

2 Elementary preparation and pruning **4**

3 Dense divisor fibers and surviving reservoirs **5**

4 Synchronizing near-divisibility windows **6**

5 Joint Fourier inversion **7**

6 Rational interval filling **10**

7 A fractional Hall theorem and disjoint prescribed representations **10**

8 Quantitative entropy for disjoint representations **12**

9 Modular coverage at a second scale **16**

10 Large-prime blocks and the exact loss budget **17**

11 Packing whole blocks and simultaneous completion **19**

12 The matching obstruction **20**

13 Consequences for unit-target thresholds and exceptional heights **20**

14 Lattice structure and scope **22**

1 Introduction and main results

For a finite set S of positive integers put $R(S) = \sum_{n\in S} 1/n$ and $H_N = R([1, N])$. All intervals of denominators are intersected with the integers, and logarithms are natural. Write

$$L = \log N, \quad \lambda = \log L, \quad \delta_N = \lambda/L, \quad g(N) = \lambda^2/L, \quad e_N = \lambda \log \lambda/L.$$

The arithmetic ingredients of the method have their roots in work of Croot [3, 7], Bloom [1], and Liu and Sawhney [4]. Martin's prescribed-length theorem [5] motivates the integer coordinates. The unrestricted one-target exponential counting rate was established by Conlon et al. [2], and related fixed-vector entropy and critical-support counting appear in [12]. Our objective is different: a uniform local limit on a protected rational grid, with quantitative stability under deletion, growing dimension, and simultaneous exact length constraints. We use mutually exclusive categorical sampling; we do not replace it by independent Bernoulli experiments for different colours.

Theorem 1.1 (Protected-reservoir multivariate local limit)**.** *Fix $a, d > 0$, $1/2 < \kappa < 1$, a positive integer b_0, and $D_0 > 0$. There are constants $\eta > 0$, $B_0 > 0$ and an absolute constant $C_2 > 0$ such that, for any fixed $B \geq B_0$, there are $C_1 > 0$ and a threshold N_0 for which the following holds for every integer $N \geq N_0$. The fixed parameters may enter η, B_0, C_1, N_0, except for the stated absolute constant and the stated independence of B_0 from b_0. The exponent threshold B_0 can be chosen independently of b_0. Put $Y = N/L^B$ and $m = N^\kappa$. If $F \subseteq [1, N]$ and $h_H = |F \cap (N/2, N]| \leq \eta N$, there exists $G \subseteq [1, N]$ with*

$$|[1, X] \setminus G| \leq C_1 X \delta_N + C_2 X h_H/N \quad (1 \leq X \leq N). \tag{1.1}$$

Let $H = G \cap (N/2, N] \setminus F$ and $Q = \operatorname{lcm}(H)$. Then $|H| \gg N$, $b_0 \mid Q$, $n \mid Q$ for all $n \in G$, and $\log Q \ll Y$.

For $1 \leq t \leq L^d$ and $A \subseteq G \cap [m, N/2] \setminus F$, choose independent labels $C_n \in \{0, 1, \dots, t\}$ for $n \in A \cup H$, with probabilities $p_{n,i}$. On A these probabilities are arbitrary, and on H require $p_{n,i} \geq L^{-a}$ for $0 \leq i \leq t$ (in particular $(t + 1)L^{-a} \leq 1$; only feasible probability vectors are considered). Define

$$Z_i = \sum_{n\in A\cup H} \frac{\mathbf{1}_{C_n=i}}{n}, \qquad T_i = \sum_{n\in A\cup H} \mathbf{1}_{C_n=i}.$$

For any $J \subseteq [t]$, including the empty set, put $s = |J|$, $\mathbf{Y} = (\mathbf{Z}, (T_i)_{i\in J})$, $\mu = \mathbb{E}\mathbf{Y}$ and $\Sigma = \operatorname{Cov} \mathbf{Y}$. Then Σ is positive definite. Uniformly for lattice points $\mathbf{y} \in Q^{-1}\mathbb{Z}^t \times \mathbb{Z}^s$ satisfying

$(\mathbf{y}-\mu)^{\top}\Sigma^{-1}(\mathbf{y}-\mu) \le D_0^2$,

$$\mathbb{P}(\mathbf{Y}=\mathbf{y}) = \frac{1+o(1)}{(2\pi)^{(t+s)/2}Q^t\sqrt{\det\Sigma}} \exp\left[-\tfrac{1}{2}(\mathbf{y}-\mu)^{\top}\Sigma^{-1}(\mathbf{y}-\mu)\right]. \tag{1.2}$$

The set G is independent of A, J, the probabilities and the target. More explicitly, for fixed initial parameters and fixed B, the relative error in (1.2) *is bounded in absolute value by a single function $\epsilon_N \to 0$ independent of F, A, t, J, the feasible probability arrays and the admissible target. In particular, if the mean belongs to the indicated lattice, the prescribed sums and selected lengths can be attained simultaneously by disjoint subsets.*

For $t = 1$, $J = \varnothing$, the theorem gives the weighted scalar local limit with probabilities allowed to be zero or one on the main pool. In particular it applies to any main pool starting at N/L^c for fixed $c > 0$, since $N/L^c > N^{\kappa}$ eventually. For $J = \varnothing$ and $t \le L^2$, $a = 4$, it includes growing-colour reciprocal completion. For $J = [t]$ it gives the full joint local limit. The use of real frequencies over a full period of length Q eliminates any need to remove wrap-around by a separate concentration argument. The threshold $\kappa > 1/2$ follows from a variance-weighted Taylor estimate, rather than a sum of worst-case cubic errors. We do not claim this threshold is necessary. For background on multivariate local limits for independent vectors see Dolgopyat [13]; our proof addresses the growing rational period, growing dimension and the particular protected-reservoir aperiodicity conditions directly.

Corollary 1.2 (Auxiliary events)**.** *Under Theorem 1.1 with $t = 1$, $J = \varnothing$, write $Z = Z_1$, $V = \operatorname{Var} Z$ and $\mu = \mathbb{E}Z$. If $Q\mu \in \mathbb{Z}$ and $\mathbb{P}(\mathcal{E}^c) = o(Q^{-1})$, then $\mathbb{P}(Z = \mu, \mathcal{E}) > 0$ for large N.*

Proof. Since $V \le N/(4m^2) = o(1)$, the central Gaussian probability is at least $(1 - o(1))/Q$. Subtract the failure probability. □

Theorem 1.3 (Matching-blocking threshold)**.** *Fix t positive rational targets r_i and integers $K_i \asymp K$ with fixed comparison constants. Let $\mathcal{B}_{\mathbf{r},\mathbf{K}}(M)$ be the minimum number of deletions from $[2, M]$ forbidding a family of pairwise disjoint S_i satisfying $|S_i| = K_i$ and $R(S_i) = r_i$. Put*

$$H_* = \min_{\varnothing \ne J \subseteq [t]} \frac{K_J}{e^{r_J} - 1}, \qquad K_J = \sum_{i \in J} K_i, \quad r_J = \sum_{i \in J} r_i.$$

For sufficiently large fixed D (depending on these fixed data), uniformly as $K \to \infty$,

$$\mathcal{B}_{\mathbf{r},\mathbf{K}}(\lfloor DK \rfloor) = H_* - \Theta\left(K \frac{\log\log K}{\log K}\right). \tag{1.3}$$

For identical targets $r_i = r$ and lengths $K_i = K$, the minimum occurs at $J = [t]$, since $s/(e^{sr} - 1)$ decreases for $s > 0$. Hence the threshold is $tK/(e^{tr} - 1)$ up to the stated deficit. The matching theorem is only asserted for fixed t; the growing-dimension local limit theorem does not automatically yield the corresponding uniform extremal result when t grows. For a fixed positive rational r and $F \subseteq [1, N]$, put $h_H = |F \cap (N/2, N]|$ and let $k_r(N; F)$ be the largest number of disjoint representations of r using denominators in $[1, N] \setminus F$. Write $k(N) = k_1(N; \varnothing)$ and

$$W_r(N; F) = \sum_{\substack{n \le N,\ n \notin F \\ 1/n \le r}} \frac{1}{n}.$$

Theorem 1.4 (Robust simultaneous harmonic packing)**.** *For each fixed rational $r > 0$ there is C_r such that, uniformly for $|F| \le N/L$ and sufficiently large N,*

$$k_r(N;F) \ge \left\lfloor \frac{W_r(N;F) - \frac{1}{2}g(N) - C_r(e_N + h_H L/N)}{r} \right\rfloor. \tag{1.4}$$

The trivial upper bound is $\lfloor W_r(N;F)/r \rfloor$. If $F = \varnothing$, then the upper bound strengthens to

$$k_r(N;\varnothing) \le \left\lfloor \frac{W_r(N;\varnothing) - \frac{1}{2}g(N) + C_r\lambda/L}{r} \right\rfloor. \tag{1.5}$$

In particular, the leading constant $1/2$ in the sufficient harmonic slack survives when $h_H = o(N\lambda^2/L^2)$. This is a one-sided robust guarantee, not a matching adversarial lower bound on the deficit for every F.

Organization. We first prove the reservoir theorem, then treat interval bases and fixed-length blocking by fractional allocation. The entropy section develops the simultaneous counting theory. The remaining sections apply sum-only growing-colour completion to robust harmonic packings and exceptional heights.

2 Elementary preparation and pruning

We use the standard estimates

$$\sum_{p \le x} \frac{1}{p} = \log\log x + M_1 + O(1/\log x), \quad \sum_{p \le x} \log p \ll x, \quad \psi(x) \ll x, \tag{2.1}$$

We also use the prime number theorem in the form $\#\{p : x/2 < p \le x\} \asymp x/\log x$. Here p always denotes a prime and $\psi(x) = \sum_{p^j \le x} \log p$. Partial summation gives $\sum_{p \le y} \log p/(p-1) \ll \log y$. For $Y = N/L^B$ with B fixed,

$$\sum_{\substack{Y < q \le N \\ q \text{ prime power}}} q^{-1} \ll_B \delta_N. \tag{2.2}$$

Indeed the prime contribution follows from (2.1), and $\sum_{p^j > Y, j \ge 2} p^{-j} \ll Y^{-1/2}$: for $p \le \sqrt{Y}$ the geometric tail is $O(Y^{-1})$ per prime, and for $p > \sqrt{Y}$ sum $O(p^{-2})$.

Let $\Omega(n)$ count prime factors with multiplicity, with $\Omega(1) = 0$. For $1 \le X \le N$ we have

$$\sum_{n \le X} (3/2)^{\Omega(n)} \ll X(\log(2X))^{1/2}. \tag{2.3}$$

To see this, expand $(3/2)^{\Omega(n)} = \sum_{d|n} g(d)$ with g nonnegative multiplicative and $g(p^j) = \frac{1}{2}(3/2)^{j-1}$. The divisor sum is at most $X \prod_{p \le X}(1 + (1/2)/(p - 3/2)) \ll X(\log(2X))^{1/2}$ by (2.1). Consequently the proportion of $n \le X$ with $\Omega(n) > 10\lambda$ is $O(L^{1/2 - 10\log(3/2)}) = o(\delta_N)$, uniformly in X. Also the proportion having $v_p(n) > 10\lambda$ for some prime p is $O(L^{-10\log 2}) = o(\delta_N)$, by a union bound over $p^{\lfloor 10\lambda \rfloor + 1}$.

Lemma 2.1 (Linear cost of damaged moduli)**.** *Let $E \subseteq [1, N]$ have size h. For $\varepsilon_0 > 0$ and $Y \le N$, call a prime power $q \le Y$ damaged if $|E \cap q\mathbb{N}| \ge \varepsilon_0 N/q$. For each prime with a damaged power retain only the smallest such power, forming a set $\mathcal{B}$. If $h/N \le \varepsilon_0^2/2$, then*

$$\sum_{q \in \mathcal{B}} q^{-1} \le 2\varepsilon_0^{-2} h/N. \tag{2.4}$$

Proof. Put $T = \sum_{q\in\mathcal{B}} 1/q$ and $g(n) = \sum_{q\in\mathcal{B}} \mathbf{1}_{q|n}$. Distinct elements of $\mathcal{B}$ are coprime, so $\sum_{n\le N} g(n)^2 \le N(T+T^2)$. On the other hand, $\varepsilon_0 NT \le \sum_{n\in E} g(n)$. Cauchy–Schwarz yields $\varepsilon_0^2 T^2 \le (h/N)(T+T^2)$. If $T>0$, divide by T and rearrange; $T=0$ is immediate. □

Apply Lemma 2.1 to $E = F \cap (N/2, N]$. Call a positive integer $n \le N$ good if

$$\begin{aligned} p^j \mid n \Longrightarrow p^j \le Y; \qquad & v_p(n) \le 10\lambda \ \text{ for every } p; \\ \Omega(n) \le 20\lambda; \qquad & q \nmid n \ \text{ for every } q \in \mathcal{B}. \end{aligned} \tag{2.5}$$

Let G consist of all good integers. Equations (2.2), (2.3), (2.4) and $\lfloor X/q \rfloor \le X/q$ prove (1.1) with $C_2 = 2\varepsilon_0^{-2}$. The absolute constant ε_0 will be fixed in Section 3. In particular the coefficient of h_H/N is independent of B. We keep the original forbidden set separate from G.

3 Dense divisor fibers and surviving reservoirs

For $y \ge 2$ let $s_y(k) = \prod_{p\le y} p^{v_p(k)}$. A positive integer j is called y-rough if no prime at most y divides j; 1 is allowed.

Lemma 3.1 (A positive-density range of smooth parts)**.** *There are absolute constants $J \ge 2$ (an integer) and $c_0 > 0$ such that for sufficiently large y, all $X \ge y^3$ and all primes p, at least $c_0 X$ integers $k \in (X/2, X]$ satisfy*

$$p \nmid k, \qquad \sqrt{y} \le s_y(k) \le y^J.$$

Proof. Set $f(k) = \sum_{\sqrt{y}\le t\le y,\ t \text{ prime},\ t\ne p} \mathbf{1}_{t|k}$. Counting multiples of t and pt in $(X/2, X]$ shows that

$$\sum_{X/2<k\le X,\ p\nmid k} f(k) = \frac{X}{2}(1-1/p)\sum_t \frac{1}{t} + O(y) \gg X.$$

The prime reciprocal sum tends to $\log 2$, uniformly in p. The second moment satisfies $\sum_{k\le X} f(k)^2 \le X(\sum_t 1/t + (\sum_t 1/t)^2) \ll X$. Hence $\gg X$ of the integers under consideration have $f(k) > 0$ and thus $s_y(k) \ge \sqrt{y}$. Moreover,

$$\sum_{k\le X} \log s_y(k) \le X \sum_{t\le y,\ t \text{ prime}} \frac{\log t}{t-1} \ll X \log y.$$

Markov's inequality bounds the number with $s_y(k) > y^J$ by $O(X/J)$. Choose J large enough. □

Lemma 3.2 (Smooth numbers in proportional intervals)**.** *For fixed $U \ge 1$ there exists $\kappa_U > 0$ such that for all sufficiently large y and all $\sqrt{y} \le z \le y^U$, at least $\kappa_U z$ integers in $(z/2, z]$ have all prime factors at most y.*

Proof. It suffices to prove the assertion for U a multiple of $1/4$. The case $U = 1$ follows from $\lfloor z \rfloor - \lfloor z/2 \rfloor \ge z/2 - 1$ (all integers at most y are smooth). For the induction step, $z \le y$ is immediate as well; otherwise choose primes $t \in [y^{1/4}, y^{1/2}]$. Then $\sqrt{y} \le z/t \le y^{U-1/4}$. The induction hypothesis gives $\ge \kappa_{U-1/4} z/t$ smooth integers in $(z/(2t), z/t]$. Multiplying by t gives smooth integers in the required interval. Each is counted at most $4U$ times since it is at most y^U. Sum over t, using $\sum_t 1/t \to \log 2$, and decrease the constant if necessary. □

Lemma 3.3 (Fiber extraction)**.** *Let $T \subseteq (X/2, X]$, $|T| \geq \zeta X$, and suppose $\sqrt{y} \leq s_y(k) \leq y^J$ for every $k \in T$. There exists a y-rough integer j such that, writing $z = X/j$ and $U_j = \{u : ju \in T,\ s_y(u) = u\}$,*

$$\sqrt{y} \leq z \leq 2y^J, \qquad U_j \subseteq (z/2, z], \qquad |U_j| \geq \gamma z,$$

where $\gamma = \zeta\kappa_{J+1} > 0$.

Proof. Partition T according to the unique decomposition $k = js_y(k)$, with j rough. For every nonempty fiber, $z = X/j$ lies in the indicated range. Lemma 3.2 provides at least $\kappa_{J+1}z$ smooth integers in the full fiber interval. Multiplying them by the corresponding j gives pairwise distinct integers across fibers, all at most X. Thus $\sum_j z_j \leq X/\kappa_{J+1}$. Since $\sum_j |U_j| \geq \zeta X$, some fiber has the required density. □

Fix J, c_0 from Lemma 3.1, $\varepsilon_0 < c_0/8$, and choose $\eta > 0$ so small that

$$\eta < \min\{\varepsilon_0^2/2,\ c_0/(16C_2),\ \varepsilon_0/(2b_0),\ 1/16,\ 1/(16C_2)\}, \quad C_2 = 2\varepsilon_0^{-2}. \tag{3.1}$$

Set $y = L^D$ for a fixed $D > 0$ to be chosen, and require $B > 3D$. A prime power $q = p^j$ is called admissible if $q \leq Y$, $j \leq 10\lambda$, and no member of $\mathcal{B}$ divides q.

Lemma 3.4 (Surviving local reservoir)**.** *For each admissible $q = p^j$ there is a set $T_q \subseteq (N/(2q), N/q]$ of size at least $c_0N/(2q)$ such that every $k \in T_q$ satisfies $qk \in H$, $p \nmid k$, and $\sqrt{y} \leq s_y(k) \leq y^J$.*

Proof. Write $X = N/q \geq L^B \geq y^3$. Lemma 3.1 gives at least c_0X candidates. Since $p \nmid k$, the p-power in qk is exactly q and meets the conditions in (2.5). For all other primes, the oversized prime-power or valuation exclusions remove $O_B(X\delta_N)$ candidates by (2.2) and the valuation bound following (2.3). The condition $\Omega(k) \leq 10\lambda$ costs $o(X)$ and ensures $\Omega(qk) \leq 20\lambda$. Damaged powers with base different from p remove at most $C_2Xh_H/N \leq c_0X/16$ candidates by (2.4); any damaged power with base p is excluded by admissibility or is higher than q. Finally q itself is not damaged, since otherwise its base's smallest damaged power would divide it. Thus fewer than ε_0X candidates have $qk \in E$. The surviving number is at least $c_0X/2$ for large N, uniformly in q. □

By (1.1) and (3.1), $|H| \geq N/2 - O(N\delta_N) - (C_2 + 1)h_H \gg N$. Every prime power dividing a good integer is admissible and, by Lemma 3.4, occurs in H. Therefore $n \mid Q$ for every $n \in G$. Every prime power dividing b_0 is admissible for sufficiently large N: its damage threshold exceeds ηN by (3.1), as does that of every smaller power of the same prime. Hence $b_0 \mid Q$. Finally $Q \mid \operatorname{lcm}(1, 2, \ldots, \lfloor Y \rfloor)$, so $\log Q \ll Y$. These arguments establish the deterministic conclusions of Theorem 1.1. Moreover, by reducing η in (3.1) if necessary, we can ensure $|(N/2, N] \setminus H| < N/48$ for sufficiently large N.

4 Synchronizing near-divisibility windows

Lemma 4.1 (Common prime divisors)**.** *For fixed $J, \gamma > 0$, there is a fixed D such that, with $y = L^D$, if $U_i \subseteq (z_i/2, z_i]$, $|U_i| \geq \gamma z_i$ and $\sqrt{y} \leq z_i \leq 2y^J$ for $i = 1, 2$, then the product of the primes which divide at least one integer in each of U_1, U_2 exceeds N.*

Proof. Consider the primes $\mathcal{P} = [L^2, L^{D/8}] \cap \{\text{primes}\}$. For any such $U \subseteq (z/2, z]$, let $\mathcal{P}_0$ consist of primes in $\mathcal{P}$ dividing no member of U. Put $\nu = \sum_{p \in \mathcal{P}_0} 1/p$ and $g(n) = \sum_{p \in \mathcal{P}_0} \mathbf{1}_{p|n}$. Counting

multiples gives $\sum_{n\le z} g(n) \ge z\nu - |\mathcal{P}_0|$ and $\sum_{n\le z} g(n)^2 \le z(\nu+\nu^2)$. Expanding the variance about ν yields

$$\sum_{n\le z} (g(n)-\nu)^2 \le z\nu + 2\nu|\mathcal{P}_0| \le 2z\nu,$$

since $|\mathcal{P}_0| \le L^{D/8} \le z^{1/4}$. On U the function g vanishes. Hence $\gamma z\nu^2 \le 2z\nu$, or $\nu \le 2/\gamma$. By (2.1) the reciprocal mass of $\mathcal{P}$ is $\log(D/16)+o(1)$. Choose D so that $\log(D/16) > 4/\gamma + 2$. Primes common to the two sets then have reciprocal mass at least 1, and as each is at least L^2 there are at least L^2 of them. Their product is at least $(L^2)^{L^2} > N$. □

Complete the parameter choice by taking $\zeta = c_0/4$ in Lemma 3.3, then D in Lemma 4.1, then $w > DJ+2$ and take any fixed B sufficiently large that $B > \max(3D, 2w+2a+d+10)$. Here D is just a carrier parameter, unrelated to the height constant in Theorem 1.3. Set

$$W = N/L^w, \qquad T_* = \lceil L^{2w+2a+d+5} \rceil.$$

Thus $W < N/(2y^J)$ and $T_* = o(N/Y)$. Let $\mathcal{Q}$ be the prime powers occurring in H. For a real ξ let $\mathcal{D}_\xi$ be the powers for which fewer than T_* denominators $n \in H$ divisible by that power have $\operatorname{dist}(\xi, n\mathbb{Z}) \ge W/2$. Write $j(\xi) = |\mathcal{Q} \setminus \mathcal{D}_\xi|$.

Lemma 4.2 (Frequency windows)**.** *For fixed $\mathcal{D} \subseteq \mathcal{Q}$ the set of real ξ modulo Q having $\mathcal{D}_\xi = \mathcal{D}$ has measure at most $WQ/\operatorname{lcm}(\mathcal{D})$, where $\operatorname{lcm}(\varnothing) = 1$ (an upper bound larger than Q is harmless). If $j(\xi) = 0$, then $\operatorname{dist}(\xi, Q\mathbb{Z}) < W/2$.*

Proof. For each good q, take the reservoir in Lemma 3.4 and remove its fewer than T_* witnesses. Lemma 3.3 gives a carrier $d_q = qj_q \ge N/(2y^J) > W$ and a positive-density fiber U_q. Each $d_q u$ has a multiple in the open window of length W centered at ξ. Since $d_q > W$, all these multiples equal the same integer x_q. The product of the common primes of any two fibers exceeds N, by Lemma 4.1, and divides $x_q - x_{q'}$. Since this difference has absolute value $< W$, the points coincide. Thus the window contains a multiple of $\operatorname{lcm}(\mathcal{D})$. For the empty set the claim is immediate. Counting neighborhoods of multiples in a full period proves the measure bound. All powers in $\mathcal{Q}$ have lcm Q, so the last assertion follows. □

5 Joint Fourier inversion

Write $\theta_i = 0$ for $i \notin J$; these are fixed coordinates, not integration variables. In all expressions involving $\boldsymbol{\theta}$, extend by zeros to t coordinates. Write $e(x) = e^{2\pi i x}$ and $\|x\| = \operatorname{dist}(x, \mathbb{Z})$. Define the centered characteristic function $\phi(\boldsymbol{\xi}, \boldsymbol{\theta})$ of $\mathbf{Y}$ with these frequencies. Independence over denominators, and the identity

$$\left| \sum_{i=0}^{t} p_i e(u_i) \right|^2 = 1 - 4 \sum_{i<j} p_i p_j \sin^2\left(\pi(u_i - u_j)\right), \quad u_0 = 0,$$

yield, using only the factors from H,

$$|\phi| \le \exp\left[-cL^{-2a} \sum_{i=1}^{t} \sum_{n\in H} \|\xi_i/n + \theta_i\|^2 \right]. \tag{5.1}$$

All estimates below concern the unnormalized integral over $[-Q/2, Q/2]^t \times [-1/2, 1/2]^s$. The inversion factor is Q^{-t}.

Cardinality frequencies. There are $N/12 + O(1)$ pairs $(2j, 3j)$ with $N/4 < j \le N/3$. Each omitted denominator destroys at most two pairs, and H omits fewer than $N/48$ points of its half-interval. Thus $\gg N$ pairs survive. For each colour the identity

$$3(\xi_i/(3j) + \theta_i) - 2(\xi_i/(2j) + \theta_i) = \theta_i$$

and Cauchy–Schwarz imply a lower bound $cN\|\theta_i\|^2$ for the inner sum in (5.1). If any selected $|\theta_i| \ge W/(4N)$, then $|\phi| \le \exp(-cN/L^{2w+2a})$. Since $\log Q^t \ll tN/L^B \le N/L^{B-d}$, the entire absolute integral on this region is smaller than $\exp(-L^{d+3})$ for large N.

Reciprocal frequencies. Suppose every $|\theta_i| < W/(4N)$. For a bad power at ξ_i there are T_* witnesses with $\|\xi_i/n + \theta_i\| \ge W/(4N)$. Each denominator is counted by at most 20λ powers. Consequently

$$|\phi| \le \prod_{i=1}^{t} \exp(-L^{d+3} j(\xi_i)) \tag{5.2}$$

for all sufficiently large N, since $T_* L^{-2a}(W/N)^2/\lambda \gg L^{d+3}$. By Lemma 4.2, if $|\xi_i| > N/8$ in the fundamental period, then $j(\xi_i) \ge 1$. For a prescribed set of j bad powers, $Q/\operatorname{lcm}(\mathcal{D}) \le N^j$, by comparing the maximal prime-power valuations. There are at most N^j choices. Hence the integral of the single-coordinate envelope from (5.2) on $|\xi_i| > N/8$ is at most

$$E_N = \sum_{j \ge 1} N^{2j+1} e^{-jL^{d+3}} \le e^{-L^{d+3}/2}.$$

Its integral over the whole period is at most $N + E_N$. Thus the region with at least one large ξ_i has integral at most $tE_N(N + E_N)^{t-1} = O(e^{-L^{d+3}/4})$. No independence of colours has been assumed: the product is an upper bound on the joint characteristic function.

Covariance and central region. Let $\mathbf{p}_n = (p_{n,1}, \dots, p_{n,t})^\top$ and $V_n = \operatorname{diag}(\mathbf{p}_n) - \mathbf{p}_n \mathbf{p}_n^\top$. The full sum-and-length covariance has blocks $\sum_n V_n/n^2$, $\sum_n V_n/n$, $\sum_n V_n$. The covariance Σ is the principal submatrix retaining the selected length coordinates. All bounds for the full matrix therefore restrict to this subspace. For $n \in H$, $V_n \succeq L^{-2a} I_t$, by the same pairwise difference identity as for the variance of a categorical variable. Put $D_N = \operatorname{diag}(NI_t, I_s)$. The surviving pairs, whose first coordinates $N/(2j)$ and $N/(3j)$ differ by at least $1/2$, give

$$cNL^{-2a} I_{t+s} \preceq D_N \Sigma D_N \preceq CN(N/m)^2 I_{t+s}. \tag{5.3}$$

To see the lower bound explicitly, for any two vector components (x_i, y_i) the two values $(N/(2j))x_i + y_i$ and $(N/(3j))x_i + y_i$ have squared sum at least $c'(x_i^2 + y_i^2)$: the matrix of this pair of linear forms has bounded entries and determinant bounded away from zero. Sum over the surviving pairs, of bounded multiplicity, and use the lower covariance bound on H (put $y_i = 0$ for an omitted length). The upper bound uses $V_n \preceq I_t$ and at most N denominators. In particular $|\log\det\Sigma| = O(t \log N)$. In the region $|\xi_i| \le N/8$, $|\theta_i| < W/(4N)$, the phases on H are unwrapped: their absolute values are less than $1/2$. The same pair estimate now gives

$$|\phi| \le \exp\left[-cL^{-2a}\left(\|\boldsymbol{\xi}\|_2^2/N + N\|\boldsymbol{\theta}\|_2^2\right)\right].$$

Put $D_* = t + s$ for the integration dimension and set

$$b_N = \sqrt{N}\, L^{a+d+3}, \qquad R_N = L^{d+3}, \qquad \mathcal{K}_N = \{\|\boldsymbol{\xi}\|_\infty \le b_N,\ \|\boldsymbol{\theta}\|_\infty \le b_N/N\}.$$

Here b_N is a frequency scale, unrelated to the fixed target denominator b_0. For large N, $b_N = o(W)$, so this box lies inside the unwrapped region and the fundamental domain (as $Q \geq \max H > N/2$). Outside $\mathcal{K}_N$ but within the unwrapped region the preceding bound gives $\exp(-cL^{-2a}b_N^2/N) = \exp(-cL^{2d+6})$. That region has volume at most N^t, so its absolute integral is negligible compared with $\exp(-Ct\log N)$ for every fixed $C > 0$. Together with the two earlier minor-region bounds this leaves only $\mathcal{K}_N$.

Variance-weighted expansion. For a real random variable U with $|U| \leq M$ and M sufficiently small, let $X = U - \mathbb{E}U$ and $v = \operatorname{Var} U$. Then $|X| \leq 2M$, $\mathbb{E}|X|^3 \leq 2Mv$, and $v \leq 4M^2$. The scalar Taylor expansion of $e(X) = \exp(2\pi i X)$ gives

$$\mathbb{E}e(X) = 1 - 2\pi^2 v + O(Mv).$$

The increment from 1 is $O(v)$, so the logarithm defined by its power series about 1 is well-defined for small M. The quadratic error of this logarithm is $O(v^2) = O(M^2 v)$. It follows, with absolute constants, that

$$\log \mathbb{E}e(U - \mathbb{E}U) = -2\pi^2 \operatorname{Var} U + O(M \operatorname{Var} U), \qquad |\mathbb{E}e(U - \mathbb{E}U)| \leq e^{-c \operatorname{Var} U}. \tag{5.4}$$

These estimates also hold for deterministic U (when both sides of the logarithmic identity vanish). Thus no positive lower bound on the probabilities on A is needed.

For $\omega = (\boldsymbol{\xi}, (\theta_i)_{i \in J}) \in \mathcal{K}_N$, the individual uncentered phase at each denominator is 0 or $\xi_i/n + \theta_i$. For large N all participating n are at least $m = N^\kappa$. Thus the phases have modulus at most $M_N = 2b_N/m = o(1)$ uniformly, using $\kappa > 1/2$. Write $q(\omega) = \omega^\top \Sigma \omega$. The sum of the individual phase variances is exactly $q(\omega)$, by independence over denominators. Multiplying the individual expansions gives

$$\phi(\omega) = \exp\{-2\pi^2 q(\omega) + O(M_N q(\omega))\}, \qquad |\phi(\omega)| \leq \exp(-cq(\omega)).$$

The error here is complex and uniformly bounded in modulus by the displayed real bound. It is the sum of the logarithms of the individual factors, not an assertion about the branch of the logarithm of their product.

Covariance ellipsoid. Consider

$$\mathcal{C}_N = \{\omega : q(\omega) \leq R_N\}. \tag{5.5}$$

By (5.3), writing $\omega = D_N u$, one has

$$cNL^{-2a}\|u\|_2^2 \leq q(\omega), \qquad \|\boldsymbol{\xi}\|_\infty \leq N\|u\|_2, \qquad \|\boldsymbol{\theta}\|_\infty \leq \|u\|_2.$$

Consequently $\mathcal{C}_N \subseteq \mathcal{K}_N$ for large N: the sum-coordinate bound is $O(\sqrt{N}L^{a+(d+3)/2}) = o(b_N)$ and the length-coordinate bound is that quantity divided by N. Throughout $\mathcal{C}_N$,

$$M_N R_N \ll N^{1/2-\kappa} L^{a+2d+6} = o(1).$$

Hence $\phi(\omega) = e^{-2\pi^2 q(\omega)}(1 + O(M_N R_N))$ there. This gives a relative error in the unphased central Gaussian integral, and an absolute error of that same size after any target phase is inserted.

To control the tails, denote the full unphased Gaussian integral by

$$I_N = (2\pi)^{-D_*/2}(\det \Sigma)^{-1/2} = \int_{\mathbb{R}^{D_*}} e^{-2\pi^2 q(\omega)}\, d\omega.$$

For any fixed $c > 0$, the bound $e^{-cq} \le e^{-cR_N/2}e^{-cq/2}$ on $q > R_N$ and a linear change of variables by $\Sigma^{1/2}$ give

$$\frac{\int_{q>R_N} e^{-cq(\omega)}\,d\omega}{I_N} \le e^{-cR_N/2}(4\pi^2/c)^{D_*/2} = o(1), \tag{5.6}$$

since $D_* \le 2L^d$ and $R_N = L^{d+3}$. In particular the integral of $|\phi|$ on $\mathcal{K}_N \setminus \mathcal{C}_N$ is $o(I_N)$, as is the discarded Gaussian tail. By (5.3), $I_N \ge e^{-O(t\log N)}$, so all earlier absolute minor-region errors are $o(I_N)$ as well. For $\Delta = \mathbf{y} - \mu$ the full phased Gaussian integral is

$$\int_{\mathbb{R}^{D_*}} e^{-2\pi^2 q(\omega)} e(-\omega^\top \Delta)\,d\omega = I_N e^{-\Delta^\top \Sigma^{-1}\Delta/2} \ge I_N e^{-D_0^2/2}.$$

Thus the absolute error $o(I_N)$ is a relative $o(1)$ uniformly for the stated targets.

Finally $\mathbf{y} = (\mathbf{x}, \mathbf{k})$, with $\mathbf{x} \in Q^{-1}\mathbb{Z}^t$ and $\mathbf{k} \in \mathbb{Z}^s$, gives the exact mixed Fourier inversion identity

$$\mathbb{P}(\mathbf{Y} = \mathbf{y}) = Q^{-t} \int_{[-Q/2,Q/2]^t \times [-1/2,1/2]^s} \phi(\omega) e(-\omega^\top(\mathbf{y} - \mu))\,d\omega. \tag{5.7}$$

Gaussian integration gives (1.2), with dimension $t + s$. To justify exact inversion, $Q(Z_i - x_i)$ and the selected length differences are integers; each normalized frequency integral is the indicator that its corresponding integer is zero. This proves Theorem 1.1. The argument treats exact lattice points, not congruences modulo one.

6 Rational interval filling

Corollary 6.1 (Robust rational interval basis)**.** *For H, Q as in Theorem 1.1, every rational x in*

$$[L^{-a}R(H), (1 - L^{-a})R(H)] \cap Q^{-1}\mathbb{Z}$$

is the reciprocal sum of a subset of H for all sufficiently large N.

Proof. Take $t = 1$, $J = \varnothing$, $A = \varnothing$ and all $p_n = x/R(H)$ on H. Then $\mu = x$, so apply (1.2) with $t = 1$ and $J = \varnothing$ uniformly. The conclusion is simultaneous in the sense that each lattice point has a representation; no assertion of disjoint representations for different targets is intended. □

More generally, one may force any fixed subset of A to be used or omitted by setting its probabilities to one or zero. For a prescribed set A_* with $R(A_*)$ rational and $A_* \subseteq G \cap [m, N/2] \setminus F$, all the lattice points in the preceding interval shifted by $R(A_*)$ can be represented using A_* plus a subset of H. Also, if $h_H = o(N)$, every prime power $q \le \min\{Y, \varepsilon_0 N/(2\max\{1, h_H\})\}$ of exponent at most 10λ is admissible, by its damage threshold, and divides Q. Thus the corollary includes some growing target denominators, not just the fixed denominator b_0.

7 A fractional Hall theorem and disjoint prescribed representations

The following finite fractional allocation criterion uses classical rearrangement and convex separation, in the spirit of majorization and Rado's theorem [14]; the subsequent arithmetic completion is the additional ingredient. For weights $w_1 \ge \cdots \ge w_m > 0$ write $U(s)$ and $L(s)$ for the sum of the largest and smallest s units of capacity, with linear interpolation. These functions are defined for $0 \le s \le m$.

Lemma 7.1 (Fractional allocation)**.** *Given $k_i \geq 0$ with $\sum k_i \leq m$ and target masses u_i, there exist $p_{j,i} \geq 0$, $\sum_i p_{j,i} \leq 1$, $\sum_j p_{j,i} = k_i$, $\sum_j w_j p_{j,i} = u_i$, if and only if*

$$L(k_J) \leq u_J \leq U(k_J) \qquad (\varnothing \neq J \subseteq [t]).$$

Proof. Necessity is rearrangement. For sufficiency the feasible mass vectors form a compact convex set. Add unused capacity as a dummy colour with linear coefficient zero. For any linear functional $\sum_i a_i u_i$, order the coefficients: assign the largest weights to the highest coefficients and the smallest weights to the lowest coefficients, with fractional boundary slots. An exchange of weights between two inversely ordered coefficients cannot improve this allocation. Decompose the positive coefficient functional into nonnegative multiples of nested sets of positive coefficients, and the negative part into negative multiples of nested sets of negative coefficients (order by absolute value on each side). The upper and lower inequalities in the lemma bound the support function by exactly the value achieved by this allocation, simultaneously on all these nested sets. The separation theorem implies feasibility. □

Proof of Theorem 1.3: lower bound. First fix the parameters of Theorem 1.1: $a = 2$, any fixed $d > 0$, $\kappa = 9/10$, $D_0 = 1$ and the common denominator b_0 of the targets. Fix B above its threshold. The constants η, C_1, C_2 in the reservoir theorem can thus be chosen before the height multiplier D. Write $N = \lfloor DK \rfloor$ and suppose $h = |F| \leq H_* - CK\delta_N$, where C will be chosen after D. Choose constants $D_1 < D_2 < D_3$ as follows. Put $K_{\mathrm{tot}} = \sum K_i$. Choose D_1 large enough that $[m_0, D_1 K]$, $m_0 = \lceil K\delta_N \rceil$, can accommodate K_{tot} after $h + O(K\delta_N)$ deletions. Choose D_2 so large that $K_J/(D_2 K) < r_J/2$ for every J, and $D_3 - D_2$ large enough to accommodate K_{tot} after those deletions. Then choose the fixed $D > 2D_3$ large enough that, for all large K, $h_H \leq \eta N$ follows from $h \leq H_* = O(K)$ and $C_2 D_3 K/N \leq 1/2$. Now apply Theorem 1.1 at height N; for large K its label bound includes the fixed t. The constants in the following $O(K\delta_N)$ terms are fixed before the final choice of C. In $[2, D_3 K]$ the number of forbidden or non-good denominators is at most

$$h_{\mathrm{eff}} \leq h - h_H + C_1 D_3 K\delta_N + C_2 D_3 K h_H/N \leq h + O(K\delta_N).$$

Take A_- to consist of the smallest K_{tot} good, unremoved denominators in $[m_0, D_1 K]$, and A_+ of K_{tot} such denominators in $[D_2 K, D_3 K]$. Their existence follows from the choices. On H give each nonzero colour probability L^{-2}, and unused colour the remaining probability. Put $s_H = L^{-2}|H|$, $z_H = L^{-2} R(H)$, $k_i = K_i - s_H$, $u_i = r_i - z_H$. These are positive for large N, and $s_H = O(K/L^2)$, $z_H = O(L^{-2})$. Use Lemma 7.1 with the weights on $A_\pm$. Uniformly for nonempty J, rearrangement gives

$$U(k_J) \geq \log\left(1 + \frac{\lfloor k_J \rfloor}{m_0 + h_{\mathrm{eff}}}\right) \geq r_J + \delta_N$$

for C large: use $H_* \leq K_J/(e^{r_J} - 1)$, $k_J = K_J - O(K/L^2)$ and the fixed comparisons $K_i \asymp K$. The middle interval gives $L(k_J) \leq k_J/(D_2 K) < r_J/2 < u_J$. Hence we can realize all expectations simultaneously on the main pool. The pool lies above N^κ for large K. All Qr_i and K_i are integral. Take $J = [t]$ in Theorem 1.1. By (1.2) the exact central event has positive probability. Its colours are disjoint, so F fails to block the family. □

Matching upper bound. Choose a minimizer J for H_* and put $r' = r_J$, $K' = K_J$. The union of its representations would be a K'-term representation of r'. For primes $p \in [K/\sqrt{\log K}, K/(\log K)^{1/4}]$, $p^2 > N$ and p is coprime to the fixed reduced denominator of r'. If a representation uses multiples pm_i of p, reduction after multiplication by p requires $\sum m_i^{-1} = 0$ modulo p. All m_i are at most $D\sqrt{\log K}$. Its positive numerator after taking a product denominator is at most

$s(D\sqrt{\log K})^{s-1}$ with $s \le D\sqrt{\log K}$, smaller than p for large K. Thus these prime multiples are unusable, and distinct such primes have disjoint multiples. Delete an initial segment $F_h = [2, h+1]$ with $h = \lfloor H_* - cK\delta_N \rfloor$. The maximal harmonic mass of K' remaining terms, before arithmetic exclusions, is $r' + O(c\delta_N + K^{-1})$. In the lower half of that maximizing interval, the unusable multiples number $\gg K\sum_p 1/p \gg K\delta_N$. Replacing any of them by larger denominators costs $\gg 1/K$ per point (or at least as much when fewer terms are used). This is seen by subtracting the boundary weight $1/(h + K' + 1)$ from the original and replacement terms. For sufficiently small c the available mass is strictly below r'. This forbids the union and therefore the required family. The two bounds prove (1.3). □

8 Quantitative entropy for disjoint representations

Let t now be fixed, and let $\mathbf{r} = (r_1, \dots, r_t) \in \mathbb{Q}^t_{>0}$ be fixed. Count ordered families of mutually disjoint subsets of $[1, N] \setminus F$ with reciprocal sums r_i. We consider two cases: either all lengths are unrestricted, or all lengths are prescribed to be $\lfloor \alpha_i N \rfloor$ for fixed $\alpha_i > 0$. Write $C_{\mathbf{r}}(N; F)$ or $C_{\mathbf{r},\boldsymbol{\alpha}}(N; F)$ for these counts, and $C(N; F)$ when the choice of constraints is understood. The corresponding choice of length multipliers is $\boldsymbol{\nu} = 0$ (not optimized) or $\boldsymbol{\nu} \in \mathbb{R}^t$, respectively. The entropy of a fixed target vector without deletions has a first-order formula in [12]; the additional results here are the quantitative deficit, arbitrary sublinear deletions, and the simultaneous exact length constraints.

Theorem 8.1 (Robust simultaneous entropy)**.** *In the prescribed-length case assume*

$$\alpha_{[t]} < 1, \qquad r_J > -\log(1 - \alpha_J) \quad (\varnothing \ne J \subseteq [t]), \tag{8.1}$$

where subscripts on vectors denote sums over the indicated indices. In the unrestricted case no additional condition is imposed. There are unique $\boldsymbol{\beta} \in (0, \infty)^t$ and (if lengths are prescribed) $\boldsymbol{\nu} \in \mathbb{R}^t$ for which, on $0 < u \le 1$,

$$Z(u) = 1 + \sum_{i=1}^{t} e^{-\beta_i/u - \nu_i}, \qquad p_i(u) = \frac{e^{-\beta_i/u-\nu_i}}{Z(u)}, \quad p_0(u) = \frac{1}{Z(u)},$$

$$\int_0^1 \frac{p_i(u)}{u}\, du = r_i \quad (1 \le i \le t), \qquad \int_0^1 p_i(u)\, du = \alpha_i \quad \text{if lengths are prescribed.} \tag{8.2}$$

Define

$$\begin{aligned} \Gamma &= \sum_i \beta_i r_i + \sum_i \nu_i \alpha_i + \int_0^1 \log Z(u)\, du \\ &= -\int_0^1 \sum_{i=0}^{t} p_i(u) \log p_i(u)\, du, \end{aligned}$$

with the length term omitted in the unrestricted case. Then for $h = |F| = o(N)$, uniformly over such deletion sets,

$$N\Gamma - O_{\mathbf{r},\boldsymbol{\alpha},t}(h + N\delta_N) \le \log C(N; F) \le N\Gamma - c_{\mathbf{r},\boldsymbol{\alpha},t} N\delta_N \tag{8.3}$$

for a positive constant $c_{\mathbf{r},\boldsymbol{\alpha},t}$ (with dependence on $\boldsymbol{\alpha}$ omitted in the unrestricted case). Thus the deficit from $N\Gamma$ is of order $N \log\log N / \log N$ if $h = O(N\delta_N)$. In the prescribed-length case this concerns the number of ordered disjoint representations, not the product of separate single-target counts.

The inequalities in (8.1) are the strict form of the necessary lower-capacity inequalities. Indeed, even without arithmetic restrictions, a union of $\alpha_J N + O(1)$ denominators has reciprocal mass at least $-\log(1-\alpha_J) + o(1)$ if $\alpha_J < 1$. If the total fraction is one, the union has mass tending to infinity. We make no assertion about feasibility or entropy on an equality boundary.

Lemma 8.2 (Continuum feasibility and Gibbs parameters)**.** *The parameters in Theorem 8.1 exist and are unique. Their expectation map has an invertible Jacobian at the solution, in the variables $\boldsymbol{\beta}$ or $(\boldsymbol{\beta}, \boldsymbol{\nu})$ as appropriate.*

Proof. Consider the convex objective

$$\Phi(\boldsymbol{\beta}, \boldsymbol{\nu}) = \sum_i \beta_i r_i + \sum_i \nu_i \alpha_i + \int_0^1 \log\left(1 + \sum_i e^{-\beta_i/u - \nu_i}\right) du$$

on $\boldsymbol{\beta} \geq 0$, with the length term and the optimization over $\boldsymbol{\nu}$ omitted in the unrestricted case. It is finite and continuous on this domain: on compact parameter sets the integrand is bounded by a constant, since the β_i are nonnegative. It is strictly convex there. For two different parameter vectors, the corresponding logits differ for some label on a set of positive measure in u; the pointwise log-partition function is strictly convex in the nonzero-label logits with label zero fixed. In the unrestricted case coercivity follows from $\Phi \geq \sum_i \beta_i r_i$.

For prescribed lengths we give the coercivity argument explicitly. Choose a fixed $\epsilon > 0$ small enough that $\sum_i \alpha_i < 1 - \epsilon$ and

$$-\log(1-\alpha_J) < r_J < \log(1 + \alpha_J/\epsilon) \qquad (\varnothing \neq J \subseteq [t]).$$

All inequalities remain strict in a neighborhood of $(\mathbf{r}, \boldsymbol{\alpha})$. On the interval $[\epsilon, 1]$ with Lebesgue measure, the fractional allocation criterion of Lemma 7.1 holds with the decreasing weight $w(u) = 1/u$. In detail, for prescribed length masses α_i' summing to less than $1 - \epsilon$, the possible reciprocal-mass vectors are the image of the functions $a_i(u) \geq 0$, $\sum_i a_i(u) \leq 1$, $\int a_i = \alpha_i'$. This set is convex and compact in the weak-* topology of bounded measurable functions, and its image in $\mathbb{R}^t$ is compact because w is bounded on the interval. For any linear objective, allocate the largest weights to the largest coefficients, the smallest weights to the smallest coefficients, including the unused colour with coefficient zero. The rearrangement bound for this allocation is attained by consecutive intervals of the appropriate lengths. Decomposing the positive and negative coefficients into nested sums as in Lemma 7.1, the support-function inequalities are exactly the subset inequalities with

$$U_\epsilon(s) = \int_\epsilon^{\epsilon+s} \frac{du}{u} = \log(1 + s/\epsilon), \qquad L_\epsilon(s) = \int_{1-s}^1 \frac{du}{u} = -\log(1-s).$$

Separation gives feasibility for every nearby vector $(\mathbf{r}', \boldsymbol{\alpha}')$. Extend the allocating functions by zero below ϵ, with all unused probability assigned to label zero.

For any such allocation, the pointwise inequality

$$\log Z(u) \geq -\sum_i a_i(u)(\beta_i/u + \nu_i) - \sum_{i=0}^t a_i(u) \log a_i(u)$$

follows from nonnegativity of relative entropy, where $a_0 = 1 - \sum_{i \geq 1} a_i$ and $0 \log 0 = 0$. Its last term is nonnegative. Integration gives $\Phi \geq \boldsymbol{\beta} \cdot (\mathbf{r} - \mathbf{r}') + \boldsymbol{\nu} \cdot (\boldsymbol{\alpha} - \boldsymbol{\alpha}')$. A fixed Euclidean ball of nearby targets is feasible. For a nonzero parameter vector choose the displacement from the

specified target to the nearby target opposite to the parameter vector, of a fixed length contained in that ball. This bounds Φ below by a positive constant times the norm of the parameter vector. Thus Φ is coercive, including in the unrestricted length-multiplier directions.

The unique minimizer cannot have $\beta_i = 0$. With all other parameters fixed there, $p_i(u)$ is bounded below by a positive constant for $u \in (0, 1]$: its numerator is $e^{-\nu_i}$ and the denominator is at most $1 + \sum_j e^{-\nu_j}$. Increasing just β_i by $v > 0$ changes the integral term by

$$\int_0^1 \log(1 - p_i(u)(1 - e^{-v/u}))\, du \le -c \int_v^1 \frac{v}{2u}\, du.$$

This dominates the positive linear change vr_i for small v. Thus the minimum is interior in every β_i. Differentiating there is justified by exponential decay at zero and yields (8.2). Strict convexity gives uniqueness. The entropy identity in the theorem follows by expanding $-\sum_i p_i \log p_i$.

Finally the Hessian is the integral of a categorical covariance: its quadratic form in a parameter direction $(\mathbf{b}, \mathbf{v})$ is the integrated variance of the random phase taking values 0 and $b_i/u + v_i$, with all label probabilities positive (put $\mathbf{v} = 0$ in the unrestricted case). If this variance vanishes identically almost everywhere then $b_i/u + v_i = 0$ for each i almost everywhere, hence the parameter direction is zero. The Hessian is positive definite. The Jacobian of the expectation map is the negative Hessian of Φ and is therefore invertible. □

Proof of Theorem 8.1. Use Theorem 1.1 with $a = 2$, any fixed $d > 0$, $\kappa = 9/10$, and b_0 a common denominator of the r_i. Put $P = G \cap [m, N] \setminus F = A \cup H$. By (1.1) and $m = o(N\delta_N)$,

$$|[1, N] \setminus P| = O(h + N\delta_N).$$

For parameters in a small compact neighborhood of the continuum solution in Lemma 8.2, the normalized discrete expectation maps are

$$r_i^{(N)} = \frac{1}{N} \sum_{n \in P} \frac{p_i(n/N)}{n/N}, \qquad \alpha_i^{(N)} = \frac{1}{N} \sum_{n \in P} p_i(n/N)$$

(the latter only if prescribed). The functions involved and all their first parameter derivatives extend smoothly at $u = 0$, since all β_i are bounded below by a positive constant. The integrands have bounded first u-derivatives uniformly in the neighborhood, so the Riemann sums and their first parameter derivatives differ from the continuum maps by $O(h/N + \delta_N)$. The omissions cost at most this order, since their cardinality has the above bound. In the length case the required discrete length fractions differ from α_i by at most $1/N$. Invertibility from Lemma 8.2 now gives parameters $(\boldsymbol{\beta}_N, \boldsymbol{\nu}_N)$ within $O(h/N + \delta_N)$ of the limiting solution at which the exact expectations are r_i and $\lfloor \alpha_i N \rfloor$ (for the actual, not normalized, length variables). For clarity this is the uniform small-perturbation inverse-function argument: on a small ball about the continuum solution, use the inverse of the limiting Jacobian to define the fixed-point map for the discrete residual; by the C^1 convergence and continuity of the limiting Jacobian its Lipschitz constant is less than $1/2$, and its displacement at the center is $O(h/N + \delta_N)$. A ball of radius a sufficiently large constant times that error is mapped into itself.

On H, all $t + 1$ categorical probabilities are bounded away from zero uniformly in N. The means are lattice points, since $b_0 \mid Q$. Apply (1.2) with $J = [t]$ or $J = \varnothing$, as appropriate. If $\mathcal{E}$ denotes the exact event, the determinant bounds in (5.3) give

$$\log \mathbb{P}(\mathcal{E}) = -t \log Q + O_t(\log N) \ge -O(N/L^B + \log N).$$

All exact outcomes have identical probability

$$\exp\left(-N\sum_i \beta_{N,i} r_i - \sum_i \nu_{N,i}\lfloor \alpha_i N\rfloor\right) \prod_{n\in P}\left(1+\sum_i e^{-\beta_{N,i}N/n-\nu_{N,i}}\right)^{-1},$$

with the length term omitted when applicable. Thus division of the exact-event probability by this probability counts the outcomes exactly. There is no possible collision from a repair map. Estimating the logarithm by Riemann sums, the omitted-denominator count, and the parameter perturbation proves the lower bound of (8.3) (since $B > 2$).

For the upper bound, let $\mathcal{U}$ be the integers in $(N/2, N]$ divisible by a prime $p \in [N/L^{1/2}, N/L^{1/4}]$. No such denominator occurs in a representation of any of the fixed r_i for large N. Indeed p does not divide their fixed reduced denominators and $p^2 > N$. Reduction modulo p after multiplying the sum by p would require a nonempty sum of reciprocals of distinct integers at most $L^{1/2}$ to vanish modulo p. Clearing denominators by their product gives a positive numerator of logarithm $O(L^{1/2}\log L) = o(\log p)$, impossible. The prime fibers are disjoint because their primes exceed $\sqrt{N}$. Each contains $\gg N/p$ integers of the upper half, uniformly over this range. The Mertens estimate therefore gives $|\mathcal{U}| \gg N\delta_N$. With the continuum parameters, the exact-count upper bound is

$$C(N;F) \le \exp\left(N\sum_i \beta_i r_i + \sum_i \nu_i\lfloor \alpha_i N\rfloor\right) \prod_{n\in[1,N]\setminus\mathcal{U}}\left(1+\sum_i e^{-\beta_i N/n-\nu_i}\right).$$

It holds whether or not F is imposed, by summing over all label assignments to the allowed denominators and dropping the other constraints. Each omitted $n \in \mathcal{U}$ subtracts from the log partition function at least $\min_{1/2\le u\le 1}\log Z(u) > 0$. The full log partition function is $N\int_0^1 \log Z(u)\,du + O(1)$ and rounding the length fractions costs $O(1)$. This proves the upper bound, and completes the proof. □

Corollary 8.3 (Scalar counting with and without prescribed length)**.** *For a fixed positive rational r write $C_r(N;F)$ for its representation count. There is a unique $\beta > 0$ with $\int_0^1 (u(1+e^{\beta/u}))^{-1}\,du = r$. Put*

$$\Gamma_r = \beta r + \int_0^1 \log(1+e^{-\beta/u})\,du.$$

Then $\log C_r(N;F)$ lies between $N\Gamma_r - O_r(h+N\delta_N)$ and $N\Gamma_r - \Omega_r(N\delta_N)$ for $h = o(N)$. If instead the count is restricted to $\lfloor \alpha N\rfloor$ terms, where $0 < \alpha < 1 - e^{-r}$, let (β,ν) be the unique pair with $\beta > 0$ and $p(u) = (1+e^{\beta/u+\nu})^{-1}$ satisfying $\int_0^1 p(u)/u\,du = r$, $\int_0^1 p(u)\,du = \alpha$. The same bounds hold with Γ_r replaced by

$$\Gamma_{r,\alpha} = \beta r + \nu\alpha + \int_0^1 \log(1+e^{-\beta/u-\nu})\,du.$$

They also hold for denominators restricted to $[2,N]$ (add 1 to F).

Proof. Apply Theorem 8.1 with one colour, and observe that $r > -\log(1-\alpha)$ is equivalent to the stated range of α. □

For identical targets $r_i = r$ and identical length fractions $\alpha_i = \alpha$, uniqueness forces common multipliers. The strict feasible condition reduces to $t\alpha < 1 - e^{-tr}$: the function $s \mapsto (1-e^{-sr})/s$

decreases on $s > 0$. In the length-constrained case the rate is therefore computed by two real parameters and

$$\Gamma = t(\beta r + \nu\alpha) + \int_0^1 \log(1 + te^{-\beta/u-\nu})\, du.$$

The result remains a fixed-t estimate; it cannot be applied by simply substituting $t \sim \log N$ into the implicit constants. The latter regime requires different counting arguments, such as those in [12].

9 Modular coverage at a second scale

We next recover most of the denominators excluded by the large prime cutoff. The two-scale idea is closely related to Lemma 4.1 and the proofs of Theorems 1.5–1.6 in the preprint version of Liu–Sawhney [4]. The point needed here is to combine such a transfer with uniform deletion estimates and simultaneous disjoint completion.

Lemma 9.1 (Robust modular coverage)**.** *Let X be a sufficiently large integer, $\ell = \log X$, and let $E_0 \subseteq (X/2, X]$ satisfy $|E_0| \le 2X/\ell^2$. For a prime $p > X$, suppose*

$$\frac{X}{(\log X)^B} \ge C_0 \log p$$

with a sufficiently large constant C_0 (for the fixed choices of the auxiliary parameters). Then for every $c \in \mathbb{F}_p$ there is a subset $E \subseteq (X/2, X] \setminus E_0$ such that

$$\sum_{m\in E} m^{-1} = c \quad \text{in } \mathbb{F}_p, \qquad R(E) \le \log 2 + O(1/X). \tag{9.1}$$

Proof. Apply the reservoir construction of Theorem 1.1 with $N = X$, $F = E_0$, $b_0 = 1$, $a = 4$, $d = 2$, $\kappa = 9/10$ and the same fixed cutoff exponent B, chosen sufficiently large, and denote its high pool and common denominator by H_X, Q_X. Let $Y_X = X/\ell^B$. By Lemma 2.1, the reciprocal sum of the damaged powers is $O(\ell^{-2})$. Among primes in $(Y_X/2, Y_X]$, at most $O(Y_X/\ell^2)$ are damaged. There are $\asymp Y_X/\log Y_X \asymp Y_X/\ell$ primes in this interval by the prime number theorem, and all but $o(1)$ of their proportion remain. Each remaining prime is admissible for Lemma 3.4, so divides Q_X. Consequently

$$\log Q_X \gg Y_X. \tag{9.2}$$

Moreover $R(H_X) = \log 2 + o(1)$ by (1.1) at scale X and $|E_0| = o(X)$, since the loss of each upper-half denominator costs at most $2/X$. For every $u \in [R(H_X)/3, 2R(H_X)/3] \cap Q_X^{-1}\mathbb{Z}$, take one nonzero colour with constant probability $u/R(H_X)$ and use Theorem 1.1 with $A = \varnothing$, $J = \varnothing$. Its two colour probabilities lie in $[1/3, 2/3]$, and its mean is a valid grid point. Thus every such u is a subset sum of H_X.

Since $p > X$, $p \nmid Q_X$. On increasing C_0 if necessary, (9.2) makes the interval of integer numerators $Q_X[R(H_X)/3, 2R(H_X)/3]$ contain at least p consecutive integers. Reducing those numerators divided by Q_X modulo p covers every class in $\mathbb{F}_p$. Choose the subset for the required class. Its harmonic weight is at most the weight of $(X/2, X]$, as claimed. □

Remark 9.2. The modulus in Lemma 9.1 is coprime to all local denominators. The result concerns arbitrary subsets of integers in a dyadic interval; it does not require or establish that the correction terms themselves are prime. In particular this is not a claim about the distinct-prime version of the modular subset-sum problem.

10 Large-prime blocks and the exact loss budget

Fix $r > 0$ rational in lowest terms with denominator b. Assume $|F| \le N/L$ and use the construction of Theorem 1.1 at the global scale with $b_0 = b$, $a = 4$, $d = 2$, $\kappa = 9/10$ and the same B as above; for large N its condition $h_H \le \eta N$ holds. The choices of B in the carrier proof are independent of b_0, which affects only the small deletion fraction η. Write G, H, Q, Y for the resulting objects. Let

$$X_0 = \left\lceil L\lambda^{B+3} \right\rceil.$$

Then $X_0 < L^B$ for all sufficiently large N (we have $B > 1$). A prime p with $Y < p \le N/X_0$ satisfies

$$x = \lfloor N/p \rfloor \in [X_0, L^B], \qquad p > x, \qquad x/(\log x)^B \ge c_B L\lambda^3 \gg \log p. \tag{10.1}$$

The final comparison uses $\log x \le B\lambda$ and is uniform over this entire range. At this scale all prime powers of a cofactor $m \le x$ are at most Y, even though the prime p is not allowed in G.

Proposition 10.1 (Disjoint modular cancellation blocks)**.** *There is a family of pairwise disjoint nonempty blocks B_p indexed by some of the primes $Y < p \le N/X_0$ such that each $B_p \subseteq p\mathbb{N} \cap ([1, N] \setminus F)$, all blocks are disjoint from G, and*

$$QR(B_p) \in \mathbb{Z}, \qquad R(B_p) \ll_B \lambda L^B/N. \tag{10.2}$$

All their denominators satisfy $1/n < r$ for large N. Let $\mathcal{R}$ consist of these blocks together with the singleton elements $\{n\}$ for $n \in G \setminus F$ with $1/n \le r$. Then the total weight of $\mathcal{R}$ satisfies

$$\sum_{A \in \mathcal{R}} R(A) \ge W_r(N; F) - \tfrac{1}{2} g(N) - O_B(e_N + h_H L/N). \tag{10.3}$$

Proof. For each prime p in the range above, put $x = \lfloor N/p \rfloor$ and define

$$C_p = \{m \in [1, x] \cap G : pm \notin F\}, \qquad f_p = |F \cap (N/2, N] \cap p\mathbb{N}|.$$

First suppose $f_p \le x/(\log x)^2$. The high interval $(x/2, x]$ loses at most f_p elements from the condition $pm \notin F$: if $m > x/2$ is integral, then $2m \ge x + 1 > N/p$, so $pm > N/2$. (When $x = N/p$ is integral the same strict inequality follows directly.) Equation (1.1) excludes at most $O_B(x\lambda/L + xh_H/N) = O_B(x\lambda/L)$ further elements on imposing $m \in G$. This is $o(x/(\log x)^2)$ uniformly by (10.1). Thus the missing set in $(x/2, x]$ has cardinality at most $2x/(\log x)^2$ for all sufficiently large N.

By Lemma 9.1 and (10.1), there is $E_p \subseteq C_p \cap (x/2, x]$ with

$$\sum_{m \in E_p} m^{-1} = \sum_{m \in C_p} m^{-1} \quad \text{in } \mathbb{F}_p, \qquad R(E_p) \ll 1. \tag{10.4}$$

Take $B_p = \{pm : m \in C_p \setminus E_p\}$, omitting empty blocks. Let $D_p = \operatorname{lcm}(C_p)$, with the empty lcm defined to be one. By Theorem 1.1, $D_p \mid Q$, and $p \nmid D_p$ since $p > x$. Equation (10.4) says that the integer numerator $D_p R(C_p \setminus E_p)$ is divisible by p. Therefore $QR(B_p) \in \mathbb{Z}$, since $R(B_p) = R(C_p \setminus E_p)/p$. Its mass is at most $H_x/p \ll_B \lambda L^B/N$.

Distinct primes in this range exceed $Y > \sqrt{N}$, so an integer at most N cannot be divisible by two of them. Also $p > Y$ implies that none of the members of B_p belong to G. These facts establish the disjointness, not merely a bound on overlaps. Each member exceeds Y, hence has weight less than r eventually.

We now account for all the mass lost relative to $W_r(N; F)$. The direct losses of F and of integers with weight greater than r are already excluded from W_r. For the remaining losses the following upper bounds suffice.

Moderate-prime cancellations. For primes treated above, $R(pE_p) = R(E_p)/p \ll 1/p$. The total is $O_B(\lambda/L)$, since $\sum_{Y<p\le N} 1/p = O_B(\lambda/L)$.

Heavily damaged moderate-prime fibres. If $f_p > x/(\log x)^2$, retain no block for p. Dropping all its available multiples costs at most

$$H_x/p \ll_B \lambda x/N \ll_B \lambda^3 f_p/N.$$

Here $1/p < (x+1)/N \le 2x/N$. Since the fibres are disjoint, $\sum_p f_p \le h_H$. The total loss is $O_B(h_H\lambda^3/N) = O_B(h_H L/N)$.

Other arithmetic restrictions. Exclude integers divisible by a prime power $p^j > Y$ with $j \ge 2$, those with $v_p(n) > 10\lambda$ or $\Omega(n) > 20\lambda$, and those divisible by any $q \in \mathcal{B}$. For the first two types use the prime-power tail bounds of the pruning section and the estimate $\sum_{n\le N, q|n} 1/n = H_{\lfloor N/q\rfloor}/q \le (1+L)/q$. The resulting costs are $O(LY^{-1/2})$ and $O(L^{1-10\log 2})$, respectively. For the third type the uniform count bound from (2.3), followed by partial summation, gives $O(L^{3/2-20\log(3/2)})$. All three are $o(\lambda/L)$. For the last type the cost is $O(h_H L/N)$ by Lemma 2.1. These estimates cover not only excluded small-prime integers but also missing cofactors in C_p: for $m \le x \le L^B < Y$, every failure of $m \in G$ is one of the latter three restrictions, inherited by pm. This accounts for the missing cofactors without assuming that the moderate-prime multiples form an unrestricted interval.

Extreme-prime loss. The remaining loss is bounded by

$$D(N, X_0) = \sum_{p > N/X_0,\ p\le N} H_{\lfloor N/p\rfloor}/p.$$

For any $2 \le X \le L^{B+1}$ the estimates (2.1), $H_{\lfloor u\rfloor} = \log u + O(1)$ for $u \ge 1$, and partial summation imply

$$\begin{aligned} D(N, X) &= \int_{N/X}^{N} \frac{\log(N/v)}{v\log v}\, dv + O_B\big((1+\log X)/L\big) \\ &= \frac{(\log X)^2}{2L} + O_B\left(\frac{1+\log X}{L} + \frac{(\log X)^3}{L^2}\right). \end{aligned} \tag{10.5}$$

For clarity, the $O(1)$ difference between $H_{\lfloor N/p\rfloor}$ and $\log(N/p)$ costs at most $O(\sum_{p>N/X} 1/p) = O_B((1+\log X)/L)$. In partial summation the Mertens remainder is $O(1/L)$ throughout $[N/X, N]$, and the boundary and integral errors are $O((1+\log X)/L)$. The main integral becomes $\int_0^{\log X} u/(L-u)\, du$ after substitution, giving the second line. Since $\log X_0 = \lambda + (B+3)\log\lambda + o(1)$,

$$D(N, X_0) = \frac{\lambda^2}{2L} + O_B(e_N). \tag{10.6}$$

All losses are covered by these categories: an available n not rejected by the other restrictions is either in G, has a unique moderate prime $p > Y$ with $n = pm$ and $m \in C_p$, or has an extreme prime. The prime fibres are disjoint, and overcounting other losses can only increase our bound. Combining the estimates proves (10.3). □

11 Packing whole blocks and simultaneous completion

We prove the lower bound in Theorem 1.4. The special case $r = 1, F = \varnothing$ will give the lower bound in Corollary 13.1.

Apply Proposition 10.1. If the equality $1/n = r$ holds for an available denominator, use $\{n\}$ as an independent representation and omit it from the rest of the argument. For large N that fixed denominator, if it exists, is in G, by $|F| \le N/L$ and the definition of damaged powers. Its subtraction changes the target number of groups by exactly one. All other singleton weights under consideration are strictly less than r.

Let $m = \left\lceil N^{9/10} \right\rceil$ and reserve the pool

$$P = G \cap [m, N] \setminus F$$

for random completion. For large N all these weights are less than r and $H \subseteq P$. The other elements of $\mathcal{R}$ (excluding the singleton just treated, if any) are individual small denominators and whole blocks B_p. Let W be the weight of all these elements together with P, and $W_D = W - R(P)$ the weight outside P. Set

$$t = \lfloor (W - 1/L)/r \rfloor . \tag{11.1}$$

We have $t = O_r(L)$, and also $t \to \infty$. Indeed $R([1, N] \setminus G) \ll_B g(N) + h_H L/N = O_B(1)$: this follows by the same loss estimates as in Proposition 10.1, now charging all primes exceeding Y using (10.5) with $X = L^B$ and charging the other restrictions as before. By deleting the $h = |F|$ smallest denominators in $[m, N]$ one maximizes the lost mass, so

$$R(P) \ge \sum_{n=m+h}^{N} 1/n - O_B(1) \ge \log\log N - O_B(1) \longrightarrow \infty. \tag{11.2}$$

We used $m + h = O(N/L)$ for large N. In particular $t \to \infty$, $t \le L^2$ eventually, and

$$rt - W_D = R(P) - (W - rt) \ge R(P) - r - 1/L \ge 2.$$

This slack is provided by the reserved pool, not by discarding additional weight from the eventual decomposition.

Deterministic allocation. By (10.2), each block has weight $o(1/L)$, hence at most $1/t$ for large N. The only elements with weight exceeding $1/t$ are singleton denominators $n < t$, of which there are fewer than t. Assign these to separate boxes, among t boxes. Their weights are at most $r - \eta_r$, where

$$\eta_r = r - \max\{1/n : n \in \mathbb{N},\ 1/n < r\} > 0.$$

Thus each initial box has weight at most $r - 1/t$ for large N. Assign every remaining element, of weight at most $1/t$, to a currently lightest box. At every step the minimum current weight is at most $W_D/t \le r - 2/t$, so no box ever exceeds $r - 1/t$. Empty boxes are allowed. Each block is assigned intact. Denote the disjoint unions of the assigned sets by D_i, $1 \le i \le t$. They are disjoint from P and avoid F. Since all singletons divide Q and all blocks have weights on its grid,

$$s_i = r - R(D_i) \ge 1/t, \qquad Qs_i \in \mathbb{Z}. \tag{11.3}$$

The divisibility uses $b \mid Q$; one may not assume a fixed denominator for the residual targets themselves.

Exact completion. Independently assign each $n \in P$ colour i with probability $p_i = s_i/R(P)$, and colour zero with probability

$$p_0 = 1 - \sum_{i=1}^{t} p_i = (W - rt)/R(P).$$

These are valid probabilities by (11.1). Since $W - rt \geq 1/L$, $R(P) \leq H_N = O(L)$ and $t = O_r(L)$, each probability (including p_0) is at least L^{-4} for sufficiently large N. For each colour the expected sum is s_i. The vector of means belongs to $Q^{-1}\mathbb{Z}^t$ by (11.3). Apply Theorem 1.1 with $A = P \setminus H$ and $J = \varnothing$ to conclude that some colouring attains *all* these sums at once. Let A_i be its set of colour i and put $S_i = D_i \cup A_i$. Then $R(S_i) = r$, and the sets S_i are pairwise disjoint, contained in $[1, N] \setminus F$. The whole-block assignment is crucial: splitting a block after cancellation could reintroduce its prime denominator.

Using (10.3) and (11.1), and adding back the isolated singleton if present, yields (1.4); the extra $1/L$ is absorbed by $O(e_N)$. This proves the robust lower bound. In particular no independent probability argument for separate colours, no unlimited-height representation theorem, and no assumption that deletions are random have been used.

12 The matching obstruction

We prove (1.5), using the familiar large-prime obstruction from Croot's work [6]. Fix a sufficiently small absolute constant $c > 0$. If a prime $p > N/(cL)$ divides any denominator $n \leq N$ of an exact representation of r, then $p^2 > N$ and $p \nmid b$ for sufficiently large N. The terms divisible by p are $1/(pm)$ for a nonempty set of integers $m \leq cL$. Multiplying the representation by p and reducing modulo p in the localization at p gives

$$\sum_{m \in A} 1/m = 0 \quad \text{in } \mathbb{F}_p.$$

Let $D = \operatorname{lcm}(1, \ldots, \lfloor cL \rfloor)$. The numerator $D \sum_{m \in A} 1/m$ is a positive integer divisible by p but at most $DH_{\lfloor cL \rfloor}$. Chebyshev's bound in (2.1) implies $\log(DH_{\lfloor cL \rfloor}) \leq CcL + O(\log \lambda) < \log p$ for c small and N large. This is a contradiction. Thus all multiples at most N of such primes are unusable. Those prime fibres are disjoint, and their weights are less than r eventually. By (10.5) with $X = cL$, their total weight equals

$$D(N, cL) = \tfrac{1}{2} g(N) + O(\lambda/L). \tag{12.1}$$

Hence $rk_r(N; \varnothing) \leq W_r(N; \varnothing) - D(N, cL)$, proving (1.5). This completes the proof of Theorem 1.4. For $r = 1$, the same obstruction applies to any single positive integer target, even one varying with N: multiplication by p gives zero modulo p for every integer. Consequently it bounds the largest integer representable with denominators at most N as well.

13 Consequences for unit-target thresholds and exceptional heights

Corollary 13.1 (Unit-target threshold)**.** *There is an absolute constant C such that, for all sufficiently large N,*

$$\left\lfloor H_N - \tfrac{1}{2} g(N) - Ce_N \right\rfloor \leq k(N) \leq \left\lfloor H_N - \tfrac{1}{2} g(N) + C\lambda/L \right\rfloor. \tag{13.1}$$

Proof. Take $r = 1$ and $F = \varnothing$ in Theorem 1.4, noting that $W_1(N; \varnothing) = H_N$ and $e_N = o(g(N))$. □

This first-order deficit coefficient is retained as an application of the multicolour method; the companion capacity manuscripts [11, 12] give stronger undeleted height results. The packing contribution emphasized here concerns arbitrary fixed rational targets and deletions.

Corollary 13.2 (Croot's threshold and integer representations)**.** *Write $H_N = j+\delta$ with $j = \lfloor H_N \rfloor$, $0 \le \delta < 1$, and let $\mathcal{I}(N)$ be the set of positive integers representable as a sum of distinct unit fractions with denominators at most N. For every fixed $\epsilon > 0$ and sufficiently large N,*

$$\mathcal{I}(N) = \begin{cases} [1, j] \cap \mathbb{N}, & \delta > (1/2 + \epsilon)g(N), \\ [1, j-1] \cap \mathbb{N}, & \delta < (1/2 - \epsilon)g(N). \end{cases}$$

In the first case $k(N) = j$; in the second $k(N) = j - 1$. For all sufficiently large N, $\mathcal{I}(N)$ is one of these two initial intervals and $k(N)$ is one of these two integers.

Proof. The union of any j' of $k(N)$ disjoint unit-sum sets represents j', so $[1, k(N)] \cap \mathbb{N} \subseteq \mathcal{I}(N)$. The trivial harmonic bound excludes integers above j, and (12.1) excludes j when $\delta < (1/2-\epsilon)g(N)$. Apply (13.1), noting $e_N = o(g(N))$ and $g(N) \to 0$. □

This is the leading-coefficient threshold formulated in [6]. It leaves the narrower transition window unresolved. Erdős Problems 308 and 309 concern single-integer representations, whereas 296 requires disjointness; their original first-order questions already had answers [8, 9, 10]. These conclusions are retained here as consequences of the robust packing proof; stronger undeleted threshold results are in [11, 12].

Corollary 13.3 (Least height)**.** *Let $M_j = \min\{N : k(N) \ge j\}$ and let m_j be the least height of a representation of the integer j without a disjointness requirement. As $j \to \infty$, both M_j and m_j satisfy*

$$M_j = e^{j-\gamma}\left(1 + \frac{(\log j)^2}{2j} + O\left(\frac{\log j \, \log\log j}{j}\right)\right), \tag{13.2}$$

with M_j replaced by m_j for the second assertion.

Proof. Use $H_N = \log N + \gamma + O(1/N)$ and monotonicity. At integer heights obtained by rounding $\exp(j - \gamma + \frac{1}{2}(\log j)^2/j \pm A \log j \log\log j/j)$, with fixed A sufficiently large, the two arguments of the floors in (13.1) place the lower height strictly below, and the upper height strictly above, the threshold for j sets. To see this, substitute $L = j + O(1)$ into $g(N)$ and e_N at these two heights; $g(N) = (\log j)^2/j + O((\log j)^2/j^2)$ and $e_N \asymp \log j \log\log j/j$. Rounding the heights contributes exponentially small errors. Exponentiating the resulting bounds gives (13.2). For m_j, the same necessary bound follows from (12.1) for integer targets, and $m_j \le M_j$ by taking a union. □

Corollary 13.4 (Exceptional integers)**.** *Let $E(X) = \#\{N \le X : k(N) < \lfloor H_N \rfloor\}$. Then $E(X) \asymp Xg(X)$, and more precisely*

$$\liminf_{X\to\infty} \frac{E(X)}{Xg(X)} = \frac{1}{2(e-1)}, \qquad \limsup_{X\to\infty} \frac{E(X)}{Xg(X)} = \frac{e}{2(e-1)}.$$

In particular, $k(N) = \lfloor H_N \rfloor$ for a set of integers of natural density one.

Proof. Set $a_j = \min\{N : H_N \geq j\} = e^{j-\gamma} + O(1)$. For sufficiently large j, (13.1) implies $a_j \leq M_j < a_{j+1}$. Monotonicity of k shows the exceptional integers in $[a_j, a_{j+1})$ are exactly $[a_j, M_j) \cap \mathbb{N}$. By (13.2) its length ℓ_j satisfies

$$\ell_j = M_j - a_j \sim \tfrac{1}{2} a_j g(a_j).$$

As $a_{j-1}/a_j \to e^{-1}$ and $g(a_{j-1})/g(a_j) \to 1$, summing previous complete intervals gives

$$\sum_{i<j} \ell_i \sim \frac{a_j g(a_j)}{2(e-1)}.$$

For example this follows by Stolz–Cesàro on the increasing denominator $a_j g(a_j)$, since its ratio at consecutive indices tends to e. During $[a_j, M_j)$ the normalizing factor $Xg(X)$ equals $(1 + o(1))a_j g(a_j)$ uniformly, while the numerator increases by ℓ_j. From M_j to $a_{j+1} - 1$ the numerator is constant and $Xg(X)$ increases for large X. Thus the two extrema in the limit occur at the start and end of the exceptional intervals, yielding the claimed constants. The density-one statement follows because $g(X) \to 0$. □

14 Lattice structure and scope

For a finite set P of at least two distinct denominators put $Q = \operatorname{lcm}(P)$, $a_n = Q/n$, and $g = \gcd_{n \in P}(a_n - a_{n_0}) > 0$. The increments $(a_n, 1)$ generate exactly

$$\Lambda = \{(v, k) \in \mathbb{Z}^2 : v \equiv a_{n_0} k \pmod{g}\}.$$

Indeed their differences generate $(g, 0)$ by Bezout, and $(a_{n_0}, 1)$ supplies the second generator. In the coordinates (Z, T) the lattice has covolume g/Q. A local limit on this lattice would therefore have prefactor $g/(2\pi Q\sqrt{\det \Sigma})$, provided quantitative aperiodicity is proved. Algebraic generation alone is insufficient: take many odd denominators including 1, and a single even denominator 2. Then Q has valuation one at 2 and Q/n is even for all odd denominators; its parity is determined by the one Bernoulli variable for 2, even though the generated lattice is full in this example: g divides $a_1 - a_2 = Q/2 = a_2$, hence divides every a_n, so $g = 1$. For instance, taking that Bernoulli probability to be $1/3$ precludes uniform parity mixing. The present proof uses many surviving pairs, not just an lcm condition. This distinction is relevant to restricted denominator families.

The protected upper-half reservoir is essential; an arbitrary dense set need not satisfy the Fourier bounds. The blocking estimate here concerns fixed t, whereas the local limit theorem permits polylogarithmically many colours. The quantitative entropy estimates concern fixed targets (and, when imposed, fixed positive length fractions in the strict feasible region); they do not give a sharp second-order coefficient or a uniform growing-target entropy theorem. In particular the critical-support and growing-number-of-unit-target results of [12] are different from the present counting statements. The proofs of the robust packing bounds also leave a smaller-order gap at the second-scale cutoff. The manuscript has undergone internal mathematical review. The strengthened central-frequency argument has not been machine-formalized. Priority beyond the literature cited here requires further checking.

References

[1] 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), 4563–4589. doi:10.4171/JEMS/1456; arXiv:2112.03726.

[2] 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 Anal.* 2025:28, 13 pp. doi:10.19086/da.154329; arXiv:2404.16016.

[3] E. S. Croot III, On unit fractions with denominators in short intervals, *Acta Arith.* **99** (2001), 99–114. doi:10.4064/aa99-2-1; arXiv:math/9904181.

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

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

[6] E. S. Croot III, On some questions of Erdős and Graham about Egyptian fractions, *Mathematika* **46** (1999), 359–372. doi:10.1112/S0025579300007828. Author's manuscript: `https://ecroot.math.gatech.edu/erdos_graham.pdf`.

[7] E. S. Croot III, On a coloring conjecture about unit fractions, *Ann. of Math.* (2) **157** (2003), 545–556. doi:10.4007/annals.2003.157.545.

[8] T. F. Bloom, Erdős Problem #296, `https://www.erdosproblems.com/296`.

[9] T. F. Bloom, Erdős Problem #308, `https://www.erdosproblems.com/308`.

[10] T. F. Bloom, Erdős Problem #309, `https://www.erdosproblems.com/309`.

[11] P. Chen, *Sharp Two-Term Capacity Asymptotics for Egyptian Fractions via Rational Grid Representations*, Hexagon preprint 2611.01225v1 (2026), `https://hexagonmath.org/2611.01225v1`.

[12] P. Chen, *Sharp Second-Order Entropy Asymptotics for Disjoint Egyptian Fraction Representations*, Hexagon preprint 2611.01226v1 (2026), `https://hexagonmath.org/2611.01226v1`.

[13] D. Dolgopyat, A local limit theorem for sums of independent random vectors, *Electron. J. Probab.* **21** (2016), no. 39, 1–15.

[14] R. Rado, An inequality, *J. London Math. Soc.* **27** (1952), 1–6. doi:10.1112/jlms/s1-27.1.1.