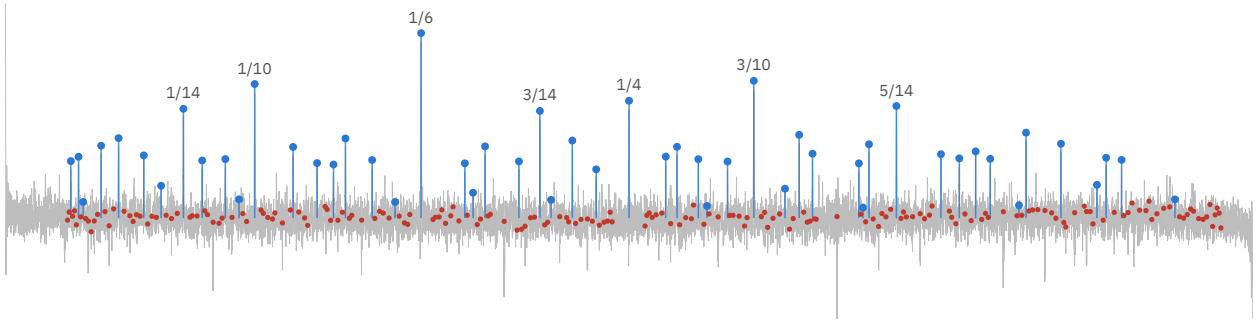


hexagon:2611.00068v2

Arcs of $\Pi(1+z^p)$

Prime races, the Gaussian core, Dirichlet L-functions and a certified solver for Subset Sum

Antonio Mennillo

Naples, Italy · ORCID 0009-0001-5978-6015

ABSTRACT

The subsets of $\{1\} \cup \{p \le N\}$ reach every integer between 0 and their total, by Bertrand's postulate, so deciding whether a target is a subset sum of the primes is trivial. Counting the solutions is not: the count carries the arithmetic of the primes. We study the generating function $F_N(z) = \Pi(1+z^w)$ on the unit circle and follow one object, the arc at a rational a/q, through three settings.

For the primes, the centre of the arc a/q rises above the sea of the minor arcs exactly when $q = 2^k (k \ge 2)$ or $q = 2p^j$, with excess $\Lambda(q/2)/\varphi(q)$, and the neighbourhoods of all the other arcs are bounded uniformly in q; Theorem E gives the expansion arc by arc. The height of an arc is a race of the primes modulo q with weights $\log|\cos(\pi ra/q)|$. By Landau's theorem a one-sided bound on it is equivalent to GRH for the even characters it sees, unconditionally at $q = 8$, 10 and 14, where the zero-free segments it needs are certified; at $q = 14$ the race changes sign infinitely often. The principal arc carries ζ in the centre and width of its Gaussian.

For arbitrary weights, the arcs become the decision rule of a Subset Sum solver that reads them on the original weights and ends with a meet-in-the-middle filtered by a random residue class. Every answer it returns is verified. The dummy weights of HURS explain why the reading is exact: they blur the counts by a flat window, and the blur is invertible on counts, though not on reachable sets. On random weights of b bits the method has rigorous guarantees, in time $2^{b/2+o(b)}$ and, with the filter, $2^{b/3+o(b)}$, resting on a local limit theorem uniform in

the size of the weights and on a concentration bound for signed subset sums. Finally, the solver is read inside a general theory of constraint propagation, in which every network factors as $T = T_{\text{comp}} \circ T_{\text{struct}}$.

---

1 Introduction

Two questions can be asked about the subset sums of a set of positive integers: whether a target is reached, and how many times. For the primes the answers live at very different depths.

The floor. Let $W_N = \{1\} \cup \{p \le N\}$. Ordered increasingly, each weight is at most one more than the sum of the previous ones, because $p_{k+1} < 2p_k$ (Bertrand's postulate). A set with this property is complete: its subset sums are all the integers from 0 to its total (§2). For the primes the decision problem is therefore trivial, and it rests on the prime number theorem at its weakest.

The ceiling. The number $r_N(T)$ of subsets with sum T is the coefficient of z^T in $F_N(z) = \prod_{w \in W_N}(1 + z^w)$. Near $z = 1$ the product gives a Gaussian; this is the local limit theorem, and in the bulk it already determines $r_N(T)$ to great accuracy (§3.6). Away from $z = 1$ the product is small, but its size along the arcs at the rationals a/q is not random. Each factor has modulus $2|\cos(\pi wa/q)|$, so $\log|F_N(e(a/q))|$ is a sum over the primes of a function of p mod q: a prime race. The question of this paper is what the arcs know about the primes, and what they can do when the weights are not primes.

The thread. The paper follows one object, the arc, through three settings.

- In §2–3 the weights are the primes. We determine which arcs rise above the sea of the minor arcs (Lemma 1), bound the neighbourhoods of all arcs uniformly in q (Lemmas P and Q), and state the arc expansion arc by arc (Theorem E).
- In §4 we read the height of an arc as a prime race. Under two conditions on the characters it sees, (R) and (N), a one-sided bound on it is equivalent to GRH for them (Proposition 3); both conditions are verified for $q = 8$, 10, 14. The principal arc carries ζ (Proposition 3_0). They are reformulations of the Knapowski–Turán type, in which the hypothesis becomes a statement about counting subsets.
- In §5–6 the weights are arbitrary. The floor fails, the dummies of HURS restore it, and the arcs, once the dummies are removed exactly, which makes them unnecessary in the solver, become the decision rule of a Subset Sum solver. We give its cost law, its benchmarks, and the rigorous form of its guarantees on random and on dense instances.
- The arc, seen from far enough, is propagation on a quotient. In §7 we develop the general theory of constraint propagation behind the algorithm: a structural part (equalities with parity, functional definitions on a complete DAG) solved in one pass, a computational part iterated to a fixpoint, the convergence of the full operator, and definitional variables that make implicit links evaluable. In §8 we read the solver of §5 inside it.

Why it matters. Prime races and Subset Sum have so far been separate subjects. The first belongs to analytic number theory (Chebyshev's bias, Knapowski–Turán, Rubinstein–Sarnak); the second is a classical NP-complete problem whose algorithms are combinatorial. The arcs join them. Counting subsets of the primes is a circle-method problem in which every arc is a prime race, so a hypothesis about the zeros of Dirichlet

L-functions becomes a statement about how many subsets of primes have a given sum, and the selection rule $q = 2^k, 2p^j$ with excess $\Lambda(q/2)/\varphi(q)$ is read off a cyclotomic value. In the other direction, the same arcs, read on arbitrary weights, are a working decision rule: they find the congruence obstructions that the structure theorem of dense Subset Sum predicts, and they steer a solver on instances where a dynamic-programming table is out of reach. The bridge is a dictionary, and both sides of it are proved.

§9 places the results in the literature and states their scope.

Notation. $e(x) = e^{2\pi i x}$; Λ is von Mangoldt's function; φ is Euler's function; Φ_q is the q-th cyclotomic polynomial; $\pi(N,\chi) = \sum_{p\le N}\chi(p)$; $\operatorname{li} x = \int_2^x dt/\log t + \operatorname{li} 2$. Constants C are absolute unless indexed.

Map of the results.

Result	What it says	Where	Inputs
Lemma 2.1, Proposition 2.3	the primes with 1 are complete; without 1 they miss six values	§2	—
Lemma 1	the arcs that emerge, $q = 2^k, 2p^j$, with excess $\Lambda(q/2)/\varphi(q)$	§3.2	—
Lemmas P, Q, Corollary Q′	the lobes, bounded uniformly in q; the highest is at $q = 3$	§3.3	Arb certificate (Q')
Propositions C, C′	the Edgeworth expansion of the core of an arc	§3.4	—
Lemmas M, L_0, L, W; Theorem E	the arc expansion: minor, non-emerging and emerging arcs	§3.5	Siegel–Walfisz, Vinogradov–Vaughan, Jackson
Proposition 3	a one-sided bound on an arc is equivalent to GRH for its characters	§4.2	(R) by Arb, (N)
Theorem S	no last sign change of the race at $q = 14$	§4.3	(R) by Arb, a zero on the line
Proposition 3_0	the principal arc carries ζ	§4.4	von Koch (one direction)
Lemmas 5.1, 5.3, blur identity	the solver and the exact de-blurring	§5	—
Lemmas A, B0′, B, B‴; Propositions B′, B⁗	the centre rule; a local limit theorem uniform in M; the endgame is non-empty	§6.1–6.3	[BR, Thm 22.1] (B'); B'''' computed
Theorems R, R′	success with probability $\ge 1-\delta$ in time $2^{b/2+o(b)}$	§6.4	—
Lemma C, Corollaries C′, C″	an exact decider on dense instances	§6.5	Bringmann–Wellnitz
Theorem LV	the endgame count has an exponential tail: Las Vegas	§6.6	certified $\rho \le 0.44495$
Lemmas J, K, V, Proposition K′, Theorem R″	the filtered endgame, time $2^{b/3+o(b)}$	§6.7	Rosser–Schoenfeld
Theorems 7.2–7.13	$T = T_{\text{comp}} \circ T_{\text{struct}}$, the information limit, convergence	§7	—
Propositions 8.1–8.5	the solver as propagation; the blur is invertible on counts	§8	—

2 The floor: completeness and linear search

Define the *omicron* of a set of positive integers $w_1 \le w_2 \le \cdots \le w_n$ as $o = \max_k w_{k+1}/(1+\sum_{i\le k} w_i)$.

Lemma 2.1 (gap filling; Brown). Let $w_1 \le w_2 \le \cdots \le w_n$ be positive integers and $S_k = \sum_{i\le k} w_i$, $S_0 = 0$. Every integer $0 \le T \le S_n$ is a subset sum if and only if $w_{k+1} \le 1 + S_k$ for every $0 \le k < n$, that is, if and only if $o \le 1$ [Brown]. Since $o \ge w_1 \ge 1$, the condition reads $o = 1$. For W_N it holds by Bertrand's postulate.

Proof. If the condition holds, the subset sums of the first k weights fill $[0, S_k]$ by induction, and adding $w_{k+1} \le S_k + 1$ fills $[0, S_{k+1}]$. Conversely, if $w_{k+1} > 1 + S_k$ for some k, the integer $S_k + 1 \le S_n$ is not a subset sum: the subsets of the first k weights sum to at most S_k, and every other subset contains a weight $w_j \ge w_{k+1} > S_k + 1$.□

Linear search. Scan the weights downwards and take w_k whenever $w_k \le T_{\text{rem}}$. Before w_k is examined, $T_{\text{rem}} \le \sum_{i\le k} w_i$. If $T_{\text{rem}} \ge w_k$, it is taken and $T_{\text{rem}} - w_k \le \sum_{i<k} w_i$. Otherwise $T_{\text{rem}} \le w_k - 1 \le \sum_{i<k} w_i$ by completeness. The scan ends at $T_{\text{rem}} = 0$. Decision is always "yes", and search costs $O(n)$ after the sieve. Two more facts complete the floor: $r_N(T) \ge 1$ for $0 \le T \le S_N$, and $r_N(T) = r_N(S_N - T)$ by complementation.

Remark 2.2 (where the primes saturate, and the grid step). In $W_N = \{1, 2, 3, 5, \dots\}$ the ratio of Lemma 2.1 is $1/1$ at the weight 1 and $2/(1+1) = 1$ at the weight 2, so $o(W_N) = 1$ is attained there. At the weight p_{m+1}, $m \ge 1$, the weights before it sum to S_{p_m}, and the ratio is

$$r_m = \frac{p_{m+1}}{1 + S_{p_m}} < 1.$$

Indeed $p_m \le 2 + \sum_{p<p_m} p$ holds for $m = 1$ and passes from m to $m+1$ by Bertrand's postulate, $p_{m+1} < 2p_m \le 2 + \sum_{p\le p_m} p$; hence $p_{m+1} < 2p_m \le 1 + S_{p_m}$. Above the weight 2 the primes are complete with a margin: $S_{p_m} \sim p_m^2/(2\log p_m) \sim m p_m/2$ gives $r_m = (2 + o(1))/m$. The same ratio is the step of the logarithmic grid $\log S_{p_m}$ of the cumulative prime sums [CPS], exactly:

$$\Delta_m = \log S_{p_{m+1}} - \log S_{p_m} = \log\Big(1 + \frac{p_{m+1}}{S_{p_m}}\Big) = \log\Big(1 + r_m\Big(1 + \frac{1}{S_{p_m}}\Big)\Big),$$

so completeness, $r_m \le 1$, reads $\Delta_m \le \log(2 + 1/S_{p_m})$ on the grid, and $\Delta_m = (2 + o(1))/m$.

Proposition 2.3 (the primes without the weight 1). Let $P_N = \{p \le N\}$ and $\Sigma_N = \sum_{p\le N} p = S_N - 1$. For $N \ge 5$ the subset sums of P_N are exactly

$$[0, \Sigma_N] \setminus \{1, 4, 6, \Sigma_N - 6, \Sigma_N - 4, \Sigma_N - 1\};$$

for $N = 2$ and $N = 3$ the missing values in $[0, \Sigma_N]$ are $\{1\}$ and $\{1, 4\}$. For $N \ge 7$ the reachable set has 7 maximal intervals: $\{0\}$, $[2, 3]$, $\{5\}$, $[7, \Sigma_N - 7]$, $\{\Sigma_N - 5\}$, $[\Sigma_N - 3, \Sigma_N - 2]$, $\{\Sigma_N\}$.

Proof. For $N = 5$, 7, 11 the statement is checked directly ($\Sigma = 10, 17, 28$). Write p_n for the n-th prime, $\Sigma_n = \sum_{k\le n} p_k$, $H = \{1, 4, 6\}$, and let R_n be the subset sums of $\{p_1, \dots, p_n\}$. Suppose $R_n = [0, \Sigma_n] \setminus (H \cup (\Sigma_n - H))$ and $7 \le p = p_{n+1} \le \Sigma_n - 13$. Then $R_{n+1} = R_n \cup (R_n + p)$, and:

- below 7 only R_n contributes, since $R_n + p \geq p \geq 7$, so 1, 4, 6 stay missing and 0, 2, 3, 5 stay reached;
- $[7, \Sigma_n - 7] \subset R_n$;
- for $j \in H$, $\Sigma_n - j - p$ lies in $[7, \Sigma_n - 7]$, because $7 - j \leq p \leq \Sigma_n - j - 7$; so $\Sigma_n - j \in R_n + p$, and the top holes of R_n are filled;
- for $\Sigma_n < x \leq \Sigma_n + p$, $x \in R_{n+1}$ if and only if $y = x - p \in R_n$; here $y > \Sigma_n - p \geq 13$, so $y \notin H$, and $y \notin R_n$ exactly when $y \in \Sigma_n - H$, that is $x \in \Sigma_{n+1} - H$.

Hence $R_{n+1} = [0, \Sigma_{n+1}] \setminus (H \cup (\Sigma_{n+1} - H))$. The hypothesis on p holds for every $n \geq 5$. Put $D_n = \Sigma_{n-1} - p_n$. By Bertrand's postulate $D_{n+1} = \Sigma_n - p_{n+1} > \Sigma_n - 2p_n = D_n$, so D is increasing, and $D_6 = \Sigma_5 - p_6 = 28 - 13 = 15 \geq 13$; the lower bound is $p_{n+1} \geq 13$. The induction therefore runs from $N = 11$ to every N. □

The proposition was also checked directly for every $N < 2 \cdot 10^4$. It gives Richert's theorem [Richert], that every integer greater than 6 is a sum of distinct primes, by taking N with $\Sigma_N \geq m + 7$; and it shows what the weight 1 does in W_N: the six missing values are 1 and its translates $4 = 3 + 1$, $6 = 5 + 1$ and their complements, so the floor of the primes needs the 1 exactly to close these six holes.

Remark 2.4 (interval representation). The reachable sets obey $R_{k+1} = R_k \cup (R_k + w_{k+1})$. The subsets of $\mathbb{Z}$, with union as addition and the Minkowski sum $A + B$ as multiplication, form a commutative idempotent semiring, the Boolean group semiring $\mathbb{B}[\mathbb{Z}]$, with zero $\emptyset$ and unit $\{0\}$; translation by w is multiplication by $\{w\}$, and $R_n = \prod_k (\{0\} \oplus \{w_k\})$ is the Boolean shadow of $F_W(z) = \prod_k (1 + z^{w_k})$. The reachable sets form a semimodule over this semiring; the structure is not tropical, since $\oplus$ is union and not a minimum. Storing R_k as its K_k maximal intervals, a step costs $O(K_k)$, with $K_k \leq \min(2^k, S_k + 1)$. $K = 1$ exactly for complete weights (Lemma 2.1), $K = 7$ for the primes alone with $N \geq 7$ (Proposition 2.3), and $K = 1$ for the dummy-augmented sets of §5.4 by construction. The last case says nothing about the original weights, which is why the dummies must be removed exactly.

The floor is blind to the zeros of ζ. A zero off the critical line would shift the gaps between primes by far less than Bertrand's margin, and completeness would survive. What the zeros change is the count, and to see them one has to look at F_N on the circle.

3 The arcs of the prime product

3.1 Factors that vanish, and the Fourier coefficients of the count

For a rational a/q with $(a, q) = 1$ and $q \geq 3$, the factors of $F_N(e(a/q))$ are $1 + e(wa/q)$, of modulus $2|\cos(\pi wa/q)|$. A factor vanishes exactly when $w \equiv q/2 \pmod{q}$; for $q \geq 3$ and $w \in W_N$ this forces q even and $w = q/2$ (an odd multiple of $q/2$ larger than $q/2$ is not a prime). So it happens for $q = 4 (w = 2)$ and $q = 2p (w = p)$. Let $\mathscr{V}(a/q)$ be the set of weights whose factor vanishes (at most one weight) and $F_N^*(a/q)$ the product of the other factors. The arc at $q = 2$, where every odd weight vanishes, is excluded throughout.

Remark 3.1 (the amplitudes are Fourier coefficients of the count). Since F_N is a polynomial,

$$F_N(e(a/q)) = \sum_{T=0}^{S_N} r_N(T)\, e(Ta/q)$$

exactly.

- If $\mathscr{V}(a/q)=\emptyset$, the amplitude $F_N^*(a/q)=F_N(e(a/q))$ is the finite Fourier coefficient of the count at frequency a/q.
- If $\mathscr{V}(a/q)=\{w\}$ with $w=q/2$, then $F_N(e(a/q))=0$, and the amplitude is a coefficient of the first moment. Differentiating $F_N=(1+z^w)F_N^*$ at $\zeta=e(a/q)$, where $\zeta^w=-1$, gives $\zeta F_N'(\zeta)=-wF_N^*(\zeta)$, that is

$$F_N^*(a/q)=-\frac{1}{w}\sum_T T\,r_N(T)\,e(Ta/q).$$

Equivalently, $F_N^*(a/q)$ is the Fourier coefficient of the count of $W_N\setminus\{w\}$.

So every arc amplitude is an exact Fourier coefficient: of $r_N(T)$ when no factor vanishes, of $T\cdot r_N(T)$ when one does. Grouping T by its class $\bmod q$, the coefficient is a discrete Fourier transform $\bmod q$ of the residue distribution of the prime subset sums, or of its first moment. Proposition 3 below is a statement about these coefficients and does not depend on the arc expansion of §3.5.

3.2 Which arcs emerge

Lemma 1. For $q\ge 3$ and $(a,q)=1$ let $\bar w(q)=\varphi(q)^{-1}\sum_{(r,q)=1}\log|\cos(\pi ra/q)|$. Then $\bar w(q)$ does not depend on a, and

$$\bar w(q)=\begin{cases}-\log 2, & q \text{ odd},\\ -\log 2+\Lambda(q/2)/\varphi(q), & q \text{ even}.\end{cases}$$

Proof. Multiplication by a permutes the residues prime to q. Since $|1+e(r/q)|=2|\cos(\pi r/q)|$, $\sum_{(r,q)=1}\log|2\cos(\pi r/q)|=\log|\Pi(1+\zeta_q^r)|=\log|\Phi_q(-1)|$. For $q>2$, $\Phi_q(-1)=p$ if $q=2p^j$ with p prime ($p=2$ allowed) and $\Phi_q(-1)=1$ otherwise [BHM, Lemma 7], so $\log|\Phi_q(-1)|=\Lambda(q/2)$ for q even and 0 for q odd. □

The lemma rests on classical facts about cyclotomic values; the reading as a level of the arcs is ours. It was checked against direct computation for every $3\le q<400$.

The sea. For almost every θ, $(1/\pi(N))\sum_{p\le N}\log|\cos(\pi p\theta)|\to\int_0^1\log|\cos\pi x|dx=-\log 2$, by the equidistribution of $p\theta \bmod 1$ [Vinogradov] together with a Borel–Cantelli control of the logarithmic singularity. This is the level of the minor arcs. By Lemma 1 the centre of the arc a/q rises above it, $\bar w(q)>-\log 2$, if and only if $q=2^k$ with $k\ge 2$ or $q=2p^j$ with p an odd prime. We call these q, together with $q=1$, the *emerging moduli*, $\mathfrak{E}$. The von Mangoldt function of $q/2$ decides which arcs emerge.

Up to $O_q(1)$ from the weights 1 and $p|q$,

$$\log|F_N^*(a/q)|=\pi(N)\big(\log 2+\bar w(q)\big)+H_{a/q}(N)+O_q(1),\qquad H_{a/q}(N)=\sum_{p\le N,\ p\nmid q}c_{p \bmod q},$$

with $c_r=\log|\cos(\pi ra/q)|-\bar w(q)$. The *height* $H_{a/q}$ is a race of the primes modulo q with weights c_r, and $\sum_r c_r=0$.

Figure 1 shows the levels at $N = 2000$: the arcs with $q \in \mathfrak{E}$ rise to the level of Lemma 1, and all the other arcs stay at the level of the sea.

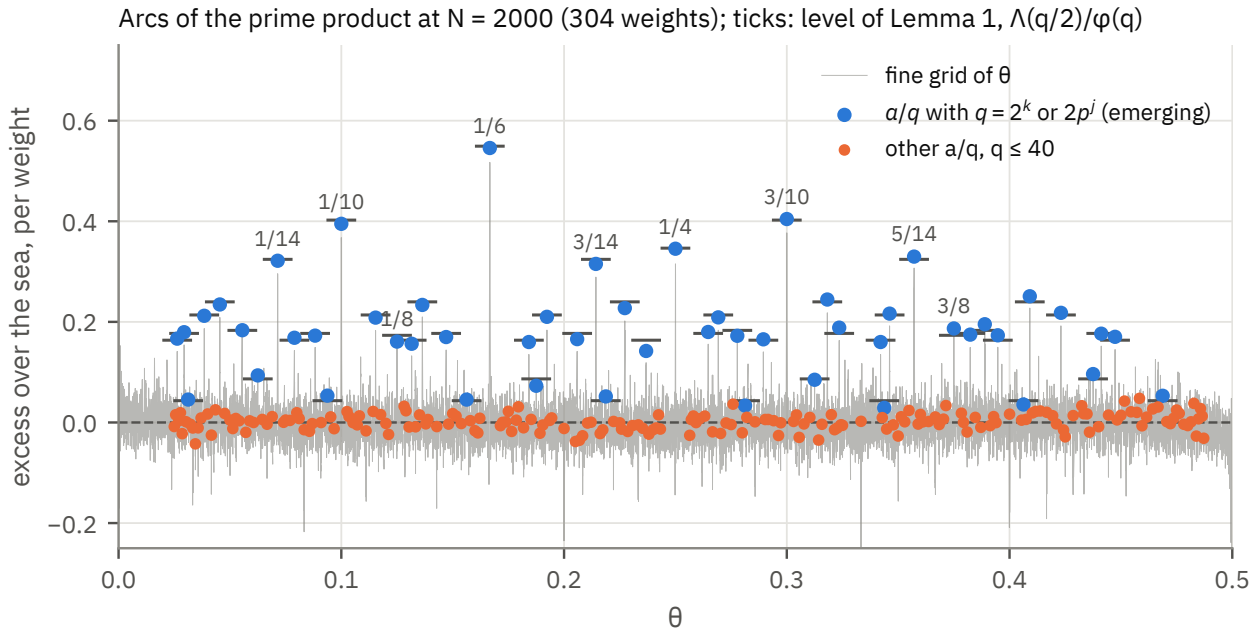


Figure 1. The arcs of the prime product at $N = 2000$ (304 weights). Grey: the per-weight excess over the sea, $(1/|W_N|)\sum_w \log|2\cos(\pi w\theta)|$, on a grid of $\theta \in (0, \frac{1}{2}]$; it fluctuates around 0. Dots: the same quantity at $\theta = a/q$, $q \le 40$, with the vanishing factor removed. The arcs with $q = 2^k$ or $2p^j$ (blue) rise to the level $\Lambda(q/2)/\varphi(q)$ of Lemma 1 (ticks); the other arcs (orange) stay at the level of the sea. The largest deviation from Lemma 1 is 0.048, the term $O_q(1)$ and the height divided by 304.

3.3 Neighbourhoods of the arcs: the lobes

Near an arc the level changes. For $q \ge 1$ and $X \ge 0$ define the neighbourhood level

$$A_q(X) = \frac{1}{X}\int_0^X \log\left|\Phi_q(-e(x))\right| dx \qquad (A_q(0) = \log|\Phi_q(-1)|).$$

By the identity $\prod_{(r,q)=1}(1+\zeta_q^r y) = \Phi_q(-y)$, $A_q(X)/\varphi(q)$ is the average excess over the sea, per prime, of $\sum_p \log|2\cos(\pi(pa/q + p\beta))|$ at $\beta = X/N$ (Lemma L below). The *core level* is $c_q = A_q(0)/\varphi(q)$: $\log 2$ for $q = 1$, $\Lambda(q/2)/\varphi(q)$ for q even, 0 for q odd ≥ 3. The *lobe level* is $\ell_q = \sup\{A_q(X)/\varphi(q) : X \ge X_q\}$, where X_q is the first local minimum of A_q, when A_q has one. We also use $\ell_q^* = \sup\{A_q(X)/\varphi(q) : X \ge 1/(2q)\}$, which needs no X_q.

q	emerging	core c_q	lobe ℓ_q	at X
1	yes	0.6931	0.1223	1.31
3	no	0	**0.1316**	0.73
4	yes	0.3466	0.0612	0.65
6	yes	0.5493	0.0782	1.24
8	yes	0.1733	0.0306	0.33
10	yes	0.4024	0.0417	1.16
14	yes	0.3243	0.0265	1.12

q	emerging	core c_q	lobe ℓ_q	at X
15	no	0	0.1088	0.20
32	yes	0.0433	0.0056	0.08

Table 1. Core and lobe levels (exact integrals of $\log|\Phi_q|$).

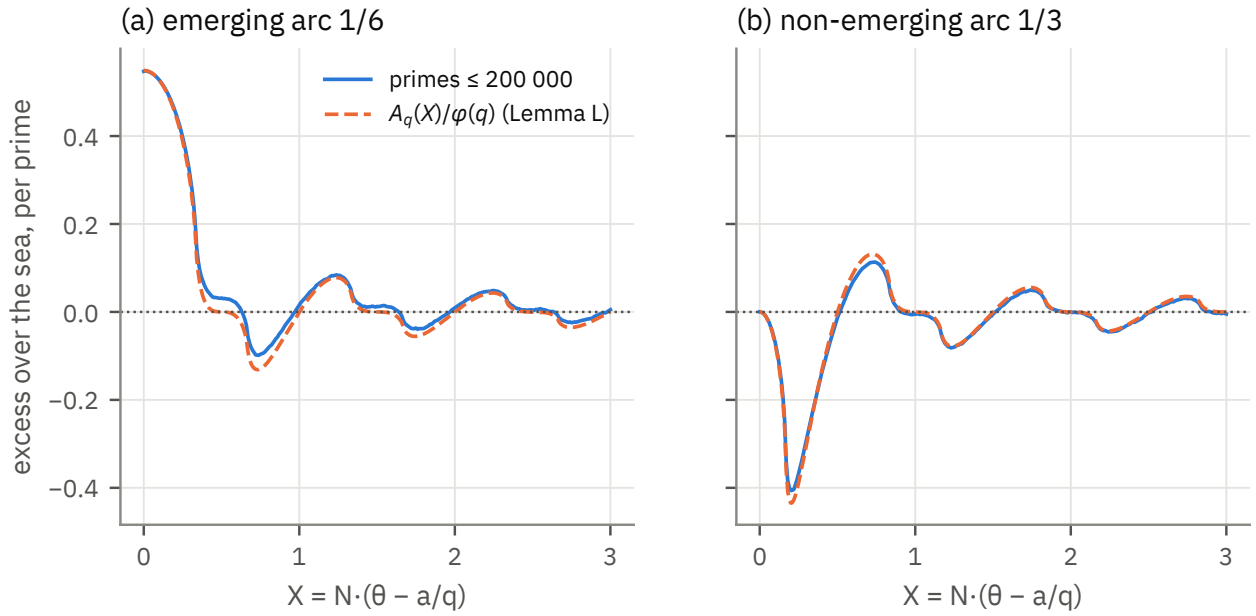


Figure 2. The neighbourhood of an arc. Blue: the per-prime excess at $\theta = a/q + X/N$, for the primes up to $N = 2 \cdot 10^5$. Dashed: the neighbourhood level $A_q(X)/\varphi(q)$, the prediction of Lemma L (§3.5). (a) The emerging arc $1/6$, with core level 0.549. (b) The non-emerging arc $1/3$, with core level 0 and the largest lobe, $\ell_3 = 0.1316$ at $X \approx 0.73$ (Table 1, Corollary Q′).

Lemma P (maximum principle for emerging arcs). For every emerging $q \geq 4$, $x \mapsto |\Phi_q(-e(x))|$ attains its maximum $\Phi_q(-1) = e^{\Lambda(q/2)}$ at $x = 0$, and only on a finite set of points per period. Hence $A_q(X) < \Lambda(q/2)$ for $X > 0$, and $\ell_q < c_q$.

Proof. With $y = -e(x)$ and $\Phi_q(z) = \prod_{d|q}(z^d - 1)^{\mu(q/d)}$:

- $q = 2^k$, $k \geq 2$: $\Phi_q(z) = z^{q/2} + 1$ and $|\Phi_q(y)| = 2|\cos(\pi q x/2)| \leq 2$;
- $q = 2p^j$, p odd: only $d \in \{q, q/2, q/p, q/(2p)\}$ contribute, and $|\Phi_q(-e(x))| = |\sin p\theta / \sin\theta|$ with $\theta = \pi p^{j-1} x$, a Dirichlet kernel bounded by p.

In both cases $\log|\Phi_q(-e(x))| \leq \Lambda(q/2)$ with equality on a null set. By Jensen's formula its mean over a period is 0, so $A_q(X) \to 0$ and the supremum over $X \geq X_q$ is attained and strictly below $\Lambda(q/2)$. □

Lemma Q (lobes of every arc). Let $q \geq 3$, $h_q(x) = \log|\Phi_q(-e(x))|$ and $\kappa = \mathrm{Cl}_2(\pi/3)/\pi = 0.32307\ldots$

1. h_q is strictly decreasing on $[0, 1/(2q))$, so A_q is strictly decreasing on $(0, 1/(2q))$, $A_q(X) < \varphi(q)c_q$ for $0 < X < 1/(2q)$, and $X_q \geq 1/(2q)$.

2. For every $X \geq 1/(2q)$,

$$A_q(X) \leq \sum_{\mu(q/d)=1} \min\Big(\log 2, \frac{\kappa}{dX}\Big) + \sum_{\mu(q/d)=-1} \beta(dX), \qquad \beta(L) = \begin{cases} \max\Big(\kappa,\ \min\big(\kappa/L,\ \log\frac{e}{2L}\big)\Big) & L \leq 1, \\ \kappa/L & L > 1, \end{cases}$$

and the right side is non-increasing in X. Writing R_q for its value at $X = 1/(2q)$ divided by $\varphi(q)$,

$$\ell_q \leq R_q \leq \frac{2^{\omega(q)-1}(1+\log 2+\log \operatorname{rad} q)}{\varphi(q)} \leq \frac{\tau(q)\log(2eq)}{2\varphi(q)}.$$

3. For $X \geq 1$, $A_q(X)/\varphi(q) \leq \kappa\tau(q)/\varphi(q)$.

Proof.

- (1) Pairing r with $q-r$, $h_q(x) = \Sigma \log|c_r + 2\cos 2\pi x|$ over $\varphi(q)/2$ pairs, with $c_r = 2\cos(2\pi r/q) \geq -2\cos(\pi/q)$ because $|2r/q-1| \geq 1/q$. On $[0, 1/(2q))$ each term is positive inside the logarithm and strictly decreasing.
- (2) By Möbius inversion $h_q(x) = \sum_{d|q} \mu(q/d) G(dx + d/2)$ with $G(u) = \log|2\sin \pi u|$. Since $\int_0^u G = -\operatorname{Cl}_2(2\pi u)/(2\pi)$ and $|\operatorname{Cl}_2| \leq \operatorname{Cl}_2(\pi/3)$, the average of $\pm G(d\cdot + d/2)$ over $[0, X]$ has modulus $\leq \kappa/(dX)$; moreover $G \leq \log 2$. For $\mu = -1$ and $L = dX \leq 1$ the average of $-G$ over an interval of length L is largest when centred at 0, and with $2|\sin \pi s| \geq 4|s|$ it is $\leq \log(e/(2L))$. At $X = 1/(2q)$ the $2^{\omega(q)-1}$ terms with $\mu = +1$ give $\leq \log 2$ each, a term with $\mu = -1$ and $d = q/e$ gives $\leq 1 + \log e$, and $\sum_{\omega(e)\text{odd}} \log e \leq 2^{\omega(q)-1} \log \operatorname{rad} q$. The max with κ makes β non-increasing and continuous at $L = 1$ (without it β would jump from $1 - \log 2$ to κ there, since $\kappa > 1 - \log 2$); at $X = 1/(2q)$ every term with $\mu = -1$ has $L \leq 1/4$, where $\beta \geq \kappa$, so R_q is unchanged.
- (3) For $X \geq 1$ every term is $\leq \kappa$, and there are $2^{\omega(q)}$ of them. □

So $\ell_q \ll \tau(q)\log q/\varphi(q) = q^{-1+o(1)}$: the lobes of far arcs sink into the sea uniformly.

Corollary Q′ (the largest lobe). $\sup_{q \notin \mathfrak{E}} \ell_q = \ell_3$, with $\ell_3 \in [0.131593, 0.132469]$ (numerically $\ell_3 = 0.1316$).

- For $q > 20000$, $2^{\omega(q)}/\varphi(q) \leq M_\theta q^{-\theta}$ with $\theta = 0.70$ and $M_{0.70} = 15.14$, and the bound of Lemma Q(2), which is decreasing in q beyond 10^4, is $< \ell_3$.
- For $q \leq 20000$, R_q is evaluated in ball arithmetic (Arb, 128 bits); it is certified $< \ell_3$ except for 145 non-emerging moduli, all ≤ 2730, $q = 3$ among them.
- For each of the other 144 moduli, $\sup_{X \geq 1/(2q)} A_q(X)/\varphi(q)$ is certified $< \ell_3$; the largest upper bound is 0.1145, at $q = 15$. Since $X_q \geq 1/(2q)$ (Lemma Q(1)), this bounds ℓ_q. The method: $H(X) = \int_0^X \log|\Phi_q(-e(x))|dx = \sum_{d|q} \mu(q/d) d^{-1}[C(dX + d/2) - C(d/2)]$ with $C(u) = -\operatorname{Cl}_2(2\pi u)/(2\pi)$, evaluated at rational points; between points, $H(X) \leq H(a) + (X-a)\cdot S$, with S a certified bound for $\sup \log|\Phi_q(-e(x))|$ from the product form $\sum_{(r,q)=1} \log|2\cos\pi(r/q + x)|$,

each term maximised exactly on the interval; beyond X^* the tail $|H(X)| \le \kappa \sum_{d|q,\mu(q/d)\neq 0} 1/d$.

- ℓ_3 itself: the lower bound is $A_3(29/40)/2$, after a certified local minimum of A_3 in $(1/5, 21/100)$; the upper bound comes from the same interval method.
- The constant $M_{0.70}$ of the first item is certified ≤ 15.1373, attained at $q = 30030$ (only $p \le 13$ contribute), and the bound at $q = 20001$ is 0.0856.

The corollary is proved: the finite part is a certified computation in interval arithmetic (`certify_lobes_arb.py`, 39 *s*).

3.4 The core of an arc: Gaussian and Edgeworth

Proposition C (core expansion). Put $t = 2\pi\beta$. A non-vanishing factor $(1+ue^{iwt})/(1+u)$, $u = e(wa/q)$, is the characteristic function of $w \cdot \mathrm{Bernoulli}(p)$ with complex $p = u/(1+u)$, with cumulants $\kappa_1 = wp$, $\kappa_2 = w^2 p(1-p)$, $\kappa_3 = w^3 p(1-p)(1-2p)$, $\kappa_4 = w^4 p(1-p)(1-6p+6p^2)$. The vanishing factor ($w = q/2$) is $1 - e^{iwt} = w(-it)e^{iwt/2}\mathrm{sinc}(wt/2)$: the factor $w(-it)$ becomes $w \cdot d/dT$ after inversion, and the rest has cumulants $w/2$, $w^2/12$, 0, $-w^4/120$. Summing gives complex cumulants $\kappa_j(a/q)$, $s = \sqrt{\kappa_2}$, $z = (T-\kappa_1)/s$, and the arc term

$$\mathscr{A}_{a/q}(T) = e(-Ta/q)\,F_N^*(a/q)\Big(\prod_{w\in\mathscr{V}} w\frac{d}{dT}\Big)\Big[\frac{\varphi(z)}{s}\Big(1 + \frac{\kappa_3}{6s^3}He_3(z) + \frac{\kappa_4}{24s^4}He_4(z) + \frac{\kappa_3^2}{72s^6}He_6(z)\Big)\Big].$$

We tested it on the exact counts at $N = 500$ (96 weights, counts up to 2^{84}), using twisted Gaussian windows $\sum_T r_N(T)e(aT/q)e^{-(T-T_0)^2/2H^2}$ with $H = 60$ at three centres T_0:

Arc	Gaussian only	with Edgeworth
0/1	0.9947–0.9956	**0.99997–0.99998**
1/8	0.980–0.985	**0.9973–1.0019**
1/6	0.964–0.974	**0.9997–1.0001**
1/10	0.955–0.974	**0.9987–0.9991**
3/10	0.969–0.983	**0.9962–0.9970**
1/4	0.983–0.988	**0.9989–0.9991**
1/14	0.961–0.984	**0.9928–0.9944**

Table 2. *Ratio exact/predicted of the arc terms.* The phases agree to $\le 2\cdot 10^{-3}$ rad. The residual, at most 0.7%, is the next Edgeworth order (Proposition C′).

Proposition C′ (the next two orders). Let a non-vanishing factor have, besides $\kappa_1, \ldots, \kappa_4$, the cumulants $\kappa_5 = w^5 p(1-p)(1-2p)(1-12p+12p^2)$ and $\kappa_6 = w^6 p(1-p)(1-30p+150p^2-240p^3+120p^4)$, and the vanishing factor $\kappa_3 = \kappa_5 = 0$, $\kappa_6 = w^6/252$. The bracket of Proposition C continues with the terms of order s^{-3},

$$\frac{\kappa_5}{120s^5}He_5(z)+\frac{\kappa_3\kappa_4}{144s^7}He_7(z)+\frac{\kappa_3^3}{1296s^9}He_9(z),$$

and of order s^{-4},

$$\frac{\kappa_6}{720s^6}He_6(z)+\Big(\frac{\kappa_4^2}{1152}+\frac{\kappa_3\kappa_5}{720}\Big)\frac{He_8(z)}{s^8}+\frac{\kappa_3^2\kappa_4}{1728s^{10}}He_{10}(z)+\frac{\kappa_3^4}{31104s^{12}}He_{12}(z).$$

Proof. These are the terms of order s^{-3} and s^{-4} of the Edgeworth series of a sum of independent factors with the given cumulants. The cumulants of $w\cdot\mathrm{Bernoulli}(p)$ are the derivatives of $\log(1-p+pe^{iwt})$; those of $e^{iwt/2}\mathrm{sinc}(wt/2)$ are those of the uniform law on $[0,w]$, whose odd cumulants beyond the first vanish and whose sixth is $w^6B_6/6=w^6/252$. □

Arc	order 1 (Proposition C)	order 2	order 3
$0/1$	0.99997–0.99998	0.99997–0.99998	**1.000000**
$1/8$	0.99953–0.99956	0.99974–0.99978	**0.99995–0.99998**
$1/6$	0.99966–1.00008	1.00015–1.00022	**0.99998–1.00000**
$1/10$	0.99867–0.99909	0.99991–1.00010	**0.99996–0.99998**
$3/10$	0.99618–0.99703	1.00021–1.00034	**0.99967–0.99975**
$1/4$	0.99894–0.99909	1.00120–1.00138	**1.00003**
$1/14$	0.99281–0.99442	1.00038–1.00052	**0.99937–0.99953**

Table 2′. *The same test with $H=120$.* The largest residual falls from 0.72% to 0.14% and 0.063%, with alternating sign, as expected of an asymptotic series. At $H=60$ the arc $1/8$ shows a residual of about 0.2% that changes with T_0 (Table 2); it is leakage of the window from neighbouring frequencies, not an Edgeworth term, and it disappears at $H\ge 120$.

3.5 The arc expansion as a theorem

Fix $B>40$ and put $\mathfrak{P}=(\log N)^B$, $M=(\log N)^{B/4}$. The major arcs are $\mathfrak{M}(a/q)=\{\theta:|\theta-a/q|\le\mathfrak{P}/(qN)\}$, $q\le\mathfrak{P}$; the minor arcs $\mathfrak{m}$ are the rest. The statements below concern $q\ge 3$; the arcs at $q=1$ and $q=2$ are the main Gaussian and its parity image. Put $R_{a/q}(T)=\int_{\mathfrak{M}(a/q)}F_N(e(\theta))e(-T\theta)d\theta$, so that $r_N(T)=\sum_{q\le\mathfrak{P}}\sum_a R_{a/q}(T)+\int_{\mathfrak{m}}F_N(e(\theta))e(-T\theta)d\theta$ exactly.

Lemma M (one-sided trigonometric majorant). For every $M\ge 2$ there is a trigonometric polynomial K_M of degree $\le M$ with $K_M(x)\ge\log|\cos\pi x|$ for all x, $\hat K_M(0)\le-\log 2+CM^{-1/2}$, and $|\hat K_M(m)|\le C$.

Proof. Let $f=\log|\cos\pi x|$ and $f_\varepsilon=\max(f,\log\varepsilon)$: $f\le f_\varepsilon$, $\hat f_\varepsilon(0)=-\log 2+O(\varepsilon)$, and f_ε is Lipschitz with constant π/ε. With J_M the Jackson kernel, $K_M=J_M*f_\varepsilon+C\pi/(\varepsilon M)$ majorizes f_ε. Take $\varepsilon=M^{-1/2}$. Since $J_M\ge 0$ has mass 1, $|\hat K_M(m)|\le\int_0^1|f_\varepsilon|\le\log 2+O(\varepsilon)$ for $m\ne 0$. □

An L^1 approximation would not do: an upper bound is needed at every prime simultaneously, and the singularities of f are at $-\infty$, so a one-sided majorant costs nothing.

Lemma L (neighbourhood of an arc). Fix B. Uniformly for $3\le q\le\mathfrak{P}$, $(a,q)=1$ and $\theta=a/q+X/N$ with $N^{-1/2}\le|X|\le\mathfrak{P}/q$,

$$\log|F_N(e(\theta))| \le n\log 2 + \frac{N}{\log N}\Big(\frac{A_q(|X|)}{\varphi(q)} - \log 2\Big) + O_B\Big(\frac{N}{\log^2 N}\Big).$$

Only this upper bound is used below.

Proof. Write $\log|1+e(w\theta)| = \log 2 + g(w\theta)$, with $g(y) = \log|\cos \pi y| \le 0$. Since $|\Phi_q(-e(-x))| = |\Phi_q(-e(x))|$, we may take $X > 0$ and put $\beta = X/N$.

- The weight 1, the primes $p \le \sqrt{N}$ and the primes dividing q contribute ≤ 0 to $\sum_w g(w\theta)$; drop them and keep the primes $\sqrt{N} < p \le N$ with $p \nmid q$.
- If $p \equiv r \pmod q$, then $pa/q - ra/q$ is an integer, so $|\cos \pi p\theta| = |\cos \pi(ra/q + p\beta)|$. Put $g_r(x) = g(ra/q + x)$ and $g_{r,\varepsilon} = \max(g_r, \log \varepsilon)$ with $\varepsilon = 1/N$. Then $g_r \le g_{r,\varepsilon} \le 0$, and on $[0,X]$ the function $g_{r,\varepsilon}$ is continuous, with at most $2X+4$ monotone pieces, each of variation $\le \log N$.
- By Siegel–Walfisz, $\pi(t;q,r) = \mathrm{li}(t)/\varphi(q) + E_r(t)$ with $|E_r(t)| \ll_B Ne^{-c_B\sqrt{\log N}}$ for $\sqrt{N} \le t \le N$ and $q \le \mathfrak{P}$ (Siegel–Walfisz with exponent $2B$, since $\mathfrak{P} \le (\log t)^{2B}$ for $t \ge \sqrt{N}$ and N large). Stieltjes integration by parts, with $|\int f dE| \le (2\sup|f| + \mathrm{Var}\, f)\cdot \sup|E|$, gives

$$\sum_{\substack{\sqrt{N}<p\le N\\ p\equiv r}} g_{r,\varepsilon}(p\beta) = \frac{1}{\varphi(q)}\int_{\sqrt{N}}^{N} g_{r,\varepsilon}(t\beta)\frac{dt}{\log t} + O_B\big((X+3)N\log N\, e^{-c_B\sqrt{\log N}}\big).$$

- Since $g_{r,\varepsilon} \le 0$ and $\log t \le \log N$, the integral is at most

$$\frac{1}{\log N}\int_{\sqrt{N}}^{N} g_{r,\varepsilon}(t\beta)\,dt = \frac{N}{X\log N}\int_{X/\sqrt{N}}^{X} g_{r,\varepsilon}(x)\,dx \le \frac{N}{X\log N}\Big[\int_0^X g_r + \int_0^X (g_{r,\varepsilon} - g_r) + \frac{X\log N}{\sqrt{N}}\Big].$$

- On each period, $g_{r,\varepsilon} - g_r$ is supported where $|\cos| < \varepsilon$ and integrates to at most 2ε, so $\int_0^X (g_{r,\varepsilon} - g_r) \le 2(X+1)/N$.
- By $\prod_{(r,q)=1}(1+\zeta_q^{ra}y) = \Phi_q(-y)$ and $|1+e(y)| = 2|\cos \pi y|$, $\sum_{(r,q)=1} g_r(x) = \log|\Phi_q(-e(x))| - \varphi(q)\log 2$, so $\sum_r \int_0^X g_r = X(A_q(X) - \varphi(q)\log 2)$.
- Summing over the $\varphi(q)$ classes,

$$\sum_w g(w\theta) \le \frac{N}{\log N}\Big(\frac{A_q(X)}{\varphi(q)} - \log 2\Big) + O\Big(\frac{\sqrt{N}}{\log N}\Big) + O(\sqrt{N}) + O_B\big(\mathfrak{P}^2 N\log N\, e^{-c_B\sqrt{\log N}}\big),$$

using $X \ge N^{-1/2}$ in the second term and $X \le \mathfrak{P}/q$, $\varphi(q) \le \mathfrak{P}$ in the last. All three error terms are $O_B(N/\log^2 N)$. □

Lemma L_0 **(the core of an arc).** For $3 \le q \le \mathfrak{P}$, $(a,q) = 1$ and $\theta = a/q + \beta$ with $|\beta| \le N^{-3/2}$,

$$\log|F_N(e(\theta))| \le \log 2 + \log|F_N^*(a/q)| + \pi q\,\pi(N)N^{-1/2}.$$

Proof. For a weight $w \notin \mathscr{V}(a/q)$, $\|wa/q - 1/2\| \geq 1/(2q)$, so $|\cos(\pi wa/q)| \geq \sin(\pi/(2q)) \geq 1/q$. Then $|\cos\pi(wa/q + w\beta)| \leq |\cos(\pi wa/q)| + \pi w|\beta| \leq |\cos(\pi wa/q)|(1 + \pi qw|\beta|)$, and $\sum_w \pi qw|\beta| \leq \pi q\pi(N)N|\beta| \leq \pi q\pi(N)N^{-1/2}$. The vanishing factor, if any, has modulus ≤ 2. □

The peak of an arc. For $q \geq 3$ let $G(\theta)$ be $F_N(e(\theta))$ with the vanishing factor removed, if there is one, so that $G(a/q) = F_N^*(a/q)$. Put $t = 2\pi(\theta - a/q)$, $S_2 = \sum_{w \notin \mathscr{V}} w^2$, and

$$\nu_{a/q} = \frac{1}{2} \sum_{w \notin \mathscr{V}} w \tan\frac{\pi wa}{q}.$$

Lemma W (the peak). Fix B. For $3 \leq q \leq \mathfrak{P}$ and $(a,q) = 1$:

1. On $|t| < \pi/(qN)$ no factor of G vanishes, and $t \mapsto \log|G|$ is concave there, with second derivative $\leq -S_2/4$.
2. Its derivative at $t = 0$ is $-\nu_{a/q}$, and $\nu_{a/q} \ll_B N^2 e^{-c_B\sqrt{\log N}}$. Moreover $S_2 \geq N^3/(64\log N)$ for N large (the primes in $(N/2, N]$, by Siegel–Walfisz at $q = 1$). Hence $\log|G|$ has a unique maximum, at a point t^* with $|t^*| \leq 4|\nu_{a/q}|/S_2 \ll_B N^{-1} e^{-c_B\sqrt{\log N}}$.
3. At $\theta^* = a/q + t^*/(2\pi)$ the cumulants of Proposition C, taken with $u = e(w\theta^*)$, have κ_1^* real, $\kappa_2^* = \sum_{w \notin \mathscr{V}} w^2/(4\cos^2(\pi w\theta^*)) \geq S_2/4$, κ_3^* purely imaginary and κ_4^* real. Moreover $\log|F_N^*(a/q)| \leq \log|G(\theta^*)| \leq \log|F_N^*(a/q)| + 4\nu_{a/q}^2/S_2$.

Proof.

- (1) The factor w vanishes where $w\theta \equiv 1/2 \pmod 1$. For $w \notin \mathscr{V}$, $\|wa/q - 1/2\| \geq 1/(2q)$, so the nearest zero is at distance $\geq 1/(2qw) \geq 1/(2qN)$ from a/q. On that interval, $d^2/dt^2 \log|\cos(\pi w\theta)| = -w^2/(4\cos^2(\pi w\theta)) \leq -w^2/4$.
- (2) With $u = e(wa/q)$ one has $u/(1+u) = 1/2 + (i/2)\tan(\pi wa/q)$, and $d/dt \log|1 + ue^{iwt}|$ at $t = 0$ equals $-w \cdot \mathrm{Im}(u/(1+u)) = -(w/2)\tan(\pi wa/q)$. For ν, group the primes by class: $\sum_{p \equiv r, p \leq N} p = \varphi(q)^{-1} \sum_{p \leq N, p \nmid q} p + O_B(N^2 e^{-c_B\sqrt{\log N}})$ by Siegel–Walfisz and partial summation (applied with $(\log N)^{2B}$ for $t \geq \sqrt{N}$, the primes $p \leq \sqrt{N}$ contributing $O(N)$). Moreover $\sum_{(r,q)=1} \tan(\pi ra/q) = 0$ (pair r with $q - r$) and $|\tan(\pi ra/q)| \leq q$; the weight 1 and the primes dividing q contribute $O(q^2 \log q)$. By concavity, the maximum lies within $4|\nu|/S_2$ of 0, which is inside the interval of (1) for N large.
- (3) On $|u| = 1$, $p(1-p) = u/(1+u)^2 = 1/(4\cos^2(\pi w\theta))$ is real and positive, and $1 - 2p = -i\tan(\pi w\theta)$ is purely imaginary. This gives the types of κ_2^*, κ_3^*, κ_4^*. κ_1^* is real because $-\mathrm{Im}\,\kappa_1^*$ is the derivative of $\log|G|$ at t^*, which vanishes. The last inequality follows from concavity: $\log|G(\theta^*)| - \log|G(a/q)| \leq |\nu||t^*| \leq 4\nu^2/S_2$. □

The offset t^* is a prime race: $\nu_{a/q}$ weights the classes $r \bmod q$ by the odd function $\tan(\pi ra/q)$, so only the odd characters $\bmod q$ enter it. The height of an arc sees the even characters (§4), its offset the odd ones. Under GRH for the odd characters $\bmod q$, $\nu_{a/q} \ll q\log q \cdot N^{3/2}\log^2(qN)$ and $|t^*| \ll q\log q \cdot N^{-3/2}\log^3(qN)$; unconditionally only the Siegel–Walfisz bound of Lemma W is available.

Theorem E (arc expansion). Fix $B > 40$ ($B = 48$ makes the error in (1) $O(1/\log N)$) and $0 < \varepsilon < 1/8$ (the statements for larger ε follow). As $N \to \infty$, for $3 \le q \le \mathfrak{P}$:

1. (minor arcs) $\sup_{\theta\in\mathfrak{m}} |F_N(e(\theta))| \le 2^n \exp((-\log 2 + O((\log N)^{5-B/8}) + O(M^{-1/2}))\pi(N))$.
2. (emerging arcs) Let $q \in \mathfrak{E}$. With the cumulants κ_j^* of Lemma W(3), $s^* = \sqrt{\kappa_2^*}$ and $z = (T-\kappa_1^*)/s^*$, put

$$\widehat{H}(T) = e^{-iTt^*} G(\theta^*) \frac{\varphi(z)}{s^*} \Big(1 + \frac{\kappa_3^*}{6s^{*3}} He_3(z) + \frac{\kappa_4^*}{24 s^{*4}} He_4(z) + \frac{\kappa_3^{*2}}{72 s^{*6}} He_6(z)\Big).$$

If $\mathscr{V}(a/q) = \emptyset$, put $\mathscr{A}^*_{a/q}(T) = e(-Ta/q)\hat{H}(T)$ and $\Omega = 1$. If $\mathscr{V}(a/q) = \{w\}$, write $1 + e(w\theta) = -iwt \cdot e^{iwt/2}\mathrm{sinc}(wt/2)$ and replace G by $\tilde{G} = G \cdot e^{iwt/2}\mathrm{sinc}(wt/2)$ in $\hat{H}$: its value $\tilde{G}(\theta^*) = G(\theta^*)e^{iwt^*/2}\mathrm{sinc}(wt^*/2)$ replaces $G(\theta^*)$, and the cumulants of $\tilde{G}$ at t^*, those of G plus those of $e^{iwt/2}\mathrm{sinc}(wt/2)$ at t^* (see the note after the theorem), replace the κ_j^*. Put $\mathscr{A}^*_{a/q}(T) = e(-Ta/q) \cdot w \cdot (d/dT)\hat{H}(T)$ and $\Omega = w(|t^*| + 1/\tilde{s}^*)$, where $\tilde{s}^*$ is the square root of the second cumulant of $\tilde{G}$ at t^*; $\tilde{s}^*$ takes the place of s^* in $\hat{H}$, in Ω and in the error term below. θ^* stays the peak of $|G|$. Then, for every T,

$$R_{a/q}(T) = \mathscr{A}^*_{a/q}(T) + O\Big(\Omega |G(\theta^*)| \frac{N^{-3/2+2\varepsilon}}{s^*}\Big) + O\big(2^n e^{(-\log 2 + \ell_q^* + o(1))\pi(N)}\big), \qquad \ell_q^* = \sup_{X \ge 1/(2q)} \frac{A_q(X)}{\varphi(q)}.$$

In particular, if $\mathscr{V}(a/q) = \emptyset$, then $R_{a/q}(T) = \mathscr{A}^*_{a/q}(T)(1 + O(N^{-1+2\varepsilon})) + O(2^n e^{(-\log 2 + \ell_q^* + o(1))\pi(N)})$ uniformly for $|z| \le \sqrt{\log N}$.

3. (other arcs) For $q \notin \mathfrak{E}$, $|R_{a/q}(T)| \le 2^n \exp((-\log 2 + \ell_q^* + o(1))\pi(N))$, and $\ell_q^* = \ell_q$ whenever A_q has a first local minimum X_q (Remark below).

Note (the cumulants of the sinc factor). At t^*, the cumulants of $e^{iwt/2}\mathrm{sinc}(wt/2)$ are $w/2 + iw^2t^*/12$, $w^2/12 + w^4t^{*2}/240$, $-iw^4t^*/120$ and $-w^4/120 - w^6t^{*2}/504$ to leading order. Since $w = q/2 \le \mathfrak{P}$ and $|t^*| \ll N^{-1}e^{-c\sqrt{\log N}}$ (Lemma W(2)), $w^2|t^*| \le 1$ for N large, and they are $w/2$, $w^2/12$, 0, $-w^4/120$ up to $O(w^2|t^*|(1+w^2))$.

Remark (the lobe exponent). Suppose A_q has a first local minimum X_q. Then $1/(2q) \le X_q$ (Lemma Q(1)), and A_q is non-increasing on $(0, X_q]$: A_q is continuous on $(0,\infty)$ and strictly decreasing on $(0, 1/(2q))$, and if $A_q(y) > A_q(x)$ for some $0 < x < y \le X_q$, the minimum of A_q on $[x', y]$, with $x' < x$ small, would be attained inside and be a local minimum before X_q. Hence $\ell_q^* = \max(A_q(1/(2q))/\varphi(q), \ell_q)$. For $q \in \mathfrak{E}$ we have $\ell_q^* < c_q$ by Lemma P, and $|G(\theta^*)| = 2^n \exp(\pi(N)(-\log 2 + c_q) + o(\pi(N)))$ by Lemma W(3) and Siegel–Walfisz. Hence, for each fixed $q \in \mathfrak{E}$, the lobe term of (2) is $\le e^{-\gamma_q \pi(N)}|G(\theta^*)|$ with $\gamma_q = (c_q - \ell_q^*)/4 > 0$, and so exponentially smaller than the arc term wherever $\varphi(z)/s^*$ is not exponentially small, for instance for $|z| \le \sqrt{\log N}$ when $\mathscr{V}(a/q) = \emptyset$. Lemma P gives no quantitative gap $c_q - \ell_q^*$, so this is not uniform in $q \le \mathfrak{P}$; numerically the gap is about 0.016, 0.009 and 0.006 for $q = 6$, 10, 14. For $q \notin \mathfrak{E}$, $A_q(1/(2q)) \le \varphi(q)c_q = 0 \le \ell_q$ (the first inequality by Lemma Q(1) and continuity) (every divisor term of $A_q(X)$ has modulus $\le \kappa/X$, so $A_q(X) \to 0$ as $X \to \infty$), so $\ell_q^* = \ell_q$.

Proof.

- (1) By Lemma M, $\sum_p \log|\cos \pi p\theta| \le \sum_p K_M(p\theta) = \pi(N)\hat K_M(0) + \sum_{1\le|m|\le M} \hat K_M(m) \sum_{p\le N} e(mp\theta)$, with $\hat K_M(0) \le -\log 2 + CM^{-1/2}$ and $|\hat K_M(m)| \le C$.
  - Let $\theta \in \mathfrak{m}$ and $1 \le m \le M$. Dirichlet's theorem with parameter $NM/\mathfrak{P}$ gives b/k with $k \le NM/\mathfrak{P}$ and $|m\theta - b/k| \le \mathfrak{P}/(kNM)$.
  - If $k \le \mathfrak{P}/M$, then $|\theta - b/(mk)| \le \mathfrak{P}/(mkNM) \le \mathfrak{P}/(mkN)$, with $mk \le \mathfrak{P}$. Reducing $b/(mk)$ to lowest terms a'/q', $q' \le mk$, this puts θ in $\mathfrak{M}(a'/q')$, a contradiction. So $\mathfrak{P}/M < k \le NM/\mathfrak{P}$.
  - Vinogradov–Vaughan [IK, Thm 13.6] then gives $\sum_{p\le N} e(mp\theta) \ll N(M/\mathfrak{P})^{1/2}\log^4 N$ (the term $N^{4/5}\log^4 N$ is smaller), after removing the prime powers, $O(\sqrt{N}\log N)$, and the weight $\log p$ by partial summation.
  - Summing over $|m| \le M$, the error is $\ll NM^{3/2}\mathfrak{P}^{-1/2}\log^4 N = N(\log N)^{4-B/8} \ll \pi(N)(\log N)^{5-B/8}$.
- (2) Split the arc $\mathfrak{M}(a/q)$ into three ranges. Put $\tau_0 = N^{-3/2+\varepsilon}$.
  - *Core*, $|t - t^*| \le \tau_0$. For $|\sigma| \le \tau_0$, each factor of G has no zero within distance $\pi/(2qw)$ of $t^* + \sigma$ (Lemma W(1), (2)). By Cauchy's estimates its logarithm satisfies $\log(\text{factor}) = \sum_{j\le 5} \kappa^*_{j,w}(i\sigma)^j/j! + O((qw\sigma)^6 \log q)$.
    - Summed over w, the remainder is $O(q^6 \log q \cdot \tau_0^6 \Sigma w^6) = O(q^7 N^{-2+6\varepsilon})$, which is $O(N^{-3/2+2\varepsilon})$ for $\varepsilon < 1/8$.
    - Let $x = \sum_{3\le j\le 5} \kappa^*_j (i\sigma)^j/j!$. Since $|\kappa^*_j| \ll q^j \Sigma w^j \ll q^j N^{j+1}$ and $s^* \gg N^{3/2}(\log N)^{-1/2}$, on the core $|\kappa^*_3\sigma^3| \ll q^3 N^{-1/2+3\varepsilon}$, $|\kappa^*_4\sigma^4| \ll q^4 N^{-1+4\varepsilon}$ and $|\kappa^*_5\sigma^5| \ll q^5 N^{-3/2+5\varepsilon}$.
    - Expanding $e^x = 1 + x + x^2/2 + O(|x|^3)$ and keeping the terms in κ^*_3, κ^*_4 and κ^{*2}_3, every dropped term is a product $\prod_i \kappa^*_{j_i}\sigma^{j_i}$ with $\sum_i (j_i - 2) \ge 3(\kappa^*_5, \kappa^*_3\kappa^*_4, \kappa^{*3}_3, \dots)$. Against the Gaussian $e^{-\kappa^*_2\sigma^2/2}$ each contributes $O(N^{-3/2+\varepsilon})$ relative to $\varphi(z)/s^*$, times a polynomial in z.
    - Extending the Gaussian integrals from $|\sigma| \le \tau_0$ to $\mathbb{R}$ costs $e^{-\kappa^*_2\tau_0^2/2} = e^{-cN^{2\varepsilon}/\log N}$. Since $d\theta = dt/(2\pi)$ and $(2\pi)^{-1}\int e^{-iT\sigma} e^{i\kappa^*_1\sigma - \kappa^*_2\sigma^2/2} d\sigma = \varphi(z)/s^*$, the core integral equals $e(-Ta/q)\hat H(T) + O(|G(\theta^*)|N^{-3/2+2\varepsilon}/s^*) + O(|G(\theta^*)|e^{-cN^{2\varepsilon}/\log N})$, uniformly in T: each dropped term integrates in absolute value to $O(N^{-3/2+\varepsilon}/s^*)$ times $|G(\theta^*)|$.
    - For $|z| \le \sqrt{\log N}$, $\varphi(z) \ge N^{-1/2}/\sqrt{2\pi}$ and the Edgeworth bracket is $\ge 1/2$ (its terms are $O(q^6 N^{-1/2}\log^3 N)$); since also $s^* \le 2qN^{3/2}$, the terms $|G(\theta^*)|e^{-cN^{2\varepsilon}/\log N}$ are absorbed, and when $\mathscr{V} = \emptyset$ the error is $O(N^{-1+2\varepsilon})$ relative to $\mathscr{A}^*_{a/q}$.
    - When $\mathscr{V} = \{w\}$, pull out $e(-Ta/q)$; the factor $-iwt$ then becomes $w \cdot d/dT$ after inversion, exactly. Here the first cumulant $\tilde\kappa^*_1$ of $\tilde G$ at t^* is not real ($\operatorname{Im}\tilde\kappa^*_1 = O(w^2|t^*|)$), so z is complex; the tilt $|e^{i\tilde\kappa^*_1\sigma}| = e^{-\sigma \operatorname{Im}\tilde\kappa^*_1}$ is $\le e$ on the core because $|\operatorname{Im}\tilde\kappa^*_1|\tau_0 \le 1$, and the Gaussian integrals hold with complex z. The error integrals acquire one more factor $|t| \le |t^*| + |\sigma|$, whose integral against the Gaussian is $O(|t^*| + 1/s^*)$. This gives the factor Ω.
  - *Shoulder*, $\tau_0 < |t - t^*|$ and $|t| < \pi/(qN)$. By Lemma W(1), $\log|G(t)| \le \log|G(\theta^*)| - S_2(t-t^*)^2/8 \le \log|G(\theta^*)| - cN^{2\varepsilon}/\log N$, and the vanishing factor, if any, is ≤ 2. This range contributes $O(|G(\theta^*)|e^{-cN^{2\varepsilon}/\log N})$, again negligible.
  - *Lobes*, $1/(2q) \le |X| \le \mathfrak{P}/q$, that is $|t| \ge \pi/(qN)$. By Lemma L, $|F_N(e(\theta))| \le 2^n \exp((N/\log N)(\ell^*_q - \log 2) + O(N/\log^2 N))$ (note that $1/(2q) \ge N^{-1/2}$, that $0 \le \ell^*_q \le \log 2$ because $A_q \le \varphi(q)\log 2$ and $A_q(X) \to 0$, and that $\pi(N) = (1 + o(1))N/\log N$), which is $2^n \exp((-\log 2 + \ell^*_q + o(1))\pi(N))$, and the range has measure $\le 2\mathfrak{P}/(qN)$.
- (3) As in (2), split at $|X| = N^{-1/2}$ and use Lemma L_0 for $|X| \le N^{-1/2}$. For $q \notin \mathfrak{E}$ there is no vanishing factor, $c_q = 0$, and $\log|F^*_N(a/q)| = n\log 2 - \pi(N)\log 2 + o(\pi(N))$ by Siegel–Walfisz. For $|X| \ge N^{-1/2}$,

Lemma L applies with $A_q(|X|)/\varphi(q) \le \ell_q^*$ for $|X| \ge 1/(2q)$, and with $A_q(|X|) < A_q(0) = 0 \le \varphi(q)\ell_q^*$ for $N^{-1/2} \le |X| < 1/(2q)$ (Lemma Q(1)); $\ell_q^* \ge 0$ because $A_q(X) \to 0$ as $X \to \infty$. □

Remark (the formula of Proposition C). If $\nu_{a/q} \ll N^{3/2+\varepsilon/2}$, as under GRH for the odd characters $\bmod q$ (for fixed q), the peak lies well inside the core of a/q, and the linear tilt $e^{-\nu\sigma}$ on the core is dominated by the Gaussian decay. The same proof, run with the expansion at a/q instead of θ^*, then gives (2) with the arc term $\mathscr{A}_{a/q}$ of Proposition C. When $\mathscr{V}(a/q) = \emptyset$ the error is relative, $O(N^{-1/2+3\varepsilon})$, uniformly for $|\operatorname{Re} z| \le \sqrt{\log N}$; z is now complex, with $|\operatorname{Im} z| \asymp |\nu|/s^*$. When $\mathscr{V}(a/q) = \{w\}$ the bound is absolute and centred at a/q: with $\tilde{s}^2$ the second cumulant of $\tilde{G}$ at a/q, for every T and every $\delta > 0$,

$$|R_{a/q}(T) - \mathscr{A}_{a/q}(T)| \le |G(a/q)| \Big(\frac{w\, e^{\nu^2/(2\tilde{s}^2)} N^{-3/2+7\varepsilon}}{\tilde{s}^2} + w\, e^{-N^{2\varepsilon}/(16384 \log N)} + 4 e^{4\nu^2/S_2} e^{-N^{2\varepsilon}/(2048 \log N)} \Big) + 2 \cdot 2^n e^{(-\log 2 + \ell_q^* + \delta)\pi(}$$

for $N \ge N_0(\delta)$. The main term $w \cdot \hat{H}' \approx -w \cdot z\varphi(z)/s^2$ vanishes at $z = 0$, so no relative bound holds near it. This is the form tested in Table 2. Unconditionally the expansion must be centred at the peak θ^*, because $\nu_{a/q}$ is only known to be $\ll N^2 e^{-c\sqrt{\log N}}$.

The argument uses no hypothesis on the zeros. It is stated arc by arc. A single expansion "all arcs plus an exponentially small error" is not available pointwise in T: the Edgeworth error of the main term is polynomially small, larger than any secondary arc. Pointwise, the arcs are invisible (Table 3); they are leading and exact in the residue-class counts.

3.6 The count itself: what the Gaussian already gives

N	n	S_N	exact DP	estimate	Gauss, $k = 0$ / 3	Edgeworth-2, $k = 0$ / 3	arcs / main term
500	96	$2.2 \cdot 10^4$	< 0.01 s	9 ms	$5 \cdot 10^{-3}$ / $6 \cdot 10^{-2}$	$1 \cdot 10^{-7}$ / $1 \cdot 10^{-4}$	10^{-10}
2000	304	$2.8 \cdot 10^5$	0.13 s	11 ms	$2 \cdot 10^{-3}$ / $2 \cdot 10^{-2}$	$6 \cdot 10^{-9}$ / $3 \cdot 10^{-6}$	10^{-23}
8000	1008	$3.7 \cdot 10^6$	6.3 s	24 ms	$5 \cdot 10^{-4}$ / $5 \cdot 10^{-3}$	$1 \cdot 10^{-10}$ / $7 \cdot 10^{-8}$	10^{-68}

Table 3. Relative error of the estimates of $r_N(T)$ at $T = S_N/2 + k\sigma$.

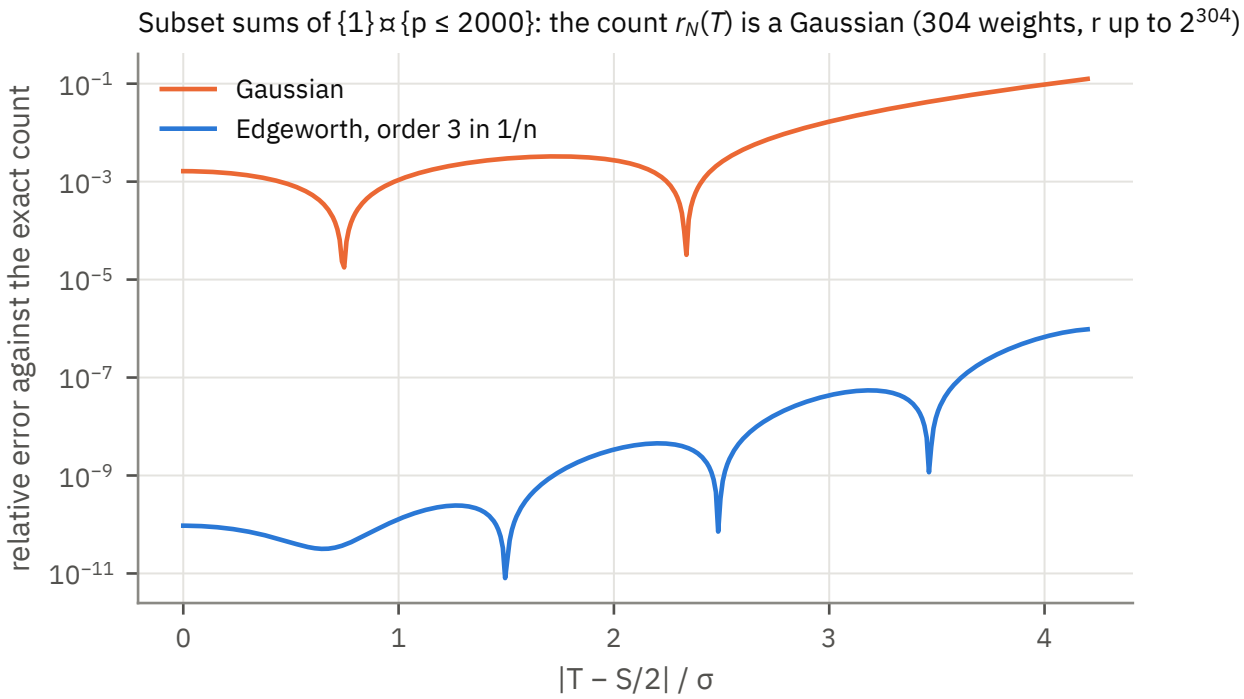


Figure 3. Relative error of the estimates of $r_N(T)$ for $W_N = \{1\} \cup \{p \le 2000\}$ (304 weights), against the exact counts of a floating-point dynamic program, for $|T - S/2| \le 4.2\sigma$. Orange: the Gaussian at lattice scale 1, about $2 \cdot 10^{-3}$ at the centre and 10^{-1} at 4σ. Blue: the Edgeworth expansion of order 3 in $1/n$ (one order beyond Table 3; the method of BENCH5, problem P4), about 10^{-10} at the centre and 10^{-6} at 4σ.

The positivity of the count does not constrain the races. The arcs enter $r_N(T)$ only through terms exponentially smaller than the main Gaussian: the strongest arc, $q = 6$, has relative height $e^{-0.144\pi(N)}$. A race changes the exponent by $O(N^\beta / \log N)$, which competes with $\pi(N)$ only if $\beta = 1$. Positivity lives at the level of the prime number theorem. The races live in the heights.

4 Prime races on the arcs

4.1 The characters seen by an arc

Expanding in Dirichlet characters, $H_{a/q}(N) = \sum_{\chi \ne \chi_0} A_\chi \pi(N, \chi)$ with $A_\chi = \varphi(q)^{-1} \sum_r c_r \bar{\chi}(r)$. The principal character drops out because $\sum_r c_r = 0$, and every odd character drops out because $c_{-r} = c_r$. Let $X(a/q) = \{\chi \ne \chi_0 : A_\chi \ne 0\}$, a set of even characters.

q	$X(a/q)$	the race
4, 6	empty	none; for $q = 6$ the measured race is identically 0
8	$\{\chi_8\}$	$\{1,7\}$ against $\{3,5\}$ mod 8
10	$\{(\cdot/5)\}$	quadratic residues against non-residues mod 5
14	the two cubic characters mod 7	$A_\chi = 0.1717 - 0.1700$i for one of them

4.2 A one-sided criterion

Proposition 3 (Landau's theorem on an arc). Fix q and a with $(a,q)=1$. Assume

- (R) $L(\sigma,\chi)\neq 0$ for $\sigma\in(\tfrac12,1]$ and $\chi\in X(a/q)$;
- (N) for every ρ with $\operatorname{Re}\rho>\tfrac12$, $\sum_{\chi\in X}A_\chi m_\chi(\rho)\neq 0$ whenever some $m_\chi(\rho)\neq 0$, where $m_\chi(\rho)$ is the multiplicity of ρ as a zero of $L(s,\chi)$.

Then the following are equivalent:

1. every $L(s,\chi)$, $\chi\in X(a/q)$, has all its zeros in the critical strip $0<\operatorname{Re}s<1$ on $\operatorname{Re}s=\tfrac12$;
2. for every $\varepsilon>0$ there is C_ε with $H_{a/q}(N)\le C_\varepsilon N^{\frac12+\varepsilon}$ for all $N\ge 2$;
3. for every $\varepsilon>0$ there is C_ε with $H_{a/q}(N)\ge -C_\varepsilon N^{\frac12+\varepsilon}$ for all $N\ge 2$.

If (2) holds for one ε, every zero of every $L(s,\chi)$, $\chi\in X$, has $\beta\le\tfrac12+\varepsilon$.

Proof. (1) $\Rightarrow$ (2), (3): under (1), $\pi(N,\chi)=O(\sqrt{N}\log N)$ [Davenport, Ch. 20], so $H=O(\sqrt{N}\log N)$.

(2) $\Rightarrow$ (1). Put $\theta=\tfrac12+\varepsilon$ and $D(s)=\sum_{p\nmid q}c_{p \bmod q}p^{-s}$. Then $D(s)=\sum_{\chi\in X}A_\chi\sum_{k\ge1}\mu(k)k^{-1}\log L(ks,\chi^k)=\sum_{\chi\in X}A_\chi\log L(s,\chi)+h(s)$, with h holomorphic on $\operatorname{Re}s>\tfrac12$. Since $L(1,\chi)\neq0$, D is regular at $s=1$, and by (R) on the real segment $(\theta,1]$. By (2), $G(x)=C_\varepsilon x^\theta-H(x)\ge0$, and $\int_1^\infty G(x)x^{-s-1}dx=C_\varepsilon/(s-\theta)-D(s)/s$ for $\operatorname{Re}s>1$. By Landau's theorem on Mellin transforms of non-negative functions [Landau; Ingham, Ch. V] the abscissa of convergence is a real singularity, so it is $\le\theta$, and D continues holomorphically to $\operatorname{Re}s>\theta$. A zero ρ of some $L(s,\chi)$ with $\operatorname{Re}\rho>\theta$ would give D a logarithmic branch point with coefficient $\Sigma_\chi A_\chi m_\chi(\rho)$, non-zero by (N): a contradiction. So $\beta\le\tfrac12+\varepsilon$ for every ε, and the functional equation of the primitive character χ^* inducing χ gives $\beta\ge\tfrac12$, since $L(s,\chi)=L(s,\chi^*)\prod_{p\mid q}(1-\chi^*(p)p^{-s})$ and the Euler factors do not vanish on $\operatorname{Re}s>0$. (3) $\Rightarrow$ (1) is the same with $G=C_\varepsilon x^\theta+H$. □

The hypotheses are verified for $q=8$, 10, 14.

- (N) is automatic when X has one element ($q=8,10$). For $q=14$, $X=\{\chi,\bar\chi\}$, $A_{\bar\chi}=\bar A_\chi$ and $A_\chi(a)=\chi(a)A_\chi(1)$ with $\chi(a)$ a cube root of unity; at $a=1$, $A_\chi/\bar A_\chi=0.0101-0.99995i$, and $\operatorname{Re}A_\chi(a)\in\{0.1717,0.0614,-0.23$ is never 0, so $A_\chi(a)/\bar A_\chi(a)$ is never real and negative and a zero of $L(s,\bar\chi)$, the conjugate of a zero of $L(s,\chi)$, cannot cancel it.
- (R) is certified by interval arithmetic (Arb, 128 bits; also checked at 64 and 256 bits). We write $L(s,\chi)=q^{-s}\sum_a\chi(a)Z(s,a/q)$, with Z the Hurwitz zeta function with its pole removed. The segment $[\tfrac12,1-2^{-8}]$ is covered by balls with a derivative enclosure, and $[1-2^{-8},1]$ is handled by a Cauchy estimate on $|s-1|=\tfrac14$. Table 4 lists the primitive characters; (R) for the characters mod 10 and 14 follows because $1-\chi^*(2)2^{-\sigma}\neq0$ for real $\sigma>0$, and for $\bar\chi$ because $L(\sigma,\bar\chi)$ is the complex conjugate of $L(\sigma,\chi)$.

Character	boxes	certified lower bound on $[\tfrac12,1]$	$L(\tfrac12)$	$L(1)$
$(\cdot/5)$	558	$L>0.23111$	0.23175	$0.43041=2\log\varphi/\sqrt{5}$
cubic mod 7, $\chi(3)=e(1/3)$	438	$\lvert L\rvert\ge0.31781$	0.31009− 0.07264i	0.53775− 0.10530i

Character	boxes	certified lower bound on $[\frac{1}{2},1]$	$L(\frac{1}{2})$	$L(1)$
χ_8	366	$L > 0.37148$	0.37369	$0.62323 = \log(1+\sqrt{2})/\sqrt{2}$

Table 4. *Certification of* (R). This agrees with Rosser [Rosser], who excludes real zeros for real characters of modulus < 227, and with Rumely [Rumely], who verifies GRH up to height 10^4 for primitive characters of conductor ≤ 13.

For $q = 8$, 10 and 14 the equivalences $(1) \Leftrightarrow (2) \Leftrightarrow (3)$ are therefore unconditional theorems. GRH for the characters seen by these arcs is a statement about counting subsets of primes: the Fourier coefficient of the subset-sum count at a/q, or of its first moment when a factor vanishes (Remark 3.1), never rises above $2^{\pi(N)}e^{\pi(N)\bar{w}(q)} \cdot O_q(1)$, with the exact $\pi(N)$, by more than a factor $e^{C_\varepsilon N^{\frac{1}{2}+\varepsilon}}$. Only one side is needed. (Replacing $\pi(N)$ by $\mathrm{li}\,N$ in this main term would add the error $\pi(N) - \mathrm{li}\,N$, a statement about ζ.)

The case of weights ± 1 on two residue classes is the one-sided theorem of Knapowski and Turán [KT], which also needs the absence of real zeros; see [MS, Example 4.3 and Theorem 4.4]. Proposition 3 is the same mechanism with general real weights; we found no statement with weights $\log|\cos(\pi ra/q)|$.

4.3 No last sign change

Theorem S. For every a coprime to 14, $H_{a/14}(N) = \Omega_\pm(N^{\frac{1}{2}-\varepsilon})$ for every $\varepsilon > 0$. In particular it changes sign infinitely often.

Proof. $X(a/14) = \{\chi, \bar{\chi}\}$, the cubic characters mod 7. Suppose that $G(x) = Cx^\theta - H_{a/14}(x) \geq 0$ for $x \geq x_0$, with $\theta < \frac{1}{2}$; the other sign is the same. The Mellin transform of G is $C/(s-\theta) - D(s)/s+$ (entire), with $D(s) = \sum_{\chi \in X} A_\chi \sum_k \mu(k) k^{-1} \log L(ks, \chi^k)$.

- $k = 1$: real zeros of $L(s,\chi)$ in $[\frac{1}{2},1]$ are excluded by Table 4, which certifies the closed segment; real zeros in $[\frac{1}{3},\frac{1}{2})$ would give, by the functional equation, real zeros of $L(s,\bar{\chi})$ in $(\frac{1}{2},\frac{2}{3}]$, also excluded.
- $k = 2$: $\chi^2 = \bar{\chi}$ is non-principal, so there is no singularity at real $s > \frac{1}{3}$ for the same reason.
- $k = 3$: $\chi^3 = \chi_0$, singular at $s = \frac{1}{3}$ only.
- $k \geq 4$: holomorphic on $\mathrm{Re}\, s > \frac{1}{4}$.

So the first real singularity is at $\max(\theta, \frac{1}{3}) < \frac{1}{2}$, and by Landau's theorem D is holomorphic on $\mathrm{Re}\, s > \max(\theta, \frac{1}{3})$. But $L(s,\chi)$ has zeros ρ on $\mathrm{Re}\, s = \frac{1}{2}$ [Rumely] (one with $\mathrm{Im}\,\rho \in [4.355, 4.357]$ is certified by Arb, see Data and code), where D has logarithmic branch points with coefficient $A_\chi m_\chi(\rho) + A_{\bar{\chi}} m_{\bar{\chi}}(\rho)$; the term $k = 2$ does not contribute there, since $L(2\rho, \bar{\chi}) \neq 0$ on $\mathrm{Re}\, s = 1$. This is non-zero because $A_\chi/\bar{A}_\chi$ is not real and negative; the argument of (N) applies to every ρ, on the line or off it. The dependence on a is only through $A_\chi(a) = \chi(a)A_\chi(1)$; this rotates $A_\chi/\bar{A}_\chi$ by $\chi(a)^2$, and the three resulting values of $\mathrm{Re}\, A_\chi(a)$ are non-zero (see (N) above). □

For $q = 8$ and 10 the single real character has $\chi^2 = \chi_0$, and the squares of the primes put a real singularity at $s = \frac{1}{2}$: the Chebyshev bias. Landau's argument does not exclude a last sign change in the direction of the

bias; Littlewood's method, through the alignment of many zeros, does this for races of a single real character [Littlewood; KT].

4.4 The principal arc carries ζ

The heights see only non-principal characters because the mean of the weights was subtracted. ζ is in what was subtracted: in the main term $\pi(N)(\log 2 + \bar{w}(q))$, whose error is $\pi(N) - \mathrm{li}\,N$, and, on the arc at $q = 1$, in the centre and width of the central Gaussian.

Proposition 3_0 **(centre and width of the prime subset sums).** For a uniformly random subset of the primes $\le N$, the sum $T = \Sigma x_p p$ has mean $\frac{1}{2}M_1(N)$ and variance $\frac{1}{4}M_2(N)$, where $M_k(N) = \sum_{p \le N} p^k$. Let $E_k(N) = M_k(N) - \mathrm{li}(N^{k+1})$; note $\mathrm{li}(N^{k+1}) = \int^N t^k dt/\log t$. For each fixed $k \ge 0$ the following are equivalent:

1. the Riemann hypothesis;
2. for every $\varepsilon > 0$ there is C_ε with $E_k(N) \le C_\varepsilon N^{k+\frac{1}{2}+\varepsilon}$ for all $N \ge 2$;
3. for every $\varepsilon > 0$ there is C_ε with $E_k(N) \ge -C_\varepsilon N^{k+\frac{1}{2}+\varepsilon}$ for all $N \ge 2$.

Proof. (1) $\Rightarrow$ (2), (3): $E_k(N) = \int^N t^k d(\pi - \mathrm{li})(t) = N^k(\pi - \mathrm{li})(N) - k\int^N t^{k-1}(\pi - \mathrm{li})(t)dt = O(N^{k+\frac{1}{2}}\log N)$, by von Koch's bound $\pi(N) - \mathrm{li}\,N = O(\sqrt{N}\log N)$ under RH [Koch]. (2) $\Rightarrow$ (1): the Mellin transform of E_k is

$$\frac{1}{s}\Big[\log\big((s-k-1)\zeta(s-k)\big) + \sum_{j\ge 2}\frac{\mu(j)}{j}\log\zeta\big(j(s-k)\big)\Big] + (\text{entire}),$$

using $P(w) = \sum_p p^{-w} = \sum_j \mu(j) j^{-1}\log\zeta(jw)$ and $\int_2^\infty t^{k-s}dt/\log t = -\log(s-k-1) + (\text{entire})$. The terms $j \ge 2$ are holomorphic for $\mathrm{Re}\, s > k + \frac{1}{2}$. $\zeta(\sigma) < 0$ on $(0, 1)$, so there is no real singularity on $(k + \frac{1}{2} + \varepsilon, \infty)$, and Landau's theorem applied to $C_\varepsilon x^{k+\frac{1}{2}+\varepsilon} - E_k(x) \ge 0$ makes $\log\zeta(s-k)$ holomorphic on $\mathrm{Re}\, s > k + \frac{1}{2} + \varepsilon$. A zero with $\mathrm{Re}\,\rho > \frac{1}{2} + \varepsilon$ would give a branch point of coefficient $m(\rho) \ne 0$; so ζ has no zero on $\mathrm{Re}\, s > \frac{1}{2}$, and the functional equation gives RH. (3) $\Rightarrow$ (1) is the same with $C_\varepsilon x^{k+\frac{1}{2}+\varepsilon} + E_k(x) \ge 0$, after increasing C_ε by $1/\log 2$ to cover the growth of $\mathrm{li}(x^{k+1})$ between consecutive integers. □

The case $k = 0$ is the classical one-sided form for $\pi(N) - \mathrm{li}\,N$. The case $k = 1$ is the centre of the Gaussian, the cumulative prime sum of [CPS]. The case $k = 2$ says that RH is equivalent to the width of the central Gaussian of the prime subset sums never exceeding its prediction $\sqrt{\mathrm{li}(N^3)}/2$ by more than a factor $1 + O(N^{-\frac{1}{2}+\varepsilon})$, on one side. With the primes up to 10^8, $E_2(N)/N^{5/2}$ lies between -0.128 and -0.0017 at 25 values of N from 10^2 to 10^8, always negative: the analogue of $\pi(N) < \mathrm{li}\,N$ in the computed range.

Figure 4 shows the three cases. It is a computation in a finite range and proves nothing about RH.

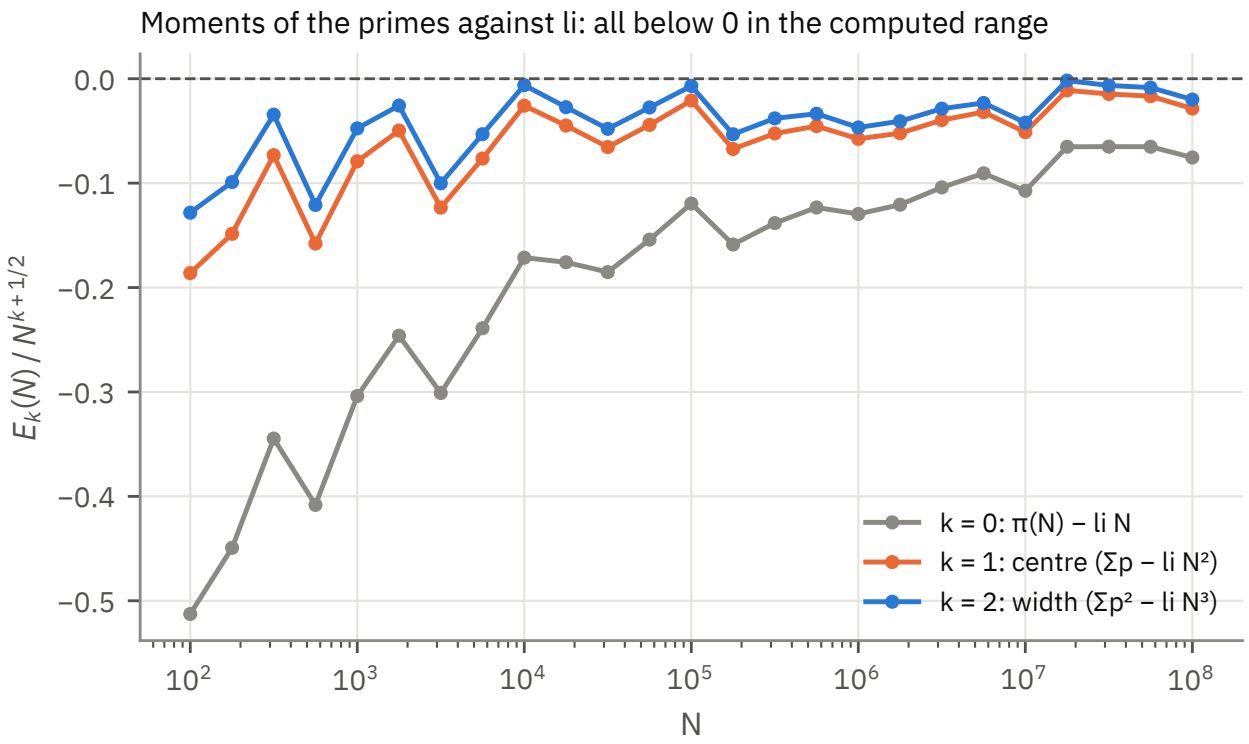


Figure 4. $E_k(N)/N^{k+\frac{1}{2}}$ for $k = 0 (\pi(N) - \mathrm{li}\, N)$, $k = 1$ (the centre) and $k = 2$ (the width of the central Gaussian), at 25 values of N from 10^2 to 10^8, log-spaced. By Proposition 3_0, RH is equivalent, for each k, to the one-sided bound $E_k(N) \le C_\varepsilon N^{k+\frac{1}{2}+\varepsilon}$. All three are negative in the computed range; for $k = 2$ the values lie between -0.128 and -0.0017.

Proposition 3_0 is a reformulation, not a route to a proof: controlling these main terms at square-root scale is the error term of the prime number theorem.

4.5 The races measured

We computed $H_{a/q}(N) \cdot \log N/\sqrt{N}$ for $N \in [10^4, 10^8]$, at 3000 points in logarithmic scale, and compared it with the Rubinstein–Sarnak prediction [RS] under GRH and the linear independence of the zeros. The mean is $\mu = -\varphi(q)^{-1} \sum_r c_r \#\{b : b^2 \equiv r \bmod q\}$, from the squares of the primes; the variance is $\Sigma_\chi |A_\chi|^2 \Sigma_\gamma (\frac{1}{4} + \gamma^2)^{-1}$, with the sums over zeros computed exactly from $L(1,\chi)$ and $L'(1,\chi)$ through Hurwitz–Stieltjes constants.

q	X	μ	sd, predicted	mean, measured	sd, measured
6	—	0	0	0 (identically)	0
8	χ_8	-0.441	0.214	-0.552	0.221
10	$(\cdot/5)$	-0.241	0.095	-0.292	0.097
14	cubic mod 7	0	0.159	-0.052	0.162

Table 5. *The races on the arcs against Rubinstein–Sarnak.* With the finite-x factor of the squares of the primes, $\frac{1}{2}\,\mathrm{li}(\sqrt{x}) \cdot \log x/\sqrt{x}$ averaged over the range (≈ 1.21; computed, not fitted), the predicted means are -0.53 and -0.29, against -0.55 and -0.29 measured. The measured means and standard deviations agree with the prediction, with no free parameter; the curves leave the ± 2 sd band only in short excursions (Figure 5).

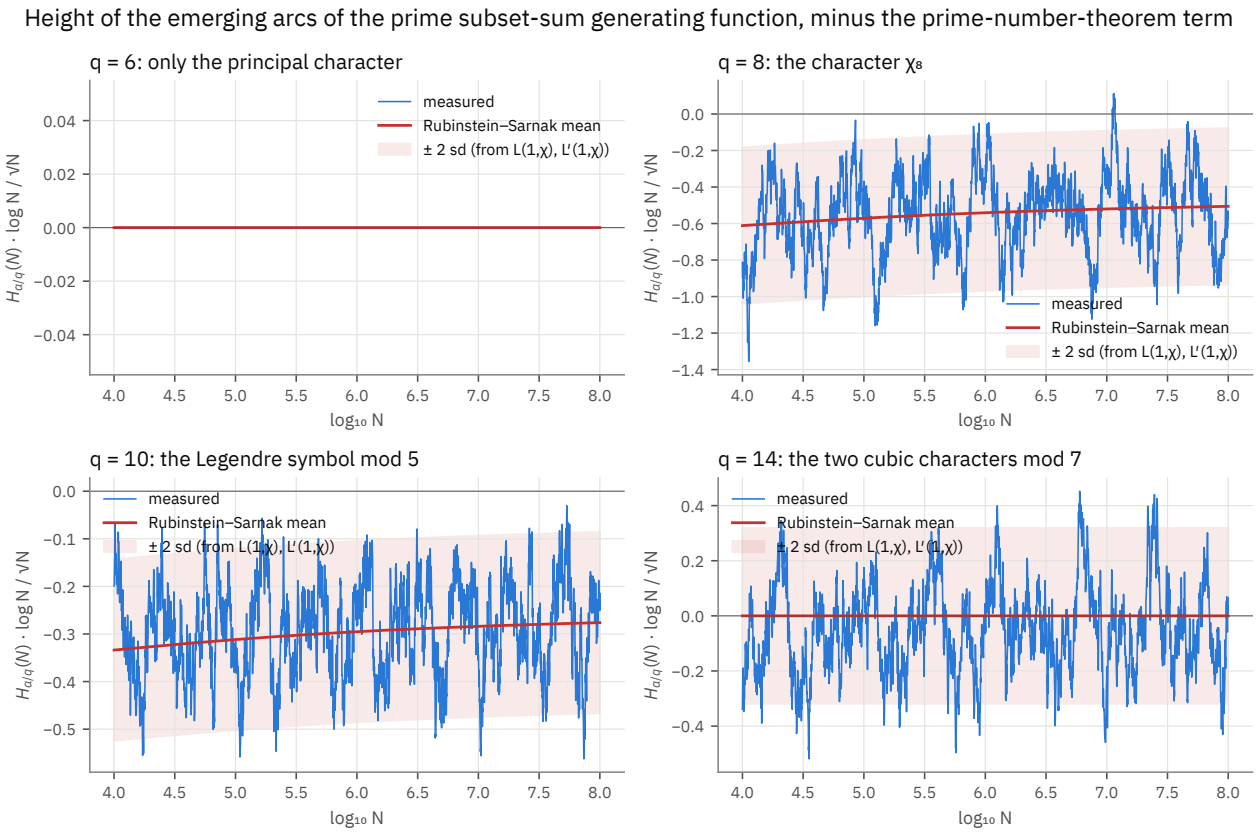


Figure 5. The heights $H_{a/q}(N) \cdot \log N/\sqrt{N}$ of the arcs $1/6$, $1/8$, $1/10$ and $1/14$, for N from 10^4 to 10^8 (blue), against the Rubinstein–Sarnak mean with the finite-x factor of the squares of the primes (red) and its ± 2 sd band, computed from $L(1,\chi)$ and $L'(1,\chi)$ (Table 5). At $q = 6$ only the principal character enters and the height vanishes identically.

5 The arcs as an algorithm

5.1 What changes for arbitrary weights

For an arbitrary set W the floor fails: $o(W) > 1$, the reachable set has gaps, and the greedy scan gets stuck. Two classical facts frame what can be hoped.

- Any polynomial-time map sending an arbitrary instance to an equivalent complete one would solve an NP-complete problem. Turning weights "into primes" is as hard as Subset Sum.
- In the dense regime $(2^n \gg S)$ the only obstructions to representability, away from the ends, are congruences modulo small almost divisors [GM; BW]. Low-density instances fall to lattice reduction [LO; CJLOSS] and have no arcs: their spectrum is all sea.

The arcs see congruences: by Remark 3.1 the class counts $N_c(q) = \#\{S \subseteq W : \Sigma S \equiv c \bmod q\}$ are the inverse discrete Fourier transform of F at the q-th roots of unity, computable exactly in $O(nq)$ per modulus. The question is whether they can also guide a construction.

5.2 A decider and a constructor

Decider. It has two parts: residue obstructions from the exact $N_c(q)$, $q \le 40$, and in the bulk a Gaussian log-count (the arc at $q = 1$), corrected by the density of live classes, required to exceed 2. Against exact reachability (bitset dynamic programming):

Instance	reachable	decider, all T	decider, bulk $\pm 3\sigma$	obstructions found
200 random weights in [1, 2000]	99.8%	99.83%	**100%**	none
198 multiples of $3+2$ arbitrary	66.5%	99.81%	**100%**	$q = 3$, class 2
150 even weights	49.9%	99.90%	**100%**	$q = 2$
mod-6 structure	49.5%	99.48%	**100%**	$q = 2$, 4, 6
30 weights in $[1, 10^5]$	77.6%	96.0%	**100%**	none

Table 6. *The arc decider.* The obstructions are found automatically by the arcs; all errors are in the tails, where the Gaussian threshold is coarse.

Constructor by self-reduction. Weights are sorted decreasingly. At step k the two branches, "take w_k" (target $T - w_k$) and "skip w_k" (target T), are scored on the suffix: $-\infty$ if a class count of the suffix vanishes, otherwise the Gaussian log-count plus the class correction. The constructor follows the larger score, which is to say the saddle point, and finishes exactly on a small block. Every output is verified.

With arcs only, down to the last weight, the success rate falls from 88% to 8% as n grows, and every failure happens in the last six weights or so: the macroscopic decisions are right, the endgame is microscopic. With the exact count on the final block, computed from the complete spectrum (all arcs $a/(S_k+1)$, which is the exact inverse transform), the success becomes 100% as soon as the final block keeps about 16 weights.

5.3 Why sixteen weights: an empirical Poisson rule

As a working model, the number of subsets of the remaining block hitting T exactly is treated as Poisson, with mean the local density $\lambda(T) = 2^m G(T; S_m/2, \sigma_m^2) \cdot (\text{class factor})$. For random weights and targets at the centre, Borgs, Chayes and Pittel prove that the count concentrates when its mean is large [BCP]; the Poisson form of the tail is a model. A branch estimated at λ is then empty with probability $\approx e^{-\lambda}$, and $\Pr(\text{success}) \approx \exp(-\sum_k e^{-\lambda_k})$, dominated by the last steps.

λ (estimated)	decisions	measured	$e^{-\lambda}$
1.3	25 285	0.161	0.264
2.4	20 814	0.074	0.093
4.2	21 404	**0.0146**	**0.0147**
7.6	20 265	0.0038	0.0006
≥ 13	$> 10^6$	0	$< 3 \cdot 10^{-6}$

Table 7. *Empty chosen branch, about* $1.3 \cdot 10^6$ *decisions.* The model is accurate near $\lambda = 4$ and optimistic beyond (0.0038 against 0.0006 at $\lambda = 7.6$); §6.6 shows that the true tail of the endgame count decays like a power of λ, not exponentially. The rule below is therefore empirical. To keep the failure below δ over n decisions, hand over to the exact endgame when the peak density of the remaining block, $2^m/\sqrt{2\pi}\sigma_m$, drops below $\lambda^* = \ln(n/\delta)$. The block size at the handover solves $2^{m^*} = \lambda^* \sqrt{2\pi}\sigma_{m^*}$, which gives $m^* \approx 15$–16 for small weights of a few hundred: the value observed.

With the filtered endgame of §5.6 a larger block costs little more, since the cost of a round grows only like $(\lambda^*)^{1/4}$, so the threshold can be raised ($\lambda^* = 5\cdot 10^4$ in Table 9′) and the Poisson misses of Table 9 disappear.

5.4 Dummies, blur and de-blurring

The solver reads the arcs directly on the original weights. This subsection explains why that reading is exact, through an augmentation that the solver itself never performs.

A simple augmentation restores the floor: insert dummy weights $d = 1 + S$ at the gaps (S the sum of the weights below the gap) until $o \le 1$. On the augmented set the reachable set is a single interval ($K = 1$ in Remark 2.4). We call the resulting forward–backward procedure HURS. On generic sets the dummies are $1, 2, \ldots, 2^{k-1}$, below $\min W$.

Blur identity. In that case

$$F_{W\cup D}(z) = F_W(z)\prod_{d\in D}(1+z^d) = F_W(z)\,\frac{1-z^M}{1-z}, \qquad M = 2^k,$$

so the dummies blur the counts with a flat window of width M: this is how o drops to 1. The dummy-free counts are recovered exactly:

$$r_W(T) - r_W(T-M) = r_{W\cup D}(T) - r_{W\cup D}(T-1).$$

Lemma 5.1 (gap lemma). Every interval of length $\Sigma_D + 1$ in $[0, S]$ contains a subset sum of W.

Proof. Every point T of the interval is $s_W + s_D$ with $s_D \le \Sigma_D$, by the completeness of $W \cup D$. □

De-blurring the arcs. The arcs are traced on $W \cup D$ and de-blurred at every step. The Gaussian cumulants of W are those of $W \cup D$ minus those of the dummies in the suffix. The class counts $N_c^W(q)$ are obtained from $N_c^{W\cup D}(q)$ by exact division by each factor $(1 + x^{d \bmod q})$ in $\mathbb{Z}[x]/(x^q - 1)$. That factor is invertible if and only if the cycle length $q/\gcd(q,d)$ is odd; where it is not, the dummy has killed those frequencies, and they are read from the original weights. The dummy 1 contributes $1 + z$, which vanishes only at $z = -1$: the frequency $\frac{1}{2}$ is lost. But exact division by $1 + x$ in $\mathbb{Z}[x]/(x^q - 1)$ fails for every even q, so the class counts at even moduli are read on the original weights.

Three findings fix the design.

- The de-blurred scores coincide exactly with the scores of the arcs on the original weights: HURS with dummies plus de-blurred arcs, and arcs on the originals, are one algorithm seen from two sides.
- Without de-blurring (arcs on $W \cup D$, dummies merely skipped) the success is 0%: the score counts completions through the dummies. A global residue check $\bmod M$ is not enough either (25, 10, 4%); the information is local.
- Marking the dummies (weights Mw, dummies $Md + 1$, target MT) makes "no dummy" the class 0 of the arc $q = M$. It works exactly but adds no information, and costs accuracy because the Gaussian is computed on $W \cup D$.

Remark 5.2 (two readings of the blur).

- The dummies $1, 2, \ldots, 2^{k-1}$ are the binary splitting of a weight 1 of multiplicity $2^k - 1$, the textbook device for bounded knapsack [Halim, Ex. 3.5.2.5]. There it acts on reachable sets. The blur identity is its form on counts: binary representations are unique, so the window is flat and the inversion exact.
- The factor $1 + x^{d \bmod q}$ acts on $\mathbb{Z}[x]/(x^q - 1)$ as $I + P^d$, with P the cyclic shift. P^d splits $\mathbb{Z}/q$ into gcd(q, d) cycles of length $L = q/\gcd(q,d)$, the Cayley graph $\mathrm{Cay}(\mathbb{Z}/q, \{d\})$, and $I + P^d$ is singular exactly when -1 is an eigenvalue of P^d, that is when L is even. A dummy kills frequencies exactly when the cycles of its Cayley graph are bipartite.

Class counts in floating point. Since the de-blurred scores coincide with the scores on the original weights, the implementation computes the suffix class counts directly on W, in floating point, with no big integers and no de-blurring. In the first version the dummies were also used in the forward pass of HURS, for a negative certificate; but $W \cup D$ is complete, so that pass certifies only $T > \Sigma W + \Sigma D$. The solver of §5.7 uses no dummies: its NO answers come from the range check and from the class counts (q, c).

- The recurrence $N_c \leftarrow N_c + N_{c-w}$ adds non-negative numbers only. The relative error is at most $n \cdot \varepsilon$, and a zero stays an exact zero, since a sum of non-negative numbers vanishes only if both terms do.
- Each row is rescaled by a power of two. A positive count lies in $[1, 2^n]$, so no positive entry can underflow for $n \leq 1020$, since 2^{-1020} is still a normal double.
- The Fourier side would be diagonal but unstable, because the inverse transform cancels down to the size of the minor arcs. The class side is stable.

Against the exact integer counts there are 0 discrepancies in the zero pattern and $|\Delta\text{score}| \leq 3 \cdot 10^{-14}$ on about 6000 scores. Treating all moduli at once costs $O(nQ^2)$: 4 ms against 27–31 ms for the exact de-blurred counts at $n = 200$.

When $o \leq 1$ already, as for the primes, every suffix block is complete, all class counts are positive, at least one branch always stays in range, and the arcs alone are exact with no endgame, at cost $O(n)$: on $\{1\} \cup \{p < 3000\}$, 100% of targets in 6 ms each.

5.5 Cost law and benchmarks

Phase	what it does	cost
0	sort	$O(n \log n)$
1 (once)	suffix class counts on W, floating point (§5.4)	$O(nQ^2)$
2 (per target)	arc decisions on the large weights	$O(nQ)$
3 (per target)	filtered endgame on m weights (§5.6)	$O(m \cdot 2^{m/4} + m \cdot 2^{m/2}/P)$ per round

Table 8. *Cost of the solver; Q is the number of moduli (24).* With the dynamic moduli of §5.7, phase 1 costs $O(n \cdot \sum_{q \in \mathscr{C}} q)$ plus $O(\sqrt{w}_{\max})$ per gcd pair for the divisors, and phase 2 costs $O(n + \sum_k |Q(k)|)$. The first version also inserted $k \leq \lceil \log_2 w_{\max} \rceil$ dummies in phase 0 and used exact integer counts on $W \cup D$ with de-blurring, $O(n^2 Q^2/64)$ word operations, and a plain meet-in-the-middle on m^* weights, $O(m^* \cdot 2^{m^*/2})$.

Cost law, first version. On generic sets σ_{m^*} is of the order of the omicron (the dummies fill $[0, \min W)$, so $o \approx \min W$), hence $m^* \approx \log_2 o + \log_2 \lambda^* + c$ and

$$T(n,o) = O(nQ) + O\big(\sqrt{\rho^* o}\,\mathrm{poly}(m^*)\big).$$

The time ratio between $o = 1.4 \cdot 10^{10}$ and $o = 2 \cdot 10^7$ is predicted as $\sqrt{700} = 26$ and measured as 21.

Cost law, filtered endgame. With $P \approx \lambda/2 \approx \lambda^*/2$ and $2^m \approx \lambda^* \sqrt{2\pi}\sigma_m$, a round costs $O(m(\lambda^*\sigma_m)^{1/4} + m(\sigma_m/\lambda^*)^{1/2})$. At $\lambda^* \asymp \sigma_m^{1/3}$ the two terms balance at $\sigma_m^{1/3}$, and

$$T(n,o) = O(nQ) + O\big(o^{1/3}\,\mathrm{poly}(m)\big).$$

On the instances of Table 9′ the measured times (15–30 ms at large λ^*) are dominated by the descent of phase 2, run in Python. The endgame alone falls from 341 ms to 3.8 ms at $m = 40$ (Table 9″).

Instance	o	dummies found		wrong outputs	ms per target
40 in $[1, 10^4]$	35	8	100%	0	0.6–0.8
30 in $[1, 10^5]$	1112	12	100% ($\lambda^* = 500$)	0	1.7–2.4
100 in $[1, 10^5]$	2588	12	100%	0	3.6–3.8
200 in $[1, 10^6]$	4443	13	100% ($\lambda^* = 500$)	0	9.6–11.1
200 in $[1, 10^{10}]$	$2 \cdot 10^7$	26	20/20	0	110
200 in $[1, 10^{12}]$	$1.4 \cdot 10^{10}$	34	20/20	0	2300
150 multiples of 7 in $[1, 7 \cdot 10^{11}]$	$3 \cdot 10^9$	32	20/20($T \equiv 0 \bmod 7$)	0	2000

Table 9. *Benchmarks, first version.* The first four rows use exact ground truth over 200 reachable targets. No wrong output occurred in more than 2400 runs. At 10^{10} and 10^{12} a dynamic-programming table would need $2 \cdot 10^{14}$ and $2 \cdot 10^{16}$ bits; meet-in-the-middle on all weights would need 2^{100} operations. The only misses are Poisson events, and raising λ^* from 50 to 500 removes them at the price of 3–4 more weights in the endgame.

The endgame in the language of arcs. On a block B of m weights, the complete spectrum resolves a single target with every frequency, at cost $m \cdot R$ with $R = \Sigma_B$. The product form $F_B = F_{B_1} F_{B_2}$ with $|B_1| = |B_2| = m/2$ is a join of two sparse half-spectra at cost $m \cdot 2^{m/2}$: this is meet-in-the-middle written as a product of generating functions. The filtered endgame of §5.6 restricts each half-spectrum to one residue class $\bmod P$ before the join. A banded spectrum (frequencies $\bmod L < R$) aliases targets and does not beat it. The $\sqrt{o}$ in the cost law is the cost of the product form.

What the arcs contribute: an ablation.

Instance	arcs	no arcs (centre rule)
200 random weights in $[1, 10^{10}]$	12/12	12/12
198 multiples of $3+2$ others, $\le 3 \cdot 10^9$	12/12	4/12
195 multiples of $7+5$ others, $\le 7 \cdot 10^9$	12/12	2/12

Table 10. On random weights the arcs are not needed: steering to the centre plus the exact endgame achieves the cost law. On sets with an almost divisor the arcs are essential: the class counts see the obstruction $\bmod d$ and keep the few non-divisible weights available to fix the residue.

Where the algorithm and the races meet. On primes restricted to a progression ($o = 3$ to 7) the arcs solve 100% of targets, and the arc centres keep the levels of Lemma 1. Restricting to $p \equiv 1 \bmod 4$ splits the arcs $1/8$ and $3/8$. The split, 0.036 per weight, about 9.7 in $\log|F|$, measures $\pi(4000;8,5) - \pi(4000;8,1)$: the Chebyshev bias, read directly in the amplitudes the algorithm uses in its scores.

Density one. We write $\rho = n/\log_2 D$ for the density of a weight set, with $D = \max W - \min W + 1$. Since $S \approx n \cdot \max W/2$, $2^n/S \geq 1$ if and only if $\rho \gtrsim 1 + \log_2(n/2)/\log_2 \max W$, so the two notions of density agree up to that correction. At density $2^n/S \approx 1$ the omicron is $\approx 2^n$ and the cost law gives $2^{n/2}$: the method becomes meet-in-the-middle, as it must; there $\lambda \approx 1$, and the filter of §5.6 has nothing to remove. There the arcs carry no usable information. The largest arc over $q \leq 24$ sits at -0.30 to -0.43 per weight, a relative deviation of the class counts of $e^{-0.3n}$ to $e^{-0.43n}$, the fluctuation of the sea. A one-level representation search [HGJ] runs at $\approx 2^{0.39n}$ on the same instances, and the full recursive versions reach $2^{0.337n}$ and $2^{0.291n}$ [HGJ; BCJ]. In a Merkle–Hellman knapsack the trapdoor $u = W^{-1} \bmod N$ gives an arc at u/N whose factors are all close to 2, an emerging arc at a very large modulus; finding it is a search over $q \tilde{} N$, which is what Shamir's attack does by other means [Shamir]. We record this as a reading, not as a result.

5.6 The filtered endgame

The endgame block E of m weights holds about λ solutions (§5.3), and the plain meet-in-the-middle enumerates all $2^{m/2}$ half-sums anyway.

- **The filter.** Split E into halves A and B. For a prime P and a residue r, keep the subsets of A whose sum is $\equiv r \pmod P$ and the subsets of B whose sum is $\equiv t - r$. A solution survives with probability $1/P$, not $1/P^2$ as with independent sampling of the two lists.
- **Building the lists.** Each filtered list comes from a meet-in-the-middle on residues over the two quarters of its half, at cost $O(m \cdot 2^{m/4} + m \cdot 2^{m/2}/P)$. The two lists are then joined on the exact sum.
- **Parameters.** The descent hands over when the peak density falls below $\lambda^* = 5 \cdot 10^4$ and the block has at most 60 weights. The filter is used when $\lambda \geq 8$ and $m \geq 16$; otherwise the block goes to the full meet-in-the-middle. These are the fixed values of the reference run BENCH5 v2; the default solver derives every one of them from the instance (§5.7, *Parameters derived from the instance*).
- **The prime.** P is uniform among the primes of $[P_0, 2P_0)$, with $P_0 = \max(3, \lfloor \lambda/2 \rfloor)$, drawn by rejection: a uniform integer of $[P_0, 2P_0)$ is drawn until it is prime. Theorem R″ (which needs $P_0 \geq 100$) proves that a round succeeds with probability at least $(1 - o(1))/(6 \ln \mathrm{mM})$, but only for the theorem mode below: in the practical solver the block is the suffix of the sorted weights, chosen after seeing them, and m and P_0 come from λ^* and the arc estimate, not from the theorem. There the bound is a guide, not a guarantee. In practice the first round almost always succeeds (§6.7). The uniform draw is the one the theorem assumes. The runs of Tables $9' - 9''$ predate it and used `sympy.randprime`, which returns the next prime after a uniform integer and so favours the primes that follow long gaps. Rerun with the uniform draw, Table $10'$ gives identical outcomes, and the reference run BENCH5 v2 of §5.7 uses it.
- **Rounds and fallback.** The implementation draws a fresh prime and residue for up to six rounds, then falls back to the full meet-in-the-middle if $m \leq 44$, so up to that size no solution is lost with respect to the plain endgame. For $m > 44$ the full lists would have more than 2^{22} elements, and the solver reports failure instead. The same bound applies when $\lambda < 8$ and the filter is skipped. These rounds run on the same block and are

correlated through its weights; Theorem R″ uses disjoint blocks, so its independence argument does not cover them.

- **Theorem mode.** A second mode runs the algorithm that Theorem R″ analyses: R disjoint blocks of m weights chosen by a random permutation, the centre rule of Lemma A on the rest, one filtered round per block with a fresh uniform prime, and no fallback. On independent uniform weights of $b = 30$ and 34 bits ($m = 42$ and 47, $P_0 = 436$ and 825, $R = 6$, targets drawn inside an estimate of the window of the theorem) it found $20/20$ targets per row, 30 of the 40 in the first round, at 4–6 ms per target. The hypotheses of the theorem held in 39 of the 40 runs; in the remaining one the realised window condition failed. These runs check the structure and the hypotheses of the theorem, not its quantitative guarantee: at $b = 30$ the bound gives $p \approx 1/164$, so with $g = 3$, $R = 6$ guarantees only about 2% success, and the R the theorem prescribes, about $300 \cdot \ln(1/\delta)$ blocks of 42 weights, would need $n \gtrsim 1.2 \cdot 10^4 \cdot \ln(1/\delta)$. The $20/20$ comes from the measured first-round rate.

- **In the language of arcs.** The list $\{y \subseteq A : S_y \equiv r \bmod P\}$ is the class r of F_A at the modulus P. The endgame is the join $[F_A]_{\equiv r} \cdot [F_B]_{\equiv t-r}$ of two class-restricted half-spectra, in place of the product $F_{B_1} F_{B_2}$ of §5.5. The probability that a round succeeds is the fraction of live classes at the modulus P (Lemma J, §6.7), the same quantity that corrects the decider of §5.2, here at a modulus of the order of λ. The block size and P are both read from the principal arc.

- **Merging.** Sorted half-lists are built by merging L and $L + w$ at each weight (MERGE-LISTS, [CLRS, §35.5]), at cost $O(2^h)$ instead of $O(h \cdot 2^h)$. On its own this gives only a factor 1.2.

Instance	first version, $\lambda^* = 50$	filtered, $\lambda^* = 50$	filtered, $\lambda^* = 500$	filtered, $\lambda^* = 5 \cdot 10^4$
200 in $[1, 10^{10}]$	143 ms	18 ms	17 ms	15 ms
200 in $[1, 10^{12}]$	4431 ms	123 ms	45 ms	30 ms
150 multiples of 7 in $[1, 7 \cdot 10^{11}]$	3302 ms	24 ms	19 ms	19 ms

Table 9′. *Time per target on the instances and targets of Table 9 (rows 5–7, same seed), first version against the filtered endgame, on one machine and one core.* Every cell found $20/20$ targets with 0 wrong outputs and 0 fallbacks. The first version needs 4.4 s at 10^{12} here, against 2.3 s in Table 9, measured on another machine. On new draws of the four shapes of the first rows of Table 9 (exact ground truth over 200 reachable targets) the filtered solver finds $200/200$ with 0 wrong outputs, at 0.9–13 ms per target.

m	weights in	λ (mean)	plain	merged	filtered
20	$[1, 10^4]$	34	0.25 ms	0.25 ms	0.92 ms
30	$[1, 10^6]$	277	5.6 ms	5.1 ms	0.56 ms
36	$[1, 10^8]$	165	69 ms	59 ms	1.3 ms
40	$[1, 10^9]$	240	341 ms	277 ms	3.8 ms

Table 9″. *The endgame alone, 20 blocks per row, targets within $M/2$ of the centre; λ is the mean over the 20 blocks of the expected number of solutions of a block.* The three methods give the same answer on $80/80$ blocks, and every solution is verified. On small blocks the filter does not pay. The table is empirical: in two rows $P_0 = \lambda/2 < 100$, outside the hypotheses of Theorem R″.

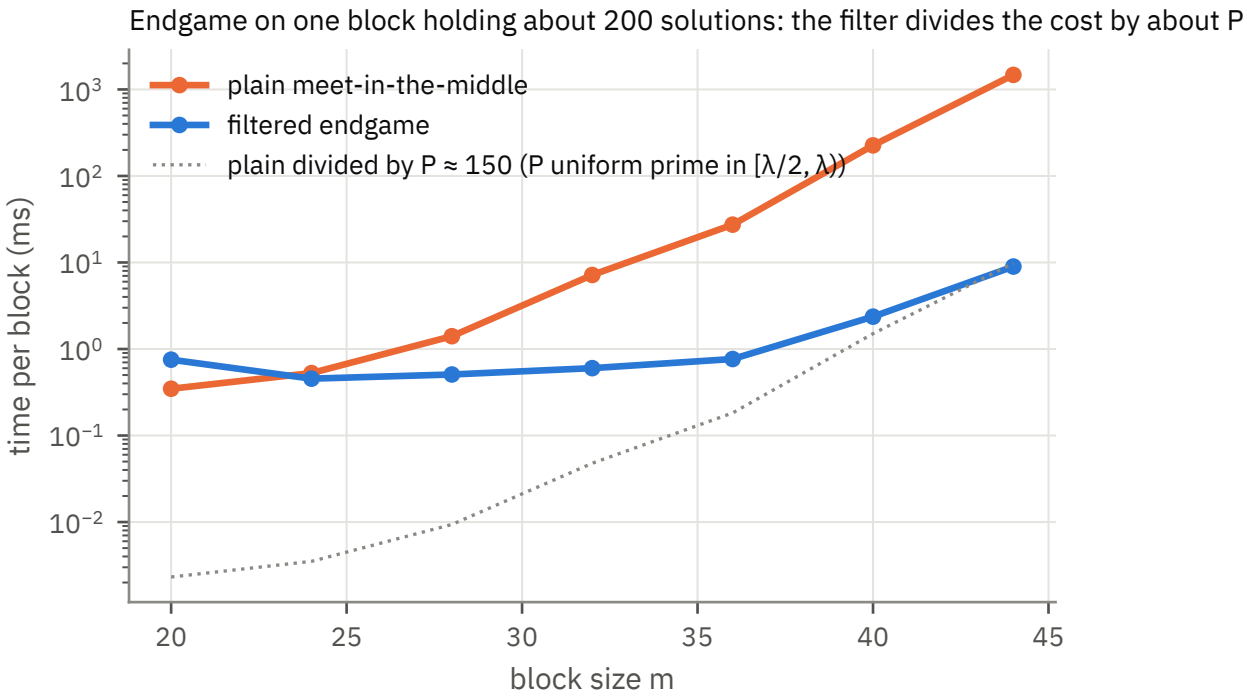


Figure 6. Time per block of the endgame on m random weights whose range makes the block hold about $\lambda \approx 200$ solutions; median over 3–5 blocks, one core. Orange: plain meet-in-the-middle. Blue: the filtered endgame. Below $m \approx 24$ the quarter lists and the overhead dominate and the filter does not pay, as in Table 9″. Beyond, with λ and therefore P fixed, both costs grow like $2^{m/2}$, and the filter divides the plain cost by about P (dotted: the plain cost divided by 150). At $m = 44$: 1.47 s against 9 ms.

5.7 Dynamic moduli

The fixed moduli $q = 2, \ldots, 24$ of §5.4 miss an almost divisor larger than 24, and on random weights they are read at every step although they carry no information. The dynamic rule chooses the moduli from the weights and, at each suffix, keeps only those whose arcs deviate.

The rule. Let W be sorted decreasingly, let $N_c^{(k)}(q)$ be the number of subsets of the suffix $W_k = (w_k, \ldots, w_n)$, of size $m = n - k + 1$, with sum $\equiv c \pmod q$, and

$$\delta_q(k) = \max_c \left| \frac{q N_c^{(k)}(q)}{2^m} - 1 \right|, \qquad \varepsilon_k = \min(1, 1/m).$$

- *Candidates.* $\mathscr{C} = \{2, \ldots, 24\} \cup \{d \le 2000 : d \mid \gcd(a, b)$ for one of 64 pairs (a, b) of weights drawn at random$\}$. If d divides all weights but a fraction f of them, a pair is missed with probability $1 - (1-f)^2$, so d escapes all 64 pairs with probability $(1-(1-f)^2)^{64}$, below 10^{-80} for $f = 5/200$ (the pairs are drawn from a fixed pseudo-random seed; drawing the two elements without replacement changes the estimate only in the third digit).
- *Selection.* At suffix k the score of §5.2 reads the moduli $Q(k) = \{q \in \mathscr{C} : \delta_q(k) \ge \varepsilon_k\}$. The class correction is $\max(0, \max_{q \in Q(k)} \ln(q N_c^{(k)}(q)/2^m))$ at $c = t \bmod q$, and an empty class gives the score $-\infty$. The negative certificate is read at the full set: if $N^{(1)}_{T \bmod q}(q) = 0$ for some $q \in Q(1)$, the answer is NO with the certificate $(q, T \bmod q)$. A NO answer carries one of three certificates: $(0,0)$ when $T < 0$ or $T > \Sigma W$; $(q, T \bmod q)$

when that class is empty mod q; and $(-3,-3)$ when an exhaustive endgame on the whole instance found no subset with sum T, which any exact search re-checks.

- *Cost.* The divisors of the 64 gcds cost $O(\sqrt{w}_{\max})$ each by trial division. The class counts of all candidates cost $O(n\cdot\sum_{q\in\mathscr{C}} q)$ once, in floating point as in §5.4, and the descent costs $O(n+\sum_k |Q(k)|)$ per target.

Lemma 5.3 (pruning). Fix k and t, and let $s_{\mathscr{C}}$ and s_Q be the scores of §5.2 with the moduli $\mathscr{C}$ and $Q(k)$.

1. If $N^{(k)}_{t \bmod q}(q) = 0$ for some $q \in \mathscr{C}$, then $q \in Q(k)$, and $s_Q = s_{\mathscr{C}} = -\infty$. In particular the NO answers with the moduli $Q(1)$ are exactly those with the moduli $\mathscr{C}$.
2. Otherwise, if $0 \le t \le S_k$ (outside this range both scores are $-\infty$), $0 \le s_{\mathscr{C}} - s_Q \le \ln(1+\varepsilon_k) \le \varepsilon_k$.
3. By Lemma C(1), applied to the suffix, $\delta_q(k) \le \sum_{d'|q,d'>1} \varphi(d') e^{-\pi^2 \kappa_k(d')/(2d'^2)}$, where $\kappa_k(d')$ is the number of weights of W_k not divisible by d'. So $q \notin Q(k)$ as soon as $\kappa_k(d') > (2d'^2/\pi^2)\ln((q-1)/\varepsilon_k)$ for every divisor $d' > 1$ of q.

Proof.

1. An empty class gives $|q\cdot 0/2^m - 1| = 1 \ge \varepsilon_k$, so $\delta_q(k) \ge \varepsilon_k$ and $q \in Q(k)$. The score with either set then returns $-\infty$ at q. At $k = 1$ this is the statement on certificates.
2. The class t mod q is non-empty for every $q \in \mathscr{C}$, so every term is finite. For $q \notin Q(k)$, $\delta_q(k) < \varepsilon_k$ gives $qN_c/2^m < 1+\varepsilon_k$ at every class, so its term is below $\ln(1+\varepsilon_k)$. The two scores share the Gaussian term, and their corrections are $b_{\mathscr{C}} = \max(b_Q, \max_{q\notin Q(k)} \ln(qN_c/2^m))$ and b_Q, with $b_Q \ge 0$. Hence $b_Q \le b_{\mathscr{C}} \le \max(b_Q, \ln(1+\varepsilon_k))$.
3. Lemma C(1) bounds $\delta_q(k)$ by the sum over its divisors, and $\sum_{d'|q,d'>1} \varphi(d') = q-1$. If every term satisfies $e^{-\pi^2\kappa_k(d')/(2d'^2)} < \varepsilon_k/(q-1)$, the sum is below ε_k. □

So the dynamic rule changes no NO answer and moves a score by at most $\ln(1+\varepsilon_k)$. A decision of the descent can change only when the scores of its two branches are within $\ln(1+\varepsilon_{k+1})$ of each other, that is, when the scores barely separate them. The endgame threshold does not use the class correction and is unchanged. The estimate $\lambda = e^s$ handed to the endgame does include it, so P_0 can move by a factor up to $1+\varepsilon_k$ and the random choices of the endgame can differ; its answers are verified either way. Part 3 is a worst-case condition: on random weights the measured $\delta_q(k)$ is far smaller than the bound, and no modulus is selected at the large suffixes. Figure 7 shows both cases.

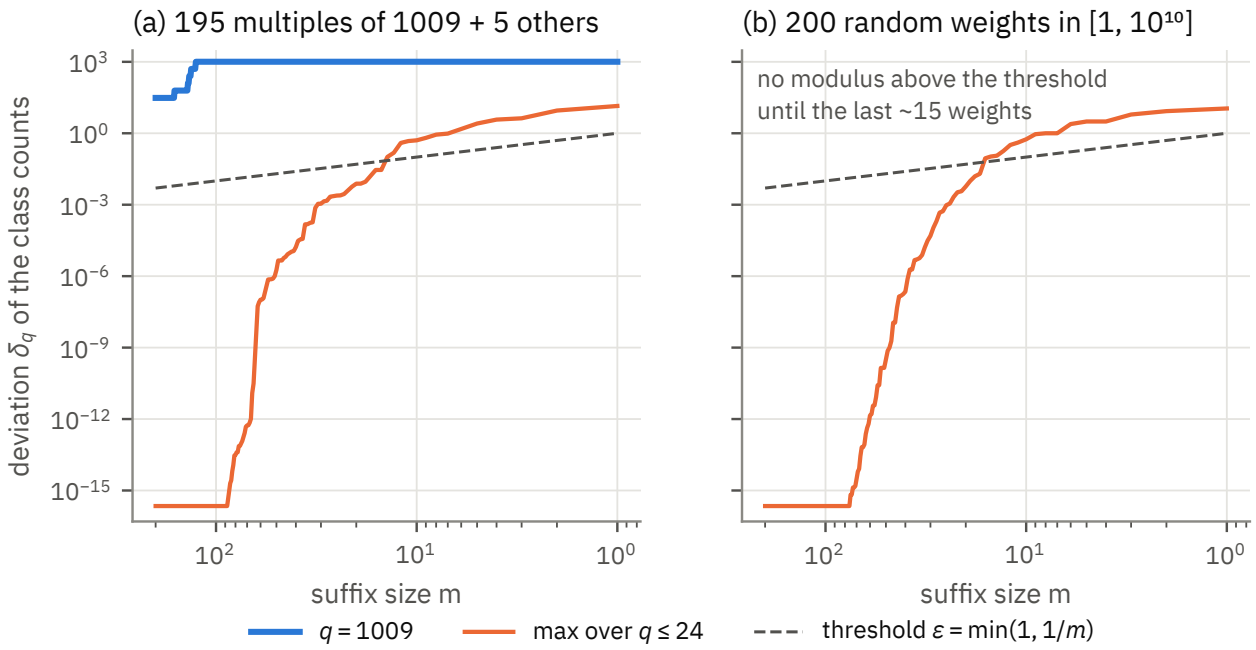


Figure 7. The selection rule of Lemma 5.3 along the suffixes (suffix size m, decreasing to the right). (a) 195 multiples of 1009 and 5 other weights. δ_{1009} stays above the threshold at every suffix, so $q = 1009$ is read throughout, while every $q \le 24$ stays below the threshold $\varepsilon = \min(1, 1/m)$ until the last 15 weights or so. (b) 200 random weights in $[1, 10^{10}]$; no modulus reaches the threshold until the last 15 weights or so. Values below 10^{-16} are rounding.

Instance	fixed $q = 2, \ldots, 24$	dynamic	no arcs	dynamic: moduli read; mean $\lvert Q(k)\rvert$
200 random in $[1, 10^{10}]$	20/20, 37.9 ms	20/20, 7.1 ms	20/20, 6.2 ms	none, 0
198 multiples of $3+2$ others, $\le 3 \cdot 10^9$	20/20, 14.9 ms	20/20, 13.7 ms	8/20	3, 6, 9, …; 22
195 multiples of $7+5$ others, $\le 7 \cdot 10^9$	20/20, 16.5 ms	20/20, 11.1 ms	7/20	7, 14, 21, …; 14
195 multiples of $101+5$ others, $\le 1.01 \cdot 10^{11}$	4/20	20/20, 11.0 ms	4/20	101, 202, 303, …; 9
195 multiples of $1009+5$ others, $\le 10^{12}$	5/20	20/20, 12.6 ms	5/20	1009; 1
190 multiples of $6+10$ others, $\le 6 \cdot 10^9$	20/20, 18.0 ms	20/20, 15.4 ms	20/20, 5.9 ms	2, 3, 4, 6, …; 22.7

Table 10′. *Fixed against dynamic moduli, 20 planted targets per row, one machine and one core; time per target where all 20 were found. The last column lists the smallest moduli read at the suffixes of more than 40 weights, and their mean number per suffix.* There were 0 wrong outputs in every cell. On random weights the dynamic rule reads no modulus and costs what the centre rule costs. The moduli 101 and 1009 are larger than 24, and only the gcd candidates find them. With five non-divisible weights at most $2^5 = 32$ residues mod 1009 are reachable, so a target in another class must be refused, and a solution must use the right subset of those five weights. The fixed moduli do not see this constraint, and fail in 15 or 16 of the 20 cases. On $\{1\} \cup \{p < 3000\}$ the rule selects no modulus at the full set, moduli enter only at the last suffixes, and 30/30 random targets are found with either rule.

Figure 8 follows one descent on an instance of this kind. Most decisions at the large weights are nearly indifferent; five are forced by the arc mod 1009.

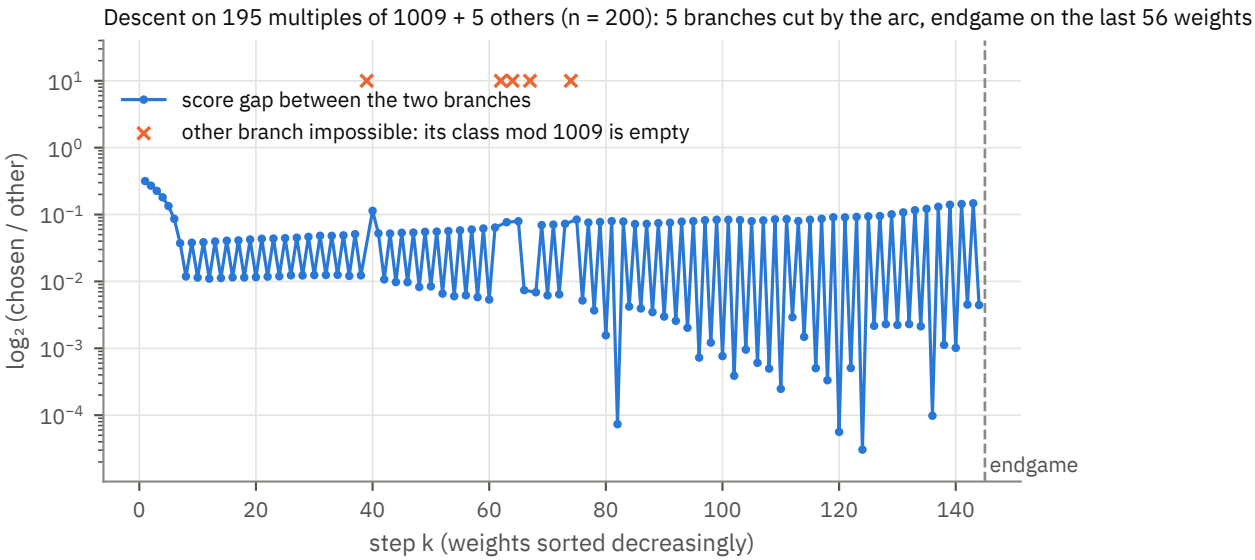


Figure 8. The descent of ARCS on one planted target of an instance with 195 multiples of 1009 and 5 other weights ($n = 200$). Blue: the gap, in bits, between the scores of the chosen branch and of the other branch at each step; it lies between 10^{-5} and 0.3 bits. Crosses: the 5 steps at which the other branch was impossible, because its class mod 1009 was empty. The descent hands over to the filtered endgame on the last 56 weights.

A preregistered check. The same comparison is part of a benchmark of five problems whose criteria were fixed, and hashed, before any instance was generated. Instance i of problem P has seed $s + 1000P + i$, where s is the integer formed by the first 64 bits of the sha256 of the criterion; there is one run, and every answer goes through an independent checker that verifies each subset sum and recomputes the reachable residues $\bmod q$ in integer arithmetic for each NO certificate. The reference run is BENCH5 v2, with the endgame of §5.6. Problem P5 has 200 weights, 195 multiples of 1009 and 5 others, with 6 planted and 6 impossible targets. The dynamic solver found $6/6$ planted targets, in a median time of 0.061 s, and certified all 6 impossible ones with the certificate $q = 1009$. The fixed moduli and the solver without arcs found $1/6$ each and certified none. CP-SAT, with one worker and a 60 s limit, decided none of the 12. There were no wrong answers. Figure 9 gives the other problems, including the one that ARCS loses.

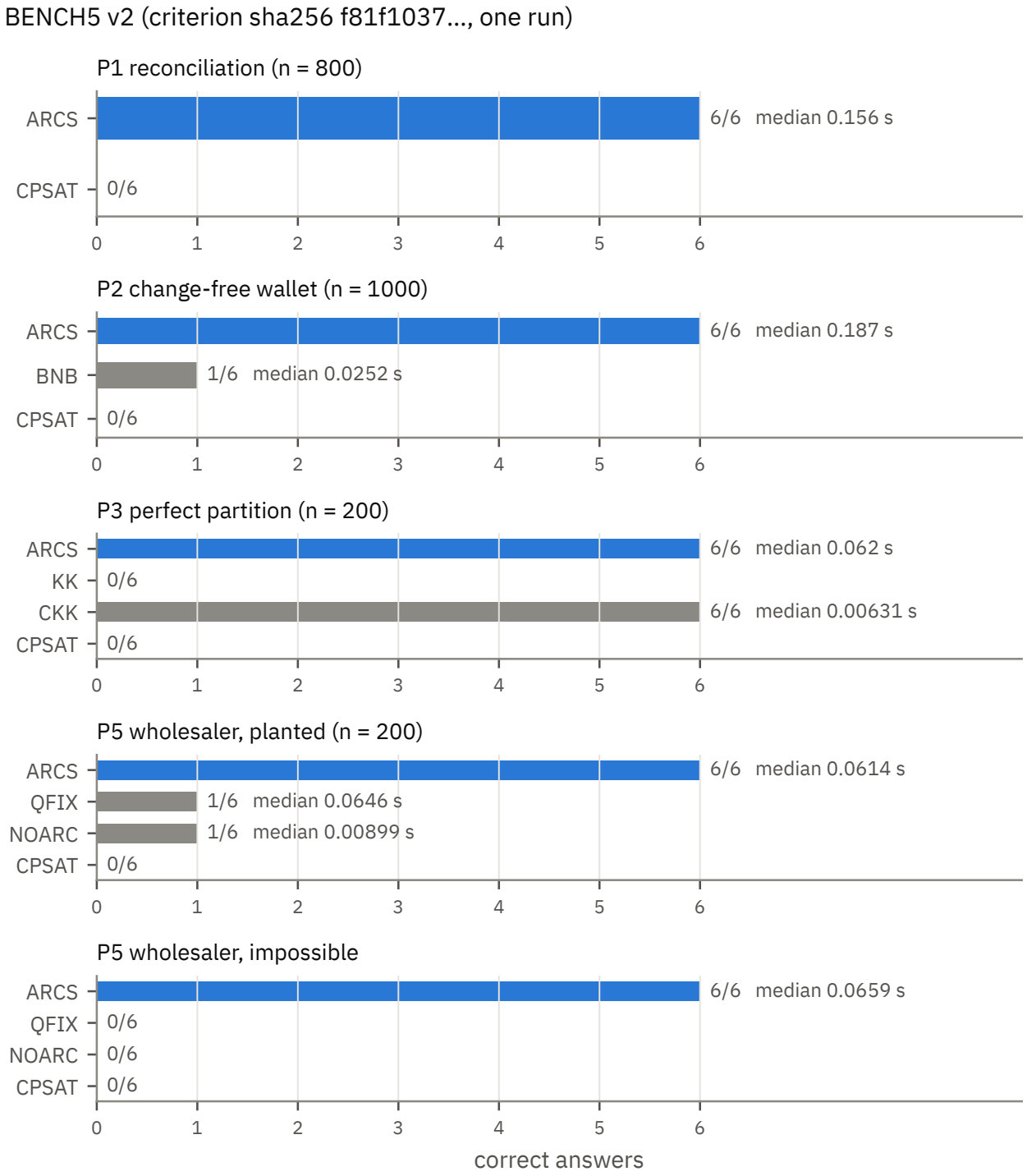


Figure 9. BENCH5 v2, the reference run (criterion sha256 f81f1037…; one run; every answer checked by the independent checker). Bars: correct answers per method; labels: correct/total and the median time of the correct answers. P1: reconciliation of 800 amounts. P2: a change-free payment from 1000 amounts. P3: a perfect partition of 200 values. P5: the instances above. ARCS (blue) solves $6/6$ in every panel and certifies all 6 impossible targets of P5 with $q = 1009$. It loses P3, where complete Karmarkar–Karp also solves $6/6$ and is 10 times faster. No method gave a wrong answer. P4 (counting the subsets of $\{1\} \cup \{p \leq N\}$ with a given sum) is not shown: the Edgeworth estimate has relative error $\leq 4.3 \cdot 10^{-8}$ in every cell and is 81, $2.7 \cdot 10^4$ and $1.2 \cdot 10^5$ times faster than the dynamic program at $N = 2000$, 8000 and 16 000; the cell at $N = 2000$ misses the preregistered factor 100.

Parameters derived from the instance. The default form of ARCS runs the steps of this section with every constant derived before the descent from the instance and from two user inputs: the accepted failure probabil-

ity δ on YES instances, and the number of operations the user will wait for, $\max_{\text{ops}}$ (with a memory bound). From the density $\rho = n/\log_2 D$ and the suffix variances V_k it computes:

- $\lambda^* = \max(400, (2.3n/\delta)^{1/3})$, from the measured tail of §6.6 with a union bound over the n decisions, and the hypothesis $P_0 \geq 100$ of Theorem R$''$;
- the suffix k^* where the peak density $2^m/(\sqrt{2\pi}\sigma_m)$ first falls below λ^*, and the cost of the cheapest endgame on the block $m^* = n - k^*$ that fits the memory (filtered rounds, full meet-in-the-middle, or the full residue scan [SS; HGJ]);
- $R = \lceil \ln(1/\delta)/2 \rceil$ filtered rounds, from the model in which each of the Poisson(λ) solutions survives the residue filter with probability $1/P$;
- the moduli: q with $\Sigma q \leq n$, and the divisors of the gcds of $\lceil \ln \delta \,/\, \ln(3/4) \rceil$ random pairs of weights, so that a divisor of at least half of the weights is missed with probability $\leq \delta$, each kept only if its class counts cost no more than the predicted endgame.

The handover rule depends on V_k only, not on the target or on the path of the descent, so k^* is the actual handover whenever an endgame fits; the cost of that endgame is a model (Table 8). If the predicted cost or memory exceeds the limits, ARCS answers FAIL with the prediction, after its NO certificates. Randomness comes from the operating system, and every prime and residue drawn is part of the output, so a run is replayable and P is independent of the weights, as Lemma K requires.

A second preregistered benchmark, with the same rules (criterion hashed before the instances, one run, independent checker), ran the default solver alone, with $\delta = 10^{-3}$, $\max_{\text{ops}} = 2^{34}$ and 2^{30} bytes. On 65 instances there were 0 wrong answers:

- P1, P2, P3 and P5 of BENCH5, with new seeds: every instance answered, the impossible targets with the certificates $q = 2$ and $q = 1009$, in median times of 0.01–0.26 *s*;
- 50 weights of 50 bits ($\rho \approx 1$), a block that the fixed caps of §5.6 refuse: $6/6$ planted targets found and $6/6$ random targets refused by an exhaustive endgame, verified by the checker's own meet-in-the-middle, in 6–11 *s*;
- 40 weights of 70 bits, beyond 64-bit integers: $6/6$ found;
- 40 weights of 60 bits ($\rho \approx 0.67$): $6/6$ found and $6/6$ refused exhaustively, in about 1.3 s;
- 100 weights of 110 bits ($\rho \approx 0.91$): $3/3$ FAIL in 4 ms, with the prediction of 2^{54} bytes and $2^{55.6}$ operations.

The predicted block equalled the actual one in all 54 runs that reached an endgame.

6 Rigorous versions of the algorithm

The outputs are verified, so correctness is never in question. What needs proof is that a solution is found, with high probability and within the stated cost. We give this for random weights (Theorems R and R$'$) and for dense structured sets (Corollary C$'$).

Setting. $W = (w_1, \ldots, w_n)$, $S = \Sigma w_i$, target T. The indices split into a head H and an endgame block E with $|E| = m$. The descent fixes x_i on H; the endgame solves $\sum_{i \in E} x_i w_i = t$ exactly by meet-in-the-middle.

6.1 *The descent never fails*

Lemma A (centre rule). Process H in any order. At step k let R_k be the sum of the head weights not yet processed, t_k the residual target, $c_k = (R_k + S_E)/2$, and choose x_k to minimize $|t_k - c_k|$. With $e_k = t_k - c_k$ and $w^* = \max_H w_i$,

$$|e_k| \le \max\Big(|e_0| - \tfrac{1}{2}\sum_{j\le k} w_j,\ \tfrac{1}{2}\max_{j\le k} w_j\Big).$$

In particular, if $|T - S/2| \le S_H/2 + w^*/2$, the descent ends with $|t - S_E/2| \le w^*/2$.

Proof. $c_k = c_{k-1} - w_k/2$, and the better of the two choices gives $|e_k| = ||e_{k-1}| - w_k/2|$. Induction. □

Writing $x_i = (1+\varepsilon_i)/2$, the endgame becomes $\sum_E \varepsilon_i w_i = u$ with $u = 2t - S_E$ and $|u| \le w^*$: a number-partitioning question. Whether it has a solution is the only probabilistic question.

6.2 *A local limit theorem uniform in the size of the weights*

Lemma B0′**.** Let $w_1, w_2, \dots$ be independent and uniform in $\{1,\dots,M\}$, $M \ge 2$, with $\mu = (M+1)/2$, $\sigma^2 = (M^2-1)/12$, $\kappa_4 = -(M^4-1)/120$, $S_k = w_1 + \cdots + w_k$, $z = (v - k\mu)/(\sigma\sqrt{k})$, $g_k(v) = (2\pi k\sigma^2)^{-1/2} e^{-z^2/2}$ and $q_M = 1/(M\sin(\pi/(2M))) \le 1/\sqrt{2}$. For $k \ge 2$ and every v,

$$\Big|\Pr(S_k = v) - g_k(v)\Big(1 + \frac{\kappa_4 He_4(z)}{24\sigma^4 k}\Big)\Big| \le \frac{0.34}{\sigma k^{5/2}} + \frac{1}{\sigma}\Big(\frac{q_M^{k-2}}{\sqrt{12}} + \Big(\frac{0.41}{k} + 0.11\Big)e^{-0.308k}\Big).$$

Consequently $|\Pr(S_k = v) - g_k(v)| \le C(k,M)/(\sigma k^{3/2})$ with

$$C(k,M) = 0.0599\,\frac{M^2+1}{M^2-1} + \frac{0.34}{k} + k^{3/2}\Big(\frac{q_M^{k-2}}{\sqrt{12}} + \Big(\frac{0.41}{k} + 0.11\Big)e^{-0.308k}\Big).$$

The constant $0.0599 = (1.2/24)\cdot 3/\sqrt{2\pi}$ is exact: it is the first Edgeworth term at $z = 0$, with $|\kappa_4|/\sigma^4 \to 6/5$.

Proof. The characteristic function of $w - \mu$ is $f(\theta) = \sin(M\theta/2)/(M\sin(\theta/2))$. Put $x = \theta/2$, $y = M\theta/2$ and $\Lambda(t) = \log(\sin t/t) + t^2/6$, so $\log f = -\sigma^2\theta^2/2 + \Lambda(y) - \Lambda(x)$.

- On $|\theta| \le \pi/M$: $\Lambda(t) = -t^4/180 + \Lambda_6(t)$ with non-positive coefficients and $|\Lambda_6(t)| \le 4.35\cdot 10^{-4} t^6$ on $(0, \pi/2]$. Since $(y^4 - x^4)/180 = -\kappa_4\theta^4/24$, $f^k = e^{-k\sigma^2\theta^2/2}(1 + k\kappa_4\theta^4/24 + E)$ with $|E| \le (k\kappa_4\theta^4/24)^2/2 + k|\Lambda_6|e^{k|\Lambda_6|}$. The linear term integrates to the Edgeworth term exactly. The two error terms integrate to $\le 0.146/(\sigma k^{5/2})$ and $\le 0.19/(\sigma k^{5/2})$, using $\kappa_4^2/\sigma^8 \le 4$ and $(M/\sigma)^6 \le 4096$.
- On $\pi/M \le |\theta| \le \pi$: $|f| \le q_M$, and $(2\pi)^{-1}\int |f|^2 = 1/M \le 1/(\sqrt{12}\sigma)$.
- The Gaussian and Edgeworth tails beyond π/M contribute $\le (0.41/k + 0.11)e^{-0.308k}/\sigma$, using $k\sigma^2\pi^2/(2M^2) \ge 0.308$k. □

Against exact convolutions ($M \in \{2,3,5,16,64,317\}, 2 \le k \le 160$) the inequality holds in every case. The

residual after the Edgeworth term is at most 0.157 against the bound's 0.34, and the measured $C(k)$ tends to 0.0599 for large M and to $0.0997 = 0.0599 \cdot 5/3$ at $M = 2$.

The constant comes from the fourth cumulant. An envelope $|f(\theta)| \le 2/(M|\theta|)$ on $[\pi/M, \pi]$ is false: $|f| \cdot M|\theta|/2$ reaches $\pi/2$ there. The true envelope is $\pi/(M|\theta|)$, and the middle range governs only small k.

6.3 The endgame is non-empty

Lemma B. Let the weights of E be independent and uniform in $\{1, \dots, M\}$, $|E| = m$, let $Z(u) = \#\{\varepsilon \in \{\pm 1\}^E : \Sigma \varepsilon_i w_i = u\}$, and

$$\lambda_m = \frac{2^{m+1}}{M\sqrt{2\pi m\, c_M}}, \qquad c_M = \tfrac{1}{3} + \tfrac{1}{2M} + \tfrac{1}{6M^2}.$$

Fix $g > 0$ and set $\eta = 1/10$. Uniformly in $M \ge 2$, in s with $|s - m\mu| \le g\sigma\sqrt{m}$ and in u with $|u| \le M$, $u \equiv s \pmod 2$:

1. $\lambda := E[Z(u) \mid S_E = s] = \lambda_m(1 + o(1))$ (the expected number of solutions of the block; F always denotes a generating function);
2. $E[Z(u)^2 | S_E = s] \le \lambda + \lambda \cdot 1[u = 0] + \lambda^2(1 + O(m^{-1/5}))$;
3. $\Pr(Z(u) = 0 | S_E = s) \le 2/\lambda + O(m^{-1/5})$.

For $u \in \{0, 1\}$ without conditioning, these are the moment estimates behind [BCP, Theorem $C(i)$]. The new point is the conditioning on S_E, which Lemma A needs.

Proof. With $\varepsilon_i = +1$ on $x \subseteq E$, $\Sigma \varepsilon_i w_i = 2S_x - S_E$, and the solutions are the x with $S_x = a := (s+u)/2$.

- (1) $\lambda = \sum_x \Pr(S_x = a) \Pr(S_{E \setminus x} = s - a) / \Pr(S_E = s)$. Restrict to $|k - m/2| \le m^{3/5}$, $k = |x|$; the other x are a fraction $\le 2e^{-2m^{1/5}}$ of all, and each of their terms is $O(\sqrt{m}/M)$, since every local probability is $\le 1/M$, so they are negligible against λ_m. On the typical range Lemma B0′ replaces each probability by its Gaussian; summed over x, the absolute errors give a relative error $O(m^{-1/2})$. The Gaussian ratio is the density at a of the conditional law of S_k given $S_m = s$, $N(ks/m, \sigma^2 k(m-k)/m)$; its variance can be frozen at $\sigma^2 m/4$, and the sum over $k = m/2 + j$ is a Riemann sum of a Gaussian, equal to $\lambda_m(1 + o(1))$ uniformly in M and in $|u| \le M$.
- (2) For a pair (x, y) put $C = x \cap y$, $A = x \setminus y$, $B = y \setminus x$, $D = E \setminus (x \cup y)$. The conditions read $S_A = S_B = a - S_C$ and $S_D = S_C - u$.
  - The diagonal gives λ. If exactly one of A, B is empty the pair is impossible. If C and D are both empty, $y = E \setminus x$, possible only for $u = 0$, giving $\lambda \cdot 1[u = 0]$.
  - For every other pair, bounding two factors by their maxima and $\sum_c \Pr(S_C = c) \Pr(S_D = c - u) = \Pr(S_C - S_D = u)$ by its maximum gives at most $Cm^{3/2}$ times the value of a typical pair. Atypical pairs are a fraction $e^{-cm^{2\eta}}$ of all.
  - For typical pairs the four class sums S_A, S_B, S_C, S_D are independent sums of $|A|$, $|B|$, $|C|$, $|D|$ weights, and Lemma B0′, uniform in M, applies to each of them. So, given S_E, the two signed sums are jointly

normal with covariance $\sigma^2(\Sigma\varepsilon_i\varepsilon_i' - (\Sigma\varepsilon_i)(\Sigma\varepsilon_i')/m)$, hence correlation $r = O(m^{-1/2+\eta})$. The joint density divided by the product of the marginals is $1 + O(m^{-1/5})$. Typical pairs have $|x|$, $|y|$, $|x \triangle y|$ within $m^{3/5}$ of $m/2$, so each class has at least $m/8$ elements; summing the four Gaussian approximations over c gives the Gaussian density of the pair, and the errors of Lemma B0′ become relative errors because $\lambda \geq \lambda_m/2$ by (1).

- (3) Chebyshev. □

Why the signed form, and the centre at $S_E/2$. For a target fixed independently of E the unsigned count does not concentrate: two overlapping subsets share the sum of their intersection, and the second moment exceeds the square of the first by a constant factor (about $1/\sqrt{1-\frac{1}{4}} \approx 1.15$ for a typical pair). With the centre at $S_E/2$ the common shift cancels.

Lemma B‴ (exact reduction to first moments). Let $\lambda_J(v,t) = E[\#\{\varepsilon \in \{\pm 1\}^J : \sum_J \varepsilon_i w_i = v\} | S_J = t]$. Then

$$E[Z(u)^2 \mid S_E = s] = \sum_{k=0}^{m} \binom{m}{k} \sum_{s_1} \frac{\Pr(S_{m-k} = s_1)\Pr(S_k = s - s_1)}{\Pr(S_m = s)} \lambda_{m-k}(u, s_1)\, \lambda_k(0, s - s_1).$$

Proof. Split a pair $(\varepsilon, \varepsilon')$ into the $m-k$ positions where the signs agree and the k where they differ. The conditions $\Sigma\varepsilon w = \Sigma\varepsilon' w = u$ become: the agreeing part sums to u and the differing part to 0. Given their weight sums, the two parts are independent. □

Proposition B′ (the constant of the second moment). Fix $M \geq 2$ and $g > 0$, and put $z = (s - m\mu)/(\sigma\sqrt{m})$. Uniformly in $|z| \leq g$ and $|u| \leq M$ with $u \equiv s \pmod 2$,

$$\frac{E[Z^2 \mid S_E = s]}{\lambda^2} = 1 + \frac{1 + \mathbf{1}[u=0]}{\lambda} + \frac{\kappa_c}{m} + \frac{\gamma_M z}{m^{3/2}} + O_{M,g}(m^{-2}),$$

and, conditioning only on the event $S_E \equiv u \pmod 2$, with $\bar{Z}_u = E[Z \mid S_E \equiv u \pmod 2]$,

$$\frac{E[Z^2 \mid S_E \equiv u \pmod 2]}{\bar{Z}_u^2} = 1 + \frac{1 + \mathbf{1}[u=0]}{\bar{Z}_u} + \frac{\kappa}{m} + \frac{c_2(M,u)}{m^2} + O_M(m^{-3}), \qquad c_2(M,u) = c_2(M,0) - \frac{2\kappa u^2}{Ew^2},$$

with

$$\kappa = \frac{\mathrm{Var}(w^2)}{4(Ew^2)^2} \to \frac{1}{5}, \qquad \kappa_c = \frac{\mathrm{Var}(w^2) - \mathrm{Cov}(w^2, w)^2/\mathrm{Var}(w)}{4(Ew^2)^2} \to \frac{1}{80},$$

$$\gamma_M = \frac{\sigma\,(B - 2\beta A)}{4(Ew^2)^2} \to -\frac{\sqrt{3}}{80}, \quad A = \mathrm{Var}(w^2) - \frac{\mathrm{Cov}(w^2,w)^2}{\mathrm{Var}(w)},\ \beta = \frac{\mathrm{Cov}(w,w^2)}{\mathrm{Var}(w)\,Ew^2},\ B = \frac{E[R^2(w-\mu)]}{\mathrm{Var}(w)},$$

where $R = w^2 - Ew^2 - (\mathrm{Cov}(w^2,w)/\mathrm{Var}(w))(w - \mu)$.

Proof. Draw signs $\varepsilon, \varepsilon' \in \{\pm 1\}^E$ uniformly and independently of the weights. Then, with $\lambda = E[Z \mid S_E = s]$

as in Lemma B, $\lambda \cdot \Pr(S_E = s) = 2^m \Pr(\Sigma Y_i = (s,u))$ with $Y_i = (w_i, \varepsilon_i w_i)$, and $E[Z^2 | S_E = s] \cdot \Pr(S_E = s) = 4^m \Pr(\Sigma X_i = (s,u,u))$ with $X_i = (w_i, \varepsilon_i w_i, \varepsilon_i' w_i)$. So the three quantities in the ratio are local probabilities of sums of m independent copies of a bounded lattice vector, in dimensions 3, 1 and 2. X has mean $(\mu, 0, 0)$, covariance $\mathrm{diag}(\sigma^2, Ew^2, Ew^2)$, and generates the lattice $\{a \equiv b \equiv c \pmod 2\}$ of index 4; Y generates $\{a \equiv b \pmod 2\}$, of index 2. The Edgeworth expansion for sums of independent lattice vectors [BR, §22], with all moments available, gives each probability as the lattice index times the Gaussian density times $1 + \sum_{r=1}^{3} m^{-r/2} P_r + O(m^{-2})$. The indices $(4 \cdot 1/2^2)$ and the Gaussian factors cancel exactly in the ratio. In standardized coordinates the cumulants that involve the sign coordinates are $\kappa_{122} = \kappa_{133} = \mathrm{Cov}(\mathrm{w}, w^2)/(\sigma E w^2)$, $\kappa_{2222} = \kappa_{3333} = Ew^4/(Ew^2)^2 - 3$, $\kappa_{1122} = \kappa_{1133}$ and $\kappa_{2233} = Ew^4/(Ew^2)^2 - 1$; the others of order three vanish because εw is symmetric and ε, ε' are independent.

- Order $m^{-1/2}$: P_1 is $-\kappa_{122} z$ for X, $-\kappa_{122} z/2$ for Y and 0 for S_E, and these cancel in $P_1(X) + P_1(S_E) - 2P_1(Y)$.
- Order m^{-1}: every fourth-cumulant term cancels except $6\kappa_{2233}/24 = \kappa$. The third-cumulant terms of P_2 give $\kappa_{122}^2 (z^2 - 1)/4$, and the expansion of the logarithm gives $-\frac{1}{2}(P_1(X)^2 + P_1(S_E)^2 - 2P_1(Y)^2) = -\kappa_{122}^2 z^2/4$. The sum is $\kappa - \kappa_{122}^2/4 = \kappa_c$, independent of z.
- Order $m^{-3/2}$: the term is odd in z, and its coefficient is γ_M: it is the linear change in z of the conditional variance of Σw^2 given S_E, $(\sigma B / \mathrm{Var}(w)) z/\sqrt{m}$ relative to mA, against the shift of the conditional mean of Σw^2, which enters squared.
- The pairs $\varepsilon' = \pm\varepsilon$ are atoms of the lattice probabilities; written out they give exactly the term $(1 + 1[u = 0])/\lambda$. For fixed M it is exponentially small in m, and we keep it because it is not when M is exponential in m.

The unconditional statement is the same computation with $X = (\varepsilon w, \varepsilon' w)$, lattice index 2 on the event $S_E \equiv u \pmod 2$, and $Y = \varepsilon w$: the term of order m^{-1} is $\kappa_{1122}/4 = \kappa$, the odd orders vanish because both laws are symmetric, and $c_2(M,u)$ collects the terms of order m^{-2}; its dependence on u comes from the cross fourth-cumulant term, $c_2(M,u) = c_2(M,0) - 2\kappa u^2/Ew^2$. □

Heuristically, the sub-block first moments are proportional to $(\Sigma w_i^2)^{-1/2}$. Without conditioning, the fluctuation of Σw^2 gives κ/m, the δ_n^2 of [BCP, Theorem 4.1]. Conditioning on S_E removes the part of Σw^2 explained by Σw: in units of M, $4/45 - 1/12 = 1/180$, and $\kappa_c = (1/180)/(4/9) = 1/80$. Measured at $m = 12$, $M = 32$ over about 10^6 instances: unconditional excess $+0.013$ (prediction 0.017 at $M = \infty$), conditional excess -0.003 ± 0.001 against a prediction of $+0.001$. The conditional value disagrees with the leading term by about 4 standard errors: at this m the next order, of size about $-0.4/m^2$, dominates. In either case, given S_E the variance of the count is close to that of a Poisson variable. An exact check, by convolution and without sampling (`b_prime_exact.py`): the unconditional remainder times m^2 is 0.25, 0.28 and 0.30 for $M = 16 (m = 80)$, 64 $(m = 80)$ and 256 $(m = 40)$, against $c_2 = 0.246$, 0.274, 0.281 computed from the series (the values $c_2(M,0)$; at $M = 16$ the exact remainder times m^2 is -0.008 at $u = 8$ and -0.771 at $u = 16$, against -0.009 and -0.772); conditionally, at the centre and $M = 16$, the remainder times m^2 is 0.017 for $m = 28$, 40, 56; at $z = 1$ and $z = 2$ the remainder is linear in z, and remainder $\cdot\, m^{3/2}/z = -0.0104$, -0.0128, -0.0138, -0.0145 at $m = 20$, 40, 80, 120, approaching $\gamma_{16} = -0.0171$ with a correction of order $m^{-1/2}$.

Proposition B‴ (the identity evaluated exactly). For moderate M the distributions of S_k are computed exactly by convolution, and Lemma B‴ gives the conditional second moment exactly over the whole window $|s - m\mu| \le 2\sigma\sqrt{m}$ and all $|u| \le M$:

M	m	$\min\lambda$	$\max \mathrm{Var}/\lambda^2$	$2/\min\lambda$
16	8	5.3	0.470	0.379
64	10	4.9	0.400	0.413
256	12	4.7	0.411	0.429
256	16	68	0.030	0.030
64	22	15 283	0.0014	0.0001

Table 11. For these M the per-round failure probability is $\le \frac{1}{2}$ as soon as $\lambda \ge 5$. Below $2/\lambda$ there is a floor that decays like κ_c/m.

6.4 The theorems

Theorem R (random weights). Let $w_1, \dots, w_n$ be independent and uniform in $\{1, \dots, M\}$, $M = 2^b$, fix $g \ge 0$, and let the target T be fixed or, more generally, a function of the weights of H and of S_E only. Choose m with $\lambda_m \ge \lambda^*$. The probability that $|T - S/2| \le S_H/2$ and the centre rule on H followed by meet-in-the-middle on E returns no solution is at most $2/\lambda^* + O(m^{-1/5}) + 2e^{-g^2/6}$, uniformly in $M \ge 2$. The time is $O(n \log n) + O(\sqrt{\lambda^* M} \cdot m^{5/4})$. Every returned solution is verified.

Proof. Lemma A places the residual within $M/2$ of $S_E/2$; since T depends on E only through S_E, so does u. S_E lies within g standard deviations of its mean with probability $\ge 1 - 2e^{-g^2/6}$ (Hoeffding). On that event, since H is independent of E, Lemma B(3) applies conditionally on S_E, and Lemma B(1) gives $\lambda \ge \lambda^*(1 - O(m^{-1/5}))$, so the failure probability is at most $2/\lambda^* + O(m^{-1/5})$. The bound is on the joint probability of the hypothesis on T and the failure. □

Theorem R′ (failure probability δ). Fix $r = \lceil \log_2(2/\delta) \rceil$ disjoint endgame blocks $E_1, \dots, E_r$ and let G be the remaining indices; assume $|T - S/2| \le S_G/2 - rmM$, with T fixed or a function of the weights of G and of the block sums only. In round j, first include every other block entirely or not at all, by a rule that looks only at the block sums, so that the target T_j left for $G \cup E_j$ satisfies $|T_j - (S_G + S_{E_j})/2| \le S_G/2 + M/2$ (adding whole blocks greedily achieves this). Then run the centre rule on G alone, whose weights are $\le M$, and meet-in-the-middle on E_j. With $g = \sqrt{6 \log(4r/\delta)}$, $\lambda_m \ge 8$ and m large enough that Lemma B(3) gives $\le \frac{1}{2}$ on the window, the probability that $|T - S/2| \le S_G/2 - rmM$ and the algorithm fails is $\le \delta$; the time is $O(\log(1/\delta) \cdot (n \log n + \sqrt{\lambda_m M} m^{5/4}))$.

Proof. By Lemma A with head G, the residual of round j satisfies $|u_j| \le M$. Given the σ-algebra generated by the weights of G and the block sums, the residual of every round is determined, and the internal configurations of the blocks are independent, each conditioned on its sum. So the failures are conditionally independent, each with probability $\le \frac{1}{2}$ on the event that all block sums are in the window, by Lemma B(1) ($\lambda \ge 6$ from $\lambda_m \ge 8$) and Lemma B(3). The complement of that event has probability $\le 2re^{-g^2/6} = \delta/2$. □

For $b = O(\log n)$ the time is polynomial. For $(\log n)^2 \ll b \ll n$ it is $2^{b/2 + o(b)}$. This is the rigorous form of the "no arcs, centre rule" variant of the ablation.

6.5 *Dense instances: the arcs and the almost divisors*

Lemma C. Let $d \geq 2$, $r(c) = \#\{x : \Sigma x_i w_i \equiv c \bmod d\}$ and, for $d' | d$, $k(d') = \#\{i : d' \nmid w_i\}$.

1. $|d \cdot r(c)/2^n - 1| \leq \sum_{a=1}^{d-1} \prod_i |\cos(\pi a w_i/d)| \leq \sum_{d'|d, d'>1} \varphi(d') e^{-\pi^2 k(d')/(2d'^2)}$.
2. If all but k weights are divisible by d and $2^k < d$, at least $d - 2^k$ classes are empty, and some arc a/d, $a \neq 0$, has $\prod_i |\cos(\pi a w_i/d)| \geq 1/(d-1)$.

Proof. (1) $r(c) = d^{-1} \sum_a e(-ac/d) \prod_i (1 + e(aw_i/d))$; for $a/d = a'/d'$ in lowest terms, a factor with $d' \nmid w_i$ has modulus $\leq \cos(\pi/d') \leq e^{-\pi^2/(2d'^2)}$. (2) The divisible weights do not move the residue; for an empty class the sum over $a \neq 0$ equals -1. □

Corollary C′ (via Bringmann–Wellnitz). Let X be a multiset of n weights with sum Σ_X and largest multiplicity μ_X. Suppose X is C_δ-dense in the sense of [BW, Def. 3.1], with $C_\delta = 1699200 \log(2n) \log^2(2\mu_X)$, and $(4 + 2C_g)\mu_X \max X \Sigma_X/n^2 \leq t \leq \Sigma_X/2$, with $C_g = 169920 \log(2\mu_X)$. For every $d' \leq D = \lfloor 4\mu_X \Sigma_X/n^2 \rfloor$ compute the exact class counts $r_{d'}(c)$ from the arcs a/d'. Then t is a subset sum if and only if $r_{d'}(t \bmod d') > 0$ for every $d' \leq D$.

Proof. Necessity is clear. By [BW, Theorems 4.1 and 4.3] there is $d \leq 4\mu_X \Sigma_X/n^2$ with $t \in \mathscr{S}_X$ if and only if $t \bmod d \in \mathscr{S}_X \bmod d$. For $t > \Sigma_X/2$ use complements. □

The decider costs $O(nD^2) = O(n^3)$ operations, against $\tilde{O}(n)$ for [BW]: it is simple and exact, not faster. [BW]'s almost divisors may have up to $\alpha \mu_X \Sigma_X/n^2$ non-divisible elements, far more than $\log_2 d$, so the reading is exact through the class counts, not through the size of the arc as in Lemma C(2).

Corollary C″ (constructor on dense instances). Under the hypotheses of Corollary C′ a solution is constructed in $O(n \cdot \Sigma_X) = O(n^4/(C_\delta \mu_X))$ operations.

Proof. C_δ-density means $n^2 \geq C_\delta \mu_X \max X$ [BW, Def. 3.1], so $\Sigma_X \leq n \cdot \max X \leq n^3/(C_\delta \mu_X)$, and the table of reachable sums over $[0, \Sigma_X]$, with backtracking, returns a subset of sum t. □

The constructor does not use the arcs: on dense instances the sums are polynomially bounded. A constructor through the class counts alone, by self-reduction, would need the hypotheses of [BW] to persist along the descent.

6.6 *Las Vegas on random instances*

The algorithm already has one-sided failure: outputs are verified, failure is reported, and Theorem R′ makes it $\leq \delta$ at a cost $\times \log(1/\delta)$. A strict Las Vegas algorithm on random instances would always answer, falling back to meet-in-the-middle on all weights after r rounds, and this would be affordable if the per-round failure were $p \leq 2^{-m/2}$. The measured tail is far below the Chebyshev bound. With $M = 64$ and S_E at the centre, $\Pr(Z = 0) = 0.0075, 0.0011, 1.2 \cdot 10^{-4}$ at $\lambda = 6.7, 12.7, 24.4$, roughly λ^{-3}; the fourth moment is $E(Z/\lambda - 1)^4 \approx (3.5 - 6.5)/\lambda^2$. A proved bound of this kind would suffice as soon as $m \geq (4/3)(b + O(\log m))$. The tail is not Poisson ($e^{-\lambda}$ would be $3 \cdot 10^{-6}$ at $\lambda = 12.7$), and it is not dominated by exact modular obstructions. Moments of fixed order around λ cannot give an exponential tail, since the relative variance has the floor

κ_c/m of Proposition B′; the proof below goes through the Fourier integral instead.

Theorem LV (exponential tail of the endgame count). In the setting of Lemma B, with $M = 2^b$, $|s - m\mu| \le g\sigma\sqrt{m}$ and $|u| \le M$, $u \equiv s \pmod 2$, there are m_0 and C_g such that for $m \ge m_0$

$$\Pr(Z(u) = 0 \mid S_E = s) \le C_g\big(m^{5/2}M^5\rho^m + 2^{-0.53m}\big), \qquad \rho = \sup_{M\ge 2}\ \sup_{\|\theta\|\ge\frac{1}{2M},\ \|\theta-\frac12\|\ge\frac{1}{4M}} E\cos^4(\pi w\theta) \le 0.446.$$

The bound $\rho \le 0.44495$, uniformly in $M \ge 2$, is computer-assisted (`lv_rate_certified.py`); it enters the proof only through the last item.

Hence the per-round failure is at most $2^{-m/2}$ as soon as $m \ge 7.6(b + O(\log m))$, and the algorithm that falls back to meet-in-the-middle on all weights after the rounds is Las Vegas on random instances.

Proof. Put $a = (s+u)/2$ and $V = \sum_E w_i^2$. Write

$$Z(u) = \int_0^1 F(\theta)e(-a\theta)\,d\theta = 2^m \int_{-1/2}^{1/2} \cos(\pi u\theta)\,P(\theta)\,d\theta, \qquad P(\theta) = \prod_{i\in E}\cos(\pi w_i\theta),$$

which holds because $s \equiv u \pmod 2$. The integrand is even. Split $[0, \frac12]$ into three parts:

- the major arc $[0, 1/(2M)]$;
- the rest $\mathfrak{B} = [1/(2M), \frac12 - 1/(4M)]$;
- the parity window $[\frac12 - 1/(4M), \frac12]$.

Put

$$\lambda_0(w) = 2^m\Big(\frac{2}{\pi V}\Big)^{1/2}, \qquad \lambda_{\min} = 2^m\Big(\frac{2}{\pi m M^2}\Big)^{1/2}.$$

Since $V \le mM^2$, $\lambda_0(w) \ge \lambda_{\min}$.

- **Good weights.** Let $k_1 = \#\{i : w_i > M/2\}$ and $j = \#\{i : w_i \text{odd}\}$, and let $G = \{k_1 \ge m/20, j \ge m/20\}$. On G, $V \ge k_1M^2/4 \ge mM^2/80$.
  - *Case $M \ge 4$.* Write $M = 2M'$ and $w_i = M'b_i + y_i$, where $b_i = 1[w_i > M/2]$ and y_i is uniform in $\{1,\dots,M'\}$, independent of b_i. Given (b_i), $S_E = M'k_1 + \sum_i y_i$ is a sum of m independent uniform variables on $M' \ge 2$ values. By Lemma B0′, every value has probability at most $13/(\sigma'\sqrt{m})$, with $\sigma'^2 = (M'^2 - 1)/12$. Hence

$$\Pr(k_1 < m/20,\ S_E = s) \le 2^{-m}\#\{b \in \{0,1\}^m :\ 20|b| < m\} \cdot \frac{13}{\sigma'\sqrt{m}}.$$

The same holds for j, with $w_i = 2y_i - 1 + p_i$, where $p_i \in \{0,1\}$ records the parity and y_i is uniform in $\{1,\dots,M'\}$. Given (p_i), $S_E = 2\sum_i y_i - \#\{i : p_i = 0\}$.

- *The count.* With $x = 23/20$,

$$\#\{b \in \{0,1\}^m : \ 20|b| < m\} \le \sum_b x^{m-20|b|} = \big(x(1+x^{-20})\big)^m \le (5/4)^m$$

So each of the two events has probability at most $(5/8)^m \cdot 13/(\sigma'\sqrt{m})$.

- *Dividing by* $\Pr(S_E = s)$. In the window, $\Pr(S_E = s) \ge e^{-g^2/2}/(2\sqrt{2\pi}\sigma\sqrt{m})$ once $17\sqrt{2\pi}e^{g^2/2}/m \le \frac{1}{2}$, by Lemma B0′. Also $\sigma \le 3\sigma'$. Hence

$$\Pr(G^c \mid S_E = s) \le 156\sqrt{2\pi}\, e^{g^2/2}(5/8)^m \le 156\sqrt{2\pi}\, e^{g^2/2} 2^{-0.53m},$$

because $5/8 \le 2^{-0.53}$.

- *Case* $M = 2$. Each w_i is 1 or 2, so $S_E + j = 2m$. In the window, $s \le 3m/2 + g\sqrt{m}/2$, so $j \ge m/4 \ge m/20$ once $m \ge 4g^2$. Also $k_1 = m - j = s - m \ge m/2 - g\sqrt{m}/2 \ge m/4$, and $V \ge m = mM^2/4$. So G holds on every instance with $S_E = s$, and $\Pr(G^c | S_E = s) = 0$.

- **Major arc, on** G**.** With $x_i = \pi w_i \theta$ we have $|x_i| \le \pi M\theta$.

  - On $0 \le \theta \le 1/(\pi M)$, $|x_i| \le 1$. There $e^{-x^2/2}(1 - x^4/4) \le \cos x \le e^{-x^2/2}$, by $1 - x^2/2 \le \cos x$ and $1 + x^2/2 \le e^{x^2/2}$. Writing $g(\theta) = e^{-\pi^2 V\theta^2/2}$, this gives

$$|g(\theta) - P(\theta)| \le \frac{\pi^4\theta^4}{4} \sum_i w_i^4 \cdot g(\theta).$$

- On $1/(\pi M) \le \theta \le 1/(2M)$, $|x_i| \le \pi/2$ and $0 \le P \le g$. Here $\cos x \le e^{-x^2/2}$ holds for $|x| \le 2$, from Taylor's bound for $|x| \le 1$ and $\cos x = 2\cos^2(x/2) - 1$.

- We compare with Gaussian integrals rather than with $\lambda(w)$. Use $\cos(\pi u\theta) \ge 1 - \pi^2 u^2\theta^2/2$, $\int_0^\infty \theta^2 g = (2\pi^2 V)^{-1}\int_0^\infty g$, $\int_0^\infty \theta^4 g = 3(\pi^2 V)^{-2}\int_0^\infty g$, and $\int_{\theta_0}^\infty g \le g(\theta_0)/(\pi^2 V\theta_0)$. With $\theta_0 = 1/(\pi M)$ this gives

$$2^{m+1}\int_0^{1/(2M)} \cos(\pi u\theta)P(\theta)\,d\theta \ \ge \ \lambda_0(w)\Big(1 - \frac{u^2}{2V} - \frac{3\sum_i w_i^4}{4V^2} - \eta_3\Big), \qquad \eta_3 = \frac{3\sqrt{2}\,e^{-V/(2M^2)}}{\sqrt{\pi V/M^2}}$$

On G, $u^2 \le M^2$ and $V \ge mM^2/80$, so $u^2/(2V) \le 40/m$. Also $\Sigma w_i^4 \le M^2 V$, so $3\Sigma w_i^4/(4V^2) \le 60/m$. Since $V/M^2 \ge m/80$ and $e^{-t/2} \le 2/t$, $\eta_3 \le 6\sqrt{2}\cdot M^2/V \le 480\sqrt{2}/m$.

- For $m \ge 12000$ the bracket is $\ge 3/4$. So the major arc, counted on both sides of 0, is at least $(3/4)\lambda_0(w) \ge (3/4)\lambda_{\min}$.

- **Parity window, on** G**.** For w even, $|\cos(\pi w(\frac{1}{2} - \delta))| \le 1$. For w odd it equals $|\sin(\pi w\delta)| \le \sin(\pi/4)$, since $0 \le w\delta \le \frac{1}{4}$. So $|P| \le 2^{-j/2} \le 2^{-m/40}$ on a window of total length $1/(2M)$. The window therefore contributes at most

$$2^m 2^{-j/2}/(2M) \le \lambda_{\min}/4.$$

This holds as soon as $2^j \ge 2\pi m$. Since $m \le 20j$, it follows from $40\pi j \le 2^j$, which is true for $j \ge 12$.

- **The rest.** Let $R(w)$ be the integral of $|P|$ over $\mathfrak{B}$, so the rest contributes at most $2^{m+1}R(w)$ in absolute value.
  - By Jensen on an interval of length $L \le \frac{1}{2}$, $R^4 \le L^3 \int_{\mathfrak{B}} P^4$.
  - By independence, $EP(\theta)^4 = (E\cos^4(\pi w\theta))^m \le \rho^m$ on $\mathfrak{B}$.
  - So $ER^4 \le \rho^m/16$. By Markov,

$$\Pr\big(2^{m+1}R(w) \ge \lambda_{\min}/4\big) \le \frac{\rho^m/16}{(\lambda_{\min}/2^{m+3})^4} = 64\pi^2 m^2 M^4 \rho^m.$$

- Dividing by $\Pr(S_E = s) \ge e^{-g^2/2}/(2\sqrt{2\pi}\sigma\sqrt{m})$, and using $\sigma \le M$, gives at most $128\pi^2\sqrt{2\pi}e^{g^2/2}m^{5/2}M^5\rho^m$ conditionally.

On G, and outside the event of the last item, we get

$$Z(u) \;\ge\; \tfrac{3}{4}\lambda_0(w) - \tfrac{1}{4}\lambda_{\min} - 2^{m+1}R(w) \;>\; \big(\tfrac{3}{4} - \tfrac{1}{4} - \tfrac{1}{4}\big)\lambda_{\min} \;>\; 0.$$

Hence, for $m \ge \max(12000, 4g^2, 34\sqrt{2\pi}e^{g^2/2})$,

$$\Pr(Z(u) = 0 \mid S_E = s) \le 2\sqrt{2\pi}\, e^{g^2/2}(78 + 64\pi^2)\big(m^{5/2}M^5\rho^m + 2^{-0.53m}\big).$$

The rate follows from $0.446^{19} \le 2^{-22}$: the per-round failure is at most $2^{-m/2}$ as soon as $m \ge 7.6(b + \log_2 m)$ and $m \ge m_0$. □

The constant 7.6 is far from the $4/3$ that the measured tail λ^{-3} would give: the proof spends the factor M^5 to turn a fourth moment into a tail. Theorem R″ below keeps the one-sided failure with a per-round success probability at least $(1 - o(1))/(6 \ln \mathrm{mM})$, at a much smaller block.

6.7 The filtered endgame: Theorem R″

Setting.

- It is that of Lemma B: a block E of m indices fixed in advance, with weights independent and uniform in $\{1, \dots, M\}$.
- $S_E = s$ lies in the window $|s - m\mu| \le g\sigma\sqrt{m}$. The residual u satisfies $|u| \le M$ and $u \equiv s \pmod 2$, and $a = (s+u)/2$.
- All probabilities and expectations are conditional on $S_E = s$.
- A search scheme is a family $\mathscr{P}$ of ordered pairs (y, z) of disjoint subsets of E, fixed in advance. The endgame of §5.6 uses $\mathscr{P}_1 = \{(x \cap A, x \cap B) : x \subseteq E\}$ for fixed halves A, B with $|A| = h = \lfloor m/2 \rfloor$.

Put

$$N = \#\{(y,z) \in \mathscr{P} : S_y + S_z = a\}, \qquad Z_r = \#\{(y,z) \in \mathscr{P} : S_y + S_z = a,\ S_y \equiv r \pmod P\}, \qquad Y_P = \sum_r Z_r^2.$$

For $\mathscr{P}_1$, $N = Z(u)$ of Lemma B. A round draws r uniformly in $\mathbb{Z}/P$, given the weights and P, and succeeds if and only if $Z_r > 0$.

Lemma J (averaging over the residues). For every finitely supported law of the weights and of P,

$$\Pr(\text{round succeeds}) \ \geq\ \frac{(\mathbb{E}N)^2}{\mathbb{E}[P \cdot Y_P]}.$$

Proof.

- Given the weights and P, the round succeeds with probability $|R_P|/P$, where $R_P = \{r : Z_r > 0\}$.
- By Cauchy–Schwarz, $N = \sum_{r \in R_P} Z_r \leq |R_P|^{1/2} Y_P^{1/2}$. Hence $|R_P|/P \geq N^2/(P \cdot Y_P)$, with $0/0 = 0$, since $N = 0$ if and only if $Y_P = 0$.
- The function $(z,t) \mapsto z^2/t$ on $t > 0$, extended by 0 at (0, 0) and by $+\infty$ elsewhere on $t = 0$, is convex and lower semicontinuous: it is the closure of the perspective of z^2.
- Jensen's inequality gives $E[N^2/(P \cdot Y_P)] \geq (EN)^2/E[P \cdot Y_P]$. □

The lemma uses nothing about the scheme or the weights.

Lemma K (a random prime). Let P be uniform among the primes of $[P_0, 2P_0)$, with $P_0 \geq 100$ and P independent of the weights. For every integer $\Delta \neq 0$ with $|\Delta| \leq X$,

$$\mathbb{E}\big[P \cdot \mathbf{1}[P \mid \Delta]\big] \leq 6 \ln X.$$

Proof.

- Δ has at most $\ln X / \ln P_0$ prime factors $\geq P_0$.
- From $\pi(y) > y/\ln y$ for $y \geq 17$ and $\pi(y) < 1.25506 \cdot y/\ln y$ [RoS] it follows that $\pi(2P_0) - \pi(P_0) \geq P_0/(3 \ln P_0)$ for $P_0 \geq 100$. The ratio is 0.483 at $P_0 = 100$ and increases; the inequality was also checked directly on $[100, 2 \cdot 10^4]$ and at 10^5, 10^6 and 10^7.
- So $\Pr(P|\Delta) \leq 3 \ln X / P_0$, and $P < 2P_0$. □

Proposition K′**.** Let $D_{eq} = E\#\{(\tau, \tau') \in \mathscr{P}^2, \tau \neq \tau'$, both hitting a, $S_y = S_{y'}\}$. Then

$$\mathbb{E}[P \cdot Y_P] \leq 2P_0\,(\mathbb{E}N + D_{\mathrm{eq}}) + 6 \ln(mM)\, \mathbb{E}N^2,$$

and hence

$$\Pr(\text{round succeeds}) \geq \Big[\frac{2P_0}{\mathbb{E}N}\Big(1+\frac{D_{\text{eq}}}{\mathbb{E}N}\Big)+6\ln(mM)\,\frac{\mathbb{E}N^2}{(\mathbb{E}N)^2}\Big]^{-1}.$$

Proof.

- $P\cdot Y_P=\Sigma P\cdot 1[S_y\equiv S_{y'}\pmod P]$, summed over ordered pairs of representations hitting a.
- Pairs with $S_y=S_{y'}$, the diagonal included, contribute at most $2P_0$ each.
- For the other pairs $|S_y-S_{y'}|\leq mM$. Lemma K applies because P is independent of the weights.
- There are at most N^2 pairs. □

Lemma V (equal half-sums). For $\mathscr{P}_1$, let $\psi=2^{m/2}/M$. Uniformly in $M\geq 2$ and in the window, for $m\geq m_0(g)$,

$$\frac{D_{\text{eq}}}{\lambda}\leq 0.72\,e^{g^2/2}\,m\,(3\psi+\psi^2)\,\big(1+O(m^{-1/5})\big),\qquad \lambda=E[Z(u)\mid S_E=s].$$

Proof. Three facts are used.

- (T1) For a non-empty class C of weights, $\max_v\Pr(S_C=v)\leq 1/M$, and $1/M<1/(\sqrt{12}\sigma)$.
- (T2) Let $X_1,\ldots,X_K$ be the sums of disjoint classes, and Λ an integer $d\times K$ matrix whose columns at d non-empty classes form a non-singular submatrix. Then $\Pr(\Lambda X=b)\leq\prod_t\max_v\Pr(X_{j_t}=v)$: given the other sums, at most one vector of pivot sums solves the system.
- (T3) In the window, $\Pr(S_E=s)\geq c_1/(\sigma\sqrt{m})$ with $c_1=e^{-g^2/2}/\sqrt{2\pi}-O(1/m)$, by Lemma B0′.

A pair (x,x') splits E into 8 classes, by membership in x and x' (patterns 11, 10, 01, 00) and by side. The event is $\{S_x=a,S_{x'}=a,S_{x_A}=S_{x'_A},S_E=s\}$. Since the weights are ≥ 1, two classes required to have equal sums cannot have exactly one of them empty.

- *Case $x_A=x'_A$, $x_B\neq x'_B$.*
  - The event is $\{S_x=S_{x'}=a,S_E=s\}$. It is impossible if exactly one of the classes $(10,B)$, $(01,B)$ is empty.
  - Otherwise take as pivots $(10,B)$, $(01,B)$ and one of $(11,A)=x_A$, $(00,A)=A\setminus x_A$. In the rows $(S_x,S_{x'},S_E)$ their columns are (1, 0, 1), (0, 1, 1) and (1, 1, 1) or (0, 0, 1), with determinant -1 or 1.
  - The probability is $\leq M^{-3}$, and there are at most $2^h\cdot 4^{m-h}$ such pairs.
- *Case $x_B=x'_B$, $x_A\neq x'_A$.* Symmetric: probability $\leq M^{-3}$, at most $2^{m-h}\cdot 4^h$ pairs.
- *Case $x_A\neq x'_A$ and $x_B\neq x'_B$.*
  - The event is $\{X_{10A}=X_{01A},X_{10B}=X_{01B},S_x=a,S_E=s\}$, impossible unless the four classes $10A$, $01A$, $10B$, $01B$ are non-empty.
  - If one of $11A$, $11B$, $00A$, $00B$ is non-empty, call it Q. The pivots $(10A,10B,01A,Q)$ have determinant ± 1: probability $\leq M^{-4}$, at most 4^m pairs.

- Otherwise $x' = E \setminus x$, which is possible only for $u = 0$. Then $S_E = s$ follows from $S_x = a$, and the pivots $(10A, 10B, 01A)$ have determinant 1: probability $\le M^{-3}$, at most 2^m pairs.

Summing. $D_{eq} \le [(2^{2m-h} + 2^{m+h} + 2^m)M^{-3} + 4^m M^{-4}]/\Pr(S_E = s)$. By Lemma B(1), $\lambda = \lambda_m(1 + O(m^{-1/5}))$ with $\lambda_m = 2^{m+1}/(M\sqrt{2\pi m c_M})$ and $c_M \le 5/8$. With (T3) and $\sigma < M/\sqrt{12}$,

$$\frac{D_{\mathrm{eq}}}{\lambda} \le \big(1 + O(m^{-1/5})\big)\, \frac{\sqrt{2\pi c_M}}{2\sqrt{12}\, c_1}\, m\Big[\frac{2^{m-h} + 2^h + 1}{M} + \frac{2^m}{M^2}\Big].$$

Finally $2^{m-h} + 2^h + 1 \le 3 \cdot 2^{m/2}$ and $\sqrt{2\pi \cdot 5/8} \cdot \sqrt{2\pi}/(2\sqrt{12}) \le 0.72$. □

Theorem R″ **(filtered endgame).** Take the setting of Theorem R′ with R disjoint blocks in place of r, so that $|T - S/2| \le S_G/2 - RmM$, and fix $g \ge 0$ with $2e^{-g^2/6} < 1$ (say $g = 3$). In round j replace the meet-in-the-middle by the filtered endgame of §5.6 on E_j. The prime P_j is uniform among the primes of $[P_0, 2P_0)$, and r_j is uniform in $\mathbb{Z}/P_j$, both drawn independently of everything else. Assume $P_0 \ge 100$.

Then, given the weights of G and the block sums, the rounds fail independently. A round whose block sum lies in the window fails with probability at most $1 - p$, where

$$p = \Big[\frac{2P_0}{\lambda_m}\big(1 + 0.72\, e^{g^2/2} m(3\psi + \psi^2)\big) + 6\ln(mM)\Big(1 + \frac{2}{\lambda_m} + O(m^{-1/5})\Big)\Big]^{-1}\big(1 - O(m^{-1/5})\big).$$

The expected cost of a round is $O(m \cdot 2^{m/4} + m \cdot 2^{m/2}/P_0)$.

A round whose block sum leaves the window counts as failed; this happens with probability at most $2e^{-g^2/6}$, independently across blocks.

Now choose $P_0 = \lfloor \lambda_m/2 \rfloor$ and m the least positive integer with $2^m \ge (\sigma\sqrt{m})^{4/3}$ (by minimality also $2^m \le 2(\sigma\sqrt{m})^{4/3}$), so $m = (4/3)b + (2/3)\log_2 b + O(1)$. Then:

- $p \ge (1 - o(1))/(6\ln(mM))$, and $2^{m/4} \asymp 2^{m/2}/P_0 \asymp (\sigma\sqrt{m})^{1/3}$;
- with $R = \lceil \ln(1/\delta)/((1 - 2e^{-g^2/6}) \cdot p) \rceil$ rounds, the probability that $|T - S/2| \le S_G/2 - RmM$ and the algorithm fails is at most δ, provided $n \ge |G| + Rm$;
- it runs in time $O(\log(1/\delta) \cdot \log(nM) \cdot (n \log n + M^{1/3} \cdot \mathrm{poly}(b)))$. For $\log n = o(b)$ and $\log(1/\delta) = 2^{o(b)}$ this is $2^{b/3 + o(b)}$.

Proof.

- As in Theorem R′, the weights of G and the block sums determine the residual of every round. The internal configurations of the blocks are independent, each conditioned on its sum, and the filters are drawn independently, so the failures are conditionally independent.
- In a round, Proposition K′ applies. By Lemma B, $EN = \lambda = \lambda_m(1 + O(m^{-1/5}))$ and $EN^2 \le \lambda + \lambda \cdot 1[u = 0] + \lambda^2(1 + O(m^{-1/5}))$; Lemma V bounds D_{eq}. This gives p.
- The filtered lists have expected size $2^{|A|} \cdot E[1/P]$ and $2^{|B|} \cdot E[1/P]$, both $O(2^{m/2}/P_0)$, since r is uniform given P. They are built from the lists of the quarters, of $O(2^{m/4})$ elements, sorted by residue, and joined after sorting, which gives the cost per round.

- With the stated m, $2^{m/4} \asymp (\sigma\sqrt{m})^{1/3}$ and $2^{m/2}/P_0 \asymp M\sqrt{m}/2^{m/2} \asymp (\sigma\sqrt{m})^{1/3}$. Also $\lambda_m \asymp M^{1/3}m^{1/6} \geq 200$ for M large, and $m\psi \asymp m^{4/3}M^{-1/3} \to 0$.
- The block sums of disjoint blocks are independent, and Lemma A needs no window. By Hoeffding each block sum lies in the window with probability at least $1-2e^{-g^2/6}$, and then the round succeeds with probability at least p. The failure probability is therefore at most $(1-(1-2e^{-g^2/6})p)^R \leq \delta$.
- Since g is fixed, Lemmas B and V apply with their fixed-g constants. In p, $6\ln(mM)\cdot O(m^{-1/5}) = o(\ln mM)$ and $e^{g^2/2}m\psi \to 0$, which gives $p \geq (1-o(1))/(6\ln mM)$. □

What is new, and what is not.

- The four-list method with a modular constraint is classical [SS; Wag]. It has rigorous analyses in the modular model [Lyu; Sha] and for k-SUM over the integers [LV], and Lyubashevsky's $2^{O(b/\log n)}$ remains better asymptotically.
- Theorem R″ adds a short proof for the filtered endgame as implemented in the theorem mode of §5.6, in the model of this paper; the practical solver, with its suffix block and repeated rounds, is outside it. Its constant $6\ln(mM)$ is the price of using no distributional input in the filter. In the runs of Table 9′ the first round almost always succeeds. With P between $\lambda/2$ and λ, between 1 and 2 survivors are expected (about 1.39 on average), and the Poisson value is about 0.75. Lemma J with the heuristic coincidence rate $1/P$ gives at least $2/3$ at $P = P_0$.

Representations. Lemma J and Proposition K′ hold for any scheme, in particular for the representation schemes of [HGJ; BCJ]. In our endgame the density is a free parameter, and part of the multiplicity that representations create at density one is already supplied by the λ solutions of the block. Representation methods can still beat $2^{b/3}$: [BCJ] runs heuristically in $2^{0.291b}$ at density one. We do not use them here because no rigorous analysis of them exists in our model; supplying the moments of Proposition K′ for those schemes would be the way to obtain one.

7 Constraint propagation: structural collapse and convergence

The arcs read a set of weights through its quotients $\mathbb{Z}/q$. Seen from far enough, this is one instance of a general mechanism: propagation of constraints on quotients and on reachable sets. The algorithm of §5 is built from a few operations: a forward pass that computes reachable sets, residue counts on quotients, injections of decisions, a backward pass, and a final search. This section develops the general theory of these operations on an arbitrary network of constraints; §8 reads the algorithm of §5 inside it.

The idea in one line. A network of constraints splits into two parts. The *structural* part (equalities with parity, functional definitions on a DAG) is solved in one pass, without search. The *computational* part (assertions, cycles, non-bijective constraints read backwards) needs iteration to a fixpoint, and that is where the whole cost lives.

7.1 The model

- Variables V, domains $W : V \to \mathscr{P}(\mathrm{Val})$, order $W' \sqsubseteq W \iff \forall v W'(v) \subseteq W(v)$. In the full state $s =$ (facts, domains, clusters) the information order is $s \preceq s' \iff F(s) \subseteq F(s') \wedge \forall v D_{s'}(v) \subseteq D_s(v)$.

- Constraints fall into three families.
  - C_{par} (binary bijective): $x = y$, $x = \neg y$, $x \oplus y = b$.
  - **Def** (functional definitions): $x := f_x(\text{pred}(x))$, with pred(x) among the nodes preceding x in a fixed order. The skeleton is the **complete DAG** (transitive tournament): every node may read any earlier node, and a concrete network is a choice of active arcs. Typical case: if-then-else, $x :=$ ite(c, a, b).
  - C_{comp} (the rest): assertions ("x must be 1"), non-bijective constraints read backwards (AND, OR, XOR of arity ≥ 3), links that close cycles.
- Free inputs I: V minus the variables defined by Def and minus those merged into a class with an input.
- Projection $\pi_v(c,W) = \{a \in W(v) | \exists \alpha(\alpha(v) = a \wedge c(\alpha) \wedge \forall u \in \text{vars}(c)\alpha(u) \in W(u))\}$; filter $T_c(W)(v) = \pi_v(c,W)$; support defect $\upsilon(c,W) = \max_v |W(v) \setminus \pi_v(c,W)|/|W(v)|$. ($\upsilon$ is a different object from the omicron $o(W)$ of the weights in §1–6, and the two are kept apart here.)

7.2 The structural part

Definition 7.1. $T_{\text{struct}} = \text{Eval} \circ \text{Par}$, where:

- Par is union-find with parity on C_{par} [Tarjan]. Every variable is written $x = r \oplus b_x$ with r the root of its class; there is one domain per class (a shared cell). The shared form is $\mathscr{S} = \{W | \forall x \sim y, W(x) = W(y) \oplus (b_x \oplus b_y)\}$.
- Eval is one pass over Def in a topological order of the parity classes (the graph of the definitions with the classes of Par contracted, assumed acyclic): $W(x) \leftarrow W(x) \cap f_x(\prod_{u \in \text{pred}(x)} W(u))$ (forward image). A topological order of the variables alone is not enough (Remark 7.2′).

Theorem 7.2 (structural collapse).

- (a) $T_{\text{struct}} \circ T_{\text{struct}} = T_{\text{struct}}$, and $T_{\text{struct}}(W) \in \mathscr{S}$.
- (b) For $W \in \mathscr{S}$ closed under Eval, $T_{\text{struct}}(W) = W$.
- (c) Injection: if $W \in \mathscr{S}$ and one variable is fixed, its parity class is updated in $O(1)$. If every input in I is fixed and no parity class contains two inputs, one pass of Eval leaves every domain with at most one value, in time $O(\sum_x \text{cost}(f_x))$. If moreover the network has a solution α with $\alpha(r)$ in the initial cell of every class, no domain becomes empty and every domain is exactly $\{\alpha(x)\}$. Without such a solution a cell can become empty even when no class contains two inputs: two definitions may write one class inconsistently ($x := u$, $y := \neg u$ with $x = y$; Lean: `Eval_blocks_counterexample`). Two inputs in one class may also be fixed inconsistently. In both cases the conflict appears as an empty cell, in $O(1)$.
- (d) For $W \in \mathscr{S}$ closed under Eval, $\upsilon(c,W) = 0$ for every $c \in C_{\text{par}}$ and every definition read forwards.

Proof. (a) On a parity class the intersection, with the bit applied, is the same for every member, and intersecting again changes nothing. In a topological order of the classes every class receives the images of parents whose classes are already final, so a second pass changes nothing. (b) Follows from (a). (c) An injection into

$\mathscr{S}$ intersects the cell of the class. With singleton inputs every image $f_x(\text{singletons})$ has at most one value, by induction on the topological order. A solution α with $\alpha(r)$ in every initial cell lies in every image and every intersection, so no cell becomes empty and each domain is $\{\alpha(x)\}$. (d) In $\mathscr{S}$ every value has support in the parity constraint; after Eval every value of x is an image of values of its parents. □

Scope. The collapse holds on the shared form, not for every $W' \geq T_{\text{struct}}(W_0)$, and it covers equalities, parities (NOT, binary XOR) and functional definitions on the complete DAG.

Remark 7.2′ **(the order of Eval).** With $x := u$, $z := w$ and the parity constraint $u = z$, the order x, z is topological on the variables, but the pass for z narrows the shared cell of u after x has read it, so a second pass changes x: T_{struct} is not idempotent. Ordering the classes, not the variables, removes the problem (Lean: `Tstruct_not_idem_counterexample`).

Theorem 7.3 (information limit). Let S be the set of solutions. If fixing one variable v determines all the others, then $|S| \leq 2$; the same holds for every extension by auxiliary variables that preserves the projection onto V. With definitions and assertions only, solutions correspond injectively to inputs: $|S| \leq 2^{|I|}$ (for Boolean Val).

Proof. If fixing v determines everything, $s \mapsto s_v$ is injective on S. An extension S' projecting onto S has $|S| \leq |S'| \leq 2$. With definitions only, the inputs determine everything (Theorem 7.2(c)). □

Reading. "Solve one, solve all" holds exactly on the components with a single free input. An auxiliary variable X built in the knowledge that the chain does not block is a certificate: checking it is forward propagation, finding it is search (the definition of NP). For Subset Sum, $o \leq 1$ (Lemma 2.1) is the case in which the certificate is built backwards in $O(n)$ (§2).

7.3 The computational part

Theorem 7.4. Let V and Val be finite, and assume every T_c, $c \in C_{\text{comp}} \cup \text{Def}$ (definitions also read backwards), is the generalized arc consistency filter.

- $(i)T_{\text{comp}}$ is decreasing and monotone: $T_{\text{comp}}(W) \sqsubseteq W$ and $W \sqsubseteq W' \to T_{\text{comp}}(W) \sqsubseteq T_{\text{comp}}(W')$.
- (ii) $T_{\text{comp}}(W) = W \iff \forall c\, \upsilon(c, W) = 0$.
- (iii) Let $\mu(W) = \sum_v (|W(v)| - 1)$, and let T_{fail} be T_{comp} with the state replaced by the failure state as soon as a domain becomes empty. If every domain of W is non-empty, then $T^m_{\text{fail}}(W) = T^{\mu(W)+1}_{\text{fail}}(W)$ for every $m \geq \mu(W) + 1$. $\mu(W)$ rounds do not suffice in general: one variable with $W = \{0\}$ and the assertion $x = 1$ has $\mu(W) = 0$ but $T_{\text{comp}}(W) \neq W$ (Lean: `Tcomp_bound_counterexample`).
- (iv) The limit W^* is the largest $W' \sqsubseteq W$ with $\upsilon(c, W') = 0$ for every c, and it does not depend on the order of the filters (chaotic iteration [Apt]).

Proof. (i) Each filter shrinks domains, and projection is monotone. (ii) $T_c(W) = W \iff \pi_v(c, W) = W(v)$. (iii) Each non-fixed step either decreases $\mu \in \mathbb{N}$ or empties a domain, after which the state is absorbing. (iv) Every $W' \sqsubseteq W$ fixed by all T_c stays below every iterate by monotonicity; the limit is fixed, hence it is the largest. □

Remark 7.5 (a natural bound that fails). A natural bound on the number of iterations is $\sum_c \upsilon(c, W_0) \cdot d$. It

fails. Take the chain $x_1 \neq \ldots \neq x_k$ with $D(x_1) = \{1\}$, $D(x_i) = \{i-1, i\}$. Then $\sum_c \upsilon(c, W_0) \cdot d = 1$, but $k-1$ iterations are needed in the worst order (simulated: $k = 3$, 5, 10, 20 give 2, 4, 9, 19). The initial defect does not see cascading prunings; the bound $\mu(W)$ of Theorem 7.4(iii) has the right order.

7.4 Factorization and algorithm

Theorem 7.6 (factorization). $T = T_{\text{comp}} \circ T_{\text{struct}}$.

- Par costs $O(|C_{\text{par}}| \cdot \alpha(n))$ with union-find; Eval costs one pass, $O(\sum_x \text{cost}(f_x))$.
- Every injection costs $O(1)$ on its class; with all inputs fixed, one pass of Eval.
- T_{comp} takes at most $\mu(W) + 1$ steps and returns W^*, independent of order.

Algorithm.

1. Union-find with parity on C_{par}.
2. Topological order of the parity classes of the definitions (complete DAG with the classes contracted, active arcs).
3. Injections: update the class, then Eval forwards.
4. AC-3-style worklist [Mackworth] on C_{comp} and on the definitions read backwards, to the fixpoint.
5. Outcome: an empty domain means unsatisfiable; all singletons means a solution; otherwise the network is locally consistent and search over the free inputs is needed (at most $2^{|I|}$ branches, Theorem 7.3).

7.5 Definitional variables

The aim is to determine as much as possible in one pass: once one variable (or a few) is fixed, how many others follow without search. The tool is the definitional variable $X := f(\text{nodes})$. It does not change the solutions, but it makes explicit a link that local propagation does not see.

Theorem 7.7. Let S be the set of solutions, with the values already fixed, and $R_{AB} = \pi_{A,B}(S)$. There is a definitional extension X such that fixing A determines B in one pass of T_{struct} if and only if R_{AB} is functional, $B = g(A)$ on S. A and B lie in the same class of Par if and only if g is a bijection.

Proof. ($\Leftarrow$)$X := g(A)$ is a definition; the link $B = X$ holds on every solution, so adding it is conservative; once A is fixed, Eval computes X and Par fixes B. ($\Rightarrow$) A definitional extension preserves the projection of the solutions: if two solutions have the same A and different B, no pass fixes B from A (Theorem 7.3). □

- X does not create links; it makes them evaluable. If R_{AB} is not functional (for instance A $\vee B = 1$), no X helps and search is needed.
- Theorem 7.7 is the network form of implicit definability: Beth's theorem turns implicit into explicit definability, Padoa's method tests it, and Craig interpolation extracts X from a proof, in the vocabulary common to the two sides [Beth; Craig]. In model counting, definable variables are detected and eliminated in [LLM].

- Finding g is hard in general: "B $= g(A)$ on every solution" is a universal statement over the solutions (coNP). The gain is where the link comes from structure that is cheap to read.

Algorithm Ext. Repeat until nothing changes:

1. **Linear X (Gauss over GF(2)).** Each XOR constraint, of any arity, and each known value is a row. The reduced form gives, for every pivot, $X := \oplus(\text{free variables}) \oplus b$. Rows with one variable are determined values; rows with two variables are parity classes of Par.
2. **Local X from definitions.** With one input known, a definition becomes linear: $\text{and}(a,1) \Longrightarrow y = a$; $\text{and}(a,0) \Longrightarrow y = 0$; dually for or; $\text{ite}(1,a,b) \Longrightarrow y = a$; $\text{ite}(c,a,a) \Longrightarrow y = a$; $\text{not}(a) \Longrightarrow y = \neg a$. These rows go to step 1.
3. **Arc consistency on the rest** (definitions not yet linear, also read backwards). New values return to step 1.

Proposition 7.8 (soundness). Every value fixed by Ext holds in every solution. *Proof.* The rows of step 2 are implied by the definitions with the known input; Gauss and arc consistency are sound; induction on the rounds. □

Proposition 7.9 (completeness in the affine case). If the network has only XOR constraints, Ext determines exactly the forced variables, and for every pair (A,B) it fixes B from A if and only if R_{AB} is functional: Theorem 7.7 becomes a polynomial algorithm. *Proof.* The solutions form an affine subspace. x is forced $\Longleftrightarrow$ e_x lies in the row space. In reduced form this happens $\Longleftrightarrow$ some row equals e_x: the only row with pivot x must appear in the combination, and no other. □

This is the affine class of Schaefer's dichotomy [Schaefer], and Gaussian elimination for XOR constraints is standard in SAT solving [SNC]. The point here is the reading: Gauss is the systematic construction of the X of Theorem 7.7, and it moves every XOR constraint from the computational to the structural part.

Probing: non-local X. For each free variable A, run Ext with $A = 0$ and with $A = 1$.

- If one branch fails, A is forced to the other value.
- If B is determined in both branches with the same value, B is forced.
- If B is determined in both branches and follows A, add X: $B = A \oplus c$ (a parity row).

The new rows return to Ext, and the procedure is repeated until nothing changes. One round costs $2n$ runs of Ext, which is polynomial. Probing (failed literals, equivalences by probing) is a known SAT preprocessing technique [LMS]; here it is read as a generator of definitional X.

Proposition 7.10 (soundness of probing). Every X found by probing holds on every solution. *Proof.* Every solution has $A = 0$ or $A = 1$. By Proposition 7.8 the values found in a branch hold on every solution of that branch, so the relation between A and B that is recorded holds on all of them. □

Proposition 7.10′ **(probing to depth ℓ).** Let $\text{Probe}_0 = \text{Ext}$, and let Probe_ℓ run $\text{Probe}_{\ell-1}$ on the network with $A = 0$ and with $A = 1$ for every variable A, already fixed or not, record what probing records, and, when one branch fails, also adopt the values fixed in the other branch; repeat until nothing changes. Then:

- (a) every X found by Probe_ℓ holds on every solution;
- (b) if each inner $\text{Probe}_{\ell-1}$ runs a single round, one round of Probe_ℓ costs at most $(2n)^\ell$ runs of Ext;
- (c) if some set B of at most ℓ variables is a strong backdoor for Ext [WGS] (for every assignment of B, Ext either fails or returns a solution), then Probe_ℓ fails if and only if the network has no solution, and otherwise fixes every variable that is constant on the solutions.

Proof. (a) Induction on ℓ: $\ell = 0$ is Proposition 7.8, and the step is the argument of Proposition 7.10 with $\text{Probe}_{\ell-1}$ in place of Ext; a failed branch has no solutions, so the values of the other branch hold on all of them. (b) A round makes $2n$ calls of $\text{Probe}_{\ell-1}$, each of one round by assumption. (c) Induction on ℓ. For $\ell = 0$, $B = \emptyset$ and Ext decides alone. For $\ell \geq 1$ take $A \in B$; in both branches $B \setminus \{A\}$ is a strong backdoor of size at most $\ell - 1$. If the network has no solution, neither branch has one, both fail by induction, and Probe_ℓ fails. Otherwise let X be constant, equal to c, on the solutions. A branch without solutions fails, and in a branch with solutions $\text{Probe}_{\ell-1}$ fixes $X = c$ by induction; in either case Probe_ℓ records $X = c$. □

One-level probing cannot be complete in general (it would decide SAT in polynomial time); probing to depth ℓ costs $(2n)^\ell$ and is complete exactly up to backdoors of size ℓ (Proposition 7.10′), and beyond them search remains. The three levels (Gauss, local linearization, probing) all produce correct definitional X, and none creates information that is not in the solutions (Theorem 7.7). The limit of what can be determined in one pass is exactly the functionality of R_{AB}, and the cost of reaching it is where the difficulty of the problem lives.

7.6 Convergence of the full operator

The propagation operator acts on states $s = (F, D, \text{Cl})$: facts, domains and clusters. The operator is $K = \text{Merge} \circ \text{Unary} \circ \text{Bridge} \circ \text{LocalDed}$:

- LocalDed applies inference rules inside each cluster;
- Bridge keeps a value of a variable only if some consistent assignment satisfies, together, all the constraints touching that variable (a filter stronger than arc consistency, close to a global solve on a complete graph);
- Unary applies the unary constraints;
- Merge merges clusters that share variables.

Its measure is $\nu(s) = |\text{Cl}(s)| + \sum_v |D_s(v)|$.

Lemma 7.11.

- L1: $F(Ks) = F(s)$. L2: $D_{Ks}(v) \subseteq D_s(v)$.
- L3: $\text{Cl}(Ks) = \text{Cl}(s)$, or $|\text{Cl}(Ks)| < |\text{Cl}(s)|$. (By induction on the merging loop: a step without over-

laps returns the same clusters; with an overlap the count drops.)

- L4: $K(s) \neq s \to \nu(Ks) < \nu(s)$.
- L5: $s \preceq s' \to K(s) \preceq K(s')$. L6: $s \preceq K(s)$.

Theorem 7.12 (termination). $\forall s_0 \exists n \leq \nu(s_0) \forall m \geq n K^m(s_0) = K^n(s_0)$. *Proof.* By L4, as long as there is no fixpoint ν drops strictly in $\mathbb{N}$; from a fixpoint the sequence stays put, by induction. □

Theorem 7.13 (closure). With $s^* = K^n(s_0)$: $\mathrm{Fix}(s^*) \wedge s_0 \preceq s^* \wedge \forall s'(\mathrm{Fix}(s') \wedge s_0 \preceq s' \to s^* \preceq s')$. The fixpoint reached is the least informed among the fixpoints above s_0: it invents nothing. *Proof.* By induction on k, $K^k(s_0) \preceq s'$ using L5 and $\mathrm{Fix}(s')$; with $k = n$, $s^* \preceq s'$. □

Remark 7.14 (a maximality statement that fails). The natural statement that a maximal fixpoint is reached is false. With one domain $\{0,1\}$, no constraints and no facts, the state with domain $\{0\}$ and fact $(t,g,0)$ is a fixpoint above s^*, but not below it. The correct statement is the closure of Theorem 7.13.

7.7 Tests

All tests compare with brute force on small networks; scripts are listed in Data and code availability.

Test	Result
Union-find with parity against brute force	0 errors on classes and bits
Eval on random if-else/AND/OR/XOR/NOT circuits on the complete DAG	0 non-singleton variables with inputs fixed; 0 violations of $\lvert S\rvert \leq 2^{\lvert I\rvert}$
Idempotence of T_{struct}	0 violations
T_{comp} under 100 filter orders	same W^* in every case; 0 cases with more than $\mu(W)$ steps
Chain of Remark 7.5, $k = 5$, 10, 20	favourable order 1 round; unfavourable order 4, 9, 19
Closure (Theorem 7.13)	0 solution values deleted
Injection, naive against shared cell	$n = 10$: 76 μs / 4.6 μs; 100: 565 μs / 4.6 μs; 1000: 14.8 ms / 5.3 μs; 10^4: 67 ms / 7.9 μs

Table 13. The generated networks have parity constraints between inputs only; the configuration of Remark 7.2′ (a parity class with members defined at different depths) does not occur in them.

Network ($n = 14$, 1–3 injected variables)	arc consistency alone	Ext	Ext+ probing	errors
XOR only	42.8%	**100%**	—	0
dense XOR	55.8%	**100%**	—	0
XOR+ definitions	53.2%	82.9–85.0%	**100%**	0
definitions + rare XOR	66.8%	61.0–70.2%	**99.9%**	0

Table 14. Fraction of the variables forced by brute force (excluding the injected ones) that each method determines.

- Proposition 7.9 in the affine case: "R_{AB} functional" $\Longleftrightarrow$ "Ext fixes B from A", 0 discrepancies over 8 778 pairs.
- Scale (XOR constraints with a planted solution, 1% of the variables injected): $n = 1000$, Ext determines 871 variables in 0.27 s against 16 for arc consistency; $n = 2000$, 1705 in 1.3 s against 45. All values agree with the planted solution.
- Where the other variables are genuinely free ($|S| > 2$ with one fixed), no method can determine them, by Theorem 7.3.

8 The Subset Sum algorithm as propagation

Model. Subset Sum is the network of §7.1 with inputs $x_i \in \{0,1\}$, definitions $s_0 = 0$, $s_i := s_{i-1} + w_i \cdot x_i$ (a chain in the complete DAG), and the assertion $s_n = T$.

Proposition 8.1 (forward pass = Eval). The forward pass $R_i = R_{i-1} \cup (R_{i-1} + w_i)$ of Remark 2.4 is exactly Eval on the chain: $R_i = W(s_i)$. *Proof.* Forward image of $s_{i-1} + w_i \cdot x_i$ with $x_i \in \{0,1\}$. □

Proposition 8.2 (one backward round suffices). After Eval, one round of the arc consistency filters from $i = n$ down to 1 reaches the fixpoint W^* of T_{comp}, and $W^*(s_i)$ is exactly the set of values reachable from 0 and extendable to T. Hence the backward choice "take w_i if $T - w_i \in R_{i-1}$" never gets stuck. *Proof.* Downward induction on i. If $W^*(s_i)$ is the set of reachable and extendable values, the filter of constraint i leaves in s_{i-1} the $v \in R_{i-1}$ with $v \in W^*(s_i)$ or $v + w_i \in W^*(s_i)$, that is, the reachable and extendable ones. Constraint i touches nothing else, and the constraints already processed stay consistent. □

This is a case of Freuder's theorem on acyclic networks [Freuder]. The bound $\mu(W)$ of Theorem 7.4(iii) is very pessimistic here: in the right order one round suffices, while in the unfavourable order up to about n rounds are needed, the same phenomenon as in Remark 7.5. The domains are unions of K_i intervals, so $T = T_{\text{comp}} \circ T_{\text{struct}}$ costs $O(n \cdot K_{\max})$. With $o \leq 1$, $K = 1$ and the cost is $O(n)$: the linear search of §2 is a corollary of Theorem 7.6.

Proposition 8.3 (the dummies are an extension that does not block but does not conserve). $W' = W \cup D$ makes $K = 1$ (the chain does not fragment), but $R(W') \supseteq R(W)$ with strict inclusion in general: the extension does not preserve the projection onto the original solutions, so it does not satisfy the hypothesis of Theorem 7.3. This is why a negative answer on W' is valid ($T \notin R(W') \Longrightarrow T \notin R(W)$, which is Theorem 7.13 applied to a relaxation), while the backward choice on the original variables alone can get stuck.

- The exact table R_i is a conservative extension that does not block, but its size is $n \cdot K$. The dummies have size $O(n + \log \max W)$, but they are not conservative. Having both properties with size polynomial in $(n, \log \max W)$ on every instance is **equivalent** to a polynomial algorithm for Subset Sum. This is an equivalence, not progress.

Proposition 8.4 (the arcs are Eval on a quotient). $N_c(q) > 0 \iff$ the class c survives Eval on the chain reduced mod q, $s_i \bmod q$. *Proof.* $N_c(q) = \#\{\text{subsets with sum} \equiv c \bmod q\}$, and Eval mod q computes the support of the same recurrence. □ An arc obstruction is a domain that empties in the quotient; by Theorem 7.13 (a relaxation deletes no solution) it is always correct.

Three semirings (Remark 2.4). The chain is evaluated in three semirings, linked by homomorphisms:

Semiring	Object	In the propagation framework
$\mathbb{B}[\mathbb{Z}]$ (union, Minkowski sum)	$R_n = \Pi(\{0\} \oplus \{w_k\})$, reachable sets	Eval $= T_{\text{struct}}$ (Proposition 8.1)
$\mathbb{N}[\mathbb{Z}/q]$	class counts $N_c(q)$	counting on the quotient; the support map $\mathbb{N} \to \mathbb{B}$ is Eval $\bmod q$ (Proposition 8.4)
$\mathbb{C}$, F_W at roots of unity	arcs a/q	Fourier transform of the counts (Remark 3.1)

Propagation lives in $\mathbb{B}$; the arcs live in $\mathbb{N}$ and $\mathbb{C}$ and project onto $\mathbb{B}$ through the support. Evaluating $\Pi(1+z^{w_i})$ at roots of unity to count subset sums is the mechanism of [LN]; the residue classes, the emerging arcs and the de-blurring are not there.

Proposition 8.5 (the blur is invertible on counts, not on Booleans). With dummies $1, 2, \ldots, 2^{k-1}$ ($M = 2^k$), $F_{W\cup D} = F_W \cdot (1-z^M)/(1-z)$, so

$$r_W(T) = \sum_{j\geq 0} \left[r_{W\cup D}(T - jM) - r_{W\cup D}(T - jM - 1) \right],$$

the blur identity of §5.4 solved by telescoping. In $\mathbb{B}[\mathbb{Z}]$ the factor $\{0, \ldots, M-1\}$ is not invertible (union has no inverses), and $R(W \cup D)$ does not determine $R(W)$.

- Proposition 8.3 (the dummies do not conserve) is a **Boolean** fact. Lifted from $\mathbb{B}$ to $\mathbb{Z}$ the extension becomes invertible, and this is what makes the de-blurring of §5.4 exact. The price is carrying the counts.

The solver of §5 in the framework.

Layer of the solver	In the framework of §7
Forward pass on $W \cup D$	Eval on a non-conservative relaxation (Proposition 8.3); negative certificate
Obstructions $N_c(q)$, $q \leq Q$	T_{comp} on quotient relaxations (Proposition 8.4), cost $O(nq)$
Gaussian estimate in the bulk	the only element outside propagation: a counting oracle, not a filter
Descent by self-reduction, Lemma A	injections (step 3) guided by the scores; Lemma A is a monovariant on $\lvert e_k \rvert$
Exact layers at the bottom	Propositions 8.1–8.2: exact Eval and one backward round, never stuck
Final meet-in-the-middle, Theorems R, R′	step 5: search over the free inputs of the block ($2^{m/2}$ instead of 2^m)
Filtered endgame (§5.6), Theorem R″	Eval on the quotient $\bmod P$ of each half (Proposition 8.4), then the search: $2^{m/4} + 2^{m/2}/P$
De-blurring (§5.4)	Proposition 8.5: inversion in $\mathbb{Z}$ of an extension that is not conservative in $\mathbb{B}$

Layer of the solver	In the framework of §7
Corollary C′	**on C_δ-dense instances, the quotient filters** $\bmod d' \le D$ **are a complete decider**: T_{comp} on quotients closes without search

Outside $o \le 1$, Corollary C′ is the one regime in which Theorem 7.6 decides without search, and it does so by the structure theorem of [BW].

Tests.

Check	Result
Proposition 8.1: Eval equals the set of subset sums (300 instances, $n \le 11$)	0 errors
Proposition 8.2: fixpoint in one backward round; W^* equals the exact projections (brute force); backward choice (1800 cases)	0 cases with more than one useful round; 0 differences; 0 stuck, 0 wrong; unfavourable order up to 12 rounds
Proposition 8.3: $R(W) \subsetneq R(W \cup D)$ (300 instances, weights ≤ 300)	$300/300$ strict; mean K of W without dummies 138; backward choice on $W \cup D$ skipping dummies fails on $979/1500$ reachable targets
Proposition 8.4: support of Eval $\bmod q$ against $N_c(q) > 0$, $q = 2, \ldots, 12$ (200 instances with divisor structure)	0 discrepancies; 36 540 obstructions out of 67 971 targets, 0 false
Proposition 8.5: telescoping reconstruction	exact in $300/300$; 303 pairs of instances with the same D and $R(W \cup D)$ but different $R(W)$

Table 15.

9 Discussion

9.1 Literature

Work	Regime	Guarantee	Cost
Galil–Margalit [GM]	dense, $w_{\max} = \tilde{O}(n^{3/2})$	worst case	$\tilde{O}(n + (w_{\max}/n)^2)$
Bringmann–Wellnitz [BW]	dense, polynomial weights	worst case	$\tilde{O}(n)$
Flaxman–Przydatek [FP]	random modular, $\log M = O((\log n)^2)$	expected, rigorous	polynomial
Lyubashevsky [Lyu]	random modular, $M = 2^{n^c}$, $c < 1$	success $1 - 2^{-\Omega(n^c)}$	$2^{O(n^c/\log n)}$

Work	Regime	Guarantee	Cost
Wan–Shi [WS]	random weights in $[1, 2^m]$	heuristic	$O(n + m2^{m/2})$
Howgrave-Graham–Joux, Becker–Coron–Joux [HGJ; BCJ]	density ≈ 1	heuristic	$2^{0.337n}$, $2^{0.291n}$
Schroeppel–Shamir, Wagner [SS; Wag]	random, density > 1, k lists	heuristic	$2^{n/2}$ (space $2^{n/4}$); $2^{b/(1+\log k)}$
Shallue [Sha]; Lin–Vasudevan [LV]	random modular; k-SUM over $\mathbb{Z}$	rigorous success bounds	k-tree
this work, Theorem R′	random integer weights, b bits	failure $\leq \delta$, rigorous	$O(\log(1/\delta)(n \log n + 2^{b/2+o(b)}))$
this work, Corollary C′	dense instances of [BW]	exact decider	$O(n^3)$
this work, Theorem R″	random integer weights, b bits	failure $\leq \delta$, rigorous	$O(\log(1/\delta) \log(nM)(n \log n + 2^{b/3+o(b)}))$
this work, arc solver	random or structured, dense to medium	outputs verified; completeness by the Poisson rule in practice, Las Vegas on random weights for $m \geq 7.6(b + O(\log m))$ (Theorem LV)	$O(nQ + m2^{m/4} + m2^{m/2}/P)$

Table 12.

- In the dense regime Bringmann–Wellnitz is rigorous and faster. Our contribution there is the arc reading of their obstructions (almost divisors are the moduli with vanishing class counts), and the measured extension of the solver to weights of $10^{10} - 10^{12}$, outside their polynomial regime.
- For $b = O((\log n)^2)$, Flaxman–Przydatek is polynomial and better. For $b = n^c$, Lyubashevsky's generalized-birthday algorithm runs in $2^{O(b/\log n)}$ on random modular instances, faster than the $2^{b/2+o(b)}$ of Theorem R′ in that window. Theorem R′ gives a proof for the simplest algorithm (centre rule and meet-in-the-middle) on integer instances, with explicit constants in its probabilistic ingredients.
- Wan–Shi's algorithm is the "no arcs, centre rule" variant of our ablation, stated heuristically and without experiments. Theorem R′ proves its guarantee. The arcs add the treatment of congruence obstructions ($12/12$ against $2 - 4/12$).
- Theorem R″ is the four-list form [SS; Wag] of Theorem R′. It gives a proof for the theorem mode of the implemented endgame, through Lemma J, with the moments of Lemma B as the only probabilistic input.
- At density ≈ 1 the endgame is a meet-in-the-middle, where the representation techniques [HGJ; BCJ] apply.

On the number-theoretic side, Richert's theorem (every integer > 6 is a sum of distinct primes) [Richert] and the Roth–Szekeres asymptotics for partitions into distinct primes [RSz] treat the arcs away from 0 as an error term. Lemma 1 rests on classical facts about cyclotomic values. The reading of the arcs of $\Pi(1 + z^p)$ as

weighted prime races, Proposition 3 with logarithmic-cosine weights, and the lobe bounds of Lemmas P and Q are, as far as we could find, new.

On the propagation side (§7–8):

- Union-find for equalities and forward evaluation of circuits are standard (congruence closure in SMT solvers, presolve in constraint programming). Labeled union-find [LUF] factors constraints by group-labelled relations (affine, XOR, modular) and refines with an AC-3 schedule; it does not treat functional definitions or DAGs. What §7 adds is the single factorization $T = T_{\text{comp}} \circ T_{\text{struct}}$ with parities and definitions on the complete DAG, the information limit (Theorem 7.3), the corrected bound (Theorem 7.4(iii) and Remark 7.5), and the closure form of convergence (Theorem 7.13).
- Theorem 7.7 is implicit definability in network form [Beth; Craig; LLM]; Gaussian elimination [SNC] and probing [LMS] are known techniques, here read as constructions of definitional variables.
- Proposition 8.2 is a case of [Freuder]; chaotic iteration is [Apt].
- Counting subset sums by evaluating the generating polynomial at roots of unity is the method of [LN]. The residue classes $\bmod q$, the emerging arcs, the prime races, the dummies and the exact de-blurring are not there.
- A deterministic navigation inside intervals of reachable sums is studied in [Nkosi], without dummies, primes or arcs.

9.2 Scope

- **The algorithmic results** concern random weights and dense instances. On them the difficulty concentrates in the endgame: $\sqrt{o}$ with a plain meet-in-the-middle, about $o^{1/3}$ with the residue filter; at density one $\lambda \approx 1$ and both are $2^{n/2}$.
- **The zeros.** Propositions 3 and 3_0 are equivalences, and Theorem S is an Ω-result. The lemmas of §6 concern independent random weights; the primes are one deterministic sequence, and carrying those lemmas over to them is the same question as the hypothesis at square-root scale. The numerical statements about the races and the moments of the primes are computed in finite ranges.
- **The count and the races.** The positivity of $r_N(T)$ lives at the level of the prime number theorem; the races live in the arcs (§3.6).
- **Propagation.** Theorem 7.6 gives a correct and terminating propagation; search acts on the free inputs and on the assertions. Definitional variables make explicit the links that are functional on the solutions (Theorem 7.7).
- **The Euler product.** The Euler product can be read as a structural collapse of the coefficients onto the primes, with the sign of $\operatorname{Re} \zeta'/\zeta$ in the computational part (the zeros), as the hybrid product of [GHK] shows.

10 Conclusion

For the primes, completeness makes the decision problem trivial and the counting problem arithmetic. The generating function of the prime subset sums has arcs at the rationals a/q. They emerge exactly at $q = 2^k$ and $q = 2p^j$, their neighbourhoods sink into the sea uniformly in q, and their heights are prime races whose one-sided bound is equivalent, under the conditions (R) and (N), to GRH for the even characters they see. That equivalence is unconditional for $q = 8$, 10 and 14, where the absence of real zeros is certified, and on the arc at 14 the race has no last sign change. The principal arc carries ζ in the centre and width of its Gaussian.

For arbitrary weights, the dummies of HURS restore completeness by blurring the counts, and the blur can be removed exactly, so the solver reads the arcs directly on the original weights. The arcs then become the decision rule of a solver whose cost is governed by its endgame: the square root of the omicron with a plain meet-in-the-middle, and about its cube root once the endgame is filtered by a residue class, which is itself an arc at a modulus of the order of the number of solutions. It returns only exact solutions, finds them on instances far beyond dynamic programming, and needs its arcs precisely where the structure theorem of dense Subset Sum puts the obstructions. On random weights of b bits the simplest form of the method has a proof: a uniform local limit theorem, and a concentration of signed subset sums conditioned on their total, give a solution with probability $1 - \delta$ in time $2^{b/2+o(b)} \cdot \log(1/\delta)$. In the modular model this is slower than generalized birthday; its interest is the proof and the explicit constants. With the residue filter the same ingredients, through an averaging over the residues, give $2^{b/3+o(b)}$: this is the four-list method, and its interest is again the proof.

The algorithm is an instance of a general propagation theory. Every constraint network factors into a structural part solved in one pass and a computational part that converges, in any order, to the least informed fixpoint; definitional variables make explicit exactly the links that are functional on the solutions. Read in this theory, the forward pass of the solver is structural evaluation, the arcs are propagation on quotients, the dummies are an extension that does not block but does not conserve, and de-blurring works because that loss, irreversible on reachable sets, is reversible on counts. Dense instances in the sense of Bringmann–Wellnitz are the regime, beyond completeness, in which propagation alone decides.

Data and code availability

All computations are reproducible with the scripts listed here (Python, NumPy, SymPy, mpmath, matplotlib for the figures, python-flint/Arb for the certifications, and OR-Tools for the CP-SAT method of BENCH5; the versions used are listed in `requirements.txt`).

- Floor (Proposition 2.3): `python/arcs_theory/holes_primes.py`.
- Arcs and races: `python/arcs_theory/arcs_fig.py`, `arcform2.py`, `edgeworth_arcs.py`, `edgeworth_arcs3.py` (Proposition C′, Table 2′), `arc_races.py`.
- Lobes: `python/arcs_theory/lobes2.py`, `lobe_proof_check.py`, `lobe_bound.py`, `lobe_tail.py`; certification of Corollary Q′ in Arb: `python/certification/certify_lobes_arb.py`, with output `python/certification/lobes_arb.txt`.
- Certification of (R): `python/certification/certify_L_real_segment.py`, with output `python/certification/out128.txt`.

- Certification of a zero on the line for the cubic character mod 7 (Theorem S): `python/certification/certify_zero_cubic7.py`, a sign change of $Z(t)$ on $[4.35546875, 4.3564453125]$, with output `python/certification/zero_cubic7_out128.txt`.
- Moments of the primes: `python/arcs_theory/width_check.py`.
- Filtered endgame and floating-point counts (§5.4–5.6, Tables $9'-9''$, Theorem R$''$): `python/algorithm/arcs_float.py`, `endgame_fast.py`, `test_123.py`, `bench_big2.py`, `bench_small2.py`.
- Theorem mode (§5.6): `python/algorithm/theorem_mode.py`, `test_theorem.py`.
- The solver: `python/algorithm/arcs_rho.py` (default ARCS, parameters derived from the instance, §5.7) and `python/algorithm/arcs_solver.py` (ARCS with the fixed constants of BENCH5 v2), with the pseudocode in `pseudocode/ARCS_pseudocode.md`.
- The benchmark of the default solver (§5.7): criterion with its sha256, runner, independent checker, results and logs, in `benchmarks/arcs_rho/`.
- Dynamic moduli (§5.7, Table 10$'$): `python/algorithm/arcs_dyn.py`, `bench_dyn.py`.
- The benchmark BENCH5 (§5.7, Figure 9), in `benchmarks/bench5_v2/`. The reference run v2: criterion (with its sha256 and an English translation), `bench5.py`, `checker5.py`, `export5.py`, results, run and check logs, and an export of every instance and answer as plain text with `verify_one.py`, which verifies the YES answers (the NO certificates are verified by `checker5.py`).
- Algorithm: `python/algorithm/decider.py`, `constructor3.py`, `poisson.py`, `deblur.py`, `bench_deblur.py`, `bench_big.py`, `ablation.py`, `nearprime.py`, `reps.py`; shared modules `hurs_arcs.py` and `hurs_hybrid.py`.
- Section 6, in `python/random_weights/`: `rig_check.py`, `b0_check.py`, `b0p_check.py`, `lemmaB_check.py`, `b_moment.py`, `b_cond.py`, `exact_B3.py`, `poisson_tail.py`, `moments4.py`, `b_prime_exact.py` (Proposition B$'$), `lv_rate.py` and `lv_rate_certified.py` (Theorem LV).
- Section 7, in `python/propagation/`: `run_tests.py` (Table 13), `ext.py` (algorithm Ext, Table 14 and the scale test), `probe.py` (probing).
- Section 8, in `python/propagation/`: `integ_hurs.py` (Propositions 8.1–8.4), `integ_algo.py` (Proposition 8.5).
- Figures, in `python/figures/`: Figure 5 comes from `python/arcs_theory/arcs_fig.py`; each of the others from its own script `fig_*.py`, with the shared style `figstyle.py` and, for Figure 6, the measured times `fig_endgame_rows.json`.
- A Lean 4 project, in `lean/`: the certified answer checker (`checkYes`, `checkNo`, `checkRange`, proved sound; the exhaustive certificate $(-3,-3)$, proved sound and complete) and `verify_bench`, which rechecks every answer of the two benchmarks; and the statements of the paper, including the analytic ones (Theorem E, Propositions C, C$'$, 3, 3_0, Theorem S, Lemmas L, W, Q, Corollary Q$'$, Lemmas B0$'$, B, B$'''$, V, Proposition B$'$, Theorems R, R$'$, R$''$, LV, and the propagation theory of §7–8): all 62 items of the paper are covered. Published theorems (Siegel–Walfisz, Vinogradov–Vaughan, Jackson, Rosser–Schoenfeld, the prime number theorem, von Koch, Bringmann–Wellnitz, [BR, Theorem 22.1], and for Proposition 3 the GRH bound for $\pi(x,\chi)$ [Davenport]) and the numerical certificates above enter as explicit hypotheses, never as axioms. Not formalized: Proposition B$''''$ (exact computations) and the running times. No `sorry`; only the standard axioms.

The scripts, the pseudocode of the solver, the benchmark files and the Lean project are in the public repository arcs-algo, `github.com/anthem87/arcs-algo` (AGPL-3.0-or-later; the version described here is commit `372ba68`). This paper is archived at doi:10.5281/zenodo.23912385 (all versions).

Declaration of AI tool use

AI tools were used to assist with computations, the drafting of text and the checking of proofs. All results were verified by the author, who takes full responsibility for the content.

References

Apt K. R. Apt, The essence of constraint propagation, Theoretical Computer Science 221 (1999) 179–210.

BCJ A. Becker, J.-S. Coron, A. Joux, Improved generic algorithms for hard knapsacks, EUROCRYPT 2011.

BCP C. Borgs, J. Chayes, B. Pittel, Phase transition and finite-size scaling for the integer partitioning problem, Random Structures & Algorithms 19 (2001), 247–288.

Beth E. W. Beth, On Padoa's method in the theory of definition, Indagationes Mathematicae 15 (1953) 330–339.

BHM B. Bzdęga, A. Herrera-Poyatos, P. Moree, Cyclotomic polynomials at roots of unity, arXiv:1611.06783.

BR R. N. Bhattacharya, R. Ranga Rao, Normal Approximation and Asymptotic Expansions, Wiley (1976); reprinted SIAM Classics in Applied Mathematics 64 (2010).

Brown J. L. Brown Jr., Note on complete sequences of integers, Amer. Math. Monthly 68 (1961) 557–560.

BW K. Bringmann, P. Wellnitz, On near-linear-time algorithms for dense Subset Sum, SODA 2021, arXiv:2010.09096.

CJLOSS M. J. Coster, A. Joux, B. A. LaMacchia, A. M. Odlyzko, C.-P. Schnorr, J. Stern, Improved low-density subset sum algorithms, Computational Complexity 2 (1992).

CPS A. Mennillo, Cumulative prime sums and Riemann zeros, Zenodo (2026).

CLRS T. H. Cormen, C. E. Leiserson, R. L. Rivest, C. Stein, Introduction to Algorithms, 3rd ed., MIT Press (2009).

Craig W. Craig, Three uses of the Herbrand–Gentzen theorem in relating model theory and proof theory, J. Symbolic Logic 22 (1957) 269–285.

Davenport H. Davenport, Multiplicative Number Theory, 3rd ed., Springer.

FP A. Flaxman, B. Przydatek, Solving medium-density subset sum problems in expected polynomial time, STACS 2005, LNCS 3404, 305–314.

Freuder E. C. Freuder, A sufficient condition for backtrack-free search, J. ACM 29 (1982) 24–32.

GHK S. M. Gonek, C. P. Hughes, J. P. Keating, A hybrid Euler–Hadamard product for the Riemann zeta function, Duke Math. J. 136 (2007) 507–549.

GM Z. Galil, O. Margalit, An almost linear-time algorithm for the dense subset-sum problem, SIAM J. Comput. 20 (1991).

HGJ N. Howgrave-Graham, A. Joux, New generic algorithms for hard knapsacks, EUROCRYPT 2010.

Halim S. Halim, F. Halim, Competitive Programming 3, Lulu (2013).

IK H. Iwaniec, E. Kowalski, Analytic Number Theory, AMS Colloquium Publications 53.

Ingham A. E. Ingham, The Distribution of Prime Numbers, Cambridge.

Koch H. von Koch, Sur la distribution des nombres premiers, Acta Math. 24 (1901) 159–182.

KT S. Knapowski, P. Turán, Comparative prime-number theory I–VIII, Acta Math. Acad. Sci. Hungar. 13–14 (1962–63).

Landau E. Landau, Über einen Satz von Tschebyschef, Math. Ann. 61 (1905).

Littlewood J. E. Littlewood, Sur la distribution des nombres premiers, C. R. Acad. Sci. Paris 158 (1914).

LLM J.-M. Lagniez, E. Lonca, P. Marquis, Improving model counting by leveraging definability, IJCAI 2016.

LMS I. Lynce, J. Marques-Silva, Probing-based preprocessing techniques for propositional satisfiability, ICTAI 2003.

LN D. Lokshtanov, J. Nederlof, Saving space by algebraization, STOC 2010.

LO J. C. Lagarias, A. M. Odlyzko, Solving low-density subset sum problems, J. ACM 32 (1985).

LV H. Lin, P. N. Vasudevan, On Wagner's k-tree algorithm over integers, arXiv:2410.06856 (2024).

LUF D. Lesbre et al., Labeled union-find for constraint factorization, NSAD 2024 (work in progress).

Lyu V. Lyubashevsky, On random high density subset sums, ECCC TR05-007 (2005).

Mackworth A. K. Mackworth, Consistency in networks of relations, Artificial Intelligence 8 (1977) 99–118.

MS G. Martin, J. Scarfy, Comparative prime number theory: a survey, arXiv:1202.3408.

Nkosi T. Nkosi, Algorithmic structure in Subset Sum: deterministic in-bound navigation and the counting complexity divide, arXiv:2506.12019 (2025).

Richert H.-E. Richert, Über Zerfällungen in ungleiche Primzahlen, Math. Z. 52 (1949).

RoS J. B. Rosser, L. Schoenfeld, Approximate formulas for some functions of prime numbers, Illinois J. Math. 6 (1962) 64–94.

Rosser J. B. Rosser, Real roots of real Dirichlet L-series, J. Res. Nat. Bur. Standards 45 (1950), 505–513.

RS M. Rubinstein, P. Sarnak, Chebyshev's bias, Experimental Mathematics 3 (1994).

RSz K. F. Roth, G. Szekeres, Some asymptotic formulae in the theory of partitions, Quart. J. Math. 5 (1954).

Rumely R. Rumely, Numerical computations concerning the ERH, Math. Comp. 61 (1993), 415–440.

Schaefer T. J. Schaefer, The complexity of satisfiability problems, STOC 1978, 216–226.

Sha A. Shallue, An improved multi-set algorithm for the dense subset sum problem, ANTS VIII, LNCS 5011 (2008).

Shamir A. Shamir, A polynomial-time algorithm for breaking the basic Merkle–Hellman cryptosystem, IEEE Trans. Inf. Theory 30 (1984).

SS R. Schroeppel, A. Shamir, A $T = O(2^{n/2})$, $S = O(2^{n/4})$ algorithm for certain NP-complete problems, SIAM J. Comput. 10 (1981) 456–464.

SNC M. Soos, K. Nohl, C. Castelluccia, Extending SAT solvers to cryptographic problems, SAT 2009.

Tarjan R. E. Tarjan, Efficiency of a good but not linear set union algorithm, J. ACM 22 (1975) 215–225.

Vinogradov I. M. Vinogradov, The Method of Trigonometrical Sums in the Theory of Numbers.

Wag D. Wagner, A generalized birthday problem, CRYPTO 2002, LNCS 2442, 288–304.

WGS R. Williams, C. P. Gomes, B. Selman, Backdoors to typical case complexity, IJCAI 2003, 1173–1178.

WS D. Wan, W. Shi, Solving medium-density subset sum problems in expected polynomial time: an enumeration approach, arXiv:0712.3203.