

LINNIK'S CONSTANT IS AT MOST 3.99

ERIC NASLUND*

ABSTRACT. Let $P(a, q)$ denote the least prime congruent to a modulo q , where $(a, q) = 1$. We give a computer-assisted proof that $P(a, q) \ll q^{3.99}$, improving the exponent 5 of Xylouris. More precisely, $P(a, q) < q^{3.99}$ for every sufficiently large q , uniformly in a .

The main new ingredient is a graded near-density estimate for zeros of Dirichlet L -functions, a weighted form of Heath-Brown's Lemma 12.1 of the kind he asked for in 1992. We combine it with the zero-location estimates and far-density method of Heath-Brown and Xylouris.

The second novelty is the scale of the case analysis. Heath-Brown and Xylouris closed their final case analyses with 14 and 21 main cases, each by a chain of inequalities evaluated in floating point. In this paper, with the help of advanced AI models, we can push this much further. The middle range of the first zero is divided into 4453 root cases and 4788 terminal cases, each closed by its own linear program. On 4,196,879 threshold boxes these programs give 29,397,336 linear relaxations, and exact integer certificates for all of them bound the normalized zero sum strictly below 1. Exceptional zeros and the remaining exterior range are treated separately. This case analysis is a kind of systematic brute force enabled by AI: an analysis of this size, with parameters tuned to each case, would be very laborious or nearly impossible to carry out by hand, and it lets the new estimate be applied separately in each case.

The near-density lemma, certificate soundness, and a conditional passage from a certified case to a prime are formalized in Lean, assuming published analytic inputs and specified facts about zeros.

***AI use:** This paper was written by Claude Opus 5.5 and GPT Astra 6 under my prompting and supervision. The research was carried out by the same models under my prompting and supervision, with some of the formalization work also done by xAI's Grok. My recommendation is for the reader to explore this paper with AI as well, and use their favorite model to ask the paper questions. The certificates are checked by three separate programs, one of them formally verified, and [Section 14](#) states precisely which parts of the proof are formally verified in Lean and which have conventional proofs only.

Date: September 30, 2026.

2020 Mathematics Subject Classification. 11N13 (primary); 11M20, 11M26, 11N35, 11Y35, 68V15, 90C05 (secondary).

Key words and phrases. Least prime in an arithmetic progression, Linnik's theorem, zeros of Dirichlet L -functions, zero density, Selberg sieve, computer-assisted proof.

CONTENTS

1. Introduction	2
2. Notation and conventions	8
3. Published inputs	12
4. Prime detection	15
5. Zero costs	17
6. Far density and representatives	24
7. Zero location	26
8. The graded near-density lemma	36
9. The near rows	43
10. Leaf programs and exact certificates	51
11. The case tree	55
12. The exterior regimes	61
13. Assembly of the proof	64
14. The computations and their verification	66
15. Concluding remarks	71
Appendix A. The complex location rows	71
Appendix B. The fixed constants	73
References	76

1. INTRODUCTION

1.1. The result. For an integer $q \geq 2$ and an integer a with $(a, q) = 1$, let $P(a, q)$ be the least prime congruent to a modulo q . Dirichlet's theorem [31] ensures that this prime exists. Linnik [77, 78] proved that there are absolute constants C and L such that

$$(1.1) \quad P(a, q) \leq Cq^L$$

for every reduced residue class. An exponent L for which such a constant C exists is called *admissible*; Linnik's constant is the infimum of the admissible exponents.

Theorem 1.1 (Least prime in a reduced residue class). *There is an absolute constant $q_0 \geq 2$ such that, for every integer $q \geq q_0$ and every integer a with $(a, q) = 1$, there is a prime p satisfying*

$$p \equiv a \pmod{q}, \quad p < q^{3.99}.$$

Consequently, there is an absolute constant $C > 0$ such that $P(a, q) \leq Cq^{3.99}$ for all $q \geq 2$ and $(a, q) = 1$. In particular, Linnik's constant is at most 3.99.

We do not compute q_0 or C ; the issue of effectivity is discussed at the end of [Section 13](#). Except in the presence of an extremely close real zero, the proof finds a prime in the smaller interval

$$q^{3.156342} < p < q^{3.99}.$$

The lower endpoint comes from the support of the weight used to detect primes. When a real zero is extremely close to 1, Heath-Brown's theorem on Siegel zeros [50] instead gives $P(a, q) \leq q^{3.5}$.

The argument combines published analytic estimates, new lemmas proved here, and exact finite computations. The theorem requires all three parts. In particular, checking a numerical certificate proves a bound for its specified configuration of zeros; the analytic argument must also show that every modulus gives a configuration covered by one of the certificates. We make this passage explicit in [Sections 11](#) and [13](#).

1.2. Earlier work. The admissible exponents in Table 1 record the main numerical developments. Heath-Brown [51] obtained 5.5, and Xylouris [143, 144, 145, 146] subsequently obtained 5.2, 5.18, and 5. Before them, Pan [98, 99] gave the first numerical values, and Chen, Jutila, Graham and Wang lowered them in a long series of papers, which includes Graham's introduction of Selberg sieve weights into zero-density estimates [39, 41]. Our proof uses the framework of these works. It rests on the three principles that Heath-Brown isolates at the start of his paper: a zero-free region with at most one exceptional zero [27, 43, 74, 97], the repulsion of other zeros by an exceptional zero, known as the Deuring–Heilbronn phenomenon [30, 53, 78], and a log-free zero-density estimate [77]. The last means that the density bound has no additional power of $\log q$, a feature needed when studying zeros at distance $O(1/\log q)$ from 1. Simpler proofs of these principles, and of Linnik's theorem, were given by Rodoski [111], Turán [133, 135], Knapowski [71], Fogels [32], Gallagher [37], Jutila [65], Motohashi [91, 92] and Bombieri [10], among others; accounts are in the books [110, 84, 62, 34].

The conjectured scale is much smaller. Chowla [24] conjectured $P(a, q) \ll_\varepsilon q^{1+\varepsilon}$ for every $\varepsilon > 0$; the maximum over reduced classes is conjecturally of order $\varphi(q)(\log q)^2$ [138, 109, 42, 76]. The Generalized Riemann Hypothesis gives admissible exponents $2 + \varepsilon$; see [4, 73, 15] for explicit conditional bounds. Explicit [9] and uniform [128] unconditional estimates for primes in progressions are also known. In the other direction, an exceptional (Siegel) zero helps: Heath-Brown [50] showed that a sufficiently strong exceptional zero forces $P(a, q) \leq q^{3+\delta}$, in an effective form that we use below, and $q^{2+\delta}$ ineffectively; see also [33, 61, 147]. Other proofs of Linnik's theorem use pretentious methods [114], sieve methods [36], or avoid L -functions [80], but give larger exponents. For moduli with special multiplicative structure there are stronger zero-free regions [7, 38, 59], and much smaller exponents are known [16].

1.3. From zeros to primes. Write $\mathcal{L} = \log q$ and express a zero of a Dirichlet L -function as

$$\rho = 1 - \frac{\lambda}{\mathcal{L}} + i \frac{\mu}{\mathcal{L}}.$$

Thus small λ means that the zero is close to the line $\sigma = 1$; μ is its height multiplied by $\log q$. We group each nonprincipal character with its complex conjugate and call this a *family*. The parameters $\lambda_1, \lambda_2, \lambda_3$ record the first zero of successive families in a specified rectangle near 1. The parameter λ' records the next zero occurrence in the first family, after removing the distinguished zero and its conjugate where appropriate. Precise definitions, including multiplicities and ties, are given in Section 2.

A nonnegative weight h , supported in $[A, L]$ with $A = L - 2T$ and $T = 0.416829$, detects primes between q^A and q^L . Let H be its Laplace transform. Xylouris's positivity criterion [145, (3.57)] gives

$$\sum_{p \equiv a(q)} \frac{\log p}{p} h\left(\frac{\log p}{\mathcal{L}}\right) \geq \frac{\mathcal{L}H(0)}{\varphi(q)}(1 - W - \eta),$$

where $\eta = 10^{-6}$ and

$$W = \frac{1}{H(0)} \sum_{\chi \neq \chi_0} \sum_{\rho \in R_\chi} |H((1 - \rho)\mathcal{L})|.$$

Here R_χ is a fixed square in the normalized coordinates, and zeros are counted with multiplicity. It is therefore enough to prove $W \leq 1 - 2\eta$. The contribution of an individual zero decays exponentially with λ , so zeros close to 1 require the most accurate estimates.

We combine three kinds of information about these zeros. First, a *per-character bound* estimates the total contribution of one L -function from a lower bound for the parameters of its zeros. Second, a *far-density bound* limits a weighted sum over characters, with weight depending on a selected zero of each character. Third, *near-density bounds* constrain the characters whose

L	Year	Author(s)	Reference
10000	1957	Pan (announced)	[98]
5448	1958	Pan	[99]
777	1965	Chen	[18]
630	1971	Jutila (reported by Turán)	[134]
550	1970	Jutila	[64]
168	1977	Chen	[19]
80	1977	Jutila	[66]
36	1977	Graham	[39]
20	1981	Graham	[41]
17	1979	Chen	[20]
16	1986	Wang	[139]
13.5	1989	Chen and Liu	[21, 22]
11.5	1991	Chen and Liu	[23]
8	1991	Wang	[140]
5.5	1992	Heath-Brown	[51]
5.2	2009	Xylouris	[143]
5.18	2011	Xylouris	[144]
5	2011	Xylouris	[145, 146]
3.99	2026	this paper	

TABLE 1. Admissible values of Linnik’s constant, following the tables in [51] and [144], in order of the bound. Graham’s value 20 was submitted before Chen’s 17 appeared.

selected zeros lie closest to 1. Zero-location estimates for λ_1 , λ' , λ_2 , and λ_3 determine which of these bounds are available in a given case.

1.4. The new near-density estimate. Heath-Brown’s near-density estimate [51, Lemma 12.1] bounds the number $N(\lambda)$ of characters with a zero in $\sigma \geq 1 - \lambda/\mathcal{L}$, $|t| \leq 1$. Its proof applies the Cauchy–Schwarz inequality to a prime sum at the single point $\beta_1 = 1 - \lambda_1/\mathcal{L}$, and compares the resulting Gram form of the characters with the diagonal. Because every character is tested at the same point, a zero at distance λ counts only through the single number $N(\lambda)$, and the bound is weaker than the density estimate of Heath-Brown’s §11 as soon as $\lambda \geq 1.1$ and void soon after. That density estimate, proved with Selberg sieve weights after Graham, has the form (11.4) of a sum over characters of a weight that depends on each character’s own distance $\lambda^{(k)}$, and it contains the count (11.5) of the characters with $\lambda^{(k)} \leq \lambda$. In his closing list of possible improvements Heath-Brown writes [51, §16, item 9]:

It would be nice to have a weighted version of Lemma 12.1, in the way that (11.4) is a weighted version of (11.5). Unfortunately, no neat way of achieving this seems available.

The graded near-density lemma (Theorem 8.4) supplies such an estimate. We test each character at its own *anchor*, a real parameter determining the point at which a smoothed prime sum is evaluated. Cauchy–Schwarz relates these prime sums to a Gram form of the characters. We estimate its off-diagonal terms by the local explicit formula and, where sieve weights are used, Burgess’s bound; Graham’s sieve asymptotic controls the diagonal terms. Allowing the anchors to vary is the grading in the name of the lemma.

After normalization, the resulting constraint has the form

$$\left(\sum_j a_j v_j\right)_+^2 \leq \sum_j D_j a_j^2 + d \left(\sum_j a_j\right)^2 \quad (a_j \geq 0).$$

Here $x_+ = \max(x, 0)$, the index j runs over character–height entries, v_j is a lower bound for the response of an entry, and $D_j > 0$ and $d > 0$ bound its diagonal and common correlation costs. Most entries correspond to one selected zero; some retain a second zero of the same character, with an explicit correction for their correlation.

Proposition 8.8 turns this inequality into a form suited to computation: there is a number $\tau \in [0, \sqrt{d}]$ such that

$$\sum_j \frac{(v_j - \tau)_+^2}{D_j} \leq 1 - \frac{\tau^2}{d}.$$

A single constraint now distinguishes zeros at different distances from 1. The sieve weighting has precedents in Graham's work [39, 41], Heath-Brown's far-density argument, and the sieve-weighted large-sieve inequalities of Motohashi [91].

Where the new lemma enters. In Heath-Brown's proof the near-density estimate enters at the very end, in the assembly of [51, §15]. There the characters with a zero below a level Λ are sorted into bins $(\Lambda_{r+1}, \Lambda_r]$, with $\Lambda_r = \Lambda - 0.025r$; each character is charged the cost of the end of its bin nearer to $s = 1$; and the numbers of characters in the bins are controlled only through the counts $N(\Lambda_r)$, bounded one level at a time by Lemma 12.1 (his Table 13). The far-density lemma controls the characters beyond Λ . Xylouris's assembly [145, §6] has the same structure, with a sharper form of Lemma 12.1 (his Lemma 5.3). A count $N(\lambda)$ treats all characters with a zero below λ alike, so it cannot tell a few zeros very close to $s = 1$ from many zeros further away. In our proof this place is taken by the *near rows* of the leaf programs described below. Each near row is one instance of **Theorem 8.4**: a single quadratic constraint in which every bin of characters enters with its own feature, computed at its own anchor. The weighted version of Lemma 12.1 that Heath-Brown asked for is thus used exactly where his count was used, in every leaf program, and it is the main source of the improvement (**Section 1.6**).

1.5. The finite case analysis. We divide the middle range $0.1 \leq \lambda_1 < 1.5$ into cases determined by the type of the first family and intervals for the first few zero parameters. There are four levels in the computation, serving different purposes.

- (1) A *root case* specifies a range of possible zero configurations. The 58 parent rows taken from published zero-location tables lead to 4453 root cases, covering the middle range.
- (2) A *leaf* is a terminal case after further subdivision and valid zero-location exclusions. There are 4788 leaves that require numerical bounds.
- (3) A *leaf program* groups characters into bins according to their selected zero parameters. Its nonnegative variables are the numbers of characters in these bins, together with weighted masses for the tails. The objective bounds W , and its constraints express the far-density bound, restrictions on the first families, and the near-density inequalities. For fixed thresholds these constraints are linear.
- (4) A *threshold box* bounds the auxiliary numbers τ in the near rows. On each box, a linear relaxation is either excluded or bounded by exact dual multipliers. Subdividing these boxes covers every possible choice of thresholds.

The case tree partitions information about zeros; the threshold boxes partition the auxiliary parameters used to bound them. Keeping these two subdivisions distinct is essential to the coverage argument.

Each leaf, or each subcase of a leaf, has its own certificate, a binary tree of threshold boxes. The 4788 leaves give 5591 such certificate trees, because some leaves are subdivided further, and these trees contain 4,196,879 threshold boxes. On each box the near rows are replaced by linear relaxations, some with two alternative forms, and every choice of alternatives, a relaxation case, is certified separately: the boxes carry 29,397,336 relaxation cases, of which 9430 are shown to be infeasible and the others are bounded by their own integer dual multipliers. The largest certified value is $0.9999998223 \dots < 1$; it bounds the zero sum together with the required error allowance. All certificate inequalities are checked in integer arithmetic. The subdivisions and zero-location arguments of the tree do not depend on the target exponent; only the leaf programs do. [Table 2](#) compares these numbers with the case analyses of Heath-Brown and Xylouris, and [Table 10](#) lists every level of the analysis.

The exterior ranges are shorter arguments. An exceptionally small λ_1 is covered by Heath-Brown’s theorem on Siegel zeros. The rest of $\lambda_1 \leq 0.1$ is handled by his quantitative Deuring–Heilbronn estimates and zero-location tables. A single further certificate treats $\lambda_1 \geq 1.5$. If the relevant rectangle contains no zero, then $W = 0$ and the positivity criterion applies immediately.

1.6. Comparison with Heath-Brown and Xylouris. Our framework is Heath-Brown’s [\[51\]](#) with Xylouris’s refinements [\[145, 146\]](#). We use the following from their work:

- the positivity criterion [\[145, \(3.57\)\]](#);
- the per-character bound [\[145, Lemma 3.10\]](#);
- the far-density lemma [\[145, Lemma 5.1\]](#);
- the zero-free regions and Deuring–Heilbronn estimates, as printed in their tables;
- the local explicit formulas [\[145, Lemmas 3.1, 3.2\]](#), which are [\[51, Lemmas 5.3, 5.2\]](#).

The case analyses compared. Both earlier proofs are computer-assisted, and both are organized, like ours, as a case analysis on the first zeros; [Table 2](#) compares the three. Heath-Brown’s paper has twelve numerical tables of zero-location and density bounds. Its final assembly [\[51, §15\]](#) treats $\lambda_1 \geq 0.348$ in fourteen intervals, mostly of length 0.02, each by a chain of inequalities evaluated numerically; the largest total it reports is 0.9943. Heath-Brown writes [\[51, §1\]](#) that “the arguments rely heavily on numerical calculations. These we do not reproduce in full, nor have we attempted any rigorous analysis of the rounding and truncation errors in the computer algorithms employed.” Xylouris’s dissertation adds eleven tables. Its final assembly [\[145, §6.4, pp. 86–87\]](#) has 21 cases: three intervals of λ_1 for a real zero of a real character, two for a complex zero of a real character, and sixteen for a nonreal character. Each case is split further according to the possible values of two zero-counting functions [\[145, \(6.46\)\]](#), and the computations were done in Maple with ten digits [\[145, p. 14\]](#); the largest value is $W < 0.998$. In both, each case is closed by a chain of inequalities in which the unknown zeros are estimated separately, one after another.

In our proof the same role is played by 4453 root cases and 4788 leaves, and each leaf is closed by a linear program with an exact certificate instead of a floating-point evaluation. The 4.2 million boxes have no counterpart in the earlier work. They do not subdivide the configurations of zeros. They subdivide the auxiliary thresholds of the near rows, and they arise because the graded lemma is a quadratic constraint rather than a count. The largest certified value, 0.99999982 against 0.9943 and 0.998, shows how much finer the analysis must be at $L = 3.99$.

Where the gain comes from. The gain from 5 to 3.99 has three sources:

- the graded near-density lemma, which replaces the counts $N(\lambda)$ and applies to every range of λ at once;
- the linear-programming combination, which lets all constraints act together on each case instead of chaining worst cases;
- a fine case analysis of the first zeros, possible because every case is certified by machine.

	Heath-Brown [51] $L = 5.5$ (1992)	Xylouris [145] $L = 5$ (2011)	This paper $L = 3.99$
Near-density input	counts $N(\lambda)$ from Lemma 12.1	counts $N(\lambda)$ from a sharper Lemma 12.1	the weighted (graded) lemma, Theorem 8.4
Numerical tables	12	11 more	58 rows from theirs, plus over 100 new
Final case analysis	14 intervals of λ_1	21 cases, each split by two zero counts	4453 roots, 4788 leaves
How a case is closed	floating point, without rounding analysis	Maple, 10-digit floating point	exact certificates on 4.2 million boxes
Largest value of W (must be below 1)	0.9943	0.998	0.99999982

TABLE 2. Three proofs of an explicit Linnik constant compared, in the middle range of the first zero ($\lambda_1 \geq 0.348$ for Heath-Brown and Xylouris, $0.1 \leq \lambda_1 < 1.5$ here). The last row is the largest bound for the normalized zero sum W over all cases of the final assembly: for Heath-Brown the largest total he reports, for Xylouris his stated bound, and for this paper the largest certified value, a bound for $W + 2\eta$. The entry “without rounding analysis” paraphrases Heath-Brown’s own remark [51, §1], quoted above.

The last source alone does not go far. Xylouris estimates that his method, with much larger computations (smaller intervals and finer grids), would give about 4.96 [146, p. 82].

1.7. The certified exponent and numerical evidence. We state the result at 3.99 because this is the exponent for which the complete cover has been certified. The tightest cases have a nonreal first character and $\lambda_1 \in [0.64, 0.80]$. In this range the available lower bounds for λ_2 are comparatively small, which restricts the anchors in the near-density estimates. Several of the corresponding leaf bounds are within 10^{-7} of 1 after the error allowance, and require fine subdivisions.

Exploratory floating-point computations suggest that a somewhat smaller exponent may be accessible. One difficult leaf, with the reserved-family columns of [Section 11.5](#), has computed values 0.992 at $L = 3.97$ and 1.020 at $L = 3.95$. These are values of particular relaxations, not certified global bounds or an obstruction to other methods. Similarly, the positive numerical margins for the exterior ranges at smaller exponents do not constitute a complete proof there. The verification reported in this paper is at $L = 3.99$.

1.8. Verification and a guide to the proof. The certificates have been checked by independent programs. Parts of the analytic argument and the soundness of the checkers have also been formalized in Lean, with the published analytic inputs stated as hypotheses. This formalization includes a theorem that a valid certified leaf yields a prime when the modulus satisfies that leaf’s hypotheses about zeros. It does not include the full passage from arbitrary moduli to leaves, or the exterior arguments. [Section 14](#) gives the verification records and separates these claims precisely.

[Sections 2](#) and [3](#) give the notation and published inputs. [Sections 4](#) to [6](#) reduce prime detection to bounds for the zero sum, and [Section 7](#) supplies the zero-location estimates. The main new analytic argument is in [Section 8](#); [Section 9](#) derives the constraints used in the computation.

Section 10 proves certificate soundness, while Section 11 shows that the zero configurations are covered and satisfy the programs. The exterior cases and the final choice of constants are treated in Sections 12 and 13. A reader interested first in the new density estimate may read Sections 2 and 3 and then Section 8, returning to the numerical construction afterwards.

Figure 1 shows how the main results depend on one another, and which of them are published inputs, new results, computations or formally verified statements. The terms of the case analysis are introduced in Section 2.4, and the fixed constants and the error budget in Section 2.7. Table 10 in Section 11 lists all levels of the finite analysis, and Table 12 in Section 14 records, for each component, its source and how it was checked.

2. NOTATION AND CONVENTIONS

We use standard notation for Dirichlet characters and their L -functions. For background on Dirichlet L -functions and their zeros we refer to [57, 131, 69, 85, 26, 62, 87, 124], and for the history of the subject to [94]. The distinctions between a zero and a zero occurrence, and between normalized and physical height, will be used throughout the proof.

2.1. Moduli, characters and zeros. Throughout, $q \geq 2$ is an integer modulus and

$$\mathcal{L} = \log q.$$

Characters are Dirichlet characters modulo q ; χ_0 is the principal character. For a nonprincipal character χ we consider the zeros $\rho = \beta + i\gamma$ of $L(s, \chi)$ with $\beta > 0$, counted with multiplicity, and we write

$$(2.1) \quad \rho = 1 - \frac{\lambda}{\mathcal{L}} + i\frac{\mu}{\mathcal{L}}, \quad \lambda = (1 - \beta)\mathcal{L}, \quad \mu = \gamma\mathcal{L}.$$

We call $\lambda = \lambda_\rho$ the *parameter* of ρ and $\mu = \mu_\rho$ its *normalized height*; $\gamma = \gamma_\rho$ is its *physical height*. Thus a height bound always concerns $|\gamma|$ or $|\mu|$. Smaller λ means a zero farther to the right. The multiplicity of ρ as a zero of the specified $L(s, \chi)$ is m_ρ . A *zero occurrence* is one copy of the pair (χ, ρ) : a zero of multiplicity m_ρ gives m_ρ occurrences. Unless stated otherwise, sums over zeros count multiplicity. A zero of an imprimitive character is a zero of the L -function of the primitive character inducing it, with the same multiplicity, as long as $\beta > 0$; all zeros considered below have $\beta > 1/2$.

2.2. Regions. Following Xylouris [145, (3.7)] we put, for $x > 0$,

$$R(x) = \left\{ \sigma + it : 1 - \frac{\log \log \mathcal{L}}{3\mathcal{L}} \leq \sigma \leq 1, |t| \leq x \right\}.$$

There is an integer $l = l(q)$ with $1 \leq l \leq \mathcal{L}/10$ such that no zero of any $L(s, \chi)$ lies in $R(10l) \setminus R(l)$ (Input 3.5). We fix one such choice, for example the least admissible integer. The distinguished zeros (Definition 2.2) and the representatives (Section 2.4) are selected in $R(l)$. The local explicit formula also involves disc zeros, which need not all belong to this rectangle. Three further regions have separate roles:

- the *prime window* $R_P = \{0 \leq \lambda \leq C_P, |\mu| \leq C_P\}$, whose zeros are the only ones that enter the zero sum;
- the *buffer* $R_B = \{0 \leq \lambda \leq C_B, |\mu| \leq C_B\}$, with $C_B > C_P + M$;
- the *height-one region* $|\gamma| \leq 1$, in which the far-density lemma operates.

The constants C_P , M and C_B are fixed independently of q , in the order described in Section 13. For sufficiently large q , $R_P \subset R_B \subset R(l)$, and the buffer has physical height $C_B/\log q < 1$. The word *inside* means that the selected first zero belongs to R_B ; *outside* means that it does not. It does not refer to $R(l)$.

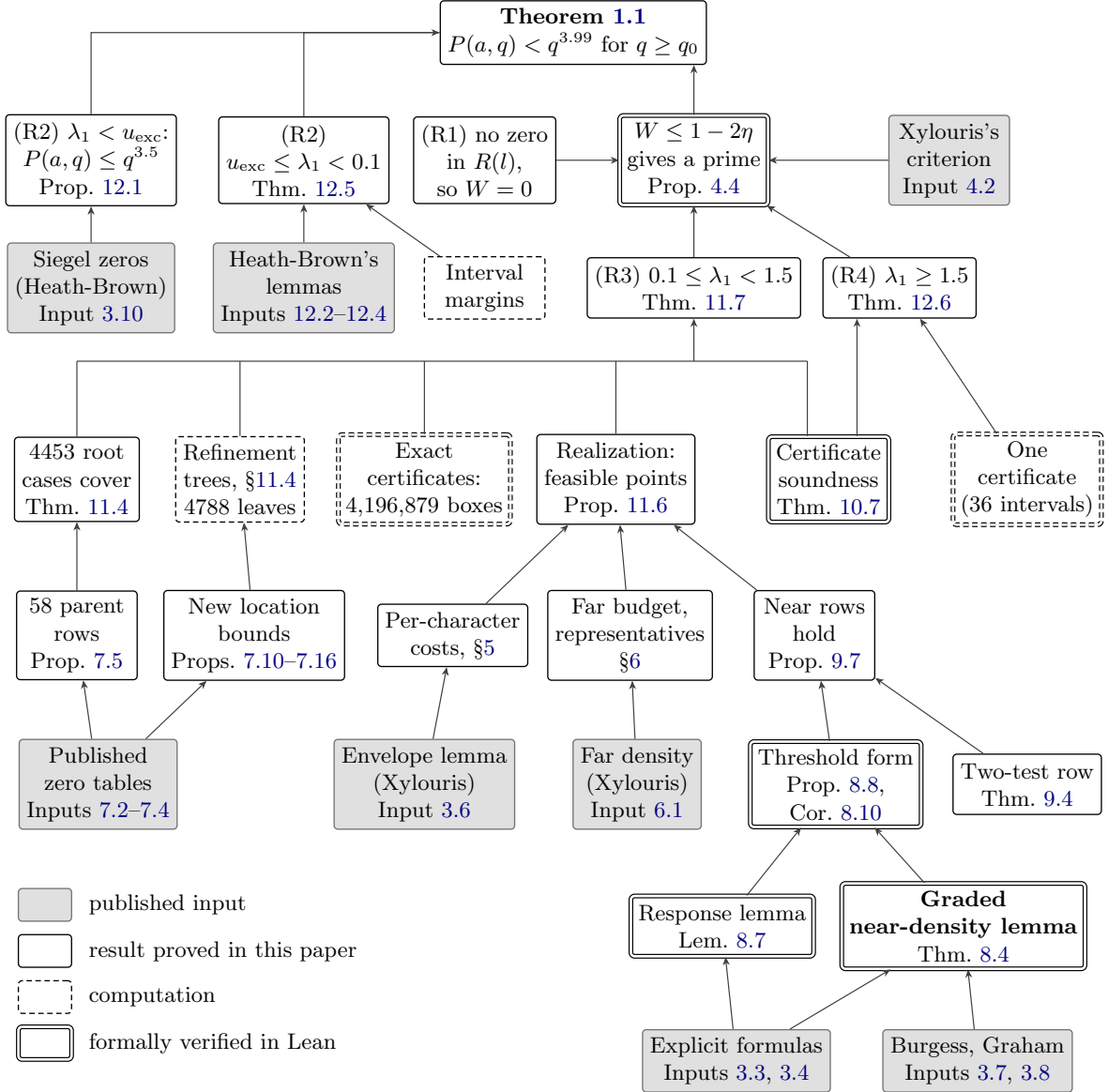


FIGURE 1. Proof map for Theorem 1.1. An arrow points from a result to a result whose proof uses it; the joined lines into (R3) are the ingredients of Theorem 11.7. The regimes (R1)–(R4) of Section 13 are defined by the first zero λ_1 , and (R2) splits at the constant u_{exc} of Section 12. A double border marks a statement formally verified in Lean, with the published inputs as hypotheses, or, for the certificates, a computation accepted by a checker proved sound in Lean. The formalization also covers the far budget and parts of Propositions 9.7 and 11.6; the zero-location results, the case tree, the costs, the two-test row, the exterior regimes and the assembly are not formally verified (Section 14.3). Some arrows are omitted, for instance from the explicit formulas to Sections 5 and 7 and into Theorem 12.6.

2.3. Families and the first zeros.

Definition 2.1 (Families and successive minima). A *family* is the set $\mathcal{F}(\chi) = \{\chi, \bar{\chi}\}$ for a nonprincipal character χ . It has one member when χ is real and two otherwise. If there is a nonprincipal zero in $R(l)$, select a pair (χ_1, ρ_1) with least parameter and write $\lambda_1 = \lambda_{\rho_1}$, $\mu_1 = \mu_{\rho_1}$, and $\mathcal{F}_1 = \mathcal{F}(\chi_1)$. Remove all occurrences belonging to $\mathcal{F}_1$ and repeat to define (χ_2, ρ_2) , λ_2 , and $\mathcal{F}_2$; continue in this way. Thus $\lambda_1 \leq \lambda_2 \leq \lambda_3 \leq \dots$. We call $\mathcal{F}_1$ the *first family*, $\mathcal{F}_2$ the *second family* and $\mathcal{F}_3$ the *third family*, and (χ_1, ρ_1) the *first zero*. If a subsequent minimum is over an empty set, its value is $+\infty$. Ties may be resolved by any admissible choice; each result used below holds for all such choices. If the initial set is empty, we use the empty-region argument of Section 12.

These are the selections of Xylouris [145, §3.1.2]. They order *families*, not the individual zeros of one L -function.

Definition 2.2 (The additional first-family zero). The *distinguished occurrences* are one copy of ρ_1 as a zero of χ_1 , together with one copy of $\bar{\rho}_1$ as a zero of χ_1 if χ_1 is real and ρ_1 is not, or as a zero of $\bar{\chi}_1$ if χ_1 is not real. We write λ' for the least parameter of a zero occurrence of χ_1 or $\bar{\chi}_1$ in $R(l)$ that is not distinguished; thus $\lambda' = \lambda_1$ when ρ_1 is a multiple zero. If no occurrence remains, put $\lambda' = +\infty$.

When λ' is finite, it can always be realized by a zero of χ_1 itself. Indeed, zeros of $L(s, \bar{\chi}_1)$ are conjugates of zeros of $L(s, \chi_1)$, with the same multiplicities. Replace any minimizing occurrence of $\bar{\chi}_1$ by its conjugate occurrence of χ_1 . We call every resulting non-distinguished occurrence ρ' with parameter λ' *admissible* (a second copy of ρ_1 is admissible when ρ_1 is a multiple zero), and write γ' for its height; as $\rho' \in R(l)$, $|\gamma'| \leq l$. The arguments below hold for every admissible choice; in the second-zero split of Section 11.6 the choice is the one used there.

The following table fixes the three types and the two counts used in the first-family bounds. Here $n = |\mathcal{F}_1|$, while α counts distinguished occurrences *per character*.

Type	First character and zero	n	α
rr	χ_1 real, ρ_1 real	1	1
rc	χ_1 real, ρ_1 nonreal	1	2
complex	χ_1 nonreal	2	1

For type rc we choose the conjugate with $\mu_1 > 0$.

2.4. The vocabulary of the case analysis. In the middle range $0.1 \leq \lambda_1 < 1.5$ the proof is a finite case analysis, defined precisely in Sections 10 and 11. Its terms are needed in the analytic sections that precede those, and we introduce them here.

- *Anchor*. The local explicit formulas of Section 3 are applied at points $s = 1 - \sigma/\mathcal{L} + i\gamma$; the real number σ is the *anchor*. An inequality anchored at σ is useful only if the zeros near s with parameter below σ are absent or are retained explicitly, so anchors are chosen at or below known lower bounds for zero parameters.
- *Specification*. A *specification* $\mathfrak{s}$ (Definition 11.2) records the type of the first family; an interval $[a, b]$ containing λ_1 , the *first-zero cell*; a lower bound $p \leq \lambda'$ and possibly an interval $[\text{lo}', \text{hi}']$ containing λ' , called a *gap*; a lower bound $l_2 \leq \lambda_2$; a lower bound r for the parameters of the zeros of height at most 1 of the families other than the first; possibly a reservation (next item); and whether ρ_1 lies inside or outside the buffer. Its *configuration set* $C(\mathfrak{s})$ consists of the zero configurations compatible with these data.
- *Reserved and ordinary characters*. A reservation concerns a family other than the first whose least parameter among zeros of height at most 1 is as small as possible. It fixes the

number $n_2 \in \{1, 2\}$ of its characters and an interval $[\text{lo}_2, \text{hi}_2]$ for that parameter. This *reserved family* need not be $\mathcal{F}_2$. The characters outside the first and the reserved family are *ordinary*.

- *Representatives.* The *height-one representative* of a nonprincipal character χ is a zero of $L(s, \chi)$ in $R(l)$ with $|\gamma| \leq 1$ and least parameter; its T^* -*representative* is defined in the same way with $|\gamma| \leq T^*$, for a height $T^* \in [\frac{1}{2}, 1]$ fixed in [Lemma 6.3](#) ([Definition 6.4](#)). The leaf programs place a character by the parameter of one of its representatives.
- *The case tree.* The *roots* are 4453 specifications whose configuration sets cover the middle range. Each root is refined by splitting intervals and by applying zero-location results. This gives a finite tree of specifications, the root's *refinement tree*; its vertices are *nodes*, and together these trees form the *case tree*. A node at which a zero-location result sharpens one of the bounds is a *location update*; a node whose configuration set is shown to be empty is *excluded*. The terminal nodes that are not excluded are the *leaves*. Some leaves are divided further into finitely many *subcases*.
- *Leaf programs, bins and columns.* Each leaf or subcase has a *leaf program* ([Definition 10.1](#)) whose value bounds W for every configuration in it. The parameters of the ordinary characters (and in some leaves those of the reserved characters) are divided into finitely many intervals, the *bins*. The *column* of a bin is the variable of the program that counts the characters whose representative zero ([Definition 6.4](#)) has parameter in the bin. The ordinary characters beyond the last bin form the *tail*, whose column is a weighted mass rather than a count. In an outside leaf the first family is represented by *hidden columns*, which bin the least parameter t_χ of the zeros in R_P of each first-family character ([Section 5.7](#)). The constant first bounds the cost of the first family in an inside leaf (0 in an outside leaf), plus any fixed charge for the reserved family, and final is an allowance for the error terms.
- *Far budget.* The far-density lemma ([Input 6.1](#)) bounds a sum over all characters of a positive decreasing function w of the parameter of a selected zero of height at most 1. We call $w(\lambda)$ the *far weight* of such a zero, and the bound ([6.1](#)) the *far budget*; in a leaf program it is one linear constraint.
- *Near rows.* The other constraints of a leaf program are conditions on counts and the *near rows*. A near row is, with one exception ([Section 9.6](#)), an instance of the graded lemma ([Theorem 8.4](#)) in the form of [Proposition 8.8](#). Its terms are *entries*, one for each character (occasionally each zero) that it includes. Each entry has a *feature*, a lower bound for how strongly its zero is detected, and a *diagonal*, what it costs on its own, and one *correlation term* bounds the interaction between distinct entries ([Proposition 8.8](#)). The entries of a family whose zeros are known are grouped into *family terms*. A row is built from two test functions, a *detector* and a *Gram test* ([Definition 8.1](#)); by the choice of its anchors and sieve weight it is a *family*, *shifted*, *graded* or *two-test* row ([Table 8](#)). It is linear in the columns once an auxiliary number, its *threshold* τ , is fixed.
- *Certificates.* A *certificate* of a leaf or subcase is a binary tree, the *certificate tree*, whose nodes are *boxes* of thresholds, starting from a box that contains every admissible threshold; its leaves are the *terminal boxes*. On a terminal box each near row is replaced by a linear relaxation, and each choice among the alternatives of these relaxations, a *relaxation case*, is either shown to be infeasible (an *exclusion*) or bounded below 1 by integer dual multipliers, in exact integer arithmetic ([Section 10](#)).

2.5. Elementary notation. For real x we write $x_+ = \max(x, 0)$, and $\mathbf{1}_E$ denotes the indicator of a condition or set E . The symbol $\varphi(q)$ is Euler's totient; the separately defined $\varphi(\chi)$ is a conductor coefficient ([Section 3](#)). A *cost* means an upper bound for a contribution to the

normalized zero sum W ; a *budget* is the right-hand side of a constraint on these contributions. All finite decimals used as parameters are exact rational numbers unless an approximation is explicitly indicated.

2.6. Asymptotic conventions. Statements about zeros below are understood for all sufficiently large q , unless a different range is stated. The threshold q_0 depends on finitely many fixed choices (the exponent, test functions, weights, constants and certificates, and the tolerances of the published results we use); the order in which these are chosen is spelled out in [Section 13](#). We write $o(1)$ for a quantity tending to 0 as $q \rightarrow \infty$, uniformly in everything except the fixed choices. The number

$$\eta = 10^{-6}$$

is a fixed tolerance used throughout, and the exponent is

$$L = 3.99.$$

Certificate coefficients use the scale $S = 10^{16}$. Counts of characters are unscaled; the precise conversion between coefficients and real quantities is given in [Section 10.1](#).

2.7. Constants and the error budget. The fixed constants of the proof, with their values, their roles and the stage of [Section 13.1](#) at which each is chosen, are listed in [Table 15](#) in [Appendix B](#). The tolerance $\eta = 10^{-6}$ is spent as follows; the allowance 5η in each leaf value covers the analytic errors with room to spare.

- (1) *Detection.* With $\varepsilon = \eta H_0$ in [Input 4.2](#), the weighted prime sum is at least $\frac{\mathcal{L}H_0}{\varphi(q)}(1 - W - \eta)$; so $W \leq 1 - 2\eta$ suffices ([Proposition 4.4](#)).
- (2) *Analytic error.* In [\(5.2\)](#) the error E has two parts ([Section 5.8](#)). The envelope errors, which include the smoothing error by [\(5.1\)](#), total less than $19\varepsilon_1$ for the ordinary characters, since $\sum_{\chi} e^{-A\lambda_{\chi}} \leq K_{\text{far}} < 19$ by [\(6.2\)](#), and at most $4\varepsilon_1$ for the at most two first-family and two reserved characters; together less than $23\varepsilon_1 = \eta$. The outside first-family bound ([Proposition 5.9](#)) adds at most $4c_{\text{out}}/(H_0 C_B^2) \leq \eta$, by the choice of C_B in (S4). Hence $E \leq 2\eta$.
- (3) *Allowance.* Every leaf value includes $\text{final}/S = 5\eta$ ([Section 11.5](#)); so does the program of [Theorem 12.6](#).
- (4) *Conclusion.* By the proof of [Proposition 11.6](#), $W \leq \mathcal{V}(x) - \text{final}/S + E \leq \mathcal{V}(x) - 3\eta$, that is, $W + 3\eta \leq \mathcal{V}(x) < 1$ by [Theorem 10.7](#). Hence $W < 1 - 3\eta$, and [Proposition 4.4](#) applies with a margin η .
- (5) *Other tolerances.* They enter the constraints rather than E : $\varepsilon = \eta J^2$ in [Input 6.1](#) gives the far budget $(1 + \eta)V$ of [\(6.1\)](#), the tolerance η' is built into every feature, diagonal and correlation term of a near row ([Corollary 8.10](#)), and every stored number is rounded in the safe direction ([Section 11.5](#)).

3. PUBLISHED INPUTS

This section records the analytic estimates used in the proof. Statements labelled *Input* are taken from the literature, with notation adapted to this paper. Our main sources are Heath-Brown [\[51\]](#) and Xylouris's dissertation [\[145\]](#); the dissertation, rather than Xylouris's shorter article, is the source of the lemma and equation numbers cited below. Any additional hypothesis needed from a source proof, or extension of a printed statement, is identified explicitly.

3.1. Test functions.

Definition 3.1 (Conditions 1 and 2). Let $x_0 > 0$ and let $f : [0, \infty) \rightarrow \mathbb{R}$ be continuous with $f(t) = 0$ for $t \geq x_0$.

- f satisfies *Condition 1* if it is twice continuously differentiable on $(0, x_0)$ with $|f''| \leq B$ there for some constant B [51, p. 280], [145, Bedingung 1, p. 17].
- f satisfies *Condition 2* if $f \geq 0$ and its Laplace transform $F(z) = \int_0^\infty f(t)e^{-zt} dt$ satisfies $\operatorname{Re} F(z) \geq 0$ for $\operatorname{Re} z \geq 0$ [51, p. 286], [145, Bedingung 2, p. 18].

Condition 1 supplies the regularity needed for the explicit formula. Condition 2 ensures that zero terms with nonnegative real argument can be discarded in an upper bound. The two conditions have different roles, so we specify each one when it is needed.

A function satisfying Condition 1 is bounded, its derivative is bounded on $(0, x_0)$, and its transform F is entire. The *conductor coefficient* of a nonprincipal character χ modulo q is [145, p. 17]

$$\varphi(\chi) = \begin{cases} \frac{1}{4}, & q \text{ cube-free, or } \operatorname{ord} \chi \leq \mathcal{L}, \\ \frac{1}{3}, & \text{otherwise.} \end{cases}$$

In particular $\varphi(\chi) = \frac{1}{4}$ for every real character when $q \geq 8$ (its order is $2 \leq \mathcal{L}$), and $\varphi(\chi) \leq \frac{1}{3}$ always.

Input 3.2 (Comparison lemma [51, Lemma 4.1], [145, Lemma 3.5, p. 22]). Let F_1, F_2 be holomorphic in $\{\operatorname{Re} z \geq 0\}$, with $\operatorname{Re} F_1(z) \geq |F_2(z)|$ on $\operatorname{Re} z = 0$, and suppose that F_1 and F_2 tend to 0 uniformly as $|z| \rightarrow \infty$ in $\operatorname{Re} z \geq 0$. Then $\operatorname{Re} F_1(z) \geq |F_2(z)|$ for all $\operatorname{Re} z \geq 0$.

This is a consequence of the maximum principle [101, 130]. All transforms below are Laplace transforms of bounded functions of compact support [142].

3.2. Explicit formulas. The following two lemmas are Heath-Brown's Lemmas 5.3 and 5.2, as stated by Xylouris [145, Lemmas 3.1 and 3.2, p. 18]. In them $s = \sigma + it$ with

$$(3.1) \quad |\sigma - 1| \leq \frac{(\log \mathcal{L})^{1/2}}{\mathcal{L}}, \quad |t| \leq \mathcal{L}.$$

Input 3.3 (Principal character). Let f satisfy Condition 1 and let s satisfy (3.1). Then

$$\sum_{n=1}^{\infty} \Lambda(n) \frac{\chi_0(n)}{n^s} f\left(\frac{\log n}{\mathcal{L}}\right) = \mathcal{L} F((s-1)\mathcal{L}) + O\left(\frac{\mathcal{L}}{\log \mathcal{L}}\right),$$

with an implied constant depending only on f .

Input 3.4 (Local explicit formula). Let $\chi \neq \chi_0$, let $s = \sigma + it$ satisfy (3.1), and let f satisfy Condition 1 with $f(0) \geq 0$. For every $\varepsilon > 0$ there are $\delta = \delta(f, \varepsilon) \in (0, 1)$ and $q_0 = q_0(f, \varepsilon)$, independent of χ and s , such that for $q \geq q_0$

$$\sum_{n=1}^{\infty} \Lambda(n) \operatorname{Re} \frac{\chi(n)}{n^s} f\left(\frac{\log n}{\mathcal{L}}\right) \leq -\mathcal{L} \sum_{|1+it-\rho| \leq \delta} \operatorname{Re} F((s-\rho)\mathcal{L}) + \frac{f(0)}{2} \varphi(\chi) \mathcal{L} + \varepsilon \mathcal{L},$$

where the sum runs over the nontrivial zeros of $L(s, \chi)$ in the disc, with multiplicity.

These are smoothed forms of the explicit formula [141, 26, 87, 62]; the conductor coefficient $\varphi(\chi)$ comes from Burgess's bounds, through Heath-Brown's growth estimate for $L(s, \chi)$ and a Jensen-type formula [51, §§2–3]. Heath-Brown's printed Lemma 5.2 omits the factor $\Lambda(n)$ on the left, a misprint that Xylouris's statement corrects. Xylouris notes [145, p. 36] that the radius δ and the threshold q_0 of Input 3.4, and the implied constant of Input 3.3, depend only on ε and

on upper bounds for x_0 , x_0^{-1} , $\sup |f|$ and $\sup_{(0,x_0)}(|f'| + |f''|)$; so both inputs hold uniformly for a family of test functions with such common bounds. In the proof of Heath-Brown's Lemma 3.1, on which Lemma 5.2 rests, the disc radius is $\delta = \min(1/(2k), \varepsilon_0/(3c_0k^2))$ with $k \geq 3$, so $\delta \leq 1/6$. The same proof works for every smaller radius, and Heath-Brown remarks after Lemma 5.2 that $\delta = 1/\log \mathcal{L}$ may be taken. We may and do assume $\delta < 1/2$.

Input 3.5 (The height l [51, Lemma 6.1], [145, Lemma 3.3, p. 19]). *There are q_0 and a number $l = l(q) \in \mathbb{N}$ with $l \leq \mathcal{L}/10$ such that for $q \geq q_0$ the function $\prod_{\chi} L(s, \chi)$ has no zeros in $R(10l) \setminus R(l)$.*

The principal L -function has no zeros in $R(l)$ for large q [145, p. 21]. All results we quote from the tables of [51, 145] concern zeros in $R(l)$ for this l ; Heath-Brown's proof of Lemma 6.1 produces l and the later sections use it only through $1 \leq l \leq \mathcal{L}/10$ and the zero-free annulus.

Input 3.6 (All-zero envelope [145, Lemma 3.10, pp. 35–36]). *Let C_0 be the size of the square in Input 4.2, and fix $\varepsilon, \lambda_{11} > 0$ and an integer $m \geq 1$. For $\lambda \in (\lambda_{11}, C_0]$, suppose that:*

- (i) *each f_{1i}^{λ} , $1 \leq i \leq m$, satisfies Condition 1 with support parameter $x_{0i} > 0$ and $f_{1i}^{\lambda}(0) \geq 0$;*
- (ii) *the quantities x_{0i} , x_{0i}^{-1} , $\sup |f_{1i}^{\lambda}|$, and $\sup_{(0,x_{0i})}(|(f_{1i}^{\lambda})'| + |(f_{1i}^{\lambda})''|)$ have common bounds depending only on λ_{11} and C_0 , and $f_1^{\lambda} = \sum_i f_{1i}^{\lambda} \geq 0$;*
- (iii) *H_2 is holomorphic in $\operatorname{Re} z \geq 0$, $|H_2(\lambda + it)| \leq \operatorname{Re} F_1^{\lambda}(it)$ for every real t , and both transforms tend uniformly to 0 as $|z| \rightarrow \infty$ in that half-plane;*
- (iv) *$\chi \neq \chi_0$ and $L(s, \chi)$ has no zero with $1 - \lambda/\mathcal{L} < \beta \leq 1$ and $|\gamma| \leq 1$.*

Here F_1^{λ} is the Laplace transform of f_1^{λ} . Then there is an effectively computable q_0 , depending on $\varepsilon, m, \lambda_{11}, C_0$ but not on λ , such that for $q \geq q_0$

$$\sum'_{\rho} |H_2((1 - \rho)\mathcal{L})| \leq F_1^{\lambda}(-\lambda) + \frac{f_1^{\lambda}(0)}{6} + \varepsilon,$$

the sum running over the zeros of $L(s, \chi)$ in the square of Input 4.2.

We have included the hypothesis $f_{1i}^{\lambda}(0) \geq 0$ in (i), although it is omitted from the printed statement. The proof [145, pp. 35–36] applies Input 3.2, then Input 3.4 at $s = 1 - \lambda/\mathcal{L}$ to each f_{1i}^{λ} , then Input 3.3. It therefore uses $f_{1i}^{\lambda}(0) \geq 0$ for each i ; every application below has this property. The zero-free hypothesis is used only to know that the zeros of $L(s, \chi)$ in the disc of Input 3.4 about 1 have $\beta \leq 1 - \lambda/\mathcal{L}$. We use the lemma in this localized form (Section 5).

3.3. Character sums and the Selberg sieve.

Input 3.7 (Burgess [51, Lemma 2.1], $k = 3$). *Let $q \geq 1$ and let χ be a primitive character modulo q . Let $N \geq 1$ and $1 \leq H \leq q$. Then for every $\varepsilon > 0$*

$$\sum_{N < n \leq N+H} \chi(n) \ll_{\varepsilon} q^{1/9+\varepsilon} H^{2/3}.$$

This is the case $k = 3$ of Heath-Brown's statement, which is Burgess's theorem [12, 11, 13, 14]; see [52] for an account. It improves on the Pólya–Vinogradov inequality [108, 136] for short intervals. For special moduli stronger bounds are known [7, 16, 6, 70, 100], but we do not use them.

The weights ξ_d in the next input are Selberg's sieve weights [116, 117] in the form used for zero-density estimates by Graham and Heath-Brown; see [45, 93, 34] for the Selberg sieve in general.

Input 3.8 (Graham [40, p. 84]; [51, (11.13), $U = 1$]). *Let $z \geq 2$ and $\xi_d = \mu(d) \log(z/d) / \log z$ for $d \leq z$, $\xi_d = 0$ otherwise. Then for $N \geq z$*

$$\sum_{n \leq N} \left(\sum_{d|n} \xi_d \right)^2 = \frac{N}{\log z} + O\left(\frac{N}{\log^2 z}\right).$$

3.4. Log-free zero density.

Input 3.9 (Log-free density [66, Theorem 1], as quoted in the proof of [51, Lemma 6.1]). *For $T \geq 1$, $\frac{4}{5} \leq \sigma \leq 1$ and every $\varepsilon > 0$,*

$$\sum_{\chi \bmod q} N(\sigma, T, \chi) \ll_{\varepsilon} (qT)^{(2+\varepsilon)(1-\sigma)},$$

where $N(\sigma, T, \chi)$ counts the nontrivial zeros $\beta + i\gamma$ of $L(s, \chi)$, with multiplicity, in $\beta \geq \sigma$, $|\gamma| \leq T$.

Estimates of this kind go back to Linnik [77, 78]; see [133, 32, 37, 84, 66, 91, 10] and [51, (1.4)], [145, Prinzip 3, p. 11]; explicit versions are in [127], and related estimates in [58, 56, 65, 49, 120, 105]. We use it only to see that the number of zeros of all $L(s, \chi)$ modulo q , counted with multiplicity, with $\lambda \leq 3$ and $|\gamma| \leq T$ is bounded independently of q , for $T = 2$ and for $T = \mathcal{L}$: take $1 - \sigma = 3/\mathcal{L}$; then $(qT)^{(2+\varepsilon)(1-\sigma)} \ll (q\mathcal{L})^{7/\mathcal{L}} \ll 1$.

3.5. Exceptional zeros.

Input 3.10 (Siegel zeros [50, Corollary 1, p. 406]). *Suppose $L(\beta_0, \chi) = 0$ for a real character χ modulo q , not necessarily primitive, and a real $\beta_0 \geq 1 - 1/(3 \log q)$, and put $\eta_0 = ((1 - \beta_0) \log q)^{-1}$. For every $\delta > 0$ there is an effectively computable constant $\eta_{\text{HB}}(\delta)$ such that $\eta_0 \geq \eta_{\text{HB}}(\delta)$ implies $P(a, q) \leq q^{3+\delta}$ for every a coprime to q .*

The zero-location results of Heath-Brown and Xylouris that we use are stated in [Section 7](#), where they are applied, and the inputs for the small first-zero regime in [Section 12](#).

4. PRIME DETECTION

Our goal is to make a nonnegative weighted sum over primes in a reduced residue class strictly positive. The explicit formula expresses this as a comparison between a main term and a sum over zeros. We first fix the weight, then state the precise zero-sum bound that will suffice.

4.1. The kernel. Fix

$$T = 0.416829, \quad \kappa = T/16,$$

and the sixteen step heights $b_0, \dots, b_{15}$ of [Table 3](#). Put

$$\psi(t) = \sum_{i=0}^{15} b_i \mathbf{1}_{[i\kappa, (i+1)\kappa)}(t), \quad \Psi(z) = \int_0^T \psi(t) e^{-zt} dt.$$

For the target exponent L put $A = L - 2T$; at $L = 3.99$ we have $A = 3.156342$. The *prime weight* is

$$h(t) = (\psi * \psi)(t - A),$$

which is supported in $[A, L]$. Its Laplace transform is

$$(4.1) \quad H(z) = \int h(t) e^{-zt} dt = e^{-Az} \Psi(z)^2, \quad H_0 = H(0) = \Psi(0)^2.$$

Numerically $\Psi(0) = \kappa \sum_i b_i = 0.193642\dots$ and $H_0 = 0.037497\dots$

i	b_i	i	b_i	i	b_i	i	b_i
0	0.02318741	4	0.23237155	8	0.47877833	12	0.75172165
1	0.07159337	5	0.29081583	9	0.54505502	13	0.82236428
2	0.12265085	6	0.35147264	10	0.61279917	14	0.89340300
3	0.17627704	7	0.41418551	11	0.68177389	15	0.96451862

TABLE 3. The step heights of the kernel ψ . They are exact decimal numbers.

Xylouris [145, (3.51)], following Heath-Brown [51, §13], uses the triangular weights

$$h_{L',K}(t) = \begin{cases} t - (L' - 2K), & L' - 2K \leq t \leq L' - K, \\ L' - t, & L' - K \leq t \leq L', \\ 0, & \text{otherwise,} \end{cases}$$

for $L' > 2K > 0$; the transform of $h_{L',K}$ is $e^{-(L'-2K)z}((1 - e^{-Kz})/z)^2$.

Lemma 4.1 (Decomposition into positive triangular weights). *For the step function ψ and shifted convolution h defined above,*

$$h = \sum_{k=0}^{30} c_k h_{A+(k+2)\kappa, \kappa}, \quad c_k = \sum_{\substack{0 \leq i, j \leq 15 \\ i+j=k}} b_i b_j > 0.$$

Every triangle in this sum has left endpoint $A + k\kappa \geq A$.

Proof. For $i, j \geq 0$ the convolution $\mathbf{1}_{[i\kappa, (i+1)\kappa]} * \mathbf{1}_{[j\kappa, (j+1)\kappa]}$ is the tent function of height κ supported on $[(i+j)\kappa, (i+j+2)\kappa]$, which is $h_{(i+j+2)\kappa, \kappa}$. Expanding $\psi * \psi$ and shifting by A gives the formula. Each c_k is a sum of products of positive numbers, and the sum over $i+j=k$ is nonempty for $0 \leq k \leq 30$. $\square$

4.2. The positivity criterion. The following is Xylouris's form of the explicit formula for a finite combination of triangles [145, (3.56)–(3.58) and the lines between (3.56) and (3.57), pp. 34–35]; it refines Heath-Brown's Lemmas 13.1 and 13.2 [51].

Input 4.2 (Xylouris's criterion). *Let $h = \sum_i \alpha_i h_{L_i, K_i}$ be a finite real combination of triangles with $L_i > 2K_i + 3$ for every i , and let H be its Laplace transform. For every $\varepsilon > 0$ there are $C_0 = C_0(\varepsilon)$ and $q_0 = q_0(\varepsilon)$ such that for every $q \geq q_0$ and every a coprime to q ,*

$$\sum_{p \equiv a(q)} \frac{\log p}{p} h\left(\frac{\log p}{\mathcal{L}}\right) \geq \frac{\mathcal{L}}{\varphi(q)} \left[H(0) - \sum_{\chi \neq \chi_0} \sum_{\rho}' |H((1-\rho)\mathcal{L})| - \varepsilon \right],$$

where $\sum_{\rho}'$ runs over the zeros $\rho = \beta + i\gamma$ of $L(s, \chi)$, with multiplicity, in the square $1 - C_0/\mathcal{L} \leq \beta \leq 1$, $|\gamma| \leq C_0/\mathcal{L}$.

Xylouris states the extension to combinations without a separate proof, and it follows from Heath-Brown's proof of his Lemma 13.2 [51, §13]. That proof starts from the explicit formula [51, Lemma 13.1], which is linear in the weight, and bounds the zeros outside the square by summing over rectangles $1 - \frac{m+1}{\mathcal{L}} \leq \beta \leq 1 - \frac{m}{\mathcal{L}}$, $\frac{n}{2\mathcal{L}} \leq |\gamma| \leq \frac{n}{\mathcal{L}}$ with $\max(m, n) \geq R$, using a zero-density bound $\ll e^{3m}(1 + n/\mathcal{L})^{3/2}$ for each rectangle and $|F((1-\rho)\mathcal{L})| \ll e^{-(L'-2K)m}n^{-2}$ for each zero. For a combination with positive coefficients c_k one has $|H| \leq \sum_k c_k |H_k|$, so the same estimate holds with $L' - 2K$ replaced by the smallest left endpoint, which exceeds 3; and the zeros in the square contribute at most $\sum |H((1-\rho)\mathcal{L})|$.

In the normalized coordinates (2.1) the square is $\{0 \leq \lambda \leq C_0, |\mu| \leq C_0\}$. Enlarging it only adds nonnegative terms to the zero sum, so the conclusion holds for every square $\{\lambda \leq C_P, |\mu| \leq C_P\}$ with $C_P \geq C_0$. We fix

$$(4.2) \quad C_P \geq \max(C_0(\eta H_0), 3),$$

which fixes the prime window $R_P = \{0 \leq \lambda \leq C_P, |\mu| \leq C_P\}$ of Section 2. Since C_P is fixed and $\log \log \mathcal{L} \rightarrow \infty$, for large q every zero in R_P lies in the rectangle $R(l)$ of Section 2.

Definition 4.3 (Normalized zero sum). For the kernel H in (4.1) and the prime window fixed by (4.2), put

$$W = W(q) = \frac{1}{H_0} \sum_{\chi \neq \chi_0} \sum_{\rho \in R_P} |H((1 - \rho)\mathcal{L})|,$$

where each zero is counted with its multiplicity.

Proposition 4.4 (A zero-sum bound that guarantees a prime). *Let $L > 3 + 2T$. There is q_0 such that for every $q \geq q_0$ with $W(q) \leq 1 - 2\eta$ and every a coprime to q there is a prime p with*

$$p \equiv a \pmod{q}, \quad q^A < p < q^L, \quad A = L - 2T.$$

Here $W(q)$ is formed from the kernel and prime window for this fixed L ; the threshold q_0 is independent of a and of the zero configuration of q .

Proof. By Lemma 4.1, h is a finite positive combination of triangles $h_{L_i, \kappa}$ with $L_i - 2\kappa = A + k\kappa \geq A > 3$. So Input 4.2 applies with $\varepsilon = \eta H_0$; it gives a q_0 such that for every $q \geq q_0$ with $W(q) \leq 1 - 2\eta$,

$$\sum_{p \equiv a(q)} \frac{\log p}{p} h\left(\frac{\log p}{\mathcal{L}}\right) \geq \frac{\mathcal{L} H_0}{\varphi(q)} (1 - W - \eta) \geq \frac{\mathcal{L} H_0 \eta}{\varphi(q)} > 0.$$

The sum is therefore nonempty. Since h vanishes outside (A, L) , some prime $p \equiv a \pmod{q}$ satisfies $A < \log p / \mathcal{L} < L$. $\square$

At $L = 3.99$ the condition $L > 3 + 2T = 3.833658$ holds. Two features of Input 4.2 matter below.

- (i) *Occurrence form.* The zero sum charges each zero occurrence separately, through the absolute value of its own term. Every bound in Section 5 is a bound for a sum of such individual terms. In particular we never need the cancellation between zeros that a grouped form of the explicit formula would record.
 - (ii) *Size of a term.* Since $|\Psi(\lambda - i\mu)| \leq \Psi(\lambda)$ for $\lambda \geq 0$,
- $$(4.3) \quad |H((1 - \rho)\mathcal{L})| = e^{-A\lambda} |\Psi(\lambda - i\mu)|^2 \leq e^{-A\lambda} \Psi(\lambda)^2 \leq H_0 e^{-A\lambda}.$$

A zero at distance λ from $\sigma = 1$ costs at most $e^{-A\lambda}$ times H_0 . The difficulty is entirely with the number of zeros close to $\sigma = 1$.

If no nonprincipal $L(s, \chi)$ has a zero in $R(l)$, then for large q it has none in R_P , and $W = 0$. This is the first exterior regime (Section 12).

5. ZERO COSTS

The zero sum W of Section 4 is a sum over characters. This section bounds the contribution of a single character, in terms of an *anchor* (Section 2.4): a number λ_* such that no zero of $L(s, \chi)$ near $s = 1$ has parameter below λ_* , at which the explicit formula is evaluated. For most characters the anchor is the parameter of a representative zero (Definition 6.4); the first family needs a finer treatment. The argument is Xylouris's proof of his Lemma 3.10 [145, pp. 35–36], which we reproduce in the localized form we need.

5.1. The envelope functions. The function $\mathcal{G}_\phi(\lambda)$ below bounds the total cost of one character. The auxiliary functions B_ϕ , $\tilde{h}$, and C separate the common exponential decay from the correction for a distinguished zero. The parameter ϕ will be $\varphi(\chi)$, or an upper bound for it.

For real λ let $\psi_\lambda(u) = \psi(u)e^{-\lambda u}$ and

$$f_\lambda(t) = 2 \int_0^\infty \psi_\lambda(u) \psi_\lambda(u+t) du = 2 \int_0^{T-t} \psi(u) \psi(u+t) e^{-\lambda(2u+t)} du \quad (t \geq 0),$$

so that $f_\lambda(t) = 0$ for $t \geq T$. Let F_λ be its Laplace transform. For $0 \leq \phi \leq \frac{1}{3}$, $\lambda \geq 0$ and $p \geq a \geq 0$ put

$$\begin{aligned} B_\phi(\lambda) &= \frac{1}{H_0} \left[2 \int_0^T \psi(u) e^{-2\lambda u} \int_u^T \psi(v) dv du + \phi \int_0^T \psi(u)^2 e^{-2\lambda u} du \right], \\ \mathcal{G}_\phi(\lambda) &= e^{-A\lambda} B_\phi(\lambda), \quad \mathcal{G} = \mathcal{G}_{1/3}, \quad \tilde{h}(\lambda) = \frac{\Psi(\lambda)^2}{H_0}, \\ C(p, a) &= \frac{2}{H_0} \int_0^T \psi(u) e^{-2pu} \int_u^T \psi(v) e^{-a(v-u)} dv du. \end{aligned}$$

Lemma 5.1 (Properties of the envelope functions). *With the functions just defined, the following hold. Parts (i)–(iii) hold for every real λ ; in (iv)–(v), take $0 \leq \phi \leq 1/3$, $\lambda \geq 0$, and $p \geq a \geq 0$.*

- (i) $f_\lambda \geq 0$, f_λ is Lipschitz on $[0, T]$, and $f_\lambda(T) = 0$.
- (ii) $\operatorname{Re} F_\lambda(iy) = |\Psi(\lambda + iy)|^2$ for real y .
- (iii) $\operatorname{Re} F_\lambda(z) \geq |\Psi(\lambda + z)|^2$ for $\operatorname{Re} z \geq 0$. In particular $\operatorname{Re} F_\lambda(z) \geq 0$ for $\operatorname{Re} z \geq 0$.
- (iv) $H_0 B_\phi(\lambda) = F_\lambda(-\lambda) + \frac{\phi}{2} f_\lambda(0)$ and $H_0 C(p, a) = F_p(a - p)$. Moreover $C(a, a) = \tilde{h}(a)$.
- (v) B_ϕ , $\mathcal{G}_\phi$ and $\tilde{h}$ are positive and decreasing on $[0, \infty)$, and so is $\mathcal{G}_\phi/w$, where w is either far weight of [Section 6](#).

Proof. (i) Nonnegativity is clear. For $0 \leq t < t' \leq T$, $|f_\lambda(t) - f_\lambda(t')| \leq 2\|\psi_\lambda\|_\infty \int |\psi_\lambda(u+t) - \psi_\lambda(u+t')| du$, which is $O(t' - t)$ because ψ_λ has bounded variation.

(ii) Let $a_\lambda(t) = \int \psi_\lambda(u) \psi_\lambda(u+|t|) du$ for $t \in \mathbb{R}$; it is even and integrable, and its Fourier transform is $|\widehat{\psi_\lambda}(y)|^2 = |\Psi(\lambda + iy)|^2$. Hence $\operatorname{Re} F_\lambda(iy) = \int_0^\infty f_\lambda(t) \cos(yt) dt = \int_{\mathbb{R}} a_\lambda(t) e^{-iyt} dt = |\Psi(\lambda + iy)|^2$.

(iii) Apply [Input 3.2](#) to $F_1 = F_\lambda$ and $F_2(z) = \Psi(\lambda + z)^2$. Both are entire. By (i), $F_\lambda(z) = z^{-1}(f_\lambda(0) + \int_0^T f'_\lambda(t) e^{-zt} dt)$, so $|F_\lambda(z)| \leq (f_\lambda(0) + \|f'_\lambda\|_1)/|z|$ for $\operatorname{Re} z \geq 0$; similarly $|\Psi(\lambda + z)| \leq (\psi_\lambda(0^+) + V(\psi_\lambda))/|z|$ with V the total variation. So both tend to 0 uniformly, and on $\operatorname{Re} z = 0$ we have equality by (ii).

(iv) Substitute $v = u + t$: $F_\lambda(-\lambda) = 2 \iint \psi(u) \psi(v) e^{-\lambda(u+v)} e^{\lambda(v-u)} dv du$ over $0 \leq u \leq v \leq T$, which is the first term of $H_0 B_\phi$; and $f_\lambda(0) = 2 \int \psi^2 e^{-2\lambda u}$. The same substitution gives $F_p(a - p) = 2 \int_0^T \psi(u) e^{-2pu} \int_u^T \psi(v) e^{-a(v-u)} dv du$. Finally $H_0 C(a, a) = 2 \iint_{u \leq v} \psi(u) \psi(v) e^{-a(u+v)} = \Psi(a)^2$.

(v) All integrands are positive and decrease in λ . For the last claim,

$$\frac{\mathcal{G}_\phi(\lambda)}{w(\lambda)} = B_\phi(\lambda) \int_{u_0}^x w_0(t)^2 e^{(2t-A)\lambda} dt,$$

and $2t - A < 0$ on $[u_0, x]$ ([Lemma 6.2](#)). □

The number $\mathcal{G}(\lambda)$ is the cost, in units of H_0 , that the argument below assigns to a character whose zeros near $s = 1$ all have parameter at least λ . For a real character the factor $\phi = \frac{1}{3}$ may be replaced by $\frac{1}{4}$ ([Proposition 5.5](#)).

5.2. Smoothing. Since ψ is a step function, f_λ has corners at the multiples of κ and does not satisfy Condition 1. We therefore work with smooth approximations of ψ in the auxiliary functions only; the prime weight h and its transform H are never changed. For $\theta \in (0, 1)$ fix a function $\psi^\theta \in C^\infty(\mathbb{R})$ with support in $(0, T)$ such that

$$0 \leq \psi^\theta \leq 1, \quad \|\psi - \psi^\theta\|_{L^1} \leq \theta,$$

for instance a mollification of $\psi \mathbf{1}_{[\theta/4, T-\theta/4]}$. Let $\Psi^\theta, f_\lambda^\theta, F_\lambda^\theta, B_\phi^\theta, \tilde{h}^\theta$ and C^θ be defined as above with ψ^θ in place of ψ , and H_0 kept.

Lemma 5.2 (Uniform smooth approximation). *Fix $\theta \in (0, 1)$ and $\Lambda \geq 1$, and choose ψ^θ as above.*

- (i) *For every $\lambda \in [0, \Lambda]$, f_λ^θ satisfies Condition 1 and Condition 2 with $x_0 = T$. The quantities $x_0, x_0^{-1}, \sup |f_\lambda^\theta|$ and $\sup_{(0, T)} (|(f_\lambda^\theta)'| + |(f_\lambda^\theta)''|)$ are bounded independently of $\lambda \in [0, \Lambda]$; the bounds may depend on θ and Λ . Lemma 5.1(i)–(iv) hold for ψ^θ .*
- (ii) *For $\operatorname{Re} z \geq 0$, $|\Psi(z)^2 - \Psi^\theta(z)^2| \leq 3\theta$.*
- (iii) *As $\theta \rightarrow 0$, $B_\phi^\theta \rightarrow B_\phi, \tilde{h}^\theta \rightarrow \tilde{h}$ and $C^\theta \rightarrow C$ uniformly on $\{0 \leq a \leq p \leq \Lambda\}$ and $0 \leq \phi \leq \frac{1}{3}$.*

Proof. (i) The function $(t, \lambda) \mapsto f_\lambda^\theta(t)$ is smooth on $\mathbb{R}^2$, and it vanishes for $t \geq T$ because ψ^θ is supported in a compact subset of $(0, T)$; this gives Condition 1 and the uniform bounds. The proof of Lemma 5.1 applies verbatim to ψ^θ , and Condition 2 is Lemma 5.1(i) and (iii) for ψ^θ . (ii) For $\operatorname{Re} z \geq 0$ we have $|\Psi(z)| \leq \|\psi\|_1 \leq T < \frac{1}{2}$, $|\Psi^\theta(z)| \leq \|\psi\|_1 + \theta$ and $|\Psi(z) - \Psi^\theta(z)| \leq \theta$; hence $|\Psi^2 - (\Psi^\theta)^2| \leq \theta(2T + \theta) \leq 3\theta$. (iii) Each quantity is a bilinear expression in (ψ, ψ) with a kernel bounded by 1 on $[0, T]^2$ for the parameters in question, and $\|\psi\|_\infty, \|\psi^\theta\|_\infty \leq 1$. $\square$

5.3. The disc and a uniform zero count. Throughout this section $\lambda_{11} = 0.05$, the smoothing parameter θ is the one fixed in stage (S3) of Section 13.1, and $\varepsilon_1 = \eta/23$. Let $\mathcal{A} \subset [\lambda_{11}, C_P]$ be the finite set consisting of the points of the grid $\lambda_{11} + h_{\mathcal{A}}\mathbb{Z}$ in $[\lambda_{11}, C_P]$ together with the anchors of the first-family bounds of all leaves (Sections 5.6 and 5.7). For a leaf with first-zero cell $[a, b]$ and lower bound p^* for λ' (the lower end of its gap if it has one, and p otherwise; Section 2.4), these are a and p^* if the leaf is inside and $p^\circ = \max(a, p^*)$ if it is outside (Section 11.5); the mesh $h_{\mathcal{A}} > 0$ is chosen so that $|B_\phi(x) - B_\phi(y)| \leq \varepsilon_1/4$ for $\phi \in \{\frac{1}{4}, \frac{1}{3}\}$ and $x, y \in [0, C_P]$ with $|x - y| \leq h_{\mathcal{A}}$. In stage (S5) we fix a radius $\delta_0 \in (0, \frac{1}{3}]$ such that Input 3.4 holds with radius δ_0 for each of the finitely many pairs of a test function and a tolerance used in this paper: the function $f^{(1)}$ below with tolerance 1, the functions f_σ^θ ($\sigma \in \mathcal{A}$) with tolerance $H_0\varepsilon_1/8$, and the detectors (Definition 8.1) of the near rows with the tolerances of Lemma 8.7. This is possible because Input 3.4 remains true for every smaller radius (see the remark after it). The *disc* of a character χ is the set of zeros ρ of $L(s, \chi)$ with $|1 - \rho| \leq \delta_0$. For fixed C and large q the square $\{\lambda \leq C, |\mu| \leq C\}$ lies in the disc, since there $|1 - \rho| \leq \sqrt{2}C/\mathcal{L}$.

Lemma 5.3 (A uniform bound for zeros in a fixed window). *Let $C \geq 1$. There are $N = N(C)$ and q_0 such that for $q \geq q_0$ the following holds. If $\chi \neq \chi_0$ and every zero in the disc of χ has parameter at least λ_{11} , then $L(s, \chi)$ has at most N zeros, counted with multiplicity, with $\lambda \leq C$ and $|\mu| \leq C$.*

Proof. Let $K_1 = 1/(4C)$ and $\vartheta(z) = (1 - e^{-K_1 z})/z$, the transform of $\mathbf{1}_{[0, K_1]}$. Let $f^{(1)}(t) = 2 \int_0^{K_1-t} e^{-\lambda_{11}(2u+t)} du$ for $0 \leq t \leq K_1$ and $f^{(1)}(t) = 0$ for $t \geq K_1$. This is the function $f_{\lambda_{11}}$ of Section 5.1 for the step $\mathbf{1}_{[0, K_1]}$; it equals $(e^{-\lambda_{11}t} - e^{-\lambda_{11}(2K_1-t)})/\lambda_{11}$ on $[0, K_1]$, so it satisfies Condition 1, and $f^{(1)}(0) \geq 0$. As in Lemma 5.1(iii), $\operatorname{Re} F^{(1)}(z) \geq |\vartheta(\lambda_{11} + z)|^2 \geq 0$ for $\operatorname{Re} z \geq 0$.

Apply Input 3.4 to $f^{(1)}$ at $s = 1 - \lambda_{11}/\mathcal{L}$, with tolerance 1 and a radius at most δ_0 . For large q , the fixed normalized square lies in this smaller disc. All its zeros still have parameter at least

λ_{11} . In the remainder of this proof, “the disc” refers to this smaller disc. The point s lies in the range (3.1) for large q , and $(s - \rho)\mathcal{L} = (\lambda_\rho - \lambda_{11}) - i\mu_\rho$. Since $-\operatorname{Re} \chi(n) \leq \chi_0(n)$ for every n and $f^{(1)} \geq 0$, Input 3.3 gives

$$-\sum_n \Lambda(n) \operatorname{Re} \frac{\chi(n)}{n^s} f^{(1)}\left(\frac{\log n}{\mathcal{L}}\right) \leq \sum_n \Lambda(n) \frac{\chi_0(n)}{n^s} f^{(1)}\left(\frac{\log n}{\mathcal{L}}\right) = \mathcal{L} F^{(1)}(-\lambda_{11}) + o(\mathcal{L}).$$

Hence $\sum_{\rho \text{ in the disc}} \operatorname{Re} F^{(1)}((\lambda_\rho - \lambda_{11}) - i\mu_\rho) \leq F^{(1)}(-\lambda_{11}) + \frac{1}{6} f^{(1)}(0) + 2 =: c_1$ for large q . Every term is nonnegative, because $\lambda_\rho \geq \lambda_{11}$. For a zero with $\lambda \leq C$, $|\mu| \leq C$ put $z = \lambda - i\mu$; then $|K_1 z| \leq \sqrt{2}/4 < \frac{1}{2}$ and $|1 - e^{-w}| \geq |w| - |w|^2 e^{|w|}/2 \geq |w|/2$ for $|w| \leq \frac{1}{2}$, so the term is at least $|\vartheta(z)|^2 \geq K_1^2/4$. The number of such zeros is therefore at most $N = \lfloor 4c_1/K_1^2 \rfloor$. $\square$

5.4. The anchored envelope. The following inequality is the common source of all the costs below. For $\sigma \in \mathcal{A}$, a character $\chi \neq \chi_0$ and a zero ρ put

$$\Phi_\sigma^\theta(\rho) = \operatorname{Re} F_\sigma^\theta((\lambda_\rho - \sigma) - i\mu_\rho).$$

By Lemma 5.1(iii), $\Phi_\sigma^\theta(\rho) \geq |\Psi^\theta(\lambda_\rho - i\mu_\rho)|^2 \geq 0$ whenever $\lambda_\rho \geq \sigma$.

Lemma 5.4 (Anchored inequality). *Fix the smoothing, finite anchor set $\mathcal{A}$, and common disc radius as in Section 5.3. There is q_0 , uniform over $\chi \neq \chi_0$ and $\sigma \in \mathcal{A}$, such that for $q \geq q_0$,*

$$\sum_{\rho \text{ in the disc of } \chi} \Phi_\sigma^\theta(\rho) \leq H_0 B_{\varphi(\chi)}^\theta(\sigma) + \frac{1}{4} H_0 \varepsilon_1.$$

Proof. Apply Input 3.4 to $f = f_\sigma^\theta$ at $s = 1 - \sigma/\mathcal{L}$ (so $t = 0$), with tolerance $H_0 \varepsilon_1/8$ and radius δ_0 , and bound the prime sum as in the proof of Lemma 5.3, using Input 3.3:

$$\mathcal{L} \sum_\rho \Phi_\sigma^\theta(\rho) \leq \frac{f_\sigma^\theta(0)}{2} \varphi(\chi) \mathcal{L} + \mathcal{L} F_\sigma^\theta(-\sigma) + \frac{1}{8} H_0 \varepsilon_1 \mathcal{L} + O\left(\frac{\mathcal{L}}{\log \mathcal{L}}\right).$$

By Lemma 5.1(iv) for ψ^θ the right side is at most $\mathcal{L}(H_0 B_{\varphi(\chi)}^\theta(\sigma) + \frac{1}{4} H_0 \varepsilon_1)$ for large q . Only the finitely many functions f_σ^θ , $\sigma \in \mathcal{A}$, occur, so one q_0 serves them all. $\square$

Proposition 5.5 (Localized envelope). *For $q \geq q_0$ the following holds. Let $\chi \neq \chi_0$ and $\lambda_* \in [\lambda_{11}, C_P]$, and suppose that every zero in the disc of χ has parameter at least λ_* . Then*

$$\frac{1}{H_0} \sum_{\rho \in R_P} |H((1 - \rho)\mathcal{L})| \leq \mathcal{G}_{\varphi(\chi)}(\lambda_*) + \varepsilon_1 e^{-A\lambda_*},$$

the sum running over the zeros of $L(s, \chi)$ in R_P with multiplicity. In particular the left side is at most $\mathcal{G}(\lambda_) + \varepsilon_1 e^{-A\lambda_*}$, and at most $\mathcal{G}_{1/4}(\lambda_*) + \varepsilon_1 e^{-A\lambda_*}$ if χ is real.*

Proof. Let $N = N(C_P)$ be as in Lemma 5.3. The parameter θ of stage (S3) satisfies

$$(5.1) \quad 3(N + 6)\theta + H_0 \sup |B_\phi^\theta - B_\phi| \leq \frac{1}{4} H_0 \varepsilon_1,$$

the supremum over $[0, C_P]$ and $\phi \in \{\frac{1}{4}, \frac{1}{3}\}$ (Lemma 5.2(iii)). Let $\sigma \in \mathcal{A}$ be the largest grid point with $\sigma \leq \lambda_*$, so that $\lambda_* - \sigma \leq h_{\mathcal{A}}$. Let $\rho \in R_P$. Then ρ is in the disc, so $\lambda_\rho \geq \lambda_* \geq \sigma$, and by (4.3), Lemma 5.2(ii) and Lemma 5.1(iii) for ψ^θ ,

$$|H((1 - \rho)\mathcal{L})| = e^{-A\lambda_\rho} |\Psi(\lambda_\rho - i\mu_\rho)|^2 \leq e^{-A\lambda_*} (\Phi_\sigma^\theta(\rho) + 3\theta).$$

The zeros of the disc that are not in R_P have $\Phi_\sigma^\theta \geq 0$. Summing over the at most N zeros in R_P and using Lemma 5.4 and (5.1),

$$\sum_{\rho \in R_P} |H((1 - \rho)\mathcal{L})| \leq e^{-A\lambda_*} \left(H_0 B_{\varphi(\chi)}^\theta(\sigma) + \frac{1}{4} H_0 \varepsilon_1 + 3N\theta \right) \leq e^{-A\lambda_*} H_0 (B_{\varphi(\chi)}(\sigma) + \frac{1}{2} \varepsilon_1),$$

and $B_{\varphi(\chi)}(\sigma) \leq B_{\varphi(\chi)}(\lambda_*) + \frac{1}{4}\varepsilon_1$ by the choice of h_A . Finally $\varphi(\chi) \leq \frac{1}{3}$ always, $\varphi(\chi) = \frac{1}{4}$ for real χ (whose order is $2 \leq \mathcal{L}$), and B_ϕ increases with ϕ . $\square$

Remark 5.6. **Proposition 5.5** is a localized form of Xylouris's Lemma 3.10 [145, pp. 35–36], applied with $m = 1$, $f_1 = f_\sigma^\theta$ and $H_2 = (\Psi^\theta)^2$, for which his condition (3.59) holds with equality by **Lemma 5.1(ii)**; the prime weight H is then compared with $(\Psi^\theta)^2$ through **Lemma 5.2(ii)**. We require only that the zeros in the disc have parameter at least λ_* , where Xylouris requires that $L(s, \chi)$ have no zero with $1 - \lambda_*/\mathcal{L} < \beta \leq 1$ and $|\gamma| \leq 1$; his proof uses this hypothesis only for the zeros in the disc, which have $|\gamma| \leq \delta_0 < 1$. His printed statement allows f_1 to be a sum of m pieces; for $m > 1$ his proof needs in addition that $f_{1i}(0) \geq 0$ for each piece, and that the sum over the square be enlarged to a sum over a disc common to all pieces before **Input 3.4** is applied to them. With $m = 1$ neither point arises. Finally, we use finitely many anchors, so that no uniformity in the anchor is needed. The coefficient $\varphi(\chi)/2$ comes from **Input 3.4**; the printed Lemma 3.10 uses $\varphi(\chi) \leq \frac{1}{3}$.

5.5. Ordinary and reserved characters. Let χ be a nonprincipal character (in the case tree, one outside the first family; in **Theorem 12.6**, any), and let λ_χ be the parameter of its T^* -representative or of its height-one representative (**Definition 6.4**). If $L(s, \chi)$ has a zero in R_P , then that zero lies in $R(l)$ and has $|\gamma| \leq C_P/\mathcal{L} < T^*$, so the representative exists and $\lambda_\chi \leq C_P$. In the regimes in which we use these costs $\lambda_1 \geq 0.1$, so $\lambda_\chi \geq \lambda_{11}$.

Corollary 5.7 (Cost of a character with a representative). *Suppose $\lambda_1 \geq 0.1$. Let $\chi \neq \chi_0$ have a zero in R_P , and let λ_χ be the parameter of either its T^* -representative or its height-one representative. For sufficiently large q , its contribution to W is at most*

$$\mathcal{G}(\lambda_\chi) + \varepsilon_1 e^{-A\lambda_\chi}, \quad \varepsilon_1 = \eta/23.$$

If χ is real, $\mathcal{G}$ may be replaced by $\mathcal{G}_{1/4}$. The threshold is uniform over these choices of character and representative.

Proof. We check the hypothesis of **Proposition 5.5** with $\lambda_* = \lambda_\chi$. A zero in the disc with parameter at most C_P lies in $R(l)$ for large q , since $C_P < \frac{1}{3} \log \log \mathcal{L}$ and $|\gamma| \leq \delta_0 < 1 \leq l$; it has $|\gamma| \leq \delta_0 < \frac{1}{2} \leq T^*$, so its parameter is at least λ_χ by the minimality of the representative. A zero in the disc with parameter above C_P has parameter above λ_χ . $\square$

5.6. The first family inside the buffer. Suppose that $\rho_1 \in R_B$, that $\lambda_1 \in [a, b]$ with $a \geq \lambda_{11}$, and that every zero occurrence of the first family in $R(l)$ other than the *distinguished* ones has parameter at least p , where $p \leq C_P$; equivalently, $p \leq \lambda'$. We assume that $a, p \in \mathcal{A}$ (in the leaves they are the numbers a and p^* of **Section 5.3**). The distinguished occurrences and their number α per character are those of **Definition 2.2** and the table after it. Let $\phi_1 = \frac{1}{4}$ if χ_1 is real and $\phi_1 = \frac{1}{3}$ otherwise, and put

$$\begin{aligned} J_{\text{old}} &= n \left[e^{-Ap} (B_{\phi_1}(a) - \alpha \tilde{h}(a))_+ + \alpha e^{-Aa} \tilde{h}(a) \right], \\ J_{\text{new}} &= n \left[\alpha e^{-Aa} \tilde{h}(a) + e^{-Ap} (B_{\phi_1}(p) - \alpha C(p, a)) \right] \quad (p \geq b). \end{aligned}$$

Proposition 5.8 (Cost of the first family inside the buffer). *Suppose $\rho_1 \in R_B$, $\lambda_1 \in [a, b]$, $a, p \in \mathcal{A}$, $a \geq \lambda_{11}$, and $p \leq \min(\lambda', C_P)$. For sufficiently large q , the contribution of all n characters of the first family to W is at most $J_{\text{old}} + n\varepsilon_1$. If also $p \geq b$, it is at most $J_{\text{new}} + n\varepsilon_1$. Here J_{old} and J_{new} are the displayed functions of a, p, n, α , and ϕ_1 .*

The two bounds use different anchors. The bound J_{old} uses the lower endpoint a of the first-zero cell. The bound J_{new} uses p , the lower bound for the remaining zeros, and estimates the

distinguished zero together with its correction in the explicit formula. Their combined expression is smaller than a separate estimate of the two terms.

Proof. It suffices to treat one first-family character χ ; the bound for its conjugate is the same, because R_P and the disc are symmetric under conjugation. A zero of the disc of χ lies in $R(l)$ or has parameter above $\frac{1}{3} \log \log \mathcal{L} > C_P$. So every zero of the disc has parameter at least $\lambda_1 \geq a$, and every non-distinguished one has parameter at least p . Write $\mathcal{S}$ for the distinguished occurrences of χ ; they lie in R_B , hence in the disc.

The bound J_{old} . Every non-distinguished occurrence has parameter at least $\max(p, \lambda_1) \geq \max(p, a)$, and J_{old} decreases as p increases; so we may and do assume $p \geq a$. For a non-distinguished $\rho \in R_P$ we then have $\lambda_\rho \geq p \geq a$, and $|H((1-\rho)\mathcal{L})| \leq e^{-Ap} |\Psi(\lambda_\rho - i\mu_\rho)|^2 \leq e^{-Ap} (\Phi_a^\theta(\rho) + 3\theta)$. By Lemma 5.4 at $\sigma = a$, since every term Φ_a^θ in the disc is nonnegative and there are at most N zeros in R_P ,

$$\sum_{\rho \in R_P \setminus \mathcal{S}} |H((1-\rho)\mathcal{L})| \leq e^{-Ap} \left(H_0 B_{\phi_1}^\theta(a) + \frac{1}{4} H_0 \varepsilon_1 + 3N\theta - \sum_{\rho \in \mathcal{S}} \Phi_a^\theta(\rho) \right).$$

For $\rho \in \mathcal{S}$ put $X_\rho = |\Psi(\lambda_1 - i\mu_\rho)|^2$; then $\Phi_a^\theta(\rho) \geq X_\rho - 3\theta$ and $|H((1-\rho)\mathcal{L})| = e^{-A\lambda_1} X_\rho$. We add $e^{-A\lambda_1} X_\rho$ for every $\rho \in \mathcal{S}$; for those in R_P this is their contribution, and for the others it is a nonnegative addition. By (5.1) the total is at most

$$e^{-Ap} H_0 B_{\phi_1}(a) + \sum_{\rho \in \mathcal{S}} X_\rho (e^{-A\lambda_1} - e^{-Ap}) + H_0 \varepsilon_1.$$

If $\lambda_1 \leq p$, each summand increases with X_ρ , and $X_\rho \leq \Psi(\lambda_1)^2 \leq \Psi(a)^2 = H_0 \tilde{h}(a)$; so the total is at most $H_0 [e^{-Ap} (B_{\phi_1}(a) - \alpha \tilde{h}(a)) + \alpha e^{-A\lambda_1} \tilde{h}(a) + \varepsilon_1]$, which is at most $H_0 (J_{\text{old}}/n + \varepsilon_1)$. If $\lambda_1 > p$, each summand is ≤ 0 , and the total is at most $H_0 [e^{-Ap} B_{\phi_1}(a) + \varepsilon_1] \leq H_0 [e^{-Ap} (B_{\phi_1}(a) - \alpha \tilde{h}(a)) + \alpha e^{-Ap} \tilde{h}(a) + \varepsilon_1]$, which is again at most $H_0 (J_{\text{old}}/n + \varepsilon_1)$ because $p \geq a$.

The bound J_{new} . Now $p \geq b \geq \lambda_1$. Apply Lemma 5.4 at $\sigma = p$. The non-distinguished zeros of the disc have $\Phi_p^\theta \geq 0$, and for those in R_P , $|H((1-\rho)\mathcal{L})| \leq e^{-Ap} (\Phi_p^\theta(\rho) + 3\theta)$. Therefore the contribution of χ is at most

$$e^{-Ap} \left(H_0 B_{\phi_1}^\theta(p) + \frac{1}{4} H_0 \varepsilon_1 + 3N\theta \right) + \sum_{\rho \in \mathcal{S}} \left(e^{-A\lambda_1} |\Psi(\lambda_1 - i\mu_\rho)|^2 \mathbf{1}_{\rho \in R_P} - e^{-Ap} \Phi_p^\theta(\rho) \right).$$

The indicator may be replaced by 1, since the term it multiplies is nonnegative, and $|\Psi(\lambda_1 - i\mu_\rho)|^2 \leq |\Psi^\theta(\lambda_1 - i\mu_\rho)|^2 + 3\theta$. For a real number y consider

$$D(y) = e^{-A\lambda_1} |\Psi^\theta(\lambda_1 + iy)|^2 - e^{-Ap} \operatorname{Re} F_p^\theta(\lambda_1 - p + iy) = \int_0^\infty k(t) \cos(yt) dt,$$

where, by Lemma 5.1(ii) for ψ^θ and the definition of f_p^θ ,

$$k(t) = e^{-A\lambda_1} f_{\lambda_1}^\theta(t) - e^{-Ap} e^{(p-\lambda_1)t} f_p^\theta(t) = 2 \int \psi^\theta(u) \psi^\theta(u+t) e^{-\lambda_1 t} (e^{-\lambda_1(A+2u)} - e^{-p(A+2u)}) du.$$

Since $\lambda_1 \leq p$, the integrand is nonnegative, so $k \geq 0$ and $D(y) \leq D(0) = \int_0^\infty k(t) dt$. Moreover $k(t)$ decreases as λ_1 increases, for each t , because both $e^{-\lambda_1 t}$ and the bracket decrease. Hence

$$D(y) \leq \int_0^\infty k(t) \big|_{\lambda_1=a} dt = e^{-Aa} \Psi^\theta(a)^2 - e^{-Ap} F_p^\theta(a-p) = H_0 (e^{-Aa} \tilde{h}^\theta(a) - e^{-Ap} C^\theta(p, a)),$$

by Lemma 5.1(iv). By Lemma 5.2(ii),(iii), $H_0 |\tilde{h}^\theta(a) - \tilde{h}(a)| \leq 3\theta$ and $H_0 |C^\theta(p, a) - C(p, a)| \leq 3\theta$. So each of the $\alpha \leq 2$ distinguished occurrences contributes at most $H_0 (e^{-Aa} \tilde{h}(a) - e^{-Ap} C(p, a)) + 9\theta$, and with (5.1) the contribution of χ is at most $H_0 (J_{\text{new}}/n + \varepsilon_1)$. $\square$

We write $J_1 = \min(J_{\text{old}}, J_{\text{new}})$, where J_{new} is omitted when $p < b$. In every leaf of the case tree the anchor p is at most $2.293 < 3 \leq C_P$, so the hypothesis $p \leq C_P$ holds. In J_{new} the quantity $B_{\phi_1}(p) - \alpha C(p, a)$ may be negative; the argument does not require it to be positive.

5.7. The first family outside the buffer. Suppose now that $\rho_1 \notin R_B$. Since $\lambda_1 < C_B$ in the regimes where this occurs, this means $|\mu_1| > C_B$. The global zero ρ_1 is not in R_P and contributes nothing to W , but the first-family characters may have other zeros in R_P . As before, the distinguished occurrences are ρ_1 (and $\bar{\rho}_1$) for χ_1 and $\bar{\rho}_1$ for $\bar{\chi}_1$, and every other zero occurrence of the first family in $R(l)$ has parameter at least $p \geq a$, where $p \in \mathcal{A}$ (in the leaves, $p = p^\circ$).

For a first-family character χ with a zero in R_P , let t_χ be the least parameter of its zeros in R_P . Then $t_\chi \geq p$, and the corresponding zero has $|\gamma| \leq C_P/\mathcal{L} \leq 1$.

Proposition 5.9 (Cost of the first family outside the buffer). *Suppose $\rho_1 \notin R_B$, $\lambda_1 \geq a \geq \lambda_{11}$, and $p \in \mathcal{A}$ satisfies $a \leq p \leq \lambda'$. For a first-family character with a zero in R_P , let t_χ be the least parameter of its zeros in R_P . Let α be the number of distinguished occurrences per character, and put $\phi_1 = 1/4$ for a real χ_1 and $\phi_1 = 1/3$ otherwise. There is a constant c_{out} , depending only on the fixed choices of stages (S1)–(S3) of Section 13.1 (in particular on θ), such that for $q \geq q_0$ the following holds. Each first-family character χ with a zero in R_P contributes at most*

$$e^{-At_\chi} B_{\phi_1}(p) + \varepsilon_1 + \frac{\alpha c_{\text{out}}}{H_0 C_B^2}$$

to W , and a first-family character without a zero in R_P contributes nothing.

Proof. Apply Lemma 5.4 at $\sigma = p \in \mathcal{A}$. For $\rho \in R_P$ we have $\lambda_\rho \geq t_\chi \geq p$ and $|H((1-\rho)\mathcal{L})| \leq e^{-At_\chi}(\Phi_p^\theta(\rho) + 3\theta)$. The non-distinguished zeros of the disc have $\Phi_p^\theta \geq 0$, and so does a distinguished occurrence with $\lambda_\rho \geq p$. A distinguished occurrence ρ in the disc with $\lambda_\rho < p$ has $\lambda_\rho - p \in [a-p, 0] \subseteq [-C_P, 0]$ and $|\mu_\rho| > C_B$. Since f_p^θ is smooth on $[0, T]$ and vanishes near T , two integrations by parts give, for $x \in [-C_P, 0]$ and real $Y \neq 0$,

$$|\operatorname{Re} F_p^\theta(x + iY)| = \left| \operatorname{Re} \left(\frac{f_p^\theta(0)}{x + iY} + \frac{(f_p^\theta)'(0)}{(x + iY)^2} + \frac{1}{(x + iY)^2} \int_0^T (f_p^\theta)''(t) e^{-(x+iY)t} dt \right) \right| \leq \frac{c_{\text{out}}}{Y^2},$$

where $c_{\text{out}} = C_P f_p^\theta(0) + |(f_p^\theta)'(0)| + e^{C_P T} \|(f_p^\theta)''\|_1$, maximized over the finitely many anchors p in use; the first term is bounded by $f_p^\theta(0)|x|/Y^2$. So the distinguished terms subtracted in Lemma 5.4 total at least $-\alpha c_{\text{out}}/C_B^2$, and summing as in Proposition 5.5, with (5.1), gives the claim. $\square$

In the leaf programs these contributions are the *hidden columns*: at most one per first-family character, carrying the value t_χ , the objective $e^{-At_\chi} B_{\phi_1}(p)$ and, because the zero realizing t_χ has $|\gamma| \leq 1$, the far weight $w(t_\chi)$ (Section 11.5).

5.8. The objective and the error allowance. In a leaf of the case tree (Section 2.4) the unknown characters are described as follows (Section 11.5): the first family, possibly a *reserved* second family of $n_2 \in \{1, 2\}$ characters with height-one representatives in a given interval $[lo_2, hi_2]$, and the *ordinary* characters, described by their T^* -representatives. By Corollary 5.7 and Propositions 5.8 and 5.9,

$$(5.2) \quad W \leq \mathcal{J}_1 + \sum_{\chi \text{ reserved}} \mathcal{G}_{\phi_2}(\nu_\chi) + \sum_{\chi \text{ ordinary}} \mathcal{G}(\lambda_\chi) + E,$$

where $\phi_2 = \frac{1}{4}$ if the reserved family is a single real character and $\phi_2 = \frac{1}{3}$ otherwise, ν_χ is the parameter of the height-one representative of a reserved character. The first-family term $\mathcal{J}_1$ is J_1 for an inside leaf and $\sum_\chi e^{-At_\chi} B_{\phi_1}(p)$ for an outside leaf. In the latter case it is represented by hidden columns, rather than by the constant part of the leaf objective (Section 11.5).

There are two sources of error in (5.2). The envelope errors total less than $19\varepsilon_1$ for the ordinary characters, by (6.2), and at most $4\varepsilon_1$ for the first and reserved families. The outside first-family bound has the additional error $4c_{\text{out}}/(H_0C_B^2)$. We choose

$$\varepsilon_1 = \frac{\eta}{23}, \quad \frac{4c_{\text{out}}}{H_0C_B^2} \leq \eta,$$

so $E \leq 2\eta$. The error η in the prime-detection criterion is accounted for separately in Proposition 4.4.

Every leaf objective includes an allowance of 5η . Consequently its value bounds $W + 3\eta$, and therefore also $W + 2\eta$, as required by the positivity criterion. These choices are made uniformly over the finite collection of leaves in Section 13.1.

6. FAR DENSITY AND REPRESENTATIVES

The per-character costs must be summed over many characters. A weighted density estimate supplies a common budget for that sum. Its zeros have physical height at most 1, so we also specify how to choose representatives within this height range.

6.1. Xylouris's weighted density lemma. The following is Xylouris's improvement [145, Lemma 5.1, (5.19), p. 66] of Heath-Brown's Lemma 11.1 [51]; the latter is the case $J = 1$, $w_0 \equiv 1$. (Xylouris writes M for our J ; we reserve M for a separation constant.)

Input 6.1 (Far density). *Let $\varepsilon, c_1, c_2 > 0$, $J \in \mathbb{N}$, and $\alpha_1, \dots, \alpha_J \geq 0$ with $\sum_i \alpha_i = 1$. Put*

$$x = \frac{2}{3} + 3c_1 + c_2, \quad u_i = \frac{1}{3} + 2c_1 + \frac{ic_2}{J} \quad (0 \leq i \leq J),$$

and let $w_0 : [u_0, x] \rightarrow \mathbb{R}$ be continuous and continuously differentiable except at finitely many points, with $1 \ll w_0 \ll 1$ and $w'_0 \ll 1$. Choose for each nonprincipal character χ modulo q at most one zero ρ_χ of $L(s, \chi)$ with $|\text{Im } \rho_\chi| \leq 1$ and $\lambda_\chi \leq \lambda_0 = \frac{1}{3} \log \log \mathcal{L}$. Then for $q \geq q_0$, with q_0 depending on all the parameters,

$$\sum_{\chi} \left(\int_{u_0}^x w_0(t)^2 e^{2\lambda_\chi t} dt \right)^{-1} \leq \frac{J^2 + \varepsilon}{c_1 c_2^2} \sum_{i=1}^J \alpha_i^2 \int_{u_{i-1}}^x w_0(t)^{-2} \min\{t - u_{i-1}, u_i - u_{i-1}\} dt.$$

In [145] the statement is made for the zeros $\rho^{(k)}$ that are chosen for the characters counted by $N(\lambda_0)$, one zero per character [145, §3.2.2, pp. 25–26], [51, §11]; any such choice is allowed, and the estimates in the proof [145, pp. 67–72] are uniform over these choices. This uniformity is needed because the selected zeros, including the $T^*(q)$ -representatives below, depend on q . A nonreal character and its conjugate are different characters, and each contributes its own term; the principal character is excluded, since $L(s, \chi_0)$ has no zero in the relevant region for large q [145, p. 21]. The hypotheses on w_0 are satisfied by our profiles, which are those of Xylouris's own application [145, (6.20)–(6.21), p. 80].

6.2. The two weight profiles. We use two profiles of the shape [145, (6.20)], both with $J = 10$ and $\varepsilon_0 = 10^{-7}$:

$$w_0(t)^2 = e^{-\theta_{\text{far}} t} \sqrt{\min(t - u_0 + \varepsilon_0, c_2 + \varepsilon_0)} \quad (u_0 \leq t \leq x).$$

Profile I has $c_1 = 0.09035$, $c_2 = 0.235968$ and $\theta_{\text{far}} = 1.28683$, and *profile II* has $c_1 = 0.0821922$, $c_2 = 0.2170903$ and $\theta_{\text{far}} = 1.4964274$. The weights α_i are the exact decimals of Table 4; in each profile they sum to 1. Both profiles satisfy the hypotheses of Input 6.1: w_0 is continuous and positive on $[u_0, x]$, and it is smooth except at $t = u_0 + c_2$.

i	profile I α_i	profile II α_i
1	0.0788827218	0.0806195583
2	0.0849386148	0.0862705300
3	0.0895629779	0.0905489307
4	0.0938231516	0.0944644867
5	0.0979710491	0.0982543287
6	0.1021284916	0.1020315591
7	0.1063732939	0.1058668825
8	0.1107651869	0.1098131140
9	0.1153565492	0.1139152197
10	0.1201979632	0.1182153903

TABLE 4. The weights α_i of the two far profiles.

Put

$$w(\lambda) = \left(\int_{u_0}^x w_0(t)^2 e^{2\lambda t} dt \right)^{-1},$$

a positive decreasing function of λ . Let V be the right-hand side of [Input 6.1](#) with $\varepsilon = 0$. Rigorous enclosures give

$$V = 175.26640331 \dots \text{ (profile I),} \quad V = 243.32098105 \dots \text{ (profile II).}$$

Taking $\varepsilon = \eta J^2$ in [Input 6.1](#), we obtain for $q \geq q_0$ and every choice of zeros as in [Input 6.1](#):

$$(6.1) \quad \sum_{\chi} w(\lambda_{\chi}) \leq (1 + \eta)V.$$

Each leaf of the case tree uses one of the two profiles, recorded in its data. Both are fixed finite choices within [Input 6.1](#).

Lemma 6.2 (Comparison of the prime weight and the far weight). *For both profiles, $x \leq 1.1737$. Hence $A > 2x$ for $L = 3.99$, and for every $\lambda \geq 0$ the function $\lambda \mapsto e^{-A\lambda}/w(\lambda)$ is decreasing. In particular $e^{-A\lambda} \leq w(\lambda)/w(0)$.*

Proof. The values are $x = 1.1736846 \dots$ and $x = 1.1303335 \dots$, so $2x < 2.348 < A = 3.156342$. Now $e^{-A\lambda}/w(\lambda) = \int_{u_0}^x w_0(t)^2 e^{(2t-A)\lambda} dt$, and $2t - A < 0$ on $[u_0, x]$. $\square$

With $w(0) = 10.7054 \dots$ (profile I) and $w(0) = 13.1730 \dots$ (profile II), [Lemma 6.2](#) and (6.1) give, for any selection satisfying [Input 6.1](#),

$$(6.2) \quad \sum_{\chi} e^{-A\lambda_{\chi}} \leq \frac{(1 + \eta)V}{w(0)} = K_{\text{far}}, \quad K_{\text{far}} < 16.38 \text{ (profile I),} \quad K_{\text{far}} < 18.48 \text{ (profile II).}$$

This bound controls errors that are summed over unboundedly many characters.

6.3. Representatives. The leaf programs describe an ordinary character through one parameter, the parameter of a *representative* zero. For the near-density argument of [Section 8](#) the representative must be separated in height from any zero with a smaller parameter, and we arrange this by choosing the height of the strip in which representatives are taken.

Fix $s_{\text{max}} = 3$; every anchor used in the near rows is at most 2.3 ([Section 9](#)), and every ordinary entry uses an anchor at most s_{max} . Let K_0 be an upper bound for the number of zeros, of all nonprincipal characters modulo q and counted with multiplicity, with $\lambda \leq s_{\text{max}}$ and $|\gamma| \leq 2$. By a log-free zero-density estimate ([Input 3.9](#)), K_0 may be taken independent of q . Next fix M as in [Section 13](#); it depends on K_0 .

Lemma 6.3 (A strip boundary separated from low-parameter zeros). *Let K_0 bound the number of zeros with $\lambda \leq s_{\max} = 3$ and $|\gamma| \leq 2$, as above, and fix $M > 0$. If $\mathcal{L} > 4M(K_0 + 1)$, there is $T^* \in [1/2, 1]$ such that no nonprincipal $L(s, \chi)$ has a zero with $\lambda \leq s_{\max}$ and $||\mu| - T^*\mathcal{L}| < M$. We take $T^* = T^*(q)$ to be the least such number.*

Proof. The at most K_0 zeros in question exclude at most K_0 open intervals of T of length $2M/\mathcal{L}$ each. The interval $[1/2, 1]$ has length $1/2 > 2MK_0/\mathcal{L}$, so the remaining set is a nonempty finite union of closed intervals, and it has a least element. $\square$

Definition 6.4 (Representatives). For a nonprincipal character χ , its T^* -representative is a zero of $L(s, \chi)$ in $R(l)$ with $|\gamma| \leq T^*$ and least parameter, if there is one; its parameter is λ_χ . Representatives are used for the characters outside the first family; in [Theorem 12.6](#) all characters use their height-one representatives. The *height-one representative* is defined in the same way with the bound $|\gamma| \leq 1$. The reserved family of a leaf ([Section 2.4](#) and [Definition 11.2](#)) and an inside first family use height-one representatives.

Every representative has $|\gamma| \leq 1$, so the far bound (6.1) applies to any selection of representatives, one per character. Representatives for conjugate characters may be chosen as conjugate zeros, with the same parameter, because every selection region is symmetric in height. The T^* -representative of χ satisfies $\lambda_\chi \geq$ the parameter of its height-one representative. Shrinking the strip can only increase the minimum parameter; in terms of real parts, the selected zero can only move to the left.

6.4. The far constraint of a leaf. In a leaf program ([Section 2.4](#)), the characters other than the first family are placed in bins by the parameters of their representatives, and each bin has a column ([Sections 10](#) and [11](#)). The column of a bin $[\text{lo}, \text{hi}]$ has far weight $W_c = \lfloor Sw(\text{hi}) \rfloor$; since w is decreasing, each character in the bin contributes at least $w(\text{hi})$ to the left side of (6.1). The *tail column* collects the ordinary characters χ with a zero in R_P and $\lambda_\chi \geq R$, where $R = \max(3, r)$ and r is the leaf's lower bound for ordinary representatives. Its value is the mass $x_{\text{tail}} = \sum w(\lambda_\chi)$ over these characters, with far weight S . The far budget of the leaf is

$$F = \lceil S(1 + \eta)V \rceil - \mathbf{1}_{\text{inside}} n \lfloor Sw(b) \rfloor.$$

Here the last term is present for an inside leaf, that is when $\rho_1 \in R_B$. Then $|\gamma_1| \leq C_B/\mathcal{L} \leq 1$, so ρ_1 and $\overline{\rho_1}$ are the height-one representatives of χ_1 and $\overline{\chi_1}$, and they contribute $n w(\lambda_1) \geq n w(b)$ to (6.1). A reserved second family is charged either through its own columns, which carry its far weight, or by subtracting the fixed amount $n_2 \lfloor Sw(\text{hi}_2) \rfloor$ from F ([Section 11.5](#)).

7. ZERO LOCATION

This section collects the lower bounds for λ_1 , λ'_1 , λ_2 and λ_3 that the case tree uses. Most are printed results of Heath-Brown [51] and Xylouris [145]; we state each in the form used, with its scope. The others are new consequences of the same explicit formula, proved here by the positivity method; one of them uses a sharpened conductor coefficient, which we prove in [Section 7.6](#). None of the results of this section depends on the exponent L .

7.1. Conventions. We fix once and for all an admissible height function $l(q)$ as in [Input 3.5](#), for instance the least one. All the zero-location results we quote are proved for an arbitrary integer l with $1 \leq l \leq \mathcal{L}/10$ for which $R(10l) \setminus R(l)$ contains no zero; their proofs use l only through these two properties, and their thresholds depend only on fixed data. So they hold simultaneously for our $l(q)$.

The zeros ρ_1, ρ_2, ρ_3 are selected as in [145, (3.11)–(3.12), p. 21] (which follows [51, §6]). At step k , remove the characters in the families $\mathcal{F}_1, \dots, \mathcal{F}_{k-1}$ and choose a zero of maximal real

part in $R(l)$ among the remaining nonprincipal L -functions. Each real character is removed once. This is the successive-minimum convention of [Definition 2.1](#). The additional zero ρ' of the first family is selected as in [\[145, pp. 21–22\]](#): $\rho' = \rho_1$ if ρ_1 is multiple, and otherwise ρ' has maximal real part among the zeros of $L(s, \chi_1)$ in $R(l)$ other than ρ_1 (and other than $\bar{\rho}_1$ if χ_1 is real). Its parameter is the λ' of [Section 2](#) (Heath-Brown also writes λ' ; Xylouris indexes this zero by 0). When several zeros qualify (ties in the real part, or several characters attaining a minimum), any of them may be selected: the proofs of the quoted results use only the extremal property of the selection, so they hold for every such choice. This is why the configuration sets of [Definition 11.2](#) quantify over the admissible choices. All the results of this section concern zeros in $R(l)$ at any height up to l , so they apply whether ρ_1 is inside or outside the buffer.

Printed tables. Every printed bound is used as an exact implication for $q \geq q_0$. Each of them was obtained by showing that a certain inequality fails with a positive margin, which persists when the error term ε is small enough [\[145, p. 43, p. 44 and footnote 2 on p. 45\]](#); for Heath-Brown's Tables 2 and 3 the printed values lie a little below the computed roots [\[51, §8\]](#). Heath-Brown states no convention for his Tables 4 and 7, and we have recomputed their roots from his printed parameters; every recomputed root exceeds the printed value, and each printed value is the root rounded down. The closest case in Table 4 is row 1.05, with root 1.43904 against 1.439, and the closest in Table 7 are rows 0.20 and 0.55, with 2.01015 against 2.01 and 1.00036 against 1.00. The rows of Table 4 are derived under the side conditions [\[51, \(8.6\), \(8.8\)\]](#); we have checked from the printed parameters that these hold for every row from 0.35 on (the tightest is (8.6), with slacks $3.8 \cdot 10^{-6}$, $3.8 \cdot 10^{-6}$ and $2.3 \cdot 10^{-6}$ for the rows 1.05, 1.15 and 1.294). For the row (0.3, 2.293) they need not hold, but that row also follows from Heath-Brown's Table 3, which gives $\lambda' \geq 2.43$ for $\lambda_1 \leq 0.30$ [\[51, §8\]](#). These checks, and the recomputation of the roots, were made in floating point, separately from the replay. A separate program compares every printed number that the proof uses, in the code and in this paper, with the text of the source pages and with the printed scope of its row; each agrees with the printed value or is weaker than it in the safe direction. Results stated in the form $c - \varepsilon$ are used with an explicit reduction: $0.857 < \frac{6}{7}$ in [Input 7.4](#), and $1.09 < \frac{12}{11}$ and $\varepsilon = 0.001$ in [Section 12](#).

7.2. Types and the first zero.

Proposition 7.1 (Lower bounds for a nonreal first zero). *For all sufficiently large q , the first family satisfies*

$$\lambda_1 > 0.440 \quad \text{if } \chi_1 \text{ is nonreal}, \quad \lambda_1 > 0.628 \quad \text{if } \chi_1 \text{ is real and } \rho_1 \text{ is nonreal}.$$

Thus $\lambda_1 \leq 0.44$ forces both χ_1 and ρ_1 to be real. These bounds are [\[145, Lemma 4.5 and Table 11, pp. 60–61\]](#).

Xylouris's lemma gives $\lambda_1 > 0.440, 0.493, 0.478, 0.498$ and 0.628 for characters of order $\geq 6, = 5, = 4, = 3$ and $= 2$ respectively; the minimum over orders at least 3 is 0.440 .

7.3. The test function. The location inequalities use one family of test functions. We define it here so that all parameters in the statements below are specified before they occur. For $\gamma > 0$, put

$$(7.1) \quad f_\gamma(t) = \begin{cases} \frac{16\gamma^5}{15}(1-u)^3(1+3u+u^2), & 0 \leq t \leq 2\gamma, \quad u = t/(2\gamma), \\ 0, & t \geq 2\gamma, \end{cases}$$

The function f_γ is the autocorrelation of $x \mapsto (\gamma^2 - x^2)_+$, the test function of [\[145, \(3.14\), \(3.17\)–\(3.18\)\]](#); it satisfies Conditions 1 and 2 ([Lemma 9.1](#)), and its transform F is positive and decreasing on $\mathbb{R}$.

7.4. Printed zero-location inputs.

Input 7.2 (Real first zero). *Let χ_1 and ρ_1 be real. For $q \geq q_0$:*

- (a) [51, Tables 3 and 4, Lemmas 8.2 and 8.3]. *If $\lambda_1 \leq t$ then $\lambda' \geq c$, for the 24 pairs (t, c) from (0.3, 2.293) to (1.294, 1.294) of the table. For a real ρ' this follows from $\lambda' \geq 2.427$ (Lemma 8.2).*
- (b) [51, Lemma 8.4]. *For every $\varepsilon > 0$: $\lambda' \geq (2 - \varepsilon) \log \lambda_1^{-1}$ if $\lambda_1 \leq 0.2$ and $q \geq q(\varepsilon)$; and $\lambda' \geq \max(\frac{3}{2} \log \lambda_1^{-1}, 1.294)$ for every λ_1 .*
- (c) [51, Table 7 and Lemma 8.7]. *If $\lambda_1 \leq t$ then $\lambda_2 \geq c$, for the 16 pairs (t, c) from (0.12, 2.56) to (0.745, 0.745) of the table; these hold whether or not $\chi_2^4 = \chi_0$ (the row $t = 0.10$ is not used).*
- (d) [51, Lemma 8.8]. $\lambda_2 \geq \max((\frac{12}{11} - \varepsilon) \log \lambda_1^{-1}, 0.745)$ for $q \geq q(\varepsilon)$.

Heath-Brown remarks that the estimates of his §8 have not been proved for extremely small values of λ_1 [51, §8]. We use (a)–(d) only for λ_1 bounded below by a fixed positive constant: $\lambda_1 \geq 0.1$ in the case tree, and, in Theorem 12.5, λ_1 at least the fixed constant $u_{\text{exc}} > 0$ of Section 12.

Input 7.3 (Nonreal first character or zero). *Let χ_1 or ρ_1 be nonreal. For $q \geq q_0$:*

- (a) [145, Table 2', p. 62]. *If $\lambda_1 \leq t$ then $\lambda' > c$, for the pairs (0.46, 1.85), ..., (0.827, 0.827) of the table.*
- (b) [145, Table 3, p. 45]. *If $\text{ord } \chi_1 \in \{2, 3, 4\}$ and $\lambda_1 \leq t$, then $\lambda' > c$, for the pairs (0.38, 2.53), ..., (1.099, 1.099); in particular $\lambda_1 \leq 0.86$ implies $\lambda' > 1.35$.*
- (c) [145, Tables 6 and 7, p. 53]. *If $\lambda_1 \leq t$ then $\lambda_2 > c$; Table 6 concerns a real χ_1 (Fall 7), Table 7 all cases.*
- (d) [51, Lemma 9.4 and Table 10]. $\lambda_2 \geq 0.702$ always, and $\lambda_1 \leq 0.70$ implies $\lambda_2 \geq 0.704$.

Input 7.4 (The third family). *For $q \geq q_0$:*

- (a) [51, Lemma 10.3]. $\lambda_3 \geq \frac{6}{7} - \varepsilon$; we use $\lambda_3 > 0.857$.
- (b) [145, Table 8, column “alle Fälle”, p. 55]. *If χ_1 or ρ_1 is nonreal and $\lambda_1 \leq t$, then $\lambda_3 > c$, for $(t, c) = (0.52, 1.320), (0.54, 1.243), (0.56, 1.160), (0.58, 1.079), (0.60, 1.001), (0.62, 0.933)$.*
- (c) [145, Lemma 4.4 and Table 10, p. 55]. *If χ_1 and ρ_1 are real, then $\lambda_1 \in [0.44, 0.60], [0.60, 0.70], [0.70, 0.80]$ imply $\lambda_3 > 1.176, 1.055, 0.952$.*
- (d) [145, (4.28)–(4.29), pp. 56–57]. *Let χ_1 be nonreal and $\lambda_1 \in [0.44, 0.85]$, and let f be the test function (7.1) with $\gamma = \frac{5}{4}$. Then*

$$0 \leq F(-\lambda_1) - F(\lambda_3 - \lambda_1) - F(\lambda_2 - \lambda_1) - F(0) + \frac{7}{6}f(0) + \varepsilon,$$

provided that $\sup_{y \in \mathbb{R}} \text{Re}\{F(-\lambda_1 + iy) - 2F(iy)\} \leq \frac{1}{6}f(0)$.

- (e) [145, (4.31) and (4.34), pp. 58–59]. *Let χ_1 and ρ_1 be real, $\lambda_1 \in [0.44, 0.80]$ and $\lambda_2 \in [0.44, 1.176]$, and let f be (7.1) with $\gamma = 1.04$. Then*

$$0 \leq F(-\lambda_2) - F(\lambda_3 - \lambda_2) - F(0) - F(\lambda_1 - \lambda_2) + \frac{9}{8}f(0) + \varepsilon,$$

provided that $\sup_{t \in \mathbb{R}} \text{Re}\{F(-\lambda_2 + it) - F(\lambda_1 - \lambda_2 + it) - F(it)\} \leq \frac{5}{48}f(0)$ on this box.

In (d), the side condition is [145, (4.29)], needed when a product character in his argument is principal; he verified it numerically, and we have verified it with interval arithmetic for all $\lambda_1 \in [0.44, 0.85]$: the supremum is at most 0.00335, while $\frac{1}{6}f(0) = 0.5425 \dots$. In (e), Xylouris's argument [145, (4.32)–(4.34), p. 59] needs

$$\sup_{t \in \mathbb{R}} \text{Re}\{F(-\lambda_2 + it) - F(\lambda_1 - \lambda_2 + it) - F(it)\} \leq \frac{5}{48}f(0)$$

on the whole box, and he states that the supremum is below 0.10, while $\frac{5}{48}f(0) = 0.1351\dots$. We have verified this with interval arithmetic as well: over $\lambda_1 \in [0.44, 0.80]$, $\lambda_2 \in [0.44, 1.18]$ and all real t the supremum is at most 0.0031. (Both checks enclose the transforms on a grid of heights with step $\frac{1}{100}$ up to 30, with second-derivative interpolation errors in all variables, and use a closed-form bound beyond height 30.) (Xylouris derives (4.28) from Heath-Brown's (10.2)–(10.6), whose derivation does not use the assumption $\lambda_3 \leq \frac{6}{7}$ made in [51, §10]; he himself applies it to prove bounds above $\frac{6}{7}$.)

Proposition 7.5 (Parent rows). *The 58 parent rows $(\mathbf{t}, A_j, B_j, p_j, r_j)$ of Table 5, from which the roots of the case tree are built (Section 11.3), are valid implications: for $q \geq q_0$, every configuration of type $\mathbf{t}$ with $\lambda_1 \in [A_j, B_j]$ has $\lambda' \geq p_j$ and $\lambda_2 \geq r_j$.*

Proof. Each entry is at most the largest of the following bounds, each valid on the whole interval $[A_j, B_j]$ for the type $\mathbf{t}$: (i) the applicable rows (t, c) of Inputs 7.2 and 7.3 with $t \geq B_j$; (ii) the bound c of an applicable row (t, c) with $c = t$ (for $\lambda_1 \leq t$ the row gives it, and for $\lambda_1 > t$ the trivial bound gives $\lambda', \lambda_2 \geq \lambda_1 > t = c$), as for the rows (0.827, 0.827) of Table 2' and (1.099, 1.099) of Table 3 of [145] and (0.745, 0.745) of Table 7 of [51]; (iii) the bounds of Input 7.2(b),(d) and Input 7.3(d), which hold for every λ_1 ; and (iv) the trivial bounds $\lambda', \lambda_2 \geq \lambda_1 \geq A_j$. $\square$

7.5. The positivity method. The use of nonnegative trigonometric polynomials to locate zeros goes back to de la Vallée Poussin [28]; it underlies all zero-free regions for Dirichlet L -functions [43, 74, 97, 82, 68] and for ζ [67, 90], and, combined with smoothed explicit formulas, the tables of Heath-Brown and Xylouris and their refinements [79]. All the new zero-location results come from the following consequence of [145, Lemmas 3.1 and 3.4]. Lemma 3.4 there is a “working version” of Input 3.4: if f satisfies Conditions 1 and 2, $\chi \neq \chi_0$ and $s \in R(9l)$, and A_1 is the set of zeros $\rho \in R(l)$ of $L(s, \chi)$ with $\operatorname{Re} \rho > \operatorname{Re} s$ (with multiplicity, at most N of them) and A_2 is any set of at most N zeros $\rho \in R(l)$ with $\operatorname{Re} \rho \leq \operatorname{Re} s$, then for every $\varepsilon > 0$ and $q \geq q_0(f, \varepsilon, N)$

$$(7.2) \quad \sum_n \Lambda(n) \operatorname{Re} \frac{\chi(n)}{n^s} f\left(\frac{\log n}{\mathcal{L}}\right) \leq -\mathcal{L} \sum_{\rho \in A_1 \cup A_2} \operatorname{Re} F((s - \rho)\mathcal{L}) + \frac{\varphi(\chi)}{2} f(0)\mathcal{L} + \varepsilon \mathcal{L}.$$

Lemma 7.6 (Positivity method). *Let f satisfy Conditions 1 and 2, with Laplace transform F . Fix a nonnegative anchor a , and suppose $a \leq \lambda_1$. Let finitely many triples $(c_j, \chi^{(j)}, t_j)$ be given, with $c_j > 0$, $\chi^{(j)}$ a character modulo q and $|t_j| \leq 9l$, such that*

$$\sum_j c_j \operatorname{Re}(\chi^{(j)}(n)n^{-it_j}) \geq 0 \quad \text{for all } n \text{ coprime to } q.$$

For each nonprincipal $\chi^{(j)}$ let $A_2(j)$ be a set of at most N zeros of $L(s, \chi^{(j)})$ in $R(l)$, with multiplicity. Then for every $\varepsilon > 0$ and $q \geq q_0$,

$$\sum_{\chi^{(j)} = \chi_0} c_j \operatorname{Re} F(-a + it_j \mathcal{L}) - \sum_{\chi^{(j)} \neq \chi_0} c_j \sum_{\rho \in A_2(j)} \operatorname{Re} F((1 + it_j - \rho)\mathcal{L} - a) + \frac{f(0)}{2} \sum_{\chi^{(j)} \neq \chi_0} c_j \varphi(\chi^{(j)}) + \varepsilon \geq 0.$$

Proof. Multiply the hypothesis by $\Lambda(n)n^{-(1-a/\mathcal{L})}f(\log n/\mathcal{L}) \geq 0$ and sum over n (for $(n, q) > 1$ every term vanishes, the principal ones included). For principal $\chi^{(j)}$ use Input 3.3 at $1 - a/\mathcal{L} + it_j$. For the others use (7.2) at $s_j = 1 - a/\mathcal{L} + it_j$, which lies in $R(9l)$; there $A_1 = \emptyset$, because every zero in $R(l)$ of a nonprincipal L -function has parameter at least $\lambda_1 \geq a$. Divide by $\mathcal{L}$. $\square$

In every application the retained zeros satisfy $\lambda_\rho \geq a$, the transform F is decreasing on $\mathbb{R}$, and the terms that we cannot evaluate are bounded over all real heights with interval arithmetic. A necessary inequality between the parameters of the retained zeros results, and a case is excluded

Type	$[A_j, B_j]$	p_j	r_j	Type	$[A_j, B_j]$	p_j	r_j
rr	[0.1, 0.12]	2.293	2.56	rc	[0.628, 0.66]	1.67	0.93
rr	[0.12, 0.14]	2.293	2.39	rc	[0.66, 0.7]	1.59	0.93
rr	[0.14, 0.16]	2.293	2.25	rc	[0.7, 0.74]	1.52	0.93
rr	[0.16, 0.18]	2.293	2.12	rc	[0.74, 0.78]	1.46	0.93
rr	[0.18, 0.2]	2.293	2.01	rc	[0.78, 0.82]	1.40	0.82
rr	[0.2, 0.25]	2.293	1.77	rc	[0.82, 0.86]	1.35	0.82
rr	[0.25, 0.3]	2.293	1.58	rc	[0.86, 0.9]	1.30	0.82
rr	[0.3, 0.348]	2.195	1.42	rc	[0.9, 0.94]	1.25	0.82
rr	[0.348, 0.35]	2.195	1.42	rc	[0.94, 0.98]	1.21	0.82
rr	[0.35, 0.4]	2.108	1.29	rc	[0.98, 1.02]	1.17	0.82
rr	[0.4, 0.45]	2.03	1.18	rc	[1.02, 1.06]	1.13	0.82
rr	[0.45, 0.5]	1.958	1.08	rc	[1.06, 1.099]	1.099	0.82
rr	[0.5, 0.55]	1.893	1.0	rc	[1.099, 1.5]	1.099	0.82
rr	[0.55, 0.6]	1.832	0.92	complex	[0.44, 0.58]	1.36	1.04
rr	[0.6, 0.65]	1.776	0.85	complex	[0.58, 0.64]	1.15	0.85
rr	[0.65, 0.7]	1.724	0.79	complex	[0.64, 0.66]	1.08	0.79
rr	[0.7, 0.745]	1.676	0.745	complex	[0.66, 0.68]	1.02	0.74
rr	[0.745, 0.75]	1.676	0.745	complex	[0.68, 0.70]	0.96	0.704
rr	[0.75, 0.8]	1.63	0.745	complex	[0.70, 0.72]	0.93	0.702
rr	[0.8, 0.85]	1.587	0.8	complex	[0.72, 0.74]	0.91	0.72
rr	[0.85, 0.9]	1.547	0.8	complex	[0.74, 0.76]	0.89	0.74
rr	[0.9, 0.95]	1.509	0.8	complex	[0.76, 0.78]	0.86	0.74
rr	[0.95, 1]	1.473	0.8	complex	[0.78, 0.80]	0.84	0.78
rr	[1, 1.05]	1.439	0.8	complex	[0.80, 0.82]	0.83	0.78
rr	[1.05, 1.1]	1.406	0.8	complex	[0.82, 1.5]	0.827	0.82
rr	[1.1, 1.15]	1.375	0.8				
rr	[1.15, 1.175]	1.36	0.8				
rr	[1.175, 1.2]	1.346	0.8				
rr	[1.2, 1.225]	1.331	0.8				
rr	[1.225, 1.25]	1.318	0.8				
rr	[1.25, 1.275]	1.304	0.8				
rr	[1.275, 1.294]	1.294	0.8				
rr	[1.294, 1.5]	1.294	0.8				

TABLE 5. Parent implications: for the stated type and $\lambda_1 \in [A_j, B_j]$, $\lambda' \geq p_j$ and $\lambda_2 \geq r_j$. Each panel is read downwards, and each row is valid on the whole closed interval. The left panel gives the 33 real-first-zero rows; the right gives the 13 real-character/nonreal-zero rows and 12 nonreal-character rows.

when that inequality fails with a positive margin. Since there are finitely many such inequalities, each with a fixed positive margin, one q_0 serves all of them.

7.6. A sharper conductor coefficient. Heath-Brown remarks [51, §16, item 4] that his estimate for λ_2 in the case of a real first character can be improved by replacing the conductor coefficient $\frac{11}{24}$ by $\frac{97}{216}$; he describes the ingredients but gives no proof. We need a weighted form of this refinement, and we prove it here. For a modulus q put

$$v = \prod_{p^e \parallel q, e \geq 3} p^e, \quad \theta_q = \frac{\log v}{\log q} \in [0, 1].$$

We also write $\text{rad } q = \prod_{p|q} p$.

Lemma 7.7 (Growth bounds with conductor dependence). *Let $k \geq 3$ and $\varepsilon > 0$. For every nonprincipal character χ modulo q ,*

$$|L(\sigma + it, \chi)| \ll_{\varepsilon, k} q^{\phi_q(\chi)(1-\sigma)(1+k^{-1})+\varepsilon} (1 + |t|) \quad \left(1 - \frac{1}{k} \leq \sigma \leq 1 + \frac{\log \mathcal{L}}{\mathcal{L}}\right),$$

where $\phi_q(\chi) = \frac{1}{4} \frac{\log(4 \operatorname{rad} q)}{\log q}$ if χ is real and $\phi_q(\chi) = \min(\frac{1}{3}, \frac{1}{4} + \frac{3}{4} \theta_q)$ otherwise.

Proof. Let χ be induced by the primitive character χ^* modulo q^* ; then $q^* \mid q$ and $|L(s, \chi)| \ll q^\varepsilon |L(s, \chi^*)|$ for $\sigma \geq \frac{1}{2}$ [51, p. 10]. Heath-Brown proves ([51, pp. 9–10], (2.2)–(2.4)) that a bound $\sum_{N < n \leq N+H} \chi^*(n) \ll_{\varepsilon, k} Q^{(k+1)/(4k^2)} q^{*\varepsilon} H^{1-1/k}$ for $1 \leq H \leq q^*$ gives, by partial summation with the cut $q_1 = Q^{(k+1)/(4k)}$ and the Pólya–Vinogradov inequality beyond q^* , $L(s, \chi^*) \ll Q^{(1-\sigma)(1+k^{-1})/4} q^{*\varepsilon} (1 + |t|)$ in the stated range (if $q_1 > q^*$ the trivial bound up to q^* is already stronger). For $\sigma \leq 1$ a bound of this form with a smaller Q is stronger; for $1 < \sigma \leq 1 + \log \mathcal{L}/\mathcal{L}$ every factor $Q^{(1-\sigma)c}$ with $1 \leq Q \leq q^4$ and $0 \leq c \leq 1$ lies between $\mathcal{L}^{-4}$ and 1, so all such bounds agree up to a factor q^ε .

If χ is real, then χ^* is a real primitive character, so q^* is the absolute value of a fundamental discriminant and $q^* \leq 4 \operatorname{rad}(q^*) \leq 4 \operatorname{rad}(q)$. Its order is 2, so by [51, Lemma 2.4], which bounds the sums with the factor $l^{3/(2k)}$ for a character of order l , the character-sum bound holds with $Q = q^*$ for every k . This gives the real case.

In general, [51, Lemma 2.3] gives the bound with $Q^{(k+1)/(4k^2)} = v^{*(3k-1)/(4k^2)} q^{*(k+1)/(4k^2)} \leq (v^* q^*)^{(k+1)/(4k^2)}$, where v^* is the cube-full part of q^* ; since $q^* \mid q$ we have $v^* \mid v$ and $v^* q^* \leq v^3 q$. This gives the exponent $\frac{1}{4}(1 + \frac{3 \log v}{\log q}) = \frac{1}{4} + \frac{3}{4} \theta_q$. The exponent $\frac{1}{3}$ is [51, Lemma 2.5]. Taking the smaller bound gives the general case. (Bounds that are uniform in q and t together are studied in [48]; here the factor $1 + |t|$ is harmless, since $|t| \leq \mathcal{L} + 1$.) $\square$

Lemma 7.8 (The refined coefficient in the explicit formula). *For every $\varepsilon > 0$, [Input 3.4](#) and (7.2) remain true, with a suitable q_0 , when $\varphi(\chi)$ is replaced by $\phi_q(\chi) + \varepsilon$, with the radius δ that the proof of [Input 3.4](#) provides for $\varphi = \frac{1}{3}$ (which serves all characters).*

Proof. Both are deduced from Heath-Brown's Lemma 3.1 [51, pp. 12–14, 23–24] [145, pp. 19–21], whose proof uses the bound (2.5) of his Lemma 2.5 only through the inequality

$$\log |L(s_0 + Re^{i\vartheta}, \chi)| \leq \{\phi(1 + k^{-1})(1 - \sigma_0 - R \cos \vartheta) + 2\varepsilon\} \mathcal{L}$$

for $\frac{\pi}{2} \leq \vartheta \leq \frac{3\pi}{2}$, $R \leq 1/k$ and large q . By [Lemma 7.7](#) this holds with $\phi = \phi_q(\chi)$, with an implied constant independent of q and χ . The remaining choices in that proof ($k = 3 + [3\phi/(2\varepsilon_0)]$, $R = 1/k$, and the radius) may be made with the upper bound $\phi \leq \frac{1}{3}$ in place of ϕ , so they do not depend on q . The deduction of Lemma 5.2 of [51] ([Input 3.4](#)) and of (7.2) from Lemma 3.1 is unchanged. $\square$

Proposition 7.9 (A coupled saving in the conductor term). *Let χ_1 be a real nonprincipal character and let χ_2 be a character with $\chi_2^k \neq \chi_0$ and $\chi_1 \chi_2^k \neq \chi_0$ for $k \in K$, a finite set of positive integers. Let $c_k > 0$ ($k \in K$) with $\Sigma = \sum_k c_k \geq \frac{1}{9}$. Suppose the polynomial $1 + \sum_{k \in K} c_k \cos(k\vartheta)$ is nonnegative for every real ϑ . In an application of [Lemma 7.6](#) to $(1 + \chi_1(n))(1 + \sum_{k \in K} c_k \operatorname{Re}((\chi_2(n)n^{-i\gamma_2})^k))$, the conductor term may, for every fixed $\varepsilon > 0$ and sufficiently large q , be taken to be*

$$\left(\frac{1}{8} + \frac{\Sigma}{3} - \frac{1}{108}\right) f(0) + \varepsilon.$$

Proof. The nonprincipal characters are χ_1 with coefficient 1, and χ_2^k and $\chi_1 \chi_2^k$, each with coefficient c_k . By [Lemma 7.8](#) the conductor term is at most $\frac{f(0)}{2} (\phi_q(\chi_1) + 2\Sigma \max_\chi \phi_q(\chi)) + \varepsilon$, the

maximum being over the characters χ_2^k and $\chi_1\chi_2^k$. Since $p \leq (p^e)^{1/3}$ when $e \geq 3$, $\text{rad } q \leq qv^{-2/3}$, so $\phi_q(\chi_1) \leq \frac{1}{4} - \frac{\theta_q}{6} + o(1)$, and $\phi_q(\chi) \leq \min(\frac{1}{3}, \frac{1}{4} + \frac{3}{4}\theta_q) + o(1)$ for every nonprincipal χ (for a real χ because $\frac{1}{4} - \frac{\theta_q}{6} \leq \frac{1}{4}$). The conductor term is therefore at most $\mathcal{K}(\theta_q)f(0) + 2\varepsilon$ for large q , with

$$\mathcal{K}(x) = \frac{1}{8} - \frac{x}{12} + \Sigma \min\left(\frac{1}{3}, \frac{1}{4} + \frac{3x}{4}\right) \quad (0 \leq x \leq 1).$$

On $[0, \frac{1}{9}]$ the slope of $\mathcal{K}$ is $\frac{3}{4}\Sigma - \frac{1}{12} \geq 0$, and on $[\frac{1}{9}, 1]$ it is $-\frac{1}{12}$. Hence $\mathcal{K}(\theta_q) \leq \mathcal{K}(\frac{1}{9}) = \frac{1}{8} + \frac{\Sigma}{3} - \frac{1}{108}$. $\square$

For $\Sigma = 1$ this is Heath-Brown's $\frac{97}{216} = \frac{11}{24} - \frac{1}{108}$.

7.7. Lower bounds for λ_2 when the first zero is real.

Proposition 7.10 (Second-family bounds for a real first zero). *Let χ_1, ρ_1 be real with $\lambda_1 \in [a, b]$, where $0 \leq a \leq b$, and let $h \geq b$. Let $P(x) = 1 + c_1x + c_2(2x^2 - 1) + c_5(16x^5 - 20x^3 + 5x)$, that is $P(\cos \vartheta) = 1 + c_1 \cos \vartheta + c_2 \cos 2\vartheta + c_5 \cos 5\vartheta$, with*

$$(c_1, c_2, c_5) = \frac{(1.41866466, 0.43287503, 0.01421037)}{1.000001},$$

so that $P \geq 0$ on $[-1, 1]$ and $\Sigma = c_1 + c_2 + c_5 \geq \frac{1}{9}$. Choose $f = f_\gamma$ from (7.1), with transform F and fixed $\gamma > 0$. If $\lambda_2 \leq h$, then for every $\varepsilon > 0$ and all sufficiently large q , the inequality corresponding to the character relation below must hold. The generic case means that χ_2 is nonreal and none of the relations in (i)–(iii) holds.

- (i) (generic) $F(b-a) + c_1F(h-a) \leq F(-a) + (\frac{1}{8} + \frac{\Sigma}{3} - \frac{1}{108})f(0) + \varepsilon$;
- (ii) ($\chi_2^2 = \chi_1$) $F(b-a) + c_1F(h-a) \leq F(-a) + (\frac{1+c_2}{8} + \frac{c_1+c_5}{3})f(0) + \sup_{\Delta} R(\Delta) + \varepsilon$, where $R(\Delta) = c_2\{\text{Re } F(-a + i\Delta) - \text{Re } F(\lambda_1 - a + i\Delta)\} - c_1 \text{Re } F(\lambda_2 - a + i\Delta)$, the supremum over $\Delta \in \mathbb{R}$, $\lambda_1 \in [a, b]$ and $\lambda_2 \in [a, h]$;
- (iii) ($\chi_2^5 \in \{\chi_0, \chi_1\}$) $F(b-a) + c_1F(h-a) \leq (1 + c_5)F(-a) + (\frac{1+c_5}{8} + \frac{c_1+c_2}{4})f(0) + \varepsilon$;
- (iv) (χ_2 real) $F(b-a) + F(h-a) \leq F(-a) + \frac{3}{8}f(0) + \varepsilon$, with a second test function (7.1).

Here χ_2 is a character of the second family. The test parameter in (iv) may be chosen independently of the one in (i)–(iii). If all four possible cases are excluded with a positive margin at $\varepsilon = 0$, then $\lambda_2 > h$ for all sufficiently large q .

Proof. Put $z_n = \chi_2(n)n^{-i\gamma_2}$ and apply Lemma 7.6 with anchor a to $(1 + \chi_1(n))P(\text{Re } z_n) \geq 0$ (with $\text{Re } z_n^k = T_k(\text{Re } z_n)$ for $|z_n| = 1$), retaining ρ_1 in the term of χ_1 and ρ_2 in the term $c_1 \text{Re } z_n$. The heights are at most $5l \leq 9l$. The nonzero frequencies are 1, 2 and 5, and $\chi_2 \neq \chi_1$. A product character is principal only if χ_2 is real, or $\chi_2^2 = \chi_1$ (order 4), or $\chi_2^5 \in \{\chi_0, \chi_1\}$ (order 5 or 10); the last two cannot occur together, since together they would give $\chi_2^2 = \chi_0$. Generic case: there is no principal product, and Proposition 7.9 gives (i), using that F is decreasing, $\lambda_1 \leq b$ and $\lambda_2 \leq h$. Case $\chi_2^2 = \chi_1$: then $\chi_1\chi_2^2 = \chi_0$ at height $2\gamma_2$, $\chi_2^2 = \chi_1$ at height $2\gamma_2$ retains ρ_1 , and $\chi_1\chi_2 = \bar{\chi}_2$ retains $\bar{\rho}_2$; with $\Delta = 2\mu_2$ these give $R(\Delta)$, while the characters of frequencies 1 and 5 are nonprincipal and are charged $\frac{1}{3}$ (conservatively, since some of them have order 4). Case $\chi_2^5 \in \{\chi_0, \chi_1\}$: exactly one term is principal, bounded by $c_5F(-a)$ since $f \geq 0$, and every other character has order at most $10 \leq \mathcal{L}$, so $\varphi = \frac{1}{4}$. Real χ_2 : apply the method to $(1 + \chi_1(n))(1 + \text{Re } z_n) \geq 0$, whose three nonprincipal characters are real. $\square$

The rows used are given in Table 6: 22 cells of width 0.0025 covering $[0.700, 0.755]$. For each cell $[a, b]$ we give h and the test parameter γ of the generic case; each of the four inequalities fails by the margin shown (normalized by $f(0)$), so that $\lambda_2 > h$ throughout the cell. The margins were computed with interval arithmetic: the correlation term $R(\Delta)$ is enclosed on a grid in $\Delta \in [0, 30]$ with second-derivative interpolation errors in all variables, and by a closed-form

bound for $|\Delta| \geq 30$. The nonnegativity of P was verified at the points $x = j/2000$ with a second-derivative bound, and independently by exact Bernstein expansion. In addition, on cells inside $[0.700, 0.7025]$ we use the degree-2 row $P(\cos \vartheta) = (\frac{7}{8} + \cos \vartheta)^2 / \frac{81}{64}$ without the conductor refinement, which gives $\lambda_2 > 0.762$ (Proposition 7.11).

Proposition 7.11 (A second-family bound near 0.70). *If χ_1, ρ_1 are real and $0.700 \leq \lambda_1 \leq 0.7025$, then $\lambda_2 > 0.762$ for $q \geq q_0$.*

Proof. As in Proposition 7.10 with $c_1 = \frac{112}{81}$, $c_2 = \frac{32}{81}$, $c_5 = 0$, the printed conductor coefficients and $\gamma = 1.09$ (and $\gamma = 0.82$ for a real χ_2); the normalized margins of the generic, order-4 and real cases are at least $3.6 \cdot 10^{-4}$, 0.082 and 0.033. $\square$

cell $[a, b]$	h	γ	generic	order 4	order 5, 10	real
[.7000, .7025]	.7793	1.09624	$1.02 \cdot 10^{-4}$	.0804	.1280	.0289
[.7025, .7050]	.7783	1.09522	$5.86 \cdot 10^{-5}$	.0803	.1279	.0287
[.7050, .7075]	.7772	1.09419	$8.66 \cdot 10^{-5}$	.0804	.1280	.0285
[.7075, .7100]	.7762	1.09317	$4.92 \cdot 10^{-5}$	.0803	.1279	.0284
[.7100, .7125]	.7751	1.09216	$8.35 \cdot 10^{-5}$	.0804	.1280	.0282
[.7125, .7150]	.7741	1.09115	$5.21 \cdot 10^{-5}$	.0803	.1279	.0280
[.7150, .7175]	.7730	1.09014	$9.25 \cdot 10^{-5}$	.0804	.1279	.0278
[.7175, .7200]	.7720	1.08914	$6.71 \cdot 10^{-5}$	.0804	.1279	.0276
[.7200, .7225]	.7710	1.08814	$4.46 \cdot 10^{-5}$	.0803	.1279	.0275
[.7225, .7250]	.7699	1.08715	$9.41 \cdot 10^{-5}$	.0804	.1279	.0273
[.7250, .7275]	.7689	1.08615	$7.75 \cdot 10^{-5}$	.0804	.1279	.0271
[.7275, .7300]	.7679	1.08517	$6.39 \cdot 10^{-5}$	.0804	.1279	.0269
[.7300, .7325]	.7669	1.08418	$5.31 \cdot 10^{-5}$	.0804	.1279	.0267
[.7325, .7350]	.7659	1.08321	$4.53 \cdot 10^{-5}$	.0804	.1279	.0265
[.7350, .7375]	.7649	1.08223	$4.03 \cdot 10^{-5}$	.0804	.1278	.0263
[.7375, .7400]	.7639	1.08126	$3.83 \cdot 10^{-5}$	.0804	.1278	.0261
[.7400, .7425]	.7629	1.08029	$3.91 \cdot 10^{-5}$	.0804	.1278	.0260
[.7425, .7450]	.7619	1.07933	$4.28 \cdot 10^{-5}$	.0804	.1278	.0258
[.7450, .7475]	.7609	1.07837	$4.93 \cdot 10^{-5}$	.0804	.1278	.0256
[.7475, .7500]	.7599	1.07741	$5.87 \cdot 10^{-5}$	.0804	.1278	.0254
[.7500, .7525]	.7589	1.07645	$7.09 \cdot 10^{-5}$	.0804	.1278	.0252
[.7525, .7550]	.7579	1.07550	$8.60 \cdot 10^{-5}$	.0804	.1278	.0250

TABLE 6. The degree-5 rows of Proposition 7.10: on each cell, $\lambda_2 > h$. The values of γ are rounded to five decimals (the exact rational values are in the data). The margins are normalized by $f(0)$. The real case uses $\gamma = 0.82$.

7.8. Lower bounds for λ_2 and λ' when the first character is nonreal. For $t > 0$ put

$$P_t(e^{i\vartheta}) = \frac{(t + \cos \vartheta)^2}{t^2 + \frac{1}{2}} = 1 + c_1 \cos \vartheta + c_2 \cos 2\vartheta, \quad c_1 = \frac{2t}{t^2 + \frac{1}{2}}, \quad c_2 = \frac{1/2}{t^2 + \frac{1}{2}},$$

so that $c_1/c_2 = 4t$.

Proposition 7.12 (Second-family bounds for a nonreal first character). *Let χ_1 be nonreal, let $\lambda_1 \in [a, b]$ with $0 \leq a \leq b$, and let $h \geq b$. Fix $t > 0$, let c_1, c_2 be the coefficients of P_t above, and let $f = f_\gamma$ with $\gamma > 0$ and transform F . Suppose $\lambda_2 \leq h$. For every $\varepsilon > 0$ and sufficiently large q , the following necessary inequalities hold:*

(i) *if χ_2 is nonreal and*

$$(7.3) \quad \sup_{\lambda \in [a, h], y \in \mathbb{R}} \operatorname{Re}\{F(-a + iy) - 4tF(\lambda - a + iy)\} \leq \frac{1}{6}f(0),$$

then $c_1\{F(b - a) + F(h - a)\} \leq F(-a) + \frac{(1+c_1+c_2)^2-1}{6}f(0) + \varepsilon$;

(ii) if χ_2 is real, choose $t' > 0$ and let c'_1, c'_2 be the coefficients of $P_{t'}$. If (7.3) holds with t' in place of t and $[a, b]$ in place of $[a, h]$, then $c'_1 F(b - a) + F(h - a) \leq F(-a) + (\frac{1}{8} + \frac{c'_1 + c'_2}{3})f(0) + \varepsilon$.

Proof. (i) Apply Lemma 7.6 to $P_t(\chi_1(n)n^{-i\gamma_1})P_t(\chi_2(n)n^{-i\gamma_2}) \geq 0$, expanded as

$$\sum_{w \in \{-2, \dots, 2\}^2} a_{|w_1|} a_{|w_2|} \chi^w(n) n^{-iw \cdot \gamma}, \quad a_0 = 1, \quad a_1 = \frac{c_1}{2}, \quad a_2 = \frac{c_2}{2}, \quad \chi^w = \chi_1^{w_1} \chi_2^{w_2}.$$

The indices $(\pm 1, 0)$ retain ρ_1 (or $\overline{\rho_1}$) and $(0, \pm 1)$ retain ρ_2 (or $\overline{\rho_2}$). If no nonzero index is principal, the conductor term is at most $\frac{1}{6}((1 + c_1 + c_2)^2 - 1)f(0)$. A nonzero principal index w is not on an axis, since χ_1 and χ_2 are nonreal, and does not have $|w_1|, |w_2| \leq 1$, since $\chi_2 \notin \{\chi_1, \overline{\chi_1}\}$. So some $|w_i| = 2$; choose such i , with $i = 1$ in case of a tie, and put $T(w) = w - \text{sgn}(w_i)e_i$. Then $\chi^{T(w)} = \chi_i^{-\text{sgn } w_i}$ is nonprincipal and not an axis index of length one, its coefficient is $4t$ times that of w , and the zero of $\chi_i^{-\text{sgn } w_i}$ selected at the axis index sits at relative height $w \cdot \mu$. The map T is injective: a collision could only come from $(2s, \sigma)$ and $(s, 2\sigma)$ with $s, \sigma = \pm 1$, and if both were principal then $\chi_1^s = \chi_2^\sigma$, which is excluded. Retaining at $T(w)$ the zero of $\chi_i^{-\text{sgn } w_i}$, the principal term at w and the extra retained term together change the inequality by at most $a_w \{\text{Re } F(-a + iy) - \frac{1}{6}f(0) - 4t \text{Re } F(\lambda_i - a + iy)\} \leq 0$, by (7.3). (ii) Use $P_{t'}(\chi_1(n)n^{-i\gamma_1})(1 + \text{Re } \chi_2(n)n^{-i\gamma_2}) \geq 0$. Now χ_2 is real, with coefficient 1 and $\varphi = \frac{1}{4}$; the only possible principal indices are $(\pm 2, \pm 1)$, which are matched to $(\pm 1, \pm 1)$ with target $\overline{\rho_1}$ or ρ_1 , as before. $\square$

Proposition 7.13 (Additional-zero bounds for a nonreal first character). *Let χ_1 be nonreal of order at least 5, let $\lambda_1 \in [a, b]$ with $0 \leq a \leq b$, and suppose $\lambda' \leq h$. Fix $t, \gamma > 0$, let c_1, c_2 be the coefficients of P_t , and let F be the transform of $f = f_\gamma$. Choose an admissible ρ' of χ_1 with parameter λ' and height γ' . Put $x = \lambda_1 - a$, $y = \lambda' - a$ and $\Delta = \mathcal{L}(\gamma_1 - \gamma')$, and*

$$M_t = (1 + c_1 + c_2)^2 - 1 - \frac{1}{2}(c_1^2 + c_2^2), \quad p_1 = \frac{1}{2}c_1^2, \quad p_2 = \frac{1}{2}c_2^2, \quad q_1 = c_1(1 + \frac{1}{2}c_2), \quad q_2 = \frac{1}{2}c_1c_2.$$

Then for every $\varepsilon > 0$ and all sufficiently large q ,

$$c_1[F(x) + F(y)] \leq F(-a) + \frac{M_t}{6}f(0) + R(\Delta) + \varepsilon,$$

where

$$\begin{aligned} R(\Delta) = & p_1 \text{Re } F(-a + i\Delta) + p_2 \text{Re } F(-a + 2i\Delta) \\ & - q_1 \text{Re}\{F(x + i\Delta) + F(y + i\Delta)\} \\ & - q_2 \text{Re}\{F(x + 2i\Delta) + F(y + 2i\Delta)\}. \end{aligned}$$

Proof. Apply Lemma 7.6 to $P_t(\chi_1(n)n^{-i\gamma_1})P_t(\chi_1(n)n^{-i\gamma'}) \geq 0$, whose terms are χ_1^{k+l} at heights $k\gamma_1 + l\gamma'$, $|k|, |l| \leq 2$. Since the order of χ_1 is at least 5 and $|k + l| \leq 4$, the principal terms are exactly those with $k + l = 0$; they have total mass p_1 at height $\pm\Delta$ and p_2 at $\pm 2\Delta$. In every term with $k + l = \pm 1$ retain both ρ_1 and ρ' (or their conjugates); a multiple zero is retained with its multiplicity. The terms $(1, 0), (0, 1)$ and their conjugates give $-c_1[F(x) + F(y)] - c_1 \text{Re}[F(x + i\Delta) + F(y + i\Delta)]$, and the terms $(2, -1), (-1, 2)$ and their conjugates give $-\frac{1}{2}c_1c_2 \text{Re}[F(x + i\Delta) + F(y + i\Delta) + F(x + 2i\Delta) + F(y + 2i\Delta)]$. $\square$

For χ_1 of order 3 or 4, Input 7.3(b) gives $\lambda' > 1.35$ whenever $\lambda_1 \leq 0.86$, which is stronger than every bound in Proposition 7.13 that we use.

The rows used are the following, each on a cell of width 0.0025 and each with a positive normalized margin. There are 25 rows of Proposition 7.12, on cells covering $[0.66, 0.7225]$, with conclusions from $\lambda_2 > 0.763$ (on $[0.66, 0.6625]$) to $\lambda_2 > 0.725$ (on $[0.72, 0.7225]$), parameters $\gamma \in \{1.25, 1.3\}$, $t \in \{0.9, 0.95\}$ and, for a real χ_2 , test parameter $\gamma_r \in \{1.1, 1.15\}$ and $t' = 0.85$;

their generic margins are at least $8.2 \cdot 10^{-4}$, and the margins in (7.3) at least 0.163. There are 69 rows of Proposition 7.13, on 69 of the 70 cells of width 0.0025 in $[0.68, 0.855]$ (all except $[0.6975, 0.70]$), with conclusions from $\lambda' > 0.975$ (on $[0.68, 0.6825]$) to $\lambda' > 0.855$ (on $[0.8525, 0.855]$), parameters $\gamma \in \{1.05, 1.1, 1.15\}$ and $t \in \{0.9, 0.95\}$, and margins at least $1.5 \cdot 10^{-3}$. For these rows, if $\lambda' \leq h$, the left side of Proposition 7.13 is at least $c_1[F(b-a) + F(h-a)]$, since F is decreasing, and $R(\Delta)$ is replaced by its supremum over all real Δ , $x \in [0, b-a]$ and $y \in [p-a, h-a]$, where $p = \max(p_j, a) \leq \lambda'$ comes from the bound p_j of the parent row of the cell (Proposition 7.5; it comes from [145, Table 2']) and the trivial bound $\lambda' \geq \lambda_1 \geq a$: $p = 0.96, 0.93, 0.91, 0.89, 0.86, 0.84, 0.83$ on the cells in $[0.68, 0.70], [0.70, 0.72], \dots, [0.80, 0.82]$, $p = 0.827$ on $[0.82, 0.8275]$, and $p = a$ on the cells beyond. So each row proves that $\lambda_1 \in [a, b]$ and $\lambda' \geq p$ imply $\lambda' > h$, and since $\lambda' \geq p$ holds for every configuration with $\lambda_1 \in [a, b]$, $\lambda_1 \in [a, b]$ implies $\lambda' > h$. The complete list is in Appendix A. On 11 further cells the inequalities of Proposition 7.12 are evaluated on the actual cell of the case tree.

7.9. A positivity exclusion.

Proposition 7.14 (Excluding a nearby real second character). *Let χ_1, ρ_1 be real with $\lambda_1 \in [a, b]$, and let $\chi_2 \neq \chi_1$ be a real nonprincipal character with a zero of height at most 1 and parameter $\nu \leq h$. If*

$$F(b-a) + F(h-a) - F(-a) - \frac{3}{8}f(0) > 0$$

for a test function (7.1), then this situation does not occur for $q \geq q_0$.

Proof. Apply Lemma 7.6 to $(1 + \chi_1(n))(1 + \operatorname{Re} \chi_2(n)n^{-i\gamma_2}) \geq 0$. Its three nonprincipal characters $\chi_1, \chi_2, \chi_1\chi_2$ are real, so $\varphi = \frac{1}{4}$; retain ρ_1 and the given zero of χ_2 , which lies in $R(l)$. This gives $F(\lambda_1 - a) + F(\nu - a) \leq F(-a) + \frac{3}{8}f(0) + \varepsilon$. $\square$

It excludes 197 nodes of the case tree (Section 2.4), all of type rr with a reserved real family; the smallest normalized margin is $7.8 \cdot 10^{-5}$.

7.10. Applying the rows.

Lemma 7.15 (Preservation of configurations under location updates). *In the notation of Definition 11.2, suppose that a result of this section proves $\lambda_2 > h$ for every configuration of type $\mathfrak{t}$ with $\lambda_1 \in [a', b']$, and let $\mathfrak{s}$ be a specification of type $\mathfrak{t}$ with cell $[a, b] \subseteq [a', b']$. Then replacing l_2 and r by $\max(l_2, h)$ and $\max(r, h)$ (and lo_2 by the new r) does not change $C(\mathfrak{s})$, and $C(\mathfrak{s}) = \emptyset$ if $\mathfrak{s}$ is reserved with $\operatorname{hi}_2 \leq h$. Similarly a proof of $\lambda' > h$ allows p and lo' to be replaced by $\max(\cdot, h)$, and makes $C(\mathfrak{s})$ empty if $\operatorname{hi}' \leq h$.*

Proof. Every family other than the first has all its zeros in $R(l)$ at parameter at least $\lambda_2 > h$; in particular $\nu(\mathcal{F}) > h$. $\square$

7.11. The third family.

Proposition 7.16 (Lower bounds for the third family). *Consider a configuration satisfying a specification of type $\mathfrak{t}$ with $\lambda_1 \in [a, b]$. Put $h = \operatorname{hi}_2$ if a family is reserved and $h = \infty$ otherwise; then $\lambda_2 \leq h$ by Lemma 11.3. Each applicable rule below gives a strict lower bound for λ_3 , for all sufficiently large q . Let $\lambda_3^{\operatorname{lo}}$ be the maximum of the finitely many bounds selected for the leaf. Then $\lambda_3 > \lambda_3^{\operatorname{lo}}$.*

- (1) 0.857 (Input 7.4(a));
- (2) for $\mathfrak{t} = \operatorname{rr}$ and $[a, b]$ inside $[0.44, 0.60], [0.60, 0.70]$ or $[0.70, 0.80]$: 1.176, 1.055, 0.952 (Input 7.4(c));
- (3) for $\mathfrak{t} \neq \operatorname{rr}$ and $b \leq t$: the value c of the smallest applicable row of Input 7.4(b);

(4) for $\mathfrak{t} = \text{complex}$ and $0.44 \leq a \leq b \leq 0.85$: any t with

$$F(t-b) + F(\min(t, h) - a) - F(-b) + F(0) - \frac{7}{6}f(0) > 10^{-5}f(0) \quad (\gamma = \frac{5}{4});$$

(5) for $\mathfrak{t} = \text{rr}$, $0.44 \leq a \leq b \leq 0.80$ and $t \in [0.952, 1.176]$: any t such that $[a, \min(t, h)]$ is covered by intervals $[c, c']$ with

$$F(-c') - F(t - c') - F(0) - F(b - c) + \frac{9}{8}f(0) < -10^{-5}f(0) \quad (\gamma = \frac{26}{25}).$$

Proof. (1)–(3) are printed. For (4), suppose $\lambda_3 \leq t$; then $\lambda_2 \leq \min(t, h)$. The function $\lambda \mapsto F(-\lambda) - F(t - \lambda) = \int f(u)e^{\lambda u}(1 - e^{-tu}) du$ is increasing, so $F(-\lambda_1) - F(\lambda_3 - \lambda_1) \leq F(-\lambda_1) - F(t - \lambda_1) \leq F(-b) - F(t - b)$, and $F(\lambda_2 - \lambda_1) \geq F(\min(t, h) - a)$. So the right side of [Input 7.4\(d\)](#) is below $-10^{-5}f(0) + \varepsilon < 0$, a contradiction; the side condition [\[145, \(4.29\)\]](#) holds by our verification. For (5), suppose $\lambda_3 \leq t \leq 1.176$. Then $\lambda_2 \in [a, \min(t, h)] \subseteq [0.44, 1.176]$ and $\lambda_1 \in [0.44, 0.80]$, so [Input 7.4\(e\)](#) applies. If $\lambda_2 \in [c, c']$, the same monotonicity gives $F(-\lambda_2) - F(\lambda_3 - \lambda_2) \leq F(-c') - F(t - c')$ and $F(\lambda_1 - \lambda_2) \geq F(b - c)$, and the right side of (e) is negative for small ε , a contradiction; the side condition [\[145, \(4.34\)\]](#) holds by our verification. $\square$

For example, on the rr cell $[0.7025, 0.705]$ rule (5) gives $\lambda_3 > 1.0502$ without a reservation, and $\lambda_3 > 1.0888$ with a reserved family below 0.85; Table 10 gives 0.952. Rules (4) and (5) are evaluated with interval arithmetic, and a value t is accepted only with a normalized margin exceeding 10^{-5} .

8. THE GRADED NEAR-DENSITY LEMMA

The purpose of this section is to constrain many characters at once while retaining the location of each selected zero. We first bound a quadratic form in smoothed prime sums, called *responses* ([Theorem 8.4](#)). We then bound each response from below using selected zeros ([Lemma 8.7](#)). Finally, an elementary optimization argument introduces the single threshold used in a near row ([Proposition 8.8](#)). The analytic inputs are stated in [Section 3](#); the scope of the formalization is described in [Section 14.3](#). [Section 8.3](#) explains the idea of the proof, and [Example 8.9](#) shows what the resulting constraint can do that separate counts cannot.

8.1. Data. The detector determines the response to a zero. The Gram test and sieve weight control correlations between characters. The anchors may vary between entries, but all test functions and allowed offsets are fixed before q tends to infinity.

Definition 8.1 (Near-density data). A *near datum* consists of the following fixed objects; Conditions 1 and 2 are those of [Definition 3.1](#).

- A *detector* f satisfying Condition 1, with transform F .
- A *Gram test* g satisfying Conditions 1 and 2, with transform G .
- Real numbers s_1 and s , called the *safe anchor* and the *reference anchor*.
- A *sieve weight*: cells $t_0 < t_1 < \dots < t_m$ with $t_0 > \frac{1}{3} + 2\varepsilon'$ for a fixed $\varepsilon' > 0$, and heights $h_0, \dots, h_{m-1} \geq 0$. We put

$$\Xi(t) = \sum_{k < m} h_k \mathbf{1}_{[t_k, t_{k+1})}(t), \quad \omega(t) = g(t)e^{s_1 t} + \Xi(t),$$

and require that ω be bounded below by a positive constant on the support of f .

- A finite set $\Delta \subset [0, \infty)$ of *offsets*.

We put $\vartheta_k = \frac{1}{2}(t_k - \frac{1}{3}) - \varepsilon'$ (the *sieve levels*), so $0 < \vartheta_k < \frac{1}{2}t_k$, and

$$I_B = \int_0^\infty \frac{e^{2st} f(t)^2}{\omega(t)} dt, \quad D(\delta) = \int_0^\infty g(t)e^{(s_1 - 2\delta)t} dt + \sum_{k < m} h_k e^{-2\delta t_k} \frac{t_{k+1}^2 - t_k^2}{2\vartheta_k}, \quad d_1 = \frac{g(0)}{6}.$$

Here the integrand of I_B is taken to be 0 where $f = 0$. We require $I_B > 0$.

The number I_B is the first Cauchy–Schwarz factor, $D(\delta)$ bounds a diagonal entry at offset δ , and d_1 bounds the correlation between distinct characters. Heath-Brown's Lemma 12.1 is the formal analogue of the case $g = f$, $s = s_1 = \lambda_1$, $\Xi = 0$ and all offsets 0 (in that case $\omega = fe^{s_1 t}$ is not bounded below on the support of f , and Heath-Brown writes $f = \sqrt{f} \cdot \sqrt{f}$ instead).

Definition 8.2 (Entries and responses). An *entry* is a triple $(\chi_j, \gamma_j, \delta_j)$ of a nonprincipal character modulo q , a real height γ_j and an offset $\delta_j \in \Delta$. Its *response anchor* is $s_j = s - \delta_j$, and its *response* is

$$R_j = -\frac{1}{\mathcal{L}} \sum_{n \geq 1} \Lambda(n) \operatorname{Re} \left(\chi_j(n) n^{-(1-s_j/\mathcal{L}+i\gamma_j)} \right) f\left(\frac{\log n}{\mathcal{L}}\right).$$

For a family of entries, the *pair excess* of entry j is

$$e_j = \sum_{\substack{k \neq j \\ \chi_k = \chi_j}} \left(\operatorname{Re} G(-\sigma_{jk} + i(\gamma_j - \gamma_k)\mathcal{L}) - d_1 \right)_+, \quad \sigma_{jk} = s_1 - \delta_j - \delta_k.$$

Definition 8.3 (Safe anchor). For $\Gamma' > 0$, the *safe-anchor hypothesis* $\text{SA}(s_1, \Gamma')$ says that no nonprincipal $L(s, \chi)$ modulo q has a zero ρ with $|\operatorname{Im} \rho| \leq \Gamma'$ and $\operatorname{Re} \rho > 1 - s_1/\mathcal{L}$.

The pair excess e_j is needed only when two entries use the same character. It records the part of their correlation that exceeds the bound d_1 for distinct characters. The safe-anchor hypothesis ensures that discarded zero terms in the Gram estimate have the required sign.

In the application $s_1 \leq \lambda_1$. Then $\text{SA}(s_1, 2l+1)$ holds for large q , by the definition of λ_1 and **Input 3.5**: a zero with $|\gamma| \leq 2l+1 \leq 10l$ and $\operatorname{Re} \rho > 1 - s_1/\mathcal{L}$ lies in $R(10l)$, hence in $R(l)$, so its parameter is at least $\lambda_1 \geq s_1$.

8.2. Statement.

Theorem 8.4 (Graded near-density lemma). Assume **Inputs 3.3, 3.4, 3.7 and 3.8**, *Mertens' estimate*, and the *divisor bound* (the number of divisors of q is $O_\varepsilon(q^\varepsilon)$). Fix a near datum as in **Definition 8.1** and a tolerance $\eta' > 0$. There is a threshold q_0 , depending only on this fixed data, with the following property.

For $q \geq q_0$, let J be a finite index set and let $(\chi_j, \gamma_j, \delta_j)_{j \in J}$ be entries satisfying

- (i) $|\gamma_j| \leq \Gamma$ for some $0 \leq \Gamma \leq \mathcal{L}/3$;
- (ii) $\text{SA}(s_1, 2\Gamma + 1)$;
- (iii) the characters χ_j are pairwise distinct whenever $\Xi \neq 0$.

Then, for every choice of real coefficients $a_j \geq 0$,

$$\left(\sum_j a_j R_j \right)_+^2 \leq (1 + \eta') I_B \left[\sum_j (D(\delta_j) - d_1 + e_j) a_j^2 + (d_1 + \eta') \left(\sum_j a_j \right)^2 \right].$$

The responses R_j and pair excesses e_j are those of **Definition 8.2**. The threshold q_0 is uniform in the number of entries, their characters and heights, and the coefficients a_j .

Two special cases of the pair excess are used.

- (i) *Separated heights*. If $\chi_k = \chi_j$ implies $|\gamma_j - \gamma_k|\mathcal{L} \geq M$ for $k \neq j$, and M is so large that $\operatorname{Re} G(-\sigma + iY) \leq d_1$ for $|Y| \geq M$ and σ in the finite set of values σ_{jk} , then $e_j = 0$ (**Lemma 8.6**; this holds once $g(0) > 0$).
- (ii) *A controlled pair*. If entry j shares its character with exactly one other entry k , with $\delta_j = \delta_k = 0$ and $|(\gamma_j - \gamma_k)\mathcal{L}|$ in a range on which $\operatorname{Re} G(-s_1 + iy) \leq c^{\text{hi}}$, then $e_j \leq (c^{\text{hi}} - d_1)_+$.

8.3. The idea of the proof. Each response R_j is a smoothed sum over primes, twisted by $\chi_j(n)n^{-i\gamma_j}$ and evaluated at the point $1 - s_j/\mathcal{L} + i\gamma_j$. By the explicit formula (Lemma 8.7 below), every zero ρ of $L(s, \chi_j)$ near $1 + i\gamma_j$ contributes about $\operatorname{Re} F((\lambda_\rho - s_j) + i(\gamma_j\mathcal{L} - \mu_\rho))$ to R_j , and this is large when λ_ρ is small. So characters with zeros close to 1 have large responses.

On the other hand, all the responses are correlations of one vector, the weighted primes, with the twisted vectors $(\chi_j(n)n^{-i\gamma_j})_n$. By the Cauchy–Schwarz inequality a combination $\sum_j a_j R_j$ is at most the length of the prime vector, which gives the factor I_B , times the length of $\sum_j a_j \chi_j(n)n^{-i\gamma_j}$ measured with the weight ω . The latter is a Gram form in the characters. Its diagonal entries are the numbers $D(\delta_j)$, and distinct characters are nearly orthogonal on the primes: their correlation is at most $d_1 = g(0)/6$. Indeed, by the explicit formula the correlation is governed by the zeros of the quotient character $\chi_j \bar{\chi}_k$ to the right of the safe anchor, and the safe-anchor hypothesis says that there are none. Hence only a bounded number of characters can have large responses.

Heath-Brown’s Lemma 12.1 is the case in which all entries are tested at the same anchor. The grading gives each entry its own anchor $s_j = s - \delta_j$. An entry tested at a larger offset δ_j has a smaller diagonal $D(\delta_j)$, because of the factor $e^{-2\delta_j t}$, so characters whose zeros are further from 1 are charged less. The sieve weight Ξ adds Selberg-sieve majorants of the primes to ω on the ranges $[q^{t_k}, q^{t_{k+1}})$. This makes ω larger and the first factor I_B smaller. The price is a sieve contribution to the diagonal, computed with Graham’s asymptotic, and off-diagonal sums of the quotient characters over long intervals. These are small by Burgess’s estimate provided the quotient characters are nonprincipal, that is, provided the characters are distinct.

8.4. A Burgess bound for all nonprincipal characters.

Lemma 8.5 (Burgess bound for imprimitive characters and long intervals). *Assume Input 3.7. For every $\varepsilon > 0$ there is C such that for every nonprincipal character χ modulo q and all real $N \geq 0$ and $H \geq 1$,*

$$\left| \sum_{N < n \leq N+H} \chi(n) \right| \leq C q^{1/9+\varepsilon} \min(H, q)^{2/3}.$$

Proof. Let χ be induced by the primitive character χ^* modulo $q^* > 1$, and let r be the product of the primes dividing q but not q^* . Then $\chi(n) = \chi^*(n) \sum_{e|(n,r)} \mu(e)$, so the sum is $\sum_{e|r} \mu(e) \chi^*(e) \sum_{N/e < m \leq (N+H)/e} \chi^*(m)$. The sum only depends on the integers in the range, so we may take the endpoints real. In each inner sum remove complete periods of χ^* , over which χ^* sums to 0 because χ^* is nonprincipal; what remains is a sum over at most $\min(H/e, q^*)$ consecutive integers, which by periodicity may be shifted to start at a point ≥ 1 . Input 3.7 bounds it by $C q^{*1/9+\varepsilon} \min(H/e, q^*)^{2/3} \leq C q^{1/9+\varepsilon} \min(H, q)^{2/3}$ (for an interval of length less than 1 the bound is trivial). By the divisor bound, the number of e is $O_\varepsilon(q^\varepsilon)$. $\square$

8.5. Proof of Theorem 8.4. Write $t_n = \log n/\mathcal{L}$. For $(n, q) = 1$ put

$$Y_n = \sum_j a_j e^{-\delta_j t_n} \chi_j(n) n^{-i\gamma_j}.$$

Since $n^{-(1-s_j/\mathcal{L}+i\gamma_j)} = n^{-1} e^{st_n} e^{-\delta_j t_n} n^{-i\gamma_j}$, and $\chi_j(n) = 0$ for $(n, q) > 1$,

$$\sum_j a_j R_j = -\operatorname{Re} \frac{1}{\mathcal{L}} \sum_{(n,q)=1} \frac{\Lambda(n)}{n} e^{st_n} f(t_n) Y_n.$$

Cauchy-Schwarz. Since $\omega > 0$ on the support of f and $\omega \geq 0$ everywhere,

$$(8.1) \quad \left(\sum_j a_j R_j \right)_+^2 \leq \left(\frac{1}{\mathcal{L}} \sum_n \frac{\Lambda(n)}{n} \frac{e^{2st_n} f(t_n)^2}{\omega(t_n)} \right) \left(\frac{1}{\mathcal{L}} \sum_{(n,q)=1} \frac{\Lambda(n)}{n} \omega(t_n) |Y_n|^2 \right).$$

The first factor. The function $\varphi_B(t) = e^{2st} f(t)^2 / \omega(t)$ is bounded, because ω is bounded below on the support of f ; it vanishes outside the support of f and is continuous except at finitely many points, so it is Riemann integrable. Given $\varepsilon_B > 0$ choose a partition $0 = u_0 < \dots < u_K$ of an interval containing the support of f whose upper Riemann sum is at most $I_B + \varepsilon_B/2$. By Mertens' estimate $\sum_{n \leq x} \Lambda(n)/n = \log x + O(1)$ (see [112] for explicit forms) we have $\mathcal{L}^{-1} \sum_{q^{u_i} < n \leq q^{u_{i+1}}} \Lambda(n)/n = u_{i+1} - u_i + O(1/\mathcal{L})$, so the first factor is at most $I_B + \varepsilon_B$ for large q .

The second factor. Split it as $Q_1 + Q_2$ according to $\omega = g e^{s_1 t} + \Xi$. Expanding $|Y_n|^2$ and writing $\chi_{jk} = \chi_j \bar{\chi}_k$,

$$Q_1 = \sum_{j,k} a_j a_k X_{jk}, \quad X_{jk} = \operatorname{Re} \frac{1}{\mathcal{L}} \sum_{(n,q)=1} \Lambda(n) \chi_{jk}(n) n^{-(1-\sigma_{jk}/\mathcal{L})-i(\gamma_j-\gamma_k)} g(t_n).$$

All the points $w_{jk} = 1 - \sigma_{jk}/\mathcal{L} + i(\gamma_j - \gamma_k)$ involved satisfy $|\operatorname{Re} w_{jk} - 1| \leq C/\mathcal{L}$ and $|\operatorname{Im} w_{jk}| \leq 2\Gamma \leq \mathcal{L}$, so [Inputs 3.3](#) and [3.4](#) apply uniformly. The error terms $o(1)$ below are uniform in the characters, heights and entries.

- *Diagonal, $j = k$.* Here $\chi_{jj} = \chi_0$ and the height is 0. By [Input 3.3](#),

$$X_{jj} = \int_0^\infty g(t) e^{(s_1 - 2\delta_j)t} dt + o(1).$$

- *Distinct characters.* Here χ_{jk} is nonprincipal. By [Input 3.4](#),

$$X_{jk} \leq - \sum_{\rho} \operatorname{Re} G((w_{jk} - \rho)\mathcal{L}) + \frac{\varphi(\chi_{jk})}{2} g(0) + o(1),$$

the sum running over the zeros of $L(s, \chi_{jk})$ in a disc of radius $\delta < 1$ about $1 + i(\gamma_j - \gamma_k)$. Such a zero has $|\operatorname{Im} \rho| \leq 2\Gamma + 1$, so by [SA\(\$s_1, 2\Gamma + 1\$ \)](#) it has $\operatorname{Re} \rho \leq 1 - s_1/\mathcal{L} \leq 1 - \sigma_{jk}/\mathcal{L} = \operatorname{Re} w_{jk}$. Hence $\operatorname{Re}(w_{jk} - \rho)\mathcal{L} \geq 0$, and each term $\operatorname{Re} G(\cdot)$ is ≥ 0 by Condition 2. Dropping them and using $\varphi \leq \frac{1}{3}$, $g(0) \geq 0$, we get $X_{jk} \leq d_1 + o(1)$. When $\sigma_{jk} < 0$ the point w_{jk} lies to the right of 1, and the same argument applies.

- *The same character, $j \neq k$.* Here $\chi_{jk} = \chi_0$ at the height $\gamma_j - \gamma_k$. By [Input 3.3](#), $X_{jk} = \operatorname{Re} G(-\sigma_{jk} + i(\gamma_j - \gamma_k)\mathcal{L}) + o(1)$.

For $a \geq 0$ the off-diagonal part is therefore at most

$$\sum_{j \neq k} a_j a_k (d_1 + \epsilon_{jk}) + o(1) \left(\sum_j a_j \right)^2, \quad \epsilon_{jk} = \mathbf{1}_{\chi_j = \chi_k} (\operatorname{Re} G(-\sigma_{jk} + i(\gamma_j - \gamma_k)\mathcal{L}) - d_1)_+.$$

Since $\epsilon_{jk} = \epsilon_{kj}$ ($\operatorname{Re} G$ is even in the imaginary part, g being real) and $2a_j a_k \leq a_j^2 + a_k^2$, we have $\sum_{j \neq k} a_j a_k \epsilon_{jk} \leq \sum_j e_j a_j^2$. Also $\sum_{j \neq k} a_j a_k = (\sum_j a_j)^2 - \sum_j a_j^2$. Hence

$$(8.2) \quad Q_1 \leq \sum_j \left(\int_0^\infty g(t) e^{(s_1 - 2\delta_j)t} dt - d_1 + e_j \right) a_j^2 + (d_1 + o(1)) \left(\sum_j a_j \right)^2.$$

The sieve part Q_2 . It is present only when $\Xi \neq 0$, so the characters are distinct and all pair excesses vanish. Fix a cell k , let $z_k = q^{\vartheta_k}$ and let

$$\xi_d = \mu(d) \frac{\log(z_k/d)}{\log z_k} \quad (d \leq z_k), \quad \nu_k(n) = \left(\sum_{d|n, d \leq z_k} \xi_d \right)^2 \geq 0.$$

- *Majorant.* If p is prime with $t_p \geq t_k$ then $p \geq q^{t_k} > z_k$, so the only divisor of p up to z_k is 1 and $\nu_k(p) = \xi_1^2 = 1$. Hence $\Lambda(p) = (\log p) \nu_k(p)$. For n not a prime power, $\Lambda(n) = 0 \leq (\log n) \nu_k(n)$. Prime powers p^e with $e \geq 2$ and $t_{p^e} \geq t_k$ contribute $\ll \mathcal{L} q^{-t_k/2} (\sum_j a_j)^2 = o(1) (\sum_j a_j)^2$. So the contribution of cell k to Q_2 is at most

$$\frac{h_k}{\mathcal{L}} \sum_{t_n \in [t_k, t_{k+1})} \frac{\log n}{n} \nu_k(n) |Y_n|^2 + o(1) \left(\sum_j a_j \right)^2,$$

where now $|Y_n|^2$ is expanded without the coprimality condition, with $\chi_j(n) = 0$ for $(n, q) > 1$.

- *Diagonal.* The terms $j = k'$ of $|Y_n|^2$ are at most $a_j^2 e^{-2\delta_j t_n}$ with $\chi_j \bar{\chi}_j \leq 1$. By [Input 3.8](#), $\sum_{n \leq x} \nu_k(n) = x / \log z_k + O(x / \log^2 z_k)$ for $x \geq z_k$. Partial summation gives

$$\frac{1}{\mathcal{L}} \sum_{t_n \in [t_k, t_{k+1})} \frac{\log n}{n} \nu_k(n) e^{-2\delta_j t_n} = \int_{t_k}^{t_{k+1}} \frac{t}{\vartheta_k} e^{-2\delta_j t} dt + O\left(\frac{1}{\mathcal{L}}\right) \leq e^{-2\delta_j t_k} \frac{t_{k+1}^2 - t_k^2}{2\vartheta_k} + O\left(\frac{1}{\mathcal{L}}\right).$$

- *Off-diagonal.* For $j \neq k'$ the character $\chi_{jk'} = \chi_j \bar{\chi}_{k'}$ is nonprincipal, since the characters are distinct. Put $W(n) = (\log n) n^{-1} e^{-(\delta_j + \delta_{k'}) t_n} n^{-i(\gamma_j - \gamma_{k'})}$. Opening ν_k , the term is

$$\frac{1}{\mathcal{L}} \sum_{d, e \leq z_k} \xi_d \xi_e \chi_{jk'}([d, e]) \sum_m \chi_{jk'}(m) W([d, e]m),$$

where m runs over an interval $[M_0, M_1)$ with $M_0 = q^{t_k} / [d, e] \geq q^{t_k - 2\vartheta_k} = q^{1/3 + 2\varepsilon'}$. By [Lemma 8.5](#), $|\sum_{M_0 \leq m \leq u} \chi_{jk'}(m)| \ll u^{2/3} q^{1/9 + \varepsilon}$ for every real $u \geq M_0$. We also have $|W(u[d, e])| \ll \mathcal{L}(u[d, e])^{-1}$ and $|\partial_u W(u[d, e])| \ll \mathcal{L}(1 + |\gamma_j - \gamma_{k'}|) u^{-2} [d, e]^{-1} \ll \mathcal{L}^2 u^{-2} [d, e]^{-1}$. Abel summation, with the bound for the partial sums $\sum_{M_0 \leq m \leq u} \chi_{jk'}(m)$ at every u , therefore bounds the inner sum by

$$\ll \mathcal{L}^2 [d, e]^{-1} q^{1/9 + \varepsilon} M_0^{-1/3} = \mathcal{L}^2 q^{1/9 + \varepsilon - t_k/3} [d, e]^{-2/3}.$$

Since $\sum_{d, e \leq z_k} [d, e]^{-2/3} \leq \sum_{g \leq z_k} g^{-2/3} (\sum_{d' \leq z_k/g} d'^{-2/3})^2 \leq 9\zeta(\frac{4}{3}) z_k^{2/3}$ and $|\xi_d| \leq 1$, the term is

$$\ll \mathcal{L} q^{1/9 + \varepsilon - t_k/3 + 2\vartheta_k/3} = \mathcal{L} q^{\varepsilon - 2\varepsilon'/3},$$

because $2\vartheta_k = t_k - \frac{1}{3} - 2\varepsilon'$. With $\varepsilon < \varepsilon'/2$ this is $o(1)$ uniformly, and the sum over j, k' is $o(1) (\sum_j a_j)^2$. (A bound for the partial sums only at the end point of the interval, combined with the total variation of W , would lose a factor $(M_1/M_0)^{2/3}$, which is too large for the cells used.)

Summing over the cells,

$$(8.3) \quad Q_2 \leq \sum_j a_j^2 \sum_{k < m} h_k e^{-2\delta_j t_k} \frac{t_{k+1}^2 - t_k^2}{2\vartheta_k} + o(1) \left(\sum_j a_j \right)^2.$$

Conclusion. By (8.2) and (8.3) the second factor is at most $B + o(1) (\sum_j a_j)^2$, where $B = \sum_j (D(\delta_j) - d_1 + e_j) a_j^2 + d_1 (\sum_j a_j)^2 \geq \sum_j (D(\delta_j) + e_j) a_j^2 \geq 0$. With the first factor $I_B + o(1)$ and $I_B > 0$, the right side of (8.1) is at most $(1 + \eta') I_B [B + \eta' (\sum_j a_j)^2]$ for q large. $\square$

8.6. Responses and the threshold form.

Lemma 8.6 (Uniform decay in the imaginary direction). *If f satisfies Condition 1, then $\operatorname{Re} F(x + iY) \rightarrow 0$ as $|Y| \rightarrow \infty$, uniformly for x in compact sets.*

Proof. Integrating by parts twice, $F(z) = f(0)/z + O_K(|z|^{-2})$ for $\operatorname{Re} z$ in a compact set K . Moreover $\operatorname{Re}(f(0)/z) = f(0) \operatorname{Re} z / |z|^2 = O_K(|z|^{-2})$. $\square$

Lemma 8.7 (Response lemma). *Assume [Input 3.4](#). Let f satisfy Conditions 1 and 2, and fix $\sigma_{\max} > 0$, an integer $K_0 \geq 0$, and $\eta' > 0$. There are $M > 0$, $\delta \in (0, 1)$ and q_0 such that the following holds for $q \geq q_0$, and it remains true when δ is replaced by any smaller radius (with a suitable q_0). Let χ be nonprincipal, $|\sigma| \leq \sigma_{\max}$ and $|\gamma| \leq \mathcal{L}$. Let $\mathcal{S}$ be a finite set of distinct zeros of $L(s, \chi)$ in the disc $|1 + i\gamma - \rho| \leq \delta$, each retained with its full multiplicity m_ρ . Suppose that every other zero of $L(s, \chi)$ in the disc with $\lambda_\rho < \sigma$ satisfies $|\gamma\mathcal{L} - \mu_\rho| \geq M$, and that there are at most K_0 such zeros, counted with multiplicity. Then the response of (χ, γ, σ) , that is $R = -\mathcal{L}^{-1} \sum_n \Lambda(n) \operatorname{Re}(\chi(n) n^{-(1-\sigma/\mathcal{L}+i\gamma)}) f(t_n)$, satisfies*

$$R \geq \sum_{\rho \in \mathcal{S}} m_\rho \operatorname{Re} F((\lambda_\rho - \sigma) + i(\gamma\mathcal{L} - \mu_\rho)) - \frac{f(0)}{6} - \eta'.$$

Proof. By [Lemma 8.6](#) choose M with $|\operatorname{Re} F(x + iY)| \leq \eta'/(2K_0 + 2)$ for $|Y| \geq M$ and $x \in [-\sigma_{\max}, 0]$. Apply [Input 3.4](#) with $\varepsilon = \eta'/2$ at the point $w = 1 - \sigma/\mathcal{L} + i\gamma$, and let δ be its disc radius, or any smaller radius (see the remark after [Input 3.4](#)). We have $(w - \rho)\mathcal{L} = (\lambda_\rho - \sigma) + i(\gamma\mathcal{L} - \mu_\rho)$, so

$$R \geq \sum_{|1+i\gamma-\rho| \leq \delta} \operatorname{Re} F((w - \rho)\mathcal{L}) - \frac{\varphi(\chi)}{2} f(0) - \frac{\eta'}{2},$$

the sum counting multiplicity. Zeros in $\mathcal{S}$ are kept. A zero outside $\mathcal{S}$ with $\lambda_\rho \geq \sigma$ has $\operatorname{Re}(w - \rho)\mathcal{L} \geq 0$ and a term ≥ 0 (Condition 2), which we drop. The remaining zeros have $\lambda_\rho \in [0, \sigma)$ and $|\gamma\mathcal{L} - \mu_\rho| \geq M$; there are at most K_0 of them, and each term is at least $-\eta'/(2K_0 + 2)$. Finally $\varphi(\chi) \leq \frac{1}{3}$ and $f(0) \geq 0$. $\square$

Proposition 8.8 (Threshold form). *Let J be a finite index set, let $D_j > 0$ and $v_j \in \mathbb{R}$ for $j \in J$, and let $d > 0$. Suppose that*

$$\left(\sum_j a_j v_j \right)_+^2 \leq \sum_j D_j a_j^2 + d \left(\sum_j a_j \right)^2 \quad \text{for every } (a_j)_{j \in J} \in [0, \infty)^J.$$

Then there is $\tau \in [0, \sqrt{d}]$ with

$$(T) \quad \sum_j \frac{(v_j - \tau)_+^2}{D_j} \leq 1 - \frac{\tau^2}{d}.$$

We call the v_j the *features*, the D_j the *diagonals* and d the *correlation term* of the inequality, and τ its *threshold*. A feature measures how strongly an entry is detected, a diagonal what the entry costs on its own, and d bounds the correlation between distinct entries.

Example 8.9 (One level and two levels). (i) Suppose that N entries have the same feature $v > 0$ and the same diagonal D , with $v^2 > d$. Taking every $a_j = 1$ in the hypothesis of [Proposition 8.8](#) gives $(Nv)^2 \leq ND + dN^2$, that is,

$$N \leq \frac{D}{v^2 - d}.$$

This is the kind of bound that Heath-Brown's Lemma 12.1 provides: a bound for the number $N(\lambda)$ of characters with a zero at distance at most λ , all counted alike.

(ii) Now take $D = 1$ and $d = \frac{1}{4}$, one entry with feature $v_1 = 1$ (a zero close to 1) and two entries with feature $v_2 = \frac{3}{4}$ (zeros further away). The count of (i), applied to each level separately, allows this: it permits up to $D/(v_1^2 - d) = \frac{4}{3}$ entries with feature at least v_1 , and up to $D/(v_2^2 - d) = \frac{16}{5}$

with feature at least v_2 . But (T) fails for every τ : for $0 \leq \tau \leq \frac{1}{2}$ its left side minus its right side is

$$(1 - \tau)^2 + 2\left(\frac{3}{4} - \tau\right)^2 - (1 - 4\tau^2) = 7\tau^2 - 5\tau + \frac{9}{8},$$

whose discriminant $25 - \frac{63}{2}$ is negative. (Equivalently, taking all three $a_j = 1$ violates the hypothesis: $(1 + \frac{3}{2})^2 = \frac{25}{4} > 3 + \frac{9}{4}$.) So this configuration is excluded. A single threshold couples the near entry and the farther ones, which separate counts cannot do. In the leaf programs this coupling acts on all the bins of characters at once.

Proof. The function $\mathcal{J}(\tau) = \tau^2/d + \sum_j (v_j - \tau)_+^2/D_j$ is convex and C^1 on $[0, \infty)$, and tends to ∞ ; let τ^* minimize it. If $\tau^* = 0$ then $\mathcal{J}'(0) \geq 0$, so all $v_j \leq 0$ and $\mathcal{J}(0) = 0$. Otherwise $\mathcal{J}'(\tau^*) = 0$, that is $\tau^* = d \sum_j a_j$ with $a_j = (v_j - \tau^*)_+/D_j \geq 0$. Then $\sum_j a_j v_j = \sum_j a_j (v_j - \tau^*) + \tau^* \sum_j a_j = \sum_j D_j a_j^2 + d(\sum_j a_j)^2 = \mathcal{J}(\tau^*)$. The hypothesis gives $\mathcal{J}(\tau^*)^2 \leq \mathcal{J}(\tau^*)$, so $\mathcal{J}(\tau^*) \leq 1$. Finally (T) forces $\tau^2 \leq d$. $\square$

Corollary 8.10 (Zero form). *Fix a near datum whose detector f also satisfies Condition 2. Fix a common bound K_0 for the omitted zeros in Lemma 8.7, and choose $\sigma_{\max} \geq \max_{\delta \in \Delta} |s - \delta|$ with $\sigma_{\max} > 0$. Choose $\eta' > 0$, $I_u \geq I_B$, and $D_u > 0$, and put*

$$\eta'' = \eta' \min\{1, \sqrt{I_B D_u}, D_u\}.$$

For sufficiently large q , consider entries satisfying Theorem 8.4 with tolerance η'' . For each entry j , suppose the retained set $\mathcal{S}_j$ satisfies Lemma 8.7 at anchor s_j , with the separation and disc radius chosen for tolerance η'' . Define

$$r_j = \sum_{\rho \in \mathcal{S}_j} m_\rho \operatorname{Re} F((\lambda_\rho - s_j) + i(\gamma_j \mathcal{L} - \mu_\rho)) - \frac{f(0)}{6},$$

and choose $D_j^+ > 0$ with $D_j^+ \geq D(\delta_j) - d_1 + e_j$ for every j . Then there is $\tau \in [0, \sqrt{d}]$ satisfying (T) with

$$v_j = \frac{r_j}{\sqrt{I_u D_u}} - \eta', \quad D_j = \frac{(1 + \eta') D_j^+}{D_u}, \quad d = \frac{(1 + \eta')(d_1 + \eta' D_u)}{D_u} \quad (= (1 + \eta')(d_1/D_u + \eta')).$$

The same threshold remains valid if features v_j are decreased, diagonals D_j or the correlation term d are increased, or nonpositive features are replaced by 0. An entry with nonpositive feature may also be omitted.

Proof. Apply Theorem 8.4 and Lemma 8.7 with the tolerance η'' fixed in the statement. The response lemma gives $R_j \geq k_j := r_j - \eta''$, so for $a \geq 0$ we have $(\sum_j a_j k_j)_+ \leq (\sum_j a_j R_j)_+$. Dividing the inequality of Theorem 8.4 by $I_B D_u$, we get for all $a \geq 0$

$$\left(\sum_j a_j \frac{k_j}{\sqrt{I_B D_u}}\right)_+^2 \leq \sum_j \frac{(1 + \eta'') D_j^+}{D_u} a_j^2 + \frac{(1 + \eta'')(d_1 + \eta'')}{D_u} \left(\sum_j a_j\right)^2,$$

where we also replaced each coefficient $D(\delta_j) - d_1 + e_j$ by the larger D_j^+ . Proposition 8.8 gives τ satisfying (T) with these features, diagonals and correlation term; only the positivity of the D_j^+ is needed, not that of $D(\delta_j) - d_1 + e_j$. We compare with the displayed quantities.

- If $r_j > 0$ then $v_j \leq k_j/\sqrt{I_B D_u}$, because $I_u \geq I_B$ and $\eta''/\sqrt{I_B D_u} \leq \eta'$. If $r_j \leq 0$ then $v_j < 0$.
- $(1 + \eta') \geq (1 + \eta'')$, so the displayed diagonals are at least the ones above.
- $(1 + \eta')(d_1 + \eta' D_u) \geq (1 + \eta'')(d_1 + \eta'')$, since $\eta'' \leq \eta' D_u$.

Finally, if (T) holds at τ for features v_j , diagonals D_j and correlation term d , it also holds at τ for features $v'_j \leq \max(v_j, 0)$ with $v'_j \leq v_j$ or $v'_j \leq 0$, diagonals $D'_j \geq D_j$ and correlation term $d' \geq d$. Indeed each term $(v'_j - \tau)_+^2 / D'_j$ is at most $(v_j - \tau)_+^2 / D_j$ (the former is 0 when $v'_j \leq 0 \leq \tau$), $1 - \tau^2/d \leq 1 - \tau^2/d'$, and $\tau \leq \sqrt{d} \leq \sqrt{d'}$. $\square$

9. THE NEAR ROWS

A *near row* is a constraint obtained by applying the threshold inequality (T) to selected characters and zeros. In a leaf program (Section 2.4), characters are grouped by parameter into bins, each with its own column. Known entries from the first or reserved family may instead be recorded separately as *family terms*. There are four kinds of rows, the family, shifted, graded and two-test rows; Table 8 in Section 9.5 lists their anchors, sieve weights and entries.

We first specify the test functions and normalization. We then justify each permitted kind of entry and describe how to combine them into rows. Most rows follow from Corollary 8.10; the two-test row in Section 9.6 has a separate proof. Proposition 9.7 is the conclusion needed later: every actual zero configuration in a leaf satisfies each of its near rows at some threshold.

9.1. The parabolic test function. All detectors and Gram tests are normalized parabolic autocorrelations. For $c > 0$ put

$$P(u) = (1 - u)^3(1 + 3u + u^2) = 1 - 5u^2 + 5u^3 - u^5, \quad \hat{f}_c(t) = \begin{cases} P(t/c), & 0 \leq t \leq c, \\ 0, & t \geq c. \end{cases}$$

Thus $\hat{f}_{2\gamma} = f_\gamma / f_\gamma(0)$ with f_γ as in (7.1). Write $\hat{F}_c$ for its transform.

Lemma 9.1 (Properties of the parabolic test functions). *Let $c > 0$.*

- (i) $\hat{f}_c \geq 0$, $\hat{f}_c(0) = 1$, $\hat{f}_c$ is decreasing on $[0, c]$, and $\hat{f}_c$ extended by 0 is twice continuously differentiable on $[0, \infty)$ with $|\hat{f}_c''| \leq 10/c^2$. So Condition 1 holds.
- (ii) $\operatorname{Re} \hat{F}_c(iy) = 60c(cy \cos \frac{cy}{2} - 2 \sin \frac{cy}{2})^2 / (cy)^6 \geq 0$ for $y \neq 0$, and Condition 2 holds.
- (iii) $x \mapsto \hat{F}_c(x)$ is positive and strictly decreasing on $\mathbb{R}$.
- (iv) For $z \neq 0$, $\hat{F}_c(z) = c\Phi(cz)$, where

$$\Phi(w) = \frac{1}{w} - \frac{10}{w^3} + \frac{30(1 + e^{-w})}{w^4} + \frac{120e^{-w}}{w^5} - \frac{120(1 - e^{-w})}{w^6} \quad (w \neq 0).$$

The singularity at 0 is removable. Consequently, for $d \geq 0$ and $|y| \geq U > 0$,

$$\operatorname{Re} \hat{F}_c(-d + iy) \leq cR(cU, e^{cd}), \quad -\operatorname{Re} \hat{F}_c(-d + iy) \leq \frac{d}{U^2} + cR(cU, e^{cd}),$$

where $R(u, E) = \frac{10}{u^3} + \frac{30}{u^4} + \frac{120}{u^6} + E\left(\frac{30}{u^4} + \frac{120}{u^5} + \frac{120}{u^6}\right)$.

- (v) For $d \geq 0$ put $\mathcal{C}_c(d) = \sup_{\operatorname{Re} z \geq -d} (-\operatorname{Re} \hat{F}_c(z))$. It is finite,

$$\mathcal{C}_c(d) = \max\left(0, \sup_{y \in \mathbb{R}} (-\operatorname{Re} \hat{F}_c(-d + iy))\right),$$

and $\mathcal{C}_c(0) = 0$.

Proof. (i) is elementary: $P'(u) = -5u(2 + u)(1 - u)^2$, and $P(1) = P'(1) = P''(1) = 0$. (ii) The function $\hat{f}_c$ is a positive multiple of the autocorrelation of $x \mapsto (\gamma^2 - x^2)_+$ with $c = 2\gamma$, so its transform on the imaginary axis is a positive multiple of $(\int_0^\infty (\gamma^2 - x^2)_+ \cos(xy) dx)^2$ [145, Lemma 3.6], which gives the closed form. Condition 2 then follows from Input 3.2 with $F_2 = 0$, since $\hat{F}_c(z) \rightarrow 0$ uniformly in $\operatorname{Re} z \geq 0$. (iii) $\hat{F}'_c(x) = -\int t \hat{f}_c(t) e^{-xt} dt < 0$. (iv) Five integrations by parts, using the values of P and its derivatives at 0 and 1. For $z = -d + iy$

we have $c \operatorname{Re} \frac{1}{cz} = \frac{-d}{d^2+y^2}$, which is ≤ 0 and $\geq -d/U^2$, and each other term is bounded by the corresponding term of $cR(c|z|, e^{cd})$, which decreases in $|z| \geq U$. (v) $-\operatorname{Re} \hat{F}_c(z-d)$ is harmonic in $\operatorname{Re} z > 0$, continuous up to the boundary and tends to 0 uniformly at infinity, so its supremum over $\operatorname{Re} z \geq 0$ is attained on the boundary or is 0; $\mathcal{C}_c(0) = 0$ by Condition 2. $\square$

9.2. Row data. Every family, shifted and graded row fixes rational numbers $\gamma_f > 0$ (detector), $\gamma_g > 0$ (Gram test), $s_1 \leq s$ (safe and reference anchors), t_0 and μ_{row} , and uses the near datum of [Section 8.1](#) with $f = \hat{f}_{2\gamma_f}$, $g = \hat{f}_{2\gamma_g}$, $\varepsilon' = \frac{1}{2000}$, and 2000 equal cells $t_0 < t_1 < \dots < t_{2000} = 2\gamma_f$, where $t_0 > \frac{1}{3} + 2\varepsilon'$. The heights $h_k \geq 0$ are rational numbers. They are chosen by a heuristic that brings ω close to the Cauchy–Schwarz optimum $\omega \propto e^{st}f$: the height of cell k is the positive part of $e^{sm_k} f(m_k) / \sqrt{\mu_{\text{row}} \kappa_k} - g(m_k) e^{s_1 m_k}$ at the midpoint m_k , where $\kappa_k = (t_{k+1}^2 - t_k^2) / (2\vartheta_k(t_{k+1} - t_k))$ is the mean sieve cost per unit length of the cell, computed in binary64 arithmetic and converted to a rational number with denominator at most 10^{12} . Any nonnegative heights are admissible, so this choice plays no role in the proofs. In every family and shifted row $\mu_{\text{row}} = 10^9$ and all the computed heights are 0, so $\Xi \equiv 0$. The offsets are the finitely many numbers $s - \sigma$ that occur for the anchors σ of the entries. In every row the lower bound of ω on each cell of the support of f is checked to be positive; in the rows with $\Xi \equiv 0$ this holds because $\gamma_g > \gamma_f$.

We use the normalization of [Corollary 8.10](#): I_u is an upper Riemann sum for I_B over 4000 cells, computed with outward interval arithmetic; $D_u \geq D(0)$; each entry j receives the feature $v_j = r_j / \sqrt{I_u D_u} - \eta'$ (rounded down, and replaced by 0 if negative) and the diagonal $(1 + \eta') D_j^+ / D_u$ (rounded up), where D_j^+ is an upper bound for $D(\delta_j) - d_1 + e_j$; and the row has the correlation term $d = (1 + \eta')(d_1 / D_u + \eta')$ (rounded up). Here $\eta' = 10^{-6}$. For an entry with offset δ the value $D(\lfloor \delta \rfloor_{1/200}) \geq D(\delta)$ is used (here $\lfloor \delta \rfloor_{1/200}$ is δ rounded down to a multiple of $\frac{1}{200}$, and D is decreasing), and an entry is omitted if its diagonal numerator does not meet the positivity margin used in the construction. An omitted entry has feature 0 in the row and contributes nothing to (T). The rounding directions are justified by the last assertion of [Corollary 8.10](#).

9.3. Facts available in a leaf. A leaf specifies intervals and lower bounds for the first few zeros. The full definition of a specification $\mathfrak{s}$ and its configuration set $C(\mathfrak{s})$ is in [Definition 11.2](#); the information needed for this section is listed below. Write $\nu(\mathcal{F})$ for the least parameter in a family among zeros of physical height at most 1, and $\nu_{T^*}(\chi)$ for the corresponding minimum at height at most T^* . A reserved family minimizes ν outside the first family. It need not be the family attaining the global parameter λ_2 .

Fix a leaf with first-zero cell $[a, b]$ and a configuration in $C(\mathfrak{s})$. For all sufficiently large q , the following facts are available.

- (Z1) $\lambda_1 \in [a, b]$. The *safe anchor* of the leaf is $s_1 = \lfloor 50a \rfloor / 50 \leq \lambda_1$.
- (Z2) $\lambda' \geq p^*$, where $p^* = \text{lo}'$ if there is a gap (an interval $[\text{lo}', \text{hi}']$ known to contain λ') and $p^* = p$ otherwise; if there is a finite gap, $\lambda' \leq \text{hi}'$.
- (Z3) Every zero in $R(l)$ of a character outside the first family has parameter at least $\lambda_2 \geq l_2$.
- (Z4) Every family $\mathcal{F} \neq \mathcal{F}_1$ has $\nu(\mathcal{F}) \geq r$, and ordinary characters have $\nu_{T^*} \geq \nu \geq r$.
- (Z5) If $\mathfrak{s}$ is reserved, the reserved family has n_2 characters with common height-one parameter $\nu_* \in [\text{lo}_2, \text{hi}_2]$.
- (Z6) Inside: $|\mu_1| \leq C_B$. Outside: $|\mu_1| > C_B$.

Lemma 9.2 (No zero to the right). *Let χ be nonprincipal, let $|\gamma| \leq 2l$ and $0 < \delta \leq 1$, and fix a real anchor σ . For sufficiently large q , every zero ρ of $L(s, \chi)$ in the disc $|1 + i\gamma - \rho| \leq \delta$ has $\lambda_\rho \geq \sigma$ in either of the following cases:*

- (a) $\sigma \leq \lambda_1$;
- (b) $\chi \notin \mathcal{F}_1$ and $\sigma \leq l_2$, where $l_2 \leq \lambda_2$ is the leaf's global second-family lower bound.

Equivalently, this disc contains no zero to the right of the line $\operatorname{Re} s = 1 - \sigma/\mathcal{L}$.

Proof. Such a zero has $|\operatorname{Im} \rho| \leq 2l + \delta < 10l$. By [Input 3.5](#) it lies in $R(l)$ or has $\lambda_\rho > \frac{1}{3} \log \log \mathcal{L}$, which exceeds every fixed σ . In $R(l)$ its parameter is at least λ_1 , and at least $\lambda_2 \geq l_2$ if $\chi \notin \mathcal{F}_1$ (Z3). $\square$

Lemma 9.3 (Separation of zeros beyond the representative strip). *Let χ be an ordinary character with T^* -representative ρ_χ , of height γ_χ and parameter λ_χ , and let $\sigma \leq \min(\lambda_\chi, s_{\max})$. Then every zero ρ of $L(s, \chi)$ with $|1 + i\gamma_\chi - \rho| \leq \delta_0$ and $\lambda_\rho < \sigma$ satisfies $|\gamma_\chi \mathcal{L} - \mu_\rho| \geq M$, and there are at most K_0 such zeros, counted with multiplicity.*

Proof. Such a zero has $|\operatorname{Im} \rho| \leq T^* + \delta_0 < 2 \leq 10l$ and $\lambda_\rho < \sigma \leq s_{\max} < \frac{1}{3} \log \log \mathcal{L}$, so it lies in $R(10l)$ and hence, by [Input 3.5](#), in $R(l)$. As $\lambda_\rho < \sigma \leq \lambda_\chi$, the minimality of the representative gives $|\operatorname{Im} \rho| > T^*$; so the zero is counted in K_0 . By [Lemma 6.3](#), $|\operatorname{Im} \rho| \geq T^* + M/\mathcal{L}$, so $\mathcal{L}|\operatorname{Im} \rho - \gamma_\chi| \geq \mathcal{L}(|\operatorname{Im} \rho| - T^*) \geq M$. $\square$

9.4. Entries. We now list the entries that occur in the rows, with their kept zeros. In each case the hypotheses of [Lemma 8.7](#) hold with the stated kept set $\mathcal{S}$ and anchor σ ; we write r for the resulting lower bound for the response, before normalization, so that the entry's feature is $r/\sqrt{I_u D_u} - \eta'$. [Table 7](#) summarizes the entries. Recall that $F(x)$ is positive and decreasing on $\mathbb{R}$, and that a kept zero with $\lambda_\rho \geq \sigma$ has $m_\rho \operatorname{Re} F(\cdot) \geq \operatorname{Re} F(\cdot) \geq 0$.

Entry	Anchor σ	Kept set $\mathcal{S}$	Response bound r	Diagonal numerator	Rows
(a) ordinary χ , bin $[\text{lo}, \text{hi}]$, at its T^* -representative	$\min(\text{lo}, s)$	$\{\rho_\chi\}$	$F(\text{hi} - \sigma) - \frac{1}{6}f(0)$	$D(s - \sigma) - d_1$	all three
(b) inside first family: χ_1 at γ_1 , and $\bar{\chi}_1$ at $-\gamma_1$	$\min(a, s)$	$\{\rho_1\}$, resp. $\{\bar{\rho}_1\}$	$F(b - \sigma) - \frac{1}{6}f(0)$	$D(s - \sigma) - d_1$	family, graded
(c) type rc: χ_1 at $\pm\gamma_1$	s_1	$\{\rho_1, \bar{\rho}_1\}$	$F(b - s_1) + R_{\text{lo}} - \frac{1}{6}f(0)$	$D(0) - d_1 + (c^{\text{hi}} - d_1)_+$	family
(d) second zero, type complex: χ_1 at γ_1, γ' , $\bar{\chi}_1$ at $-\gamma_1, -\gamma'$	s_1	$\{\rho_1, \rho'\}$ and conjugates	$F(b - s_1) + R_p - \frac{1}{6}f(0)$ at γ_1 ; $F(\text{hi}' - s_1) + R_q - \frac{1}{6}f(0)$ at γ'	$D(0) - d_1 + (c^{\text{hi}} - d_1)_+$	family
(e) shifted first family: χ_1 at γ_1 , and $\bar{\chi}_1$ at $-\gamma_1$	s	$\{\rho_1\}$; $\{\rho_1, \bar{\rho}_1\}$ for rc	$F(b - s) - \mathbf{1}_{\text{rc}} C_Z - \frac{1}{6}f(0)$	$D(0) - d_1$	shifted
(f) reserved family, at height-one representatives	σ_2	the representative	$F(\text{hi}_2 - \sigma_2) - \frac{1}{6}f(0)$	$D(s - \sigma_2) - d_1$	all three
(g) outside first family: χ_1 at γ_1 (and $\bar{\chi}_1$), and one hidden entry per character	s_1	$\{\rho_1\}$; $\{\rho_h\}$	$F(b - s_1) - \frac{1}{6}f(0)$; $F(\text{hi}_h - s_1) - \frac{1}{6}f(0)$	$D(0) - d_1$	family

TABLE 7. The entries of the near rows, justified in the list below. The anchor σ enters through the offset $s - \sigma$, where s is the reference anchor of the row; σ_2 and the constants R_{lo} , R_p , R_q , c^{hi} and C_Z are defined in the list. “All three” means the family, shifted and graded rows of [Section 9.5](#).

- (a) *Ordinary character in the bin* $[\text{lo}, \text{hi}]$. Entry $(\chi, \gamma_\chi, s - \sigma)$ at its T^* -representative, with $\sigma = \min(\text{lo}, s)$; $\mathcal{S} = \{\rho_\chi\}$. If $\sigma \leq l_2$ there is no zero to the right ([Lemma 9.2](#)); otherwise [Lemma 9.3](#) gives the hypotheses. Then $r = F(\text{hi} - \sigma) - \frac{1}{6}f(0)$, and the diagonal numerator is $D(s - \sigma) - d_1$.
- (b) *Inside first family, one entry per character*. Entries $(\chi_1, \gamma_1, s - \sigma)$ and, for type complex, $(\bar{\chi}_1, -\gamma_1, s - \sigma)$, with $\sigma = \min(a, s) \leq \lambda_1$; $\mathcal{S} = \{\rho_1\}$ (resp. $\{\bar{\rho}_1\}$). No zero to the right. $r = F(b - \sigma) - \frac{1}{6}f(0)$, and the diagonal numerator is $D(s - \sigma) - d_1$.
- (c) *Conjugate pair, type rc* (rows with $s = s_1$ and $\Xi = 0$). Entries $(\chi_1, \pm\gamma_1, 0)$, a pair of entries of the same character with height difference $2\gamma_1$; $\mathcal{S} = \{\rho_1, \bar{\rho}_1\}$ for both. No zero to the right, and

$$r = F(b - s_1) + R_{\text{lo}} - \frac{1}{6}f(0), \quad R_{\text{lo}} \leq \inf \operatorname{Re} F(x + 2i\mu)$$

over $x \in [a - s_1, b - s_1]$ and μ in the height range of the leaf ($[0, 1]$ or $[1, \infty)$). Pair excess $e \leq (c^{\text{hi}} - d_1)_+$ with $c^{\text{hi}} \geq \sup \operatorname{Re} G(-s_1 + 2i\mu)$ over the range, so the diagonal numerator is $D(0) - d_1 + (c^{\text{hi}} - d_1)_+$.

- (d) *Second zero of the first family* (rows with $s = s_1$ and $\Xi = 0$; type complex with a finite gap $[p^*, \text{hi}']$). Let ρ' be an admissible zero of χ_1 realizing λ' ([Section 2](#)), of height γ' , namely the one that witnesses the piece of the second-zero split ([Section 11.6](#)) in which the configuration lies, and $y = (\gamma_1 - \gamma')\mathcal{L}$, restricted to that piece: $|y| \in [0, 1], [1, \frac{3}{2}], [\frac{3}{2}, 2], [2, 3], [3, 5]$ or $[5, \infty)$. Entries $(\chi_1, \gamma_1, 0)$, $(\chi_1, \gamma', 0)$, $(\bar{\chi}_1, -\gamma_1, 0)$, $(\bar{\chi}_1, -\gamma', 0)$, with kept sets $\{\rho_1, \rho'\}$, $\{\rho', \rho_1\}$ and their conjugates (a zero is kept only if it lies in the disc of the entry). No zero to the right. Then

$$r_{\gamma_1} = F(b - s_1) + R_p - \frac{1}{6}f(0), \quad r_{\gamma'} = F(\text{hi}' - s_1) + R_q - \frac{1}{6}f(0),$$

where $R_p \leq \inf \operatorname{Re} F(x + iy)$ over $x \in [p^* - s_1, \text{hi}' - s_1]$ and $R_q \leq \inf \operatorname{Re} F(x + iy)$ over $x \in [a - s_1, b - s_1]$, with $|y|$ in the piece. Here $x \geq 0$ (as $s_1 \leq a \leq p^*$), so $\operatorname{Re} F(x + iy) \geq 0$ by Condition 2 and we take $R_p, R_q \geq 0$; the multiplicities of the kept zeros can then only increase the sums. On the last piece $R_p = R_q = 0$, which also covers a zero outside the disc (its term is then absent). Each same-character pair is a controlled pair with excess $(c^{\text{hi}} - d_1)_+$, $c^{\text{hi}} \geq \sup \operatorname{Re} G(-s_1 + iy)$, so every diagonal numerator is $D(0) - d_1 + (c^{\text{hi}} - d_1)_+$; an entry of χ_1 and an entry of $\bar{\chi}_1$ form a pair of distinct characters, with the nonprincipal quotient χ_1^2 , and carry no mutual excess. If $\rho' = \rho_1$ is a double zero, the two entries coincide and $r \geq m_{\rho_1} F(\lambda_1 - s_1) - \frac{1}{6}f(0)$ with $m_{\rho_1} \geq 2$, which is at least both r_{γ_1} and $r_{\gamma'}$ (here $y = 0$ and $\lambda' = \lambda_1$).

- (e) *Shifted first family* (the shifted row, $s = \min(1.9, p^*, l_2)$, used only if $s > s_1$ and $s \geq a$). Entries $(\chi_1, \gamma_1, 0)$ and, for type complex, $(\bar{\chi}_1, -\gamma_1, 0)$, with anchor s . The only zeros of χ_1 with parameter below s are ρ_1 and, for type rc, $\bar{\rho}_1$, since every other occurrence has parameter at least $\lambda' \geq p^* \geq s$. With $\mathcal{S} = \{\rho_1\}$ (and $\bar{\rho}_1$ for type rc) nothing remains in the hypotheses of [Lemma 8.7](#). As conjugate zeros of a real character have the same multiplicity m ,

$$\sum_{\rho \in \mathcal{S}} m_\rho \operatorname{Re} F(\cdot) = m(F(\lambda_1 - s) + \mathbf{1}_{\text{rc}} \operatorname{Re} F(\lambda_1 - s + 2i\mu_1)) \geq m(F(b - s) - \mathbf{1}_{\text{rc}} C_Z),$$

with $C_Z = \mathcal{C}_c(s - a)$ from [Lemma 9.1\(v\)](#) (0 if $s \leq a$). If $F(b - s) - \mathbf{1}_{\text{rc}} C_Z \geq 0$ this is at least $F(b - s) - \mathbf{1}_{\text{rc}} C_Z$, and we take $r = F(b - s) - \mathbf{1}_{\text{rc}} C_Z - \frac{1}{6}f(0)$. Otherwise the proposed bound $F(b - s) - C_Z - \frac{1}{6}f(0)$ is negative and the entry receives the feature 0; this is admissible whatever the true response bound r_j is, by the last clause of [Corollary 8.10](#) (with $v'_j = 0$). The diagonal numerator is $D(0) - d_1$.

- (f) *Reserved family.* Entries $(\chi, \gamma_\chi, s - \sigma_2)$ for its n_2 characters at their height-one representatives, with $\sigma_2 = \min(\max(a, \min(l_2, lo_2)), s)$. Every zero of these characters in $R(l)$ has parameter at least $\max(\lambda_1, \lambda_2) \geq \max(a, l_2) \geq \sigma_2$, so there is no zero to the right, $r = F(hi_2 - \sigma_2) - \frac{1}{6}f(0)$, and the diagonal numerator is $D(s - \sigma_2) - d_1$. With second-family columns, the column $[lo_{2,j}, hi_{2,j}]$ uses $\sigma = \min(\max(a, \min(l_2, lo_{2,j})), s)$, $r = F(hi_{2,j} - \sigma) - \frac{1}{6}f(0)$ and the diagonal numerator $D(s - \sigma) - d_1$.
- (g) *Outside first family* (family row, $s = s_1$, $\Xi = 0$). Global entries $(\chi_1, \gamma_1, 0)$ with $\mathcal{S} = \{\rho_1\}$ and, for type complex, $(\bar{\chi}_1, -\gamma_1, 0)$ with $\mathcal{S} = \{\bar{\rho}_1\}$ (for type rc there is the single global entry $(\chi_1, \gamma_1, 0)$); $r = F(b - s_1) - \frac{1}{6}f(0)$. And a hidden entry $(\chi, \gamma_h, 0)$ for each first-family character χ with a zero in R_P , at its zero ρ_h of least parameter $t_\chi \in [lo_h, hi_h]$, with $\mathcal{S} = \{\rho_h\}$ and $r = F(hi_h - s_1) - \frac{1}{6}f(0)$. No zero to the right. Entries of the same character have normalized height differences at least $C_B - C_P > M$, so their pair excess vanishes, and every diagonal numerator is $D(0) - d_1$.

9.5. The rows. Each leaf program has one or more rows, of the kinds listed in Table 8. In the family row of an inside leaf the first family is entered by (b), by (c) for type rc (with the height split of Section 11.6), or by (d) in a piece of the second-zero split. In the shifted row every ordinary anchor is at most $s \leq l_2$. In the graded rows all characters are distinct.

Row	Reference anchor s	Sieve weights	Entries, inside	Entries, outside
family	s_1	none ($\Xi = 0$)	(a), (f) and (b), (c) or (d)	(a), (f), (g)
shifted	$\min(1.9, p^*, l_2)$	none ($\Xi = 0$)	(a), (e), (f)	not used
graded	1.5, 1.6, 1.9, 2.1 or 2.3	present	(a), (b), (f)	(a), (f)
two-test	the shift of the leaf	two test functions	Theorem 9.4, on 37 roots	not used

TABLE 8. The kinds of near rows. The shifted row is used only if $s > s_1$ and $s \geq a$.

In every family, shifted and graded row the safe-anchor hypothesis holds by (Z1) and Input 3.5, as explained after Definition 8.3, and the characters are pairwise distinct whenever $\Xi \neq 0$. All the entries have heights at most l . The first family is entered once per character in a row with sieve weights (for type rc only ρ_1 is kept); the pair and second-zero entries occur only in rows with $\Xi = 0$.

9.6. The two-test row. On 37 inside roots the graded rows alone do not certify, and the leaf program contains in addition the following near row. It is not an instance of Theorem 8.4: its Gram form is at the shift s of the leaf, which may exceed λ_1 , and it pays explicit corrections for the quotient characters $\chi_1, \bar{\chi}_1$. Its proof follows Xylouris's proof of his Lemma 5.3 [145, (5.34)–(5.40), pp. 72–76], with two test functions.

Fix $\gamma_G > \gamma_Z > 0$ and let $f = f_{\gamma_Z}$ and $g = f_{\gamma_G}$ as in (7.1), so that $g > 0$ on the support of f . Let $s \leq \min(p^*, l_2)$ be the shift of the leaf, and put

$$R_G = G(-s), \quad R_Z = \int_0^{2\gamma_Z} \frac{f(t)^2}{g(t)} e^{st} dt, \quad \mathcal{N} = \sqrt{R_G R_Z}, \quad d_0 = \frac{g(0)}{6R_G},$$

$$C_G = \sup_{\operatorname{Re} z \geq a-s} (-\operatorname{Re} G(z)), \quad C_F = \sup_{\operatorname{Re} z \geq a-s} (-\operatorname{Re} F(z)), \quad c_G = \frac{C_G}{R_G}, \quad k_\eta = 1 + \eta.$$

The function f^2/g , extended by 0, satisfies Condition 1, since it vanishes to order 6 at $2\gamma_Z$. Define the features

$$v(\lambda) = \frac{F(\lambda - s) - \frac{1}{6}f(0)}{\mathcal{N}} - \eta, \quad v_2 = \frac{F(\text{hi}_2 - s) - \frac{\phi_2}{2}f(0)}{\mathcal{N}} - \eta,$$

$$v_f = \frac{F(b - s) - \frac{\phi_1}{2}f(0) - \mathbf{1}_{\text{rc}}C_F}{\mathcal{N}} - \eta,$$

with $\phi_1 = \frac{1}{4}$ for a real χ_1 and $\phi_1 = \frac{1}{3}$ otherwise, $\phi_2 = \frac{1}{4}$ if $n_2 = 1$ and $\phi_2 = \frac{1}{3}$ otherwise, and $\mathbf{1}_{\text{rc}}$ the indicator of type rc; the correlation term $d = k_\eta(d_0 + \eta)$; and the diagonals

type	D_o/k_η	D_f/k_η
rr	$1 - d_0 + c_G$	$1 - d_0$
rc	$1 - d_0 + 2c_G$	$1 - d_0$
complex	$1 - d_0 + 2c_G$	$1 - d_0 + c_G$

Theorem 9.4 (Two-test inequality). *Use the two-test data defined above, with $\gamma_G > \gamma_Z > 0$ and $s \leq \min(p^*, l_2)$. For every sufficiently large q and every configuration of an inside leaf, form the following entries:*

- (i) any finite set of distinct ordinary characters, each at a zero $\rho_\chi \in R(l)$ with $|\text{Im } \rho_\chi| \leq 1$;
- (ii) the first character at ρ_1 , and also $\bar{\chi}_1$ at $\bar{\rho}_1$ when χ_1 is nonreal;
- (iii) the n_2 characters of the reserved family at their height-one representatives, if a family is reserved.

For type rc, the single first-family entry retains both ρ_1 and $\bar{\rho}_1$. Assign features $v(\lambda_{\rho_\chi})$, v_f , and v_2 to groups (i)–(iii), and diagonals D_o , D_f , and D_o , respectively. Then for every choice of real coefficients $a_j \geq 0$,

$$\left(\sum_j a_j v_j \right)_+^2 \leq \sum_j D_j a_j^2 + d \left(\sum_j a_j \right)^2.$$

Proof. Put $\sigma_s = 1 - s/\mathcal{L}$ and, in the Hilbert space of sequences indexed by n ,

$$A_j(n) = \sqrt{\Lambda(n)n^{-\sigma_s}g(t_n)} \chi_j(n) n^{-i\gamma_j}, \quad B(n) = \chi_0(n) \sqrt{\Lambda(n)n^{-\sigma_s}} \frac{f(t_n)}{\sqrt{g(t_n)}}.$$

Then $-\text{Re}\langle A_j, B \rangle = \mathcal{L}R_j$, where R_j is the response of the entry at the anchor s , and $\mathcal{L} \sum_j a_j R_j \leq \|\sum_j a_j A_j\| \|B\|$.

Responses. We use Lemma 8.7 (or directly Input 3.4). For an ordinary or reserved character every zero in the disc has parameter at least s (Lemma 9.2(b), as $s \leq l_2$), so $R_j \geq F(\lambda_{\rho_\chi} - s) - \frac{\phi}{2}f(0) - \varepsilon$. For the first family, the zeros of χ_1 with parameter below s are ρ_1 and, for type rc, $\bar{\rho}_1$ (the others have parameter at least $\lambda' \geq p^* \geq s$); keeping them gives $F(\lambda_1 - s) \geq F(b - s)$ and, for type rc, $\text{Re } F(\lambda_1 - s + 2i\mu_1) \geq -C_F$ (Lemma 9.1(v)). Conjugate zeros of a real character have the same multiplicity $m \geq 1$, so the kept terms total at least $m(F(b - s) - \mathbf{1}_{\text{rc}}C_F)$, which is at least $F(b - s) - \mathbf{1}_{\text{rc}}C_F$ when this is nonnegative (otherwise $v_f < 0$, see below). Since $\varphi(\chi) = \frac{1}{4}$ for real χ and $\varphi(\chi) \leq \frac{1}{3}$ always (Input 3.4), the terms $\frac{\phi}{2}f(0)$ are at most $\frac{1}{6}f(0)$, $\frac{\phi_1}{2}f(0)$ and $\frac{\phi_2}{2}f(0)$ respectively. Thus $R_j \geq \mathcal{N}(v_j + \eta) - \varepsilon$ for entries with positive feature. Entries with nonpositive feature may be omitted in estimating the positive part on the left; we justify restoring their coefficients to the right-hand side below.

Norms. By Input 3.3 applied to f^2/g , $\|B\|^2 = \mathcal{L}R_Z(1 + o(1))$, and $\|A_j\|^2 = \mathcal{L}R_G(1 + o(1))$.

Off-diagonal terms. For $j \neq k$ the characters are distinct and $\langle A_j, A_k \rangle$ is the g -weighted sum for the nonprincipal character $\chi_{jk} = \chi_j \bar{\chi}_k$ at $\sigma_s + i(\gamma_j - \gamma_k)$. By Input 3.4 its real part is at most $\mathcal{L}(d_0 R_G + \varepsilon)$ plus $\mathcal{L}$ times the sum of $-\text{Re } G(\cdot)$ over the zeros of χ_{jk} in the disc with parameter below s . If $\chi_{jk} \notin \{\chi_1, \bar{\chi}_1\}$ there are none, by Lemma 9.2(b), which applies because

$|\gamma_j - \gamma_k| \leq 2 \leq 2l$. If $\chi_{jk} = \chi_1$ (or $\bar{\chi}_1$), they are ρ_1 (resp. $\bar{\rho}_1$), and for type rc both, each contributing at most C_G ; these zeros are simple, since a multiple ρ_1 would give $\lambda_1 = \lambda' \geq s$. For a given j , the k with $\chi_j \bar{\chi}_k \in \{\chi_1, \bar{\chi}_1\}$ are $\chi_k \in \{\chi_j \bar{\chi}_1, \chi_j \chi_1\}$: at most two, and at most one if χ_1 is real; a first-family character has none if χ_1 is real ($\chi_1 \bar{\chi}_1 = \chi_0$) and at most one (χ_1^2) otherwise. With $2a_j a_k \leq a_j^2 + a_k^2$ for each such pair, these corrections give the diagonals of the table.

Assembly. For $a \geq 0$, $\|\sum_j a_j A_j\|^2 \leq \mathcal{L}R_G[\sum_j a_j^2(1 - d_0 + \text{exc}_j) + (d_0 + \varepsilon)(\sum_j a_j)^2]$, and $\|B\|^2 \leq (1 + \varepsilon)\mathcal{L}R_Z$. Dividing by $\mathcal{L}^2\mathcal{N}^2$ and absorbing the errors in η gives the claim for the positive-feature entries. To restore any omitted entries, expand the right-hand side as $\sum_j (D_j + d)a_j^2 + 2d\sum_{j < k} a_j a_k$. Its coefficients are nonnegative: the table gives $D_j + d \geq k_\eta(1 + \eta) > 0$ and $d > 0$. Restoring nonnegative a_j therefore increases this bound, while a nonpositive feature can only decrease the positive part on the left. $\square$

Proposition 9.5 (Detector mixture). *Theorem 9.4 remains true when f is replaced by $f_\epsilon = f + \epsilon g$ with $\epsilon \geq 0$, provided F , $f(0)$ and C_F are replaced by $F + \epsilon G$, $f(0) + \epsilon g(0)$ and $C_F + \epsilon C_G$, and $\mathcal{N}$ by $\mathcal{N}_\epsilon = \sqrt{R_G(R_Z + 2\epsilon F(-s) + \epsilon^2 R_G)}$; the Gram quantities are unchanged.*

Proof. f_ϵ satisfies Conditions 1 and 2, and $f_\epsilon^2/g = f^2/g + 2\epsilon f + \epsilon^2 g$ is a sum of functions satisfying Condition 1, so $\|B_\epsilon\|^2 = \mathcal{L}(R_Z + 2\epsilon F(-s) + \epsilon^2 R_G)(1 + o(1))$. The rest of the proof is unchanged, using $-\text{Re}(F + \epsilon G) \leq C_F + \epsilon C_G$ on $\text{Re } z \geq a - s$. $\square$

Proposition 9.6 (Paired first family). *Use the data and inside-leaf hypotheses of Theorem 9.4. Suppose χ_1 is nonreal and the specification $\mathfrak{s}$ has a finite gap $\lambda' \in [p^*, h]$, where $h = \text{hi}'$. Fix $\zeta > 0$ and a number $\mathcal{B}$ with*

$$(9.1) \quad \frac{\text{Re } G(-s + i\Delta)}{R_G} - \zeta \frac{\text{Re } F(\lambda_1 - s + i\Delta) + \text{Re } F(\lambda' - s + i\Delta)}{2\mathcal{N}} \leq \mathcal{B}$$

for all real Δ , $\lambda_1 \in [a, b]$ and $\lambda' \in [p^*, h]$. Let $e \geq C_F/(2\mathcal{N})$, and put

$$m = \frac{F(b-s) + F(h-s)}{2\mathcal{N}} - \frac{f(0)}{6\mathcal{N}} - \eta - e, \quad D_* = k_\eta \left\{ \frac{1 + \mathcal{B}}{2} - d_0 + c_G - \frac{\zeta e}{2} \right\}, \quad \zeta_* = \frac{k_\eta \zeta}{2}.$$

Let $m' \leq m$ and $D' \geq D_*$ with $D' > 0$ and $2D' \geq \zeta_* \max(m', 0)$. Then, for every configuration of the leaf, there is $\tau \in [0, \sqrt{d}]$ satisfying (T) for the entries of Theorem 9.4, with the first-family feature m' and diagonal D' in place of v_f and D_f .

Proof. For χ_1 use the averaged vector $A_f = \frac{1}{2}(A_{\chi_1, \gamma_1} + A_{\chi_1, \gamma'})$, with an admissible ρ' of height γ' (Section 2; so ρ' is a zero of χ_1 and $|\gamma'| \leq l$), and its conjugate for $\bar{\chi}_1$. Its response is the average of the two responses, which we bound by (7.2) at the points $1 - s/\mathcal{L} + i\gamma_1$ and $1 - s/\mathcal{L} + i\gamma'$; they lie in $R(9l)$, and (7.2) does not require the kept zeros to lie in a disc. The set A_1 of zeros of χ_1 in $R(l)$ with parameter below s is contained in $\{\rho_1\}$, since every occurrence in $R(l)$ other than the distinguished one has parameter at least $\lambda' \geq p^* \geq s$. We put ρ' into A_2 , and also ρ_1 if $\lambda_1 \geq s$, with their multiplicities (if $\rho' = \rho_1$ this is ρ_1 with multiplicity at least 2); by Input 3.9 their number is bounded independently of q . The multiplicities can only increase the response: the terms of ρ' are ≥ 0 by Condition 2 because $\lambda' \geq s$, the term $F(\lambda_1 - s)$ is positive, and if ρ_1 is a multiple zero then $\lambda_1 = \lambda' \geq s$, so its terms are ≥ 0 as well. With $\Delta = \mathcal{L}(\gamma_1 - \gamma')$, $P = \text{Re } G(-s + i\Delta)/R_G$ and Q the second fraction in (9.1), the feature of A_f is therefore at least $\frac{F(b-s) + F(h-s)}{2\mathcal{N}} - \frac{f(0)}{6\mathcal{N}} + Q - \eta$, and $\|A_f\|^2 = \mathcal{L}R_G \frac{1+P}{2}(1 + o(1))$ by Input 3.3. Each off-diagonal term $\langle A_f, A_k \rangle$ is the average of two terms of the kind estimated in the proof of Theorem 9.4, at the heights $\gamma_1 - \gamma_k$ and $\gamma' - \gamma_k$, whose absolute values are at most $l + 1 \leq 2l$; the term between A_f and its conjugate is an average of four terms, at heights $2\gamma_1$, $\gamma_1 + \gamma'$ and $2\gamma'$, which are also at most $2l$ in absolute value. So Lemma 9.2 applies to all of them, averaging creates no new quotient characters, and the off-diagonal estimates are unchanged (for a cubic χ_1 the quotient

$\chi_1^2 = \bar{\chi}_1$ of the conjugate pair is paid by the c_G in D_*). Put $U = Q + e$; then $U \geq 0$, since $\lambda' \geq s$ gives $Q \geq -C_F/(2\mathcal{N})$.

By (9.1), $P \leq \mathcal{B} + \zeta Q$, so the actual first-family feature is at least $m + U$ and the actual diagonal $k_\eta\{\frac{1+P}{2} - d_0 + c_G\}$ is at most $D_* + \zeta_*U$. Put $U' = U + (m - m') \geq U \geq 0$. Then the actual feature is at least $m' + U'$, and the actual diagonal is at most $D' + \zeta_*U'$, which is positive. Replacing the first-family features and diagonals by these bounds preserves the quadratic inequality of Theorem 9.4, and Proposition 8.8 gives $\tau \in [0, \sqrt{d}]$ satisfying (T) with the first-family data $(m' + U', D' + \zeta_*U')$. Finally, for $\tau \geq 0$ and $U' \geq 0$,

$$\frac{(m' + U' - \tau)_+^2}{D' + \zeta_*U'} \geq \frac{(m' - \tau)_+^2}{D'}.$$

Indeed the right side is 0 if $m' \leq \tau$. Otherwise put $z = m' - \tau \in (0, m']$. The derivative of $(z+u)^2/(D' + \zeta_*u)$ in $u \geq 0$ is $(z+u)(2D' - \zeta_*z + \zeta_*u)/(D' + \zeta_*u)^2 \geq 0$, because $2D' \geq \zeta_*m' \geq \zeta_*z$. So (T) holds at the same τ with (m', D') . $\square$

In the leaf program the two-test row is the threshold form (T) of Theorem 9.4 (or of one of its two variants, Propositions 9.5 and 9.6):

$$\sum_i x_i \frac{(V_i - \tau)_+^2}{D_o} + n \frac{(v_f - \tau)_+^2}{D_f} + n_2 \frac{(v_2 - \tau)_+^2}{D_o} \leq 1 - \frac{\tau^2}{d},$$

with $V_i = v(\text{hi}_i)$ the feature of the right end of the ordinary bin i (the tail and the reserved columns have feature 0 in this row), and negative features replaced by 0. The T^* -representatives are admissible entries, since Theorem 9.4 allows any zero of height at most 1, and all of them have parameter at least $l_2 \geq s$ (Z3). The reserved family appears as a fixed term with the feature v_2 at hi_2 , which bounds its feature over the whole reserved range, and with the diagonal D_o . The row is used on 37 roots, once on each: with a single test on 31 of them, with a mixture on 2, and in the paired form on 4 (all with a nonreal χ_1 , as Proposition 9.6 requires). In the paired form the certified pair (m', D') consists of a lower bound for m and an upper bound for D_* , and the condition $2D' \geq \zeta_* \max(m', 0)$ is checked for this pair; the margins $2D' - \zeta_*m'$ are between 0.18 and 0.30. The condition (9.1) is verified with interval arithmetic. Both of its terms are even in Δ , so $\Delta \geq 0$ suffices. On the grid $\Delta \in \frac{1}{200}\mathbb{Z} \cap [0, 30]$ the left side is enclosed at the four corners of the parameter box $[a, b] \times [p^*, h]$. As the left side is a sum of a function of λ_1 and a function of λ' , its maximum over the box exceeds the maximum over the corners by at most $\frac{\zeta}{2\mathcal{N}}(\frac{(b-a)^2}{8} \int t^2 f(t) e^{(s-a)+t} dt + \frac{(h-p^*)^2}{8} \int t^2 f(t) dt)$, the linear interpolation error in each variable, and this is added. Between grid points a second-derivative bound in Δ is used, and for $\Delta \geq 30$ a closed-form bound from Lemma 9.1(iv). The value $\mathcal{B}$ used is at least 0.

9.7. Validity of the rows.

Proposition 9.7 (Every configuration satisfies the near rows). *Let $\mathfrak{s}$ specify a leaf or one of its subcases. For sufficiently large q , take any configuration in $C(\mathfrak{s})$ and assign its characters to columns as in Proposition 11.6, obtaining column values x . Then, for every near row k , there is a threshold τ_k such that*

$$0 \leq \tau_k \leq \sqrt{d_k/S}, \quad \Phi_k(x, \tau_k) \leq 1 - \frac{\tau_k^2}{d_k/S}.$$

Here d_k is the stored correlation term, $S = 10^{16}$, and Φ_k is the row expression of Definition 10.1. Each row may have its own threshold.

Proof. For a graded, family or shifted row, take as entries the characters of the configuration that are placed in the columns and family terms of the row: the ordinary characters of O

at their T^* -representatives, the reserved characters, and the first family, with the entries of Section 9.4. These form an admissible family of entries for Theorem 8.4, with the kept sets verified in Section 9.4, so Corollary 8.10 gives a threshold τ at which (T) holds with the normalized features, diagonals and correlation term. Each column's feature is at most the feature of every character placed in it (features are taken at the right end of the column, and F is decreasing), or is 0; each column's diagonal is at least theirs, or the column has feature 0 (an entry whose diagonal numerator is not safely positive is omitted, and its column is given feature 0), in which case it contributes nothing for $\tau \geq 0$; each family term (n, v, D) receives n entries with feature at least v and diagonal at most D ; and the characters in the tail column and in the hidden tail are simply omitted from the row. Since each term $(v - \tau)_+^2/D$ increases with v and decreases with D for $\tau \geq 0$, the row of the leaf holds at τ . For the two-test row the same argument applies to Theorem 9.4 and its variants, through Proposition 8.8. $\square$

The numbers I_u , D_u , $D(\delta)$, the transforms and the special terms R_{lo} , c^{hi} , R_p , R_q , C_Z and C_F are computed from the rational parameters of each row with outward interval arithmetic. The transforms are enclosed with the closed form of Lemma 9.1(iv) or with its power series near 0. The infima and suprema over heights in R_{lo} , c^{hi} , R_p and R_q are enclosed on grids of step at most $\frac{1}{50}$, with second-derivative bounds $\int t^2 \hat{f}_c(t) e^{dt} dt$ between grid points and Lipschitz bounds in the real parts, and by the tail bounds of Lemma 9.1(iv) beyond the grid. The constants C_Z , C_F and C_G are enclosed on the line $\text{Re } z = -d$ (by Lemma 9.1(v)), for heights in $[0, 20]$ by adaptive linear interpolation, starting from 200 cells of width $\frac{1}{10}$ and bisecting cells until the largest cell bound is within a small fixed tolerance of the best value found, with the same second-derivative bound, and beyond height 20 by a closed-form tail bound. The features, diagonals and correlation terms of all rows other than the 37 two-test rows have been recomputed from their rational parameters by an independent program whose soundness is formally verified (Section 14).

10. LEAF PROGRAMS AND EXACT CERTIFICATES

For each case of the zero analysis, we construct a family of linear programs indexed by the thresholds in the near rows. Their common objective bounds the zero sum. We cover the threshold domain by finitely many boxes and give a linear relaxation on each box. An exact certificate then bounds every feasible point, without having to find an optimal one.

The argument uses weak linear-programming duality [25, 115] with outward rounding of the data; compare [63, 95, 1]. This section is independent of the analytic number theory. Its conclusion is the certificate soundness theorem, Theorem 10.7, which is formally verified (Section 14.3).

10.1. Leaf data. The coefficients, budgets, and objective constants of a leaf are stored as integers at scale

$$S = 10^{16}.$$

Thus a stored coefficient v represents v/S . Character counts such as n , n_g , and n_E are ordinary unscaled integers. The variables x_c and thresholds τ_k are also unscaled real numbers.

A leaf has a finite set of *columns* c and a finite list of *near rows* $k = 1, \dots, K$. A column c carries

- an objective coefficient G_c and a far weight W_c ;
- a count coefficient C_c , a hidden-count coefficient N_c and a second-family indicator E_c (stored at scale S : $C_c, E_c \in \{0, S\}$, and N_c is 0, S or, for the hidden tail, $\lfloor S/w(\text{end}) \rfloor$, with end as in Section 11.5);
- for each row k , a feature v_{ck} and a diagonal $D_{ck} > 0$.

A row k carries a correlation term $d_k > 0$ and finitely many *family terms* (n, v, D) , where $n \geq 0$ is a count, v a feature and $D > 0$ a diagonal. The leaf also has a far budget F , counts n_g (the number n of first-family characters, in outside leaves) and n_E (the number of reserved characters carried by second-family columns: the size of the reserved family if it is charged through columns, and 0 otherwise), and two constants: first, which bounds the cost of the first family of an inside leaf (0 for an outside leaf) plus any fixed charge for the reserved family, and final, the error allowance, with $\text{final}/S = 5\eta$ (Section 11.5).

Definition 10.1 (Leaf program). For real column values $x = (x_c)$ and real thresholds $\tau = (\tau_k)$, put

$$\Phi_k(x, \tau_k) = \sum_c x_c \frac{(v_{ck}/S - \tau_k)_+^2}{D_{ck}/S} + \sum_{(n,v,D) \in \text{fam}(k)} n \frac{(v/S - \tau_k)_+^2}{D/S}.$$

The pair (x, τ) is *feasible* if all of the following hold:

- $x_c \geq 0$ for every c ;
- (far) $\sum_c W_c x_c \leq F$;
- (count) $\sum_c C_c x_c \leq 2S$;
- (hidden count) $\sum_c N_c x_c \leq n_g S$;
- (second family) $\sum_c E_c x_c = n_E S$;
- (near rows) for every k : $0 \leq \tau_k$, $\tau_k^2 \leq d_k/S$ and $\Phi_k(x, \tau_k) \leq 1 - \tau_k^2/(d_k/S)$.

The *value* of x is

$$\mathcal{V}(x) = \frac{1}{S} \left(\text{first} + \text{final} + \sum_c G_c x_c \right).$$

For fixed τ , these are linear constraints on x . Allowing τ to vary gives the family of programs to be certified. The variables x_c are real and nonnegative: integer character counts are feasible values, but the relaxation also permits fractional counts and the weighted tail masses defined in Section 11.5.

In Section 11 we show that for every configuration of zeros the actual characters give a feasible pair whose value bounds the zero sum W of Section 4. It is therefore enough to show that every feasible pair has value below 1.

10.2. Relaxations of a near row. Fix a row k and write $d = d_k/S$. The near-row constraint says $Q_k(\tau) \leq 1$, where

$$Q_k(\tau) = \sum_j c_j (u_j - \tau)_+^2 + \frac{\tau^2}{d}$$

with $c_j \geq 0$ collecting the coefficients $x_c/(D_{ck}/S)$ and $n/(D/S)$ and u_j the features. The function Q_k is convex and continuously differentiable, with $Q'_k(\tau) = -2 \sum_j c_j (u_j - \tau)_+ + 2\tau/d$. In the next two lemmas, $d > 0$, all $c_j \geq 0$, and $0 \leq a \leq b$.

Lemma 10.2 (First-order relaxation). *If $a \leq \tau \leq b$ and $Q_k(\tau) \leq 1$, then*

$$\sum_j c_j (u_j - b)_+^2 \leq 1 - \frac{a^2}{d}.$$

Proof. Each $(u_j - \tau)_+^2 \geq (u_j - b)_+^2$ since $\tau \leq b$, and $\tau^2 \geq a^2$ since $0 \leq a \leq \tau$. □

Lemma 10.3 (Tangent relaxation). *Let $m = (a + b)/2$ and $h = (b - a)/2$. If $a \leq \tau \leq b$ and $Q_k(\tau) \leq 1$, then at least one of the following holds:*

$$\begin{aligned} (\text{case } 0) \quad & \sum_j c_j \left(((u_j - m)_+ + h)^2 - h^2 \right) \leq 1 - \frac{(m - h)^2 - h^2}{d}, \\ (\text{case } 1) \quad & \sum_j c_j \tilde{\theta}_j \leq 1 - \frac{(m + h)^2 - h^2}{d}, \quad \tilde{\theta}_j = \begin{cases} ((u_j - m) - h)^2 - h^2, & u_j > m, \\ 0, & u_j \leq m. \end{cases} \end{aligned}$$

Proof. By convexity, $1 \geq Q_k(\tau) \geq Q_k(m) + Q'_k(m)(\tau - m) \geq Q_k(m) - h|Q'_k(m)|$, so $Q_k(m) - hQ'_k(m) \leq 1$ or $Q_k(m) + hQ'_k(m) \leq 1$. Expanding with $(u - m)_+^2 \pm 2h(u - m)_+ = ((u - m)_+ \pm h)^2 - h^2$ and $m^2 \mp 2hm = (m \mp h)^2 - h^2$ gives the two cases. In case 1 a term with $u_j \leq m$ is 0. $\square$

The left-hand sides are linear in the c_j , hence in the column values x . In case 1 the coefficients $\tilde{\theta}_j$ can be negative.

10.3. Certificates. Thresholds are measured in ticks of $1/T_S$, where $T_S = 10^6$, and dual multipliers are scaled by $D_S = 10^{12}$. In this subsection $[a, b]$ denotes an interval of thresholds in ticks, not a first-zero cell, and m, h, p and β are local to the relaxation formulas.

Definition 10.4 (Root threshold box). For row k , put

$$e_k = \left\lceil \sqrt{\lfloor d_k T_S^2 / S \rfloor + 1} \right\rceil.$$

The *root box* is $\prod_k [0, e_k]$ in integer-tick coordinates. Its corresponding domain for real thresholds is $\prod_k [0, e_k / T_S]$.

Since $e_k^2 S \geq d_k T_S^2$, every threshold with $\tau_k^2 \leq d_k / S$ lies in $[0, e_k / T_S]$.

Definition 10.5 (Integer costs and budgets). Let $[a, b]$ be an interval of integer ticks with $0 \leq a < b$, and let v and $D > 0$ be a stored feature and diagonal. The following costs are coefficients of column variables; the family terms are subtracted from the row budget.

(a) (First order.) The cost is $\lfloor (v - bS/T_S)_+^2 / D \rfloor$. The budget of row k is

$$S - \left\lfloor \frac{S^2 a^2}{T_S^2 d_k} \right\rfloor - \sum_{(n, v, D)} \left\lfloor \frac{n(v - bS/T_S)_+^2}{D} \right\rfloor.$$

(b) (Tangent, case $\iota \in \{0, 1\}$.) Put $m = (a + b)/(2T_S)$, $h = (b - a)/(2T_S)$ and $\bar{v} = v - mS$ (mS and $2hS$ are integers, since $2T_S$ divides S). The cost is $\lfloor S\Theta/D \rfloor$, where $\Theta = 0$ if $\bar{v} \leq 0$, and otherwise $\Theta = \lfloor \bar{v}(\bar{v} + 2hS)/S \rfloor$ for $\iota = 0$ and $\Theta = \lfloor \bar{v}(\bar{v} - 2hS)/S \rfloor$ for $\iota = 1$. The budget of row k is $\lceil S\beta \rceil$, where β is the exact rational number

$$\beta = 1 - \frac{p^2 - h^2}{d_k/S} - \sum_{(n, v, D)} n \frac{\Theta_\iota(v/S)}{D/S},$$

with $p = a/T_S$ for $\iota = 0$ and $p = b/T_S$ for $\iota = 1$, and $\Theta_0(u) = ((u - m)_+ + h)^2 - h^2$, $\Theta_1(u) = ((u - m) - h)^2 - h^2$ if $u > m$ and 0 otherwise.

The costs are the exact relaxation coefficients, multiplied by S and rounded *down*. The budgets are at least the exact right-hand sides multiplied by S : the tangent budget is rounded *up*, and in the first-order budget each subtracted term is rounded *down*.

Definition 10.6 (Certificate). A certificate for a leaf is a finite binary tree starting from the root box of Definition 10.4. An internal node splits one coordinate k of its box $[a, b]$ at an integer tick $a < \text{mid} < b$, into $[a, \text{mid}]$ and $[\text{mid}, b]$. At a terminal box, each row uses either the first-order relaxation or the two alternatives of the tangent relaxation. A *relaxation case* chooses one alternative for each tangent row. All combinations must be certified, because the alternative supplied by Lemma 10.3 can depend on the feasible point.

For consistency with the stored data, the mode is encoded by $o_k = 2$ for a first-order row and $o_k = 1$ for a tangent row. The case vector has $\iota_k = 2$ when $o_k = 2$, and $\iota_k \in \{0, 1\}$ when $o_k = 1$. In the stored trees the halves of a split are listed in the order $[a, \text{mid}]$, $[\text{mid}, b]$, the rows are numbered from 0, and the relaxation cases of a terminal box are listed in lexicographic order of $(\iota_1, \dots, \iota_K)$. For every relaxation case the node records either

- (i) an *exclusion*: a row k whose costs are all ≥ 0 and whose budget is < 0 ; or
- (ii) *integer duals* $Y_F, Y_C, Y_N, Y_E^+, Y_E^- \geq 0$ and $Z_1, \dots, Z_K \geq 0$ such that for every column c

$$(10.1) \quad D_S G_c \leq Y_F W_c + Y_C C_c + Y_N N_c + (Y_E^+ - Y_E^-) E_c + \sum_k Z_k \text{cost}_{ck}.$$

The value of the relaxation case is then

$$\left\lceil \frac{Y_F F + 2SY_C + n_g SY_N + (Y_E^+ - Y_E^-) n_E S + \sum_k Z_k \text{budget}_k}{D_S} \right\rceil + \text{first} + \text{final}.$$

The value of a terminal box is the maximum of -1 and the values of its relaxation cases that carry duals (so it is -1 if every relaxation case is excluded), and the certificate is *accepted* if every terminal box has value $< S$. The data must be *well formed*: all diagonals and correlation terms are positive, and all family counts are nonnegative.

Theorem 10.7 (Soundness of the integer certificates). *Let a leaf have positive diagonals and correlation terms and nonnegative family counts. If it has an accepted certificate in the sense of Definition 10.6, then every feasible pair (x, τ) of Definition 10.1 satisfies $\mathcal{V}(x) < 1$. This conclusion holds for all real nonnegative column values, including fractional ones.*

Proof. Let (x, τ) be feasible and put $t_k = \tau_k T_S$. Since $\tau_k^2 \leq d_k/S$ we have $t_k \leq e_k$, so t lies in the root box. At every internal node t lies in one of the two halves, so there is a terminal box $\prod_k [a_k, b_k]$ that contains t .

For each row, Lemma 10.2 (if $o_k = 2$) or Lemma 10.3 (if $o_k = 1$) gives a case ι_k in which the exact relaxed inequality holds at x . Take the relaxation case ι so obtained. Multiply the relaxed inequality of row k by S . Each coefficient of x_c is then at least cost_{ck} and the right side is at most budget_k , by the rounding directions in Definition 10.5 and since $x_c \geq 0$ and $n \geq 0$. Therefore

$$(10.2) \quad \sum_c x_c \text{cost}_{ck} \leq \text{budget}_k \quad (1 \leq k \leq K).$$

Here we used $S \cdot (v/S - b/T_S)_+^2 / (D/S) = (v - bS/T_S)_+^2 / D$ and the corresponding identities for the tangent terms.

If the relaxation case ι is excluded through row k , then the left side of (10.2) is ≥ 0 and the right side is < 0 , a contradiction. So ι carries duals. Multiply (10.1) by $x_c \geq 0$ and sum over c . Then use the far, count and hidden-count constraints, with multipliers $Y_F, Y_C, Y_N \geq 0$, the second-family equality, with the multiplier $Y_E^+ - Y_E^-$ of either sign, and (10.2), with multipliers $Z_k \geq 0$. This gives

$$D_S \sum_c G_c x_c \leq Y_F F + 2SY_C + n_g SY_N + (Y_E^+ - Y_E^-) n_E S + \sum_k Z_k \text{budget}_k.$$

Hence first + final + $\sum_c G_c x_c$ is at most the value of ι , which is at most the value of the terminal box, which is $< S$. $\square$

Remark 10.8. The formal proof of [Theorem 10.7](#) brought out that the family counts n must be nonnegative. With a negative count, the first-order relaxation can fail. The leaf builders already ensure $n \geq 0$, and the checker requires it.

11. THE CASE TREE

This section describes the finite case analysis of the middle range $0.1 \leq \lambda_1 < 1.5$ and proves that it covers every configuration of zeros ([Theorem 11.7](#)). The analytic facts it uses are the zero-location results of [Section 7](#), the costs of [Section 5](#), the far budget of [Section 6](#) and the near rows of [Section 9](#). There are two tasks: showing that the cases cover every possible configuration, and showing that a configuration in a case gives a feasible point of its program. These are proved separately before being combined in [Theorem 11.7](#). [Table 10](#), after [Section 11.4](#), lists the levels of the analysis, with their numbers and the places where they are defined and checked.

11.1. Configurations.

Definition 11.1 (Zero configuration). Fix a sufficiently large q . Its *zero configuration* is the multiset $\mathcal{Z}$ of pairs (χ, ρ) with $\chi \neq \chi_0$ a character modulo q , $\rho \in R(l)$ and $L(\rho, \chi) = 0$, each pair repeated according to the multiplicity of ρ .

Minima over empty sets are $+\infty$. In the middle range the first zero exists. The quantities of [Section 2](#) are read off from $\mathcal{Z}$: the first zero (χ_1, ρ_1) , the first family $\mathcal{F}_1$ with $n = |\mathcal{F}_1|$, the parameters λ_2, λ_3 and the family $\mathcal{F}_2$ ([Definition 2.1](#)), and the distinguished occurrences, λ' and an admissible ρ' ([Definition 2.2](#)); for type rc, $\mu_1 > 0$. In addition we use the following.

- The *height-one parameter* of χ is $\nu(\chi) = \min\{\lambda_\rho : (\chi, \rho) \in \mathcal{Z}, |\gamma_\rho| \leq 1\}$, and the T^* -parameter is $\nu_{T^*}(\chi)$, defined in the same way with $|\gamma_\rho| \leq T^*$ ([Definition 6.4](#)). Since $L(\rho, \chi) = 0$ if and only if $L(\bar{\rho}, \bar{\chi}) = 0$, both are functions of the family. We put $\nu_* = \min\{\nu(\mathcal{F}) : \mathcal{F} \neq \mathcal{F}_1\}$.
- The configuration is *inside* if $\rho_1 \in R_B$ and *outside* otherwise. Since $\lambda_1 < 1.5 < C_B$ in the middle range, outside means $|\mu_1| > C_B$. Type rr is always inside.

The definitions give $\lambda_1 \leq \lambda'$, $\lambda_1 \leq \lambda_2 \leq \lambda_3$, and $\nu_{T^*}(\chi) \geq \nu(\chi) \geq \lambda_2$ for $\chi \notin \mathcal{F}_1$.

11.2. Specifications. A specification records interval information about a configuration. It may describe no actual configuration; in that event the corresponding case can be excluded. Its two lower bounds l_2 and r have different scopes: l_2 applies throughout $R(l)$, whereas r applies only to the height-one representatives.

Definition 11.2 (Specification). A *specification* is a tuple

$$\mathfrak{s} = (\mathfrak{t}, [a, b], p, \text{gap}, l_2, r, \text{res}, \text{loc})$$

with the following data:

- a type $\mathfrak{t} \in \{\text{rr}, \text{rc}, \text{complex}\}$ and a first-zero cell $[a, b]$, where $a < b$;
- a lower bound p for λ' , and optionally an interval $\text{gap} = [\text{lo}', \text{hi}']$ with $p \leq \text{lo}' < \text{hi}' \leq \infty$, called a *gap* (used only for type complex);
- a global second-family lower bound l_2 and a height-one lower bound r , with $l_2 \leq r \leq 3$;
- either no reservation, or a pair $\text{res} = (\text{hi}_2, n_2)$ with $r < \text{hi}_2 \leq 2$ and $n_2 \in \{1, 2\}$; in the latter case put $\text{lo}_2 = r$;
- a height class $\text{loc} \in \{\text{inside}, \text{outside}\}$.

The *configuration set* $C(\mathfrak{s})$ consists of the configurations for which some admissible choices of the first zero and the minimizing families satisfy all of the following:

- (C1) the type is $\mathfrak{t}$ and $\lambda_1 \in [a, b]$;
- (C2) $\lambda' \geq p$, and $\lambda' \in [\text{lo}', \text{hi}']$ if a gap is present (with $\lambda' = \infty$ allowed when $\text{hi}' = \infty$);
- (C3) $\lambda_2 \geq l_2$;
- (C4) if res is none, then $\nu(\mathcal{F}) \geq r$ for every family $\mathcal{F} \neq \mathcal{F}_1$; if $\text{res} = (\text{hi}_2, n_2)$, then $\nu_* \in [r, \text{hi}_2]$ and some family $\mathcal{F}_{\text{res}} \neq \mathcal{F}_1$ with $\nu(\mathcal{F}_{\text{res}}) = \nu_*$ has exactly n_2 characters (so it is a real character if $n_2 = 1$ and a pair of conjugate nonreal characters if $n_2 = 2$);
- (C5) the configuration is inside if $\text{loc} = \text{inside}$ and outside if $\text{loc} = \text{outside}$.

In the reserved case, (C4) implies $\nu(\mathcal{F}) \geq r$ for every $\mathcal{F} \neq \mathcal{F}_1$. We call $\mathcal{F}_{\text{res}}$ the *reserved family* and its characters *reserved*; the characters outside $\mathcal{F}_1$ and $\mathcal{F}_{\text{res}}$ are *ordinary*.

In a formula containing lo' , that argument is omitted if no gap is present. Likewise, reserved-family terms are absent when there is no reservation.

The reserved family is a family with the least height-one parameter among the families other than the first; it is not assumed to be the family $\mathcal{F}_2$ that attains λ_2 , which may have its nearest zero at a height above 1. The following three facts are used in every leaf.

Lemma 11.3 (Consequences of reserving a family). *Let $\mathfrak{s}$ be a specification and fix a configuration and admissible choices witnessing its membership in $C(\mathfrak{s})$.*

- (i) (Count.) *At most two characters $\chi \notin \mathcal{F}_1$ have $\nu_{T^*}(\chi) < \lambda_3$.*
- (ii) (Drop.) *If $\mathfrak{s}$ is reserved, then $\nu_{T^*}(\chi) \geq \nu(\chi) \geq \lambda_3$ for every ordinary character χ .*
- (iii) (Cap.) *If $\mathfrak{s}$ is reserved, then $\lambda_2 \leq \nu_* \leq \text{hi}_2$.*

Proof. (i) Such a character has a zero in $R(l)$ with parameter below λ_3 , so it lies in $\mathcal{F}_2$, and $|\mathcal{F}_2| \leq 2$. (ii) Let $\mathcal{F}$ be the family of an ordinary character. If $\mathcal{F} \neq \mathcal{F}_2$, every zero of $\mathcal{F}$ in $R(l)$ has parameter at least λ_3 . If $\mathcal{F} = \mathcal{F}_2$, then $\mathcal{F}_{\text{res}} \notin \{\mathcal{F}_1, \mathcal{F}_2\}$, so $\nu(\mathcal{F}_{\text{res}}) \geq \lambda_3$, and $\nu(\mathcal{F}) \geq \nu_* = \nu(\mathcal{F}_{\text{res}})$ by minimality. (iii) The height-one representative of $\mathcal{F}_{\text{res}}$ is a zero in $R(l)$ of a character outside $\mathcal{F}_1$. $\square$

11.3. The source cover. The roots of the case tree are 2768 specifications, each used with the height class inside, and the 1685 of them whose type is not rr used also with the height class outside: 4453 roots in all. They are constructed from the zero-location results of [Section 7](#) in four steps.

- (1) *Parent rows.* There are 58 *parent rows* $(\mathfrak{t}, A_j, B_j, p_j, r_j)$: 33 of type rr tiling $[0.1, 1.5]$, 13 of type rc tiling $[0.628, 1.5]$, and 12 of type complex tiling $[0.44, 1.5]$. Each is the statement

$$(T2) \quad \text{type } \mathfrak{t}, \lambda_1 \in [A_j, B_j] \implies \lambda' \geq p_j \text{ and } \lambda_2 \geq r_j,$$

which is proved in [Proposition 7.5](#) from the tables of Heath-Brown and Xylouris.

- (2) *Base cells.* Each parent interval is cut into cells of width 0.01, giving 340 *base cells*.
- (3) *Cells and gaps.* The base cells are tiled by 478 first-zero cells in all. A cell $[\text{lo}, \text{hi}]$ of the parent j receives numbers $p \leq \max(p_j, \text{lo})$ and $l_2 \leq \max(r_j, \text{lo})$; this is valid since $\lambda', \lambda_2 \geq \lambda_1 \geq \text{lo}$. A cell of type complex may be *partitioned* by gaps $[p, m_1], [m_1, m_2], \dots, [m_k, \infty]$ for λ' . The cells that are not partitioned, and the gaps of those that are, are the 684 *gap cases*.
- (4) *Records.* Each gap case is either *unreserved*, with one specification with $r = l_2$ and no reservation, or *reserved* with a chain $l_2 = u_0 < u_1 < \dots < u_m = e \leq 2$, with specifications $(\text{res} = (u_{j+1}, n_2), r = u_j)$ for $n_2 = 1, 2$ and $0 \leq j < m$, followed by an unreserved *tail* specification with $r = e$.

By type there are 1083 specifications of type rr (195 unreserved, 444 reserved with $n_2 = 1$ and 444 with $n_2 = 2$), 115 of type rc (all unreserved and without gap) and 1570 of type complex.

Theorem 11.4 (Coverage by the root cases). *Every configuration with $0.1 \leq \lambda_1 < 1.5$ lies in $C(\mathfrak{s})$ for at least one of the 4453 root specifications $\mathfrak{s}$.*

Proof. Let $\mathfrak{t}$ be the type. By [Proposition 7.1](#), $\lambda_1 > 0.628$ if $\mathfrak{t} = \text{rc}$ and $\lambda_1 > 0.44$ if $\mathfrak{t} = \text{complex}$, so λ_1 lies in some cell of type $\mathfrak{t}$ (the cells are closed, and a boundary value lies in two cells). By (T2) for the parent of the cell, $\lambda' \geq p$ and $\lambda_2 \geq l_2$. If the cell is partitioned, some gap contains λ' . In the corresponding gap case, if it is unreserved, every family $\mathcal{F} \neq \mathcal{F}_1$ has $\nu(\mathcal{F}) \geq \lambda_2 \geq l_2 = r$. If it is reserved with chain $u_0 < \dots < u_m$, then $\nu_* \geq \lambda_2 \geq u_0$; if $\nu_* < e$ we choose j with $\nu_* \in [u_j, u_{j+1}]$ and a minimizing family, which is a real character or a nonreal pair, and the configuration lies in the corresponding record; otherwise it lies in the tail record. Finally the height class is determined by $|\mu_1|$. $\square$

11.4. Refinement trees. Each root is the root of a finite *refinement tree*. Every internal node carries a specification (computed from the root, never read from the data) and is of one of the following kinds. In each case every configuration of the node lies in the configuration set of one of its children, or the node has no configuration at all.

- *First-zero split* at $a < m < b$: children with cells $[a, m]$ and $[m, b]$, all other data kept. Exhaustive, since $\lambda_1 \leq m$ or $\lambda_1 \geq m$.
- *Second-family split* at $r < m < \text{hi}_2$ (reserved nodes): children ($\text{hi}_2 := m$) and ($r := m$, $\text{lo}_2 := m$). Exhaustive, since $\nu_* \leq m$ or $\nu_* \geq m$, and in the second case every family $\mathcal{F} \neq \mathcal{F}_1$ has $\nu(\mathcal{F}) \geq \nu_* \geq m$.
- *Gap split* at $\text{lo}' < m < \text{hi}'$ (type complex): children with gaps $[\text{lo}', m]$ and $[m, \text{hi}']$. Exhaustive.
- *Location update*: a lower bound $\lambda_2 > h$ valid on the whole cell ([Propositions 7.10](#) and [7.12](#); [Lemma 7.15](#)) replaces l_2 and r by $\max(l_2, h)$ and $\max(r, h)$, and excludes the node if it is reserved with $\text{hi}_2 \leq h$. Sound, since every family $\mathcal{F} \neq \mathcal{F}_1$ has $\nu(\mathcal{F}) \geq \lambda_2$.
- *Exclusion* of a reserved interval or a gap by a lower bound for λ_2 or λ' ([Propositions 7.11](#) and [7.13](#)), or of a whole node by the positivity argument of [Proposition 7.14](#).

Before a leaf model is built, all applicable location updates are applied once more to its specification; this only adds valid implications. None of these nodes depends on L . The terminal nodes are the *leaves*. The inside trees have 3146 leaves. Of these, 233 (in 199 roots) are stored in a separate file and are called *supplementary leaves*. Every leaf, supplementary or not, receives the same model. The outside trees have 1642 leaves. In all there are 4788 leaves; 343 inside roots and 70 outside roots have none, because all their configurations are excluded.

Node	Inside	Outside
first-zero / second-family / gap splits	1 / 408 / 593	0 / 27 / 0
location updates, real (restrict / exclude)	225 / 75	–
location updates, complex (restrict / exclude)	11 / 0	–
exclusions by lower bounds for λ_2 or λ'	352	70
exclusions by positivity	197	0
leaves	3146	1642

TABLE 9. Nodes of the refinement trees.

Level	Number	What it partitions or represents	Where defined	How it is checked
Parent rows	58	Published implications: type t and $\lambda_1 \in [A_j, B_j]$ give $\lambda' \geq p_j$, $\lambda_2 \geq r_j$ (33 rr, 13 rc, 12 complex)	Proposition 7.5	Printed tables; Heath-Brown's Tables 4 and 7 recomputed
Base cells	340	Parent intervals cut into cells of width 0.01 (145, 89, 106)	Section 11.3	Validated in the replay
First-zero cells	478	Tiling of the base cells (195, 115, 168), with bounds $p \leq \lambda'$ and $l_2 \leq \lambda_2$	Section 11.3	Validated in the replay
Gap cases	684	422 cells without a gap; 262 gaps for λ' partitioning 56 complex cells	Section 11.3	Validated in the replay
Specifications	2768	Gap cases with a second-family record: 1083 rr (195 unreserved, 444 + 444 reserved), 115 rc, 1570 complex	Definition 11.2	Validated in the replay
Root cases	4453	2768 specifications inside, and the 1685 not of type rr also outside	Theorem 11.4	Hand proof; replay
Refinement nodes	1959	1029 splits, 236 restricting location updates, 694 exclusions (75 by real location updates, 422 by lower bounds for λ_2 or λ' , 197 by positivity); $4453 + 1029 = 4788 + 694$	Section 11.4 , Table 9	Replay; some location rows by separate programs
Leaves	4788	3146 inside (233 supplementary) and 1642 outside; 343 inside and 70 outside roots have none	Section 11.4	Replay; Lean numeric checker
Certificate trees	5591	One per leaf or subcase of a leaf: 3949 inside (349 for supplementary leaves), 1642 outside	Section 11.6	Replay; independent checker; Lean checker
Threshold boxes	4,196,879	Terminal boxes of the trees: 4,109,455 inside (1,042,363 for supplementary leaves), 87,424 outside	Definition 10.6	As above
Relaxation cases	29,397,336	2^k in a box with k tangent rows; 9430 closed by an exclusion, the rest by integer duals	Definition 10.6	As above, in exact arithmetic

Rows and exclusions

Near rows	16,842	All family, shifted and graded rows (13,528 inside, 3314 outside), with 11,135,323 entries	Section 9.5	Interval arithmetic; Lean row checker
Two-test rows	37	On 37 inside roots: 31 single, 2 mixture, 4 paired	Section 9.6	Replay; exact comparison
Complex location rows	94	25 for λ_2 and 69 for λ' (Propositions 7.12 and 7.13)	Appendix A	Interval arithmetic; separate program
Degree-5 real rows	22	Cells covering $[0.700, 0.755]$, used in 300 location nodes	Proposition 7.10 , Table 6	Interval arithmetic; replay
Positivity exclusions	197	Reserved nodes of type rr; least normalized margin $7.8 \cdot 10^{-5}$	Proposition 7.14	Interval arithmetic; replay

Triples of numbers are counts for the types rr, rc, complex; “supplementary” refers to the supplementary leaves of [Section 11.4](#). Each split adds one terminal node; hence the identity in the refinement row. The printed tables were compared with the pages of [\[51, 145\]](#), and Heath-Brown's Tables 4 and 7 were recomputed in floating point ([Section 7.1](#)). The replay does not regenerate the 94 complex rows or the conditions [\[145, \(4.29\), \(4.34\)\]](#); separate interval programs check them. The Lean row checker covers all near rows except the 37 two-test rows.

TABLE 10. The finite analysis of the middle range $0.1 \leq \lambda_1 < 1.5$ at a glance.

The case tree, from the parent rows to the leaves, partitions configurations of zeros; the boxes and relaxation cases partition the thresholds of the near rows.

The checks are described in [Section 14](#).

11.5. The leaf model. This subsection gives the real coefficients before integer scaling. An interval $[lo, hi)$ is charged at its left endpoint in the objective, where the cost is largest, and at its right endpoint in density constraints, where the available weight is smallest. These choices enlarge the feasible set and give an upper bound for every actual distribution of characters in the interval.

Let $\mathfrak{s}$ be the specification of a leaf, after the location updates. Let s denote the *shift* $\min(1.9, p^*, l_2)$, where $p^* = lo'$ if $\mathfrak{s}$ has a gap and $p^* = p$ otherwise, and let λ_3^{lo} be the lower bound for λ_3 of [Proposition 7.16](#), computed from $[a, b]$, the type and, for a reserved $\mathfrak{s}$, the cap hi_2 . The leaf program ([Definition 10.1](#)) has the columns of [Table 11](#). Every number is rounded in the safe direction at scale S : far weights, features and the hidden-tail count coefficient $1/w(\text{end})$ down; objective coefficients, diagonals, correlation terms, the far budget and the constant first up. Let O be the set of ordinary characters with a zero in R_P , and $\mathcal{H}$ the set of first-family characters with a zero in R_P .

- (a) *Ordinary bins.* Let $R = \max(3, r)$ and let $r = \xi_0 < \xi_1 < \dots < \xi_N = R$ consist of r , R and the multiples of $1/\text{den}$ between them, where $\text{den} \in \{200, 400, 500, 800, 2000\}$ is fixed per leaf. The bins are the intervals $[\xi_i, \xi_{i+1})$, except that in a reserved leaf the bins with $\xi_{i+1} \leq \lambda_3^{lo}$ are omitted. The ordinary characters beyond R form the tail.
- (b) *Reserved family.* Either (fixed terms) the family is charged $n_2 \mathcal{G}_{\phi_2}(lo_2)$ in first and $n_2 w(hi_2)$ in the far budget, and it appears in the near rows as a family term with its feature at hi_2 (there are then no second-family columns, and the count n_E of the second-family equality of [Definition 10.1](#) is 0); or (columns) the interval $[lo_2, hi_2]$ is cut at the multiples of $1/400$ into reserved columns $[lo_{2,j}, hi_{2,j}]$, with values z_j and $\sum_j z_j = n_2 = n_E$.
- (c) *Hidden columns* (outside leaves only). Put $p^\circ = \max(a, p, lo')$ and $\text{end} = \max(3, p^\circ)$. The hidden bins $[lo_h, hi_h)$ divide $[p^\circ, \text{end}]$ on the grid $1/\text{den}$, and the hidden tail collects the values $t_\chi \geq \text{end}$. The hidden count is at most n .

Column	Value x_c	G_c	W_c	C_c	N_c	E_c	Near rows
ordinary bin $[\xi_i, \xi_{i+1})$	number of $\chi \in O$ with $\nu_{T^*}(\chi)$ in the bin	$\mathcal{G}(\xi_i)$	$w(\xi_{i+1})$	0 or 1^a	0	0	entry (a)
tail	$\sum w(\nu_{T^*}(\chi))$ over $\chi \in O$ with $\nu_{T^*}(\chi) \geq R$	$\mathcal{G}(R)/w(R)$	1	0	0	0	omitted
reserved column ^b $[lo_{2,j}, hi_{2,j}]$	$z_j = n_2$ if ν_* is in the column, else 0	$\mathcal{G}_{\phi_2}(lo_{2,j})$	$w(hi_{2,j})$	0	0	1	entry (f)
hidden bin ^c $[lo_h, hi_h)$	number of $\chi \in \mathcal{H}$ with t_χ in the bin	$e^{-A lo_h} B_{\phi_1}(p^\circ)$	$w(hi_h)$	0	1	0	entry (g)
hidden tail ^c	$\sum w(t_\chi)$ over $\chi \in \mathcal{H}$ with $t_\chi \geq \text{end}$	$\frac{e^{-A \text{end}} B_{\phi_1}(p^\circ)}{w(\text{end})}$	1	0	$\frac{1}{w(\text{end})}$	0	omitted

TABLE 11. The columns of a leaf program and their real coefficients before scaling by S : objective G_c , far weight W_c , count C_c , hidden count N_c and second-family indicator E_c ([Definition 10.1](#)). The last column gives the entry of [Table 7](#) that supplies the feature and diagonal of the column in each near row. (a) 1 if $\mathfrak{s}$ is unreserved and $\xi_{i+1} \leq \lambda_3^{lo}$, and 0 otherwise. (b) Only if the reserved family is charged through columns. (c) Outside leaves only.

Before scaling and rounding, the far budget is $(1+\eta)V - \mathbf{1}_{\text{inside}} n w(b)$, with a further subtraction of $n_2 w(\text{hi}_2)$ for a fixed reserved-family term. The stored integer budget F is formed by rounding the initial budget up and each subtraction down, as in [Section 6](#). The count budget is 2, and the stored allowance satisfies $\text{final}/S = 5\eta$. For an inside leaf, first/S is rounded up from $J_1 = \min(J_{\text{old}}, J_{\text{new}})$ with anchor p^* , omitting J_{new} when $p^* < b$. For an outside leaf it starts at 0. The fixed reserved-family charge, if used, is added with upward rounding. The near rows are those of [Section 9](#).

11.6. Subdivision of a leaf. A leaf may be certified through finitely many subcases, each of which is certified with its own leaf model. Besides the second-family split above, three kinds are used. The first replaces the specification by three specifications. The other two restrict one further quantity Q of the configuration to one of finitely many intervals I whose union is the range of Q ; the configuration set of the subcase $(\mathfrak{s}, I)$ is the set $C(\mathfrak{s}, I)$ of configurations for which, for some admissible choice of (χ_1, ρ_1) , of the minimizers and of ρ' , (C1)–(C5) hold and $Q \in I$. Then $C(\mathfrak{s})$ is the union of the sets $C(\mathfrak{s}, I)$, and all the results of [Sections 9](#) and [11](#) that are stated for $C(\mathfrak{s})$ hold for $C(\mathfrak{s}, I)$ with the same proofs.

Lemma 11.5 (Reservation split). *Let $\mathfrak{s}$ be unreserved with ordinary lower bound r , and let $r < h \leq 2$. Let $\mathfrak{s}_1$ and $\mathfrak{s}_2$ be $\mathfrak{s}$ with the reservation $(h, 1)$ and $(h, 2)$ (so $\text{lo}_2 = r$ and $\text{hi}_2 = h$), and let $\mathfrak{s}_0$ be $\mathfrak{s}$ with r replaced by h . Then $C(\mathfrak{s}) \subseteq C(\mathfrak{s}_0) \cup C(\mathfrak{s}_1) \cup C(\mathfrak{s}_2)$.*

Proof. If $\nu_* < h$, a family $\mathcal{F} \neq \mathcal{F}_1$ attaining ν_* exists, and it is a real character or a pair of conjugate nonreal characters; the configuration lies in $C(\mathfrak{s}_1)$ or $C(\mathfrak{s}_2)$, since $\nu_* \geq r$. Otherwise every family $\mathcal{F} \neq \mathcal{F}_1$ has $\nu(\mathcal{F}) \geq h$, and the configuration lies in $C(\mathfrak{s}_0)$. $\square$

In a reserved child the model uses only the minimality of the reserved family ([Lemma 11.3\(ii\)](#)) and the cap $\lambda_2 \leq h$ ([Lemma 11.3\(iii\)](#)); it never uses that the reserved family is $\mathcal{F}_2$. The split of [Lemma 11.5](#) is used with $h = \min(\lambda_3^{\text{lo}}, 2)$ on four inside roots and on five roots that contain supplementary leaves.

Height split for type rc. An inside leaf of type rc is split by $Q = \mu_1$ (recall that $\mu_1 > 0$ for this type) into $\mu_1 \in [0, 1]$ and $\mu_1 \in [1, \infty)$, which changes only the family row ([Section 9](#)).

Second-zero height split. An inside leaf of type complex with a finite gap is split by $Q = |y|$, where $y = \mu_1 - \mu_{\rho'} = (\gamma_1 - \gamma')\mathcal{L}$ is the normalized height difference between ρ_1 and the non-distinguished zero ρ' of χ_1 realizing λ' (a zero of χ_1 , not of $\bar{\chi}_1$; [Section 2](#)), into six pieces, according as $|y|$ lies in $[0, 1]$, $[1, \frac{3}{2}]$, $[\frac{3}{2}, 2]$, $[2, 3]$, $[3, 5]$ or $[5, \infty)$. This changes only the family row. A piece may be certified by an exclusion ([Definition 10.6\(i\)](#)) rather than by duals: the near row alone is then infeasible.

11.7. Realization and the cover theorem.

Proposition 11.6 (Realization). *Let $\mathfrak{s}$ specify a leaf or one of its subcases, with the program constructed in [Section 11.5](#). For every sufficiently large q whose configuration lies in $C(\mathfrak{s})$, there is a feasible pair (x, τ) such that*

$$W(q) + 2\eta \leq \mathcal{V}(x).$$

The column values x are obtained from the actual character counts and weighted tail masses. The threshold q_0 may be chosen uniformly over the finite set of leaves and subcases.

Proof. Let O and $\mathcal{H}$ be as in [Section 11.5](#) ($\mathcal{H}$ is used only if the configuration is outside). For $\chi \in O$ we have $\nu_{T^*}(\chi) \leq C_P$ and $\nu_{T^*}(\chi) \geq \nu(\chi) \geq r$ by (C4). Put x_i equal to the number of $\chi \in O$ with $\nu_{T^*}(\chi) \in [\xi_i, \xi_{i+1})$, $x_{\text{tail}} = \sum w(\nu_{T^*}(\chi))$ over $\chi \in O$ with $\nu_{T^*}(\chi) \geq R$, z_j equal to n_2 for the reserved column containing ν_* and 0 otherwise, x_i^h equal to the number of $\chi \in \mathcal{H}$ whose

value t_χ (Section 5.7) lies in the hidden bin i , and $x_{\text{tail}}^h = \sum w(t_\chi)$ over the $\chi \in \mathcal{H}$ in the hidden tail. In a reserved leaf no $\chi \in O$ lies in an omitted bin, by Lemma 11.3(ii) and $\lambda_3 \geq \lambda_3^{\text{lo}}$.

Far. The characters of O , the reserved characters, the first-family characters of $\mathcal{H}$ and, for an inside configuration, χ_1 and $\bar{\chi}_1$ are distinct, and to each we assign one zero with $|\gamma| \leq 1$ and parameter at most C_P : the T^* -representative, the height-one representative, the zero realizing t_χ , and ρ_1 or $\bar{\rho}_1$, which has $|\gamma_1| \leq C_B/\mathcal{L} \leq 1$. By (6.1), and since w is decreasing, the far constraint holds.

Count. By Lemma 11.3(i), since a character in a counted bin has $\nu_{T^*}(\chi) < \xi_{i+1} \leq \lambda_3^{\text{lo}} \leq \lambda_3$.

Hidden count and second family. We have $|\mathcal{H}| \leq n$, a character in the hidden tail contributes $w(t_\chi)/w(\text{end}) \leq 1$ to the hidden count, and $\sum_j z_j = n_2 = n_E$ when the reserved family is charged through columns (otherwise there are no second-family columns and $n_E = 0$).

Near rows. By Proposition 9.7, each near row holds for some threshold, with features at least and diagonals at most those of the columns.

Value. We use (5.2). An ordinary character in a bin costs at most the objective $\mathcal{G}(\xi_i)$ at the left end of the bin, a reserved character in a column at most $\mathcal{G}_{\phi_2}$ at the left end (with fixed terms, at most $\mathcal{G}_{\phi_2}(\text{lo}_2)$), and the tail characters at most $\mathcal{G}(R)/w(R)$ per unit of far mass, by the monotonicity of $\mathcal{G}$, $\mathcal{G}_{\phi_2}$ and $\mathcal{G}/w$ (Lemma 5.1(v)). For an outside configuration we apply Proposition 5.9 with the anchor $p^\circ = \max(a, p, \text{lo}')$, which is admissible because every non-distinguished occurrence of the first family has parameter at least $\lambda' \geq p^\circ$; then $t_\chi \geq p^\circ$, a character in a hidden bin $[\text{lo}_h, \text{hi}_h)$ costs at most the objective $e^{-A \text{lo}_h} B_{\phi_1}(p^\circ)$ since $e^{-A\lambda}$ is decreasing, and a character in the hidden tail costs at most $e^{-A \cdot \text{end}} B_{\phi_1}(p^\circ)/w(\text{end})$ per unit of far mass, since $e^{-A\lambda}/w(\lambda)$ is decreasing (Lemma 6.2). Including the directions of integer rounding, we obtain

$$W \leq \mathcal{V}(x) - \frac{\text{final}}{S} + E \leq \mathcal{V}(x) - 5\eta + 2\eta \leq \mathcal{V}(x) - 2\eta.$$

□

Theorem 11.7 (The zero-sum bound in the middle range). *For the case tree and certificates described in this paper, there is a threshold q_0 such that every $q \geq q_0$ with $0.1 \leq \lambda_1 < 1.5$ satisfies*

$$W(q) < 1 - 2\eta.$$

More precisely, its configuration belongs to a leaf or subcase whose program has a feasible pair (x, τ) with $W(q) + 2\eta \leq \mathcal{V}(x) < 1$.

Proof. By Theorem 11.4 the configuration lies in the configuration set of a root. Following the refinement tree and the subdivisions of Section 11.6, each of which is exhaustive, we reach a leaf or subcase whose configuration set contains the configuration; it is not an excluded node, because excluded nodes have no configurations. Proposition 11.6 gives the feasible pair, and Theorem 10.7 and the certificates of Section 14 give $\mathcal{V}(x) < 1$. □

12. THE EXTERIOR REGIMES

The case tree covers $0.1 \leq \lambda_1 < 1.5$. We now treat the remaining possibilities: an empty zero rectangle, a first zero with $\lambda_1 < 0.1$, and a first zero with $\lambda_1 \geq 1.5$. The small-parameter case is different from the others. Its first zero is real: it is an exceptional zero in the sense of Landau and Page [74, 97]. By the theorems of Siegel and Tatzuza [118, 123] (see also [102, 103]) it cannot be extremely close to 1, but these bounds do not keep λ_1 away from 0. Although such a zero has a large cost, the Deuring–Heilbronn phenomenon [30, 53, 78] repels the remaining zeros. We use Heath-Brown’s quantitative estimates for this repulsion (see also [8] for explicit versions) and, for a sufficiently close zero, his separate theorem on Siegel zeros [50], by which a very strong exceptional zero is even helpful.

12.1. No zero in $R(l)$. If no nonprincipal $L(s, \chi)$ vanishes in $R(l)$, then for large q none vanishes in R_P , because $C_P \leq \frac{1}{3} \log \log \mathcal{L}$ and $C_P/\mathcal{L} \leq l$. So $W = 0$ and [Proposition 4.4](#) gives a prime $p \equiv a \pmod{q}$ with $q^A < p < q^L$. Xylouris notes this case himself [[145](#), p. 36].

12.2. A tiny exceptional zero. Let $\eta_{\text{HB}}(\delta)$ be the effectively computable constant of [Input 3.10](#) and put

$$u_{\text{exc}} = \min\left(\frac{1}{\eta_{\text{HB}}(1/2)}, 0.08\right).$$

Proposition 12.1 (Primes in the presence of a very close real zero). *Let u_{exc} be the positive cutoff defined above. For all sufficiently large q , if $\lambda_1 < u_{\text{exc}}$, then $P(a, q) \leq q^{3.5}$ for every $(a, q) = 1$.*

Proof. By [Proposition 7.1](#), χ_1 and ρ_1 are real, so $\beta_0 = \rho_1 = 1 - \lambda_1/\mathcal{L}$ is a real zero of the real nonprincipal $L(s, \chi_1)$. Since $\lambda_1 < u_{\text{exc}} \leq \frac{1}{3}$ we have $\beta_0 \geq 1 - 1/(3 \log q)$, and $((1 - \beta_0) \log q)^{-1} = 1/\lambda_1 > 1/u_{\text{exc}} \geq \eta_{\text{HB}}(\frac{1}{2})$. [Input 3.10](#) with $\delta = \frac{1}{2}$ gives the claim. $\square$

12.3. A small exceptional zero. For $u_{\text{exc}} \leq \lambda_1 \leq 0.1$ we use a different prime weight, a single triangle, and Heath-Brown's own estimates. The inputs are the following.

Input 12.2 (Prime detection with a single triangle [[51](#), Lemma 13.2]). *Let $L' > 2K + 3$, let $f = h_{L', K}$ and $F(z) = e^{-(L'-2K)z} F_2(z)$ with $F_2(z) = \left(\frac{1-e^{-Kz}}{z}\right)^2$. For every $\varepsilon > 0$ there is $R = R(\varepsilon)$ such that for large q*

$$\sum_{p \equiv a(q)} \frac{\log p}{p} f\left(\frac{\log p}{\mathcal{L}}\right) \geq \frac{\mathcal{L}}{\varphi(q)} \left[K^2 - \varepsilon - \sum_{\chi \neq \chi_0} \sum_{\rho}' |F((1 - \rho)\mathcal{L})| \right],$$

where $\sum'$ runs over the zeros with $1 - R/\mathcal{L} \leq \beta \leq 1$, $|\gamma| \leq R/\mathcal{L}$.

Input 12.3 (Per-character bound for the triangular kernel [[51](#), Lemma 13.3]). *Let $\chi \neq \chi_0$, $0 < \lambda \leq R$, and suppose that $L(s, \chi)$ has no zero with $1 - \lambda/\mathcal{L} < \beta \leq 1$ and $|\gamma| \leq 1$. Then for every $\eta_1 > 0$ and $q \geq q(\eta_1, \lambda)$*

$$\sum_{\rho}' |F_2((1 - \rho)\mathcal{L})| \leq \mathcal{B}_{\varphi(\chi)}(\lambda) + \eta_1, \quad \mathcal{B}_{\phi}(\lambda) = \frac{\phi}{2} \frac{1 - e^{-2K\lambda}}{\lambda} + \frac{2K\lambda - 1 + e^{-2K\lambda}}{2\lambda^2}.$$

Input 12.4 (Heath-Brown's weighted density estimate). *In the notation of [[51](#), Lemma 11.1] and [[145](#), (5.18), p. 66], the following holds. Let $\varepsilon, c_1, c_2 > 0$ and $\phi = \max_{\chi} \varphi(\chi)$. For $q \geq q(\varepsilon, c_1, c_2)$ and any choice, for each nonprincipal χ with a zero in $\beta \geq 1 - \lambda_0/\mathcal{L}$, $|\gamma| \leq 1$ ($\lambda_0 = \frac{1}{3} \log \log \mathcal{L}$), of one such zero with parameter $\lambda(\chi)$,*

$$\sum_{\chi} \frac{\lambda(\chi)}{e^{(4\phi+6c_1+2c_2)\lambda(\chi)} - e^{(2\phi+4c_1)\lambda(\chi)}} \leq \frac{2\phi + 2c_1 + c_2}{4c_1c_2} + \varepsilon.$$

The left side decreases and the right side increases with ϕ , so the inequality holds with $\phi = \frac{1}{3}$. We also use [Input 7.2\(b\),\(d\)](#) and the following two rows of Heath-Brown's tables, which hold for all $\lambda_1 \leq 0.10$ bounded below by a fixed positive constant [[51](#), §8, Table 2 and Table 5] (see the remark after [Input 7.2](#); we use them for $\lambda_1 \geq 0.08$): if χ_1 and ρ_1 are real and $\lambda_1 \leq 0.10$, then $\lambda' \geq 4.96 - \varepsilon$ and $\lambda_2 \geq 2.83 - \varepsilon$ for $q \geq q(\varepsilon)$. We have recomputed these rows from Heath-Brown's printed parameters, obtaining 4.96355 and 2.84577; we use them with $\varepsilon = 0.001$. (Table 5 is stated under $\lambda_2 \leq \lambda'$, which is redundant here since $\lambda' \geq 4.959 > 2.83$.)

Fixed data. $L = 3.99$, $K = 0.1821$, $A_s = L - 2K = 3.6258$; $c_1 = 0.057$, $c_2 = 0.1554$, $\phi = \frac{1}{3}$;

$$a_d = \frac{4}{3} + 6c_1 + 2c_2 = \frac{3724}{1875}, \quad b_d = \frac{2}{3} + 4c_1 = \frac{671}{750}, \quad V_d = \frac{2/3 + 2c_1 + c_2}{4c_1c_2} = \frac{184750}{6993};$$

$\alpha = 1.09 < \frac{12}{11}$, $B_1 = K^2 + \frac{K}{4}$, $m^* = \alpha \log 12.5 = 2.7530442 \dots$, $m_2 = 2.829$ and $m'_2 = 4.959$. Let $H_s(u) = e^{-A_s u} F_2(u)$ and $\mathcal{M}(u) = (K^2 - H_s(u))/u$, and

$$Q(m) = \frac{e^{-(A_s - a_d)m} - e^{-(A_s - b_d)m}}{m} = \int_{A_s - a_d}^{A_s - b_d} e^{-tm} dt.$$

Theorem 12.5 (The small exceptional-zero range). *For all sufficiently large q with $u_{\text{exc}} \leq \lambda_1 \leq 0.1$, and every integer a with $(a, q) = 1$, there is a prime p such that*

$$p \equiv a \pmod{q}, \quad q^{3.6258} < p < q^{3.99}.$$

The threshold is uniform over the fixed interval $[u_{\text{exc}}, 0.1]$.

Proof. Write $u = \lambda_1$. By Proposition 7.1, χ_1 and ρ_1 are real. The zero ρ_1 is simple, since a double zero would give $\lambda' = u$, whereas $\lambda' \geq (2 - \varepsilon) \log(1/u) > u$ (Input 7.2(b)) on $[u_{\text{exc}}, 0.08]$ and $\lambda' \geq 4.959 > u$ on $[0.08, 0.1]$.

The criterion. Apply Input 12.2 with $L' = L$, since $3.99 > 2K + 3$; the zero ρ_1 contributes $H_s(u)$. Enlarging R only adds nonnegative terms to $\sum'$, so the inequality of Input 12.2 remains true with R replaced by $\max(R, 3)$; we use it with $R \geq 3$, so that Input 12.3 applies with the values $\lambda \leq m_2 < 3$ below. It remains to bound the other terms. Put $m(u) = \alpha \log(1/u)$ on $\mathcal{P}_1 = [u_{\text{exc}}, 0.08]$ and $m(u) = m_2$ on $\mathcal{P}_2 = [0.08, 0.1]$, and $m_1(u) = m(u)$ on $\mathcal{P}_1$, $m_1(u) = m'_2$ on $\mathcal{P}_2$. By Input 7.2(b),(d) (with $\varepsilon = \frac{12}{11} - \alpha$) and the two table rows, every zero $\rho \neq \rho_1$ of $L(s, \chi_1)$ in $R(l)$ has $\lambda_\rho \geq m_1(u)$, and every zero of every other nonprincipal $L(s, \chi)$ in $R(l)$ has $\lambda_\rho \geq m(u)$.

The first character. $L(s, \chi_1)$ has no zero with $\lambda < u_{\text{exc}}/2$ and $|\gamma| \leq 1$, so Input 12.3 with $\lambda = u_{\text{exc}}/2$ and $\varphi(\chi_1) = \frac{1}{4}$ gives $\sum'_\rho |F_2((1 - \rho)\mathcal{L})| \leq \mathcal{B}_{1/4}(u_{\text{exc}}/2) + \eta_1 \leq B_1 + \eta_1$, as $\mathcal{B}_\phi$ is decreasing in λ , with limit $\phi K + K^2$ as $\lambda \rightarrow 0$. Hence the zeros of χ_1 other than ρ_1 contribute

$$T_1 \leq e^{-A_s m_1(u)} (B_1 + \eta_1).$$

The other characters. For $\chi \neq \chi_0, \chi_1$ with a zero in the rectangle of Input 12.2, let $\lambda(\chi)$ be the parameter of its zero of least parameter with $|\gamma| \leq 1$; then $\lambda(\chi) \geq m(u)$. Apply Input 12.3 with the fixed value $\lambda = m^*$ on $\mathcal{P}_1$ (note $m^* \leq m(u)$ there) and $\lambda = m_2$ on $\mathcal{P}_2$, and with $\varphi(\chi) \leq \frac{1}{3}$. Since every zero in the rectangle has parameter at least $\lambda(\chi)$, the zeros of χ contribute at most $e^{-A_s \lambda(\chi)} (\mathcal{B}_{1/3}(\bar{m}) + \eta_1)$, with $\bar{m} = m^*$ or m_2 . Since $A_s > a_d$, the function $e^{-A_s t} (e^{a_d t} - e^{b_d t})/t = Q(t)$ is decreasing, so $e^{-A_s \lambda(\chi)} \leq Q(m(u)) \lambda(\chi) / (e^{a_d \lambda(\chi)} - e^{b_d \lambda(\chi)})$. By Input 12.4, the characters other than χ_1 contribute

$$T_2 \leq (\mathcal{B}_{1/3}(\bar{m}) + \eta_1) (V_d + \varepsilon_2) Q(m(u)).$$

Monotonicity. We have $K^2 - H_s(u) = \int h_{L,K}(t) (1 - e^{-ut}) dt$, so $\mathcal{M}(u) = \int h_{L,K}(t) \frac{1 - e^{-ut}}{u} dt$, which decreases in u because $h_{L,K} \geq 0$ and $(1 - e^{-ut})/u$ decreases for each $t > 0$. On $\mathcal{P}_1$ we have $u = e^{-m/\alpha}$, and

$$\frac{Q(m(u))}{u} = \int_{A_s - a_d - 1/\alpha}^{A_s - b_d - 1/\alpha} e^{-tm(u)} dt, \quad \frac{e^{-A_s m(u)}}{u} = e^{-(A_s - 1/\alpha)m(u)},$$

with $A_s - a_d - 1/\alpha = \frac{236171}{327000} > 0$; both increase with u . On $\mathcal{P}_2$ the bounds for T_1 and T_2 do not depend on u , and $u \geq 0.08$. Hence on $\mathcal{P}_i$

$$K^2 - H_s(u) - T_1 - T_2 \geq \varpi_i u - \mathcal{E},$$

where

$$\begin{aligned}\varpi_1 &= \mathcal{M}(0.08) - V_d \mathcal{B}_{1/3}(m^*) \frac{Q(m^*)}{0.08} - B_1 \frac{e^{-A_s m^*}}{0.08}, \\ \varpi_2 &= \mathcal{M}(0.1) - V_d \mathcal{B}_{1/3}(m_2) \frac{Q(m_2)}{0.08} - B_1 \frac{e^{-A_s m'_2}}{0.08},\end{aligned}$$

and $\mathcal{E} = \eta_1(1 + (V_d + \varepsilon_2)(a_d - b_d)) + \varepsilon_2 \mathcal{B}_{1/3}(0)(a_d - b_d)$, using $Q \leq a_d - b_d$.

Numerical margins. The interval computations give the following values, displayed here to ten decimal places:

	$\mathcal{M}$	other characters	first character	ϖ_i
$\mathcal{P}_1$	0.1088458429	0.0783116746	0.0000454655	0.0304887029
$\mathcal{P}_2$	0.1050056698	0.0668874692	0.0000000153	0.0381181853

The certified lower bounds give $\varpi_i > 0.0304$. Choose η_1 , ε_2 and the ε of [Input 12.2](#), after u_{exc} , with $\mathcal{E} + \varepsilon < 0.03u_{\text{exc}}$. Then the bracket in [Input 12.2](#) is at least $(0.0304 - 0.03)u_{\text{exc}} > 0$, and there is a prime $p \equiv a \pmod{q}$ with $\log p/\mathcal{L} \in (A_s, L)$. $\square$

At $L = 3.99$, the margins are approximately 28% and 36% of the main term. Exploratory evaluations with the same K , c_1 , and c_2 remain positive at $L = 3.90$; the certified exterior result used here is the one at 3.99.

12.4. A large first zero.

Theorem 12.6 (The large first-zero range). *For all sufficiently large q with a nonempty zero rectangle and $\lambda_1 \geq 1.5$,*

$$W(q) \leq 0.9966814 - 2\eta.$$

Consequently, every reduced residue class contains a prime p with $q^{3.156342} < p < q^{3.99}$.

Proof. Every nonprincipal character with a zero in R_P is treated as ordinary, with its height-one representative, of parameter $\lambda_\chi \geq \lambda_1 \geq 1.5$; by [Corollary 5.7](#) it contributes at most $\mathcal{G}(\lambda_\chi) + \varepsilon_1 e^{-A\lambda_\chi}$. Consider the leaf program with the bins $[1.5 + \frac{j}{100}, 1.5 + \frac{j+1}{100})$, $0 \leq j < 150$, and the tail $[3, \infty)$; the far budget $F = (1 + \eta)V$ with profile I ([Section 6](#)); first = 0 and final = 5η ; no count, hidden or second-family constraints; and one near row. The near row is [Corollary 8.10](#) with $\Xi \equiv 0$, $s = s_1 = \frac{3}{2}$, detector $\hat{f}_2$ ($\gamma_f = 1$), Gram test $\hat{f}_3$ ($\gamma_g = \frac{3}{2}$) and no family entries, normalized as in [Section 9.2](#) with $t_0 = 1$ (so I_u is the upper Riemann sum over 2000 cells of each of $[0, 1)$ and $[1, 2)$): every zero of every nonprincipal L -function in $R(l)$ has parameter at least $1.5 = s_1$, so the safe-anchor hypothesis holds, and every entry has no zero to the right of its anchor $\min(\text{lo}, s) = s$. The actual characters give a feasible point as in [Proposition 11.6](#).

The certificate consists of 36 contiguous threshold intervals covering $[0, e]$; on each it uses the first-order relaxation and integer duals $Y_F, Z \geq 0$ satisfying [\(10.1\)](#); there are no exclusions and 5436 column inequalities. The largest value is 0.9966813270344512, on the threshold interval $[0.1625, 0.165]$, where $Y_F = 5145926349$ and $Z = 1593393341176$. By [Theorem 10.7](#), $\mathcal{V}(x) < 0.9966814$, and $W \leq \mathcal{V}(x) - 2\eta$ as in [Proposition 11.6](#). [Proposition 4.4](#) gives the prime. $\square$

In this regime one near row, together with the far-density budget, suffices. All characters are treated by the same per-character bound.

13. ASSEMBLY OF THE PROOF

Each preceding argument holds once q exceeds a threshold depending on fixed choices. We first verify that these choices can be made in a noncircular order. A single threshold then works for every leaf and every exterior case.

13.1. The order of the choices. The constants of the argument are chosen in the following order. Each stage depends only on the earlier ones, and there are finitely many choices at each stage.

- (S1) The fixed data: the exponent $L = 3.99$ and the kernel of [Section 4](#); the two far profiles of [Section 6](#); the test functions, anchors, sieve weights and grids of all near rows ([Section 9](#)); the case tree, its leaf programs and their certificates ([Sections 10 and 11](#)); the zero-location rows of [Section 7](#); the data of the exterior regimes ([Section 12](#)); $\eta = \eta' = 10^{-6}$, $\lambda_{11} = 0.05$ and $s_{\max} = 3$; and $u_{\text{exc}} = \min(1/\eta_{\text{HB}}(\frac{1}{2}), 0.08)$, with $\eta_{\text{HB}}(\delta)$ from [Input 3.10](#); and the envelope tolerance $\varepsilon_1 = \eta/23$ of [Section 5.8](#).
- (S2) The prime window: $C_P \geq \max(C_0(\eta H_0), 3)$, with C_0 from [Input 4.2](#).
- (S3) The zero count $N = N(C_P)$ of [Lemma 5.3](#); the mesh h_A and the finite set of anchors $\mathcal{A}$ of [Section 5.3](#); then the smoothing parameter θ of [Section 5.2](#), chosen so that (5.1) holds for $\varepsilon_1 = \eta/23$; then the constant c_{out} of [Proposition 5.9](#).
- (S4) The count K_0 of [Section 6.3](#), a bound, uniform in q , for the number of zeros of all nonprincipal L -functions modulo q with $\lambda \leq s_{\max}$ and $|\gamma| \leq 2$ ([Input 3.9](#)). Then the separation M , so large that for $|Y| \geq M$ we have $|\text{Re } F(x + iY)| \leq \eta''_{\min}/(2K_0 + 2)$ for every detector and $x \in [-s_{\max}, 0]$, where $\eta''_{\min} > 0$ is the least of the numbers $\eta'' = \eta' \min(1, \sqrt{I_B D_u}, D_u)$ of [Corollary 8.10](#) over the finitely many rows, and $|\text{Re } G(-\sigma + iY)| \leq d_1$ for every Gram test and every $\sigma = s_1 - \delta - \delta'$ that occurs ([Lemma 8.6](#)). Then the buffer $C_B > \max(C_P + M, 1.5)$ with $4c_{\text{out}}/(H_0 C_B^2) \leq \eta$.
- (S5) The tolerances of the published results and of the lemmas proved here, each below the finitely many positive margins it must fit: the tolerances of [Inputs 3.3 and 3.4](#) for the finitely many test functions (those of [Section 5](#), the functions f_σ^θ with tolerance $H_0 \varepsilon_1/8$, and the detectors of the near rows with tolerance $\eta''_{\min}/2$), the disc radius $\delta_0 \leq \frac{1}{3}$ of [Section 5.3](#), which is at most the radius that [Input 3.4](#) provides for each of these, and the tolerances of (7.2) and [Lemma 7.8](#), of [Input 6.1](#) (with $\varepsilon = \eta J^2$), of [Inputs 3.7 and 3.8](#) and Mertens' estimate, of [Input 4.2](#) (with $\varepsilon = \eta H_0$), of the zero tables, and of the small-branch estimates of [Theorem 12.5](#) (with $\mathcal{E} + \varepsilon < 0.03u_{\text{exc}}$).
- (S6) Finally q_0 : the maximum of the finitely many thresholds of the previous stages, together with those of [Input 3.5](#), of the inclusions $R_P \subseteq R_B \subseteq$ the discs, and $\mathcal{L} > 4M(K_0 + 1)$ for [Lemma 6.3](#).

The certificates of stage (S1) depend only on fixed data, not on C_P , M , C_B or q . The count K_0 counts zeros up to height 2, which exceeds $1 + \delta_0$ for every admissible radius; so M , which depends on K_0 , can be fixed before the radius δ_0 , and there is no circular dependence. The strip height $T^* = T^*(q) \in [\frac{1}{2}, 1]$ of [Lemma 6.3](#) is then determined by q .

Envelope errors that are summed over unboundedly many characters are weighted by $e^{-A\lambda_\chi}$, and $\sum_\chi e^{-A\lambda_\chi} < 19$ by (6.2); errors in the near rows are uniform per entry or are multiples of $(\sum_j a_j)^2$ ([Theorem 8.4](#)). Thus the error allowances remain uniform as the number of characters grows with q .

13.2. The regimes. Let $q \geq q_0$. Exactly one of the following holds.

- (R1) No nonprincipal $L(s, \chi)$ has a zero in $R(l)$.
- (R2) $\lambda_1 < 0.1$.
- (R3) $0.1 \leq \lambda_1 < 1.5$.
- (R4) $\lambda_1 \geq 1.5$.

(The first zero λ_1 is defined when (R1) fails; since $R(l)$ is compact there are finitely many zeros in it.)

13.3. Proof of Theorem 1.1. Let $q \geq q_0$ and let a be coprime to q .

In case (R1), $W = 0$ and Proposition 4.4 gives a prime $p \equiv a \pmod{q}$ with $q^A < p < q^L$.

In case (R2), if $\lambda_1 < u_{\text{exc}}$ then $P(a, q) \leq q^{3.5}$ by Proposition 12.1; if $u_{\text{exc}} \leq \lambda_1 < 0.1$ then Theorem 12.5 gives a prime $p \equiv a \pmod{q}$ with $q^{3.6258} < p < q^{3.99}$.

In case (R3), Theorem 11.7 gives a leaf of the case tree, or a subcase of a leaf, and a feasible point of its program with $W \leq \mathcal{V}(x) - 2\eta < 1 - 2\eta$. In case (R4), Theorem 12.6 gives $W < 1 - 2\eta$. In both cases Proposition 4.4 gives a prime $p \equiv a \pmod{q}$ with $q^A < p < q^L$.

Thus $P(a, q) < q^{3.99}$ for every $q \geq q_0$ and every $(a, q) = 1$. There are only finitely many moduli $2 \leq q < q_0$, and finitely many reduced residue classes for each such modulus. Dirichlet's theorem supplies a least prime in each class. Taking the maximum of 1 and the finitely many ratios $P(a, q)/q^{3.99}$ gives an absolute C valid for all $q \geq 2$. $\square$

Remark 13.1 (Effectivity). Heath-Brown's theorem on Siegel zeros [50, Corollary 1], used for the closest exceptional zeros, is effective. The other source estimates used here are also stated with effective constants. However, we have neither audited the effectivity of every step of the assembly nor computed a value of q_0 or C . Theorem 1.1 asserts their existence and does not supply an explicit numerical bound for either constant.

14. THE COMPUTATIONS AND THEIR VERIFICATION

The finite computation has three logical components. The zero-location checks justify the case analysis; enclosures of analytic quantities supply the leaf coefficients; and integer certificates bound the resulting programs. The exterior ranges require separate computations. A check of one component does not by itself verify the others. This section records the evidence for each and the scope of the Lean formalization.

Computer assistance is well established in explicit analytic number theory, for instance in the ternary Goldbach problem [54] and in verifications of the Riemann hypothesis [106, 107], and linear programs with rigorously checked bounds are central to the proof of the Kepler conjecture [96, 119, 46]. In our computations, the real coefficients and margins are enclosed by interval arithmetic [88, 89, 113, 132]. It is implemented with the interval context of mpmath [126], at a working precision of at least 50 digits, and with a vectorized binary64 interval module that widens the result of every operation outward by one unit in the last place and evaluates exp, sin and cos by Taylor polynomials with explicit remainder bounds, not by the platform's library; the Lean checkers of Section 14.3 use their own verified rational interval arithmetic. Floating-point linear programming (with the HiGHS solver [55] in SciPy [137]) is used only to propose dual multipliers, and every certificate is checked in exact integer arithmetic. Independent numerical audits use multiple-precision arithmetic [126]. The supplementary floating-point checks of certain published table entries are identified separately in Section 7.1. Table 12 summarizes, for each component of the proof, whether it is published or new, how it was obtained and how it was checked.

14.1. The data. The certificates are stored in three compressed JSON record files. The inside and outside files are indexed by root case; the supplementary file records the 233 supplementary leaves of 199 inside roots (Section 11.4). Their sizes and SHA-256 hashes are:

Component	Source	Kind	Method	Checked by
Positivity criterion and detection (Input 4.2 and Proposition 4.4)	Published [145]	Analytic	–	Lean, with the criterion as a hypothesis
Envelope and first-family bounds (Section 5)	New, after [145]	Analytic	Interval arithmetic	Hand proof; a hypothesis in Lean, which checks only the enclosures
Far density (Input 6.1)	Published [145]	Analytic	–	Lean hypothesis; profile conditions proved in Lean
Far constants V , w (Section 6)	New profiles	Numerical	Interval arithmetic	Multiple-precision audit; Lean numeric checker
Printed zero-location tables (Proposition 7.5)	Published [51, 145]	Analytic	As printed	Compared with the pages; Heath-Brown's Tables 4 and 7 recomputed in floating point
New location rows (Section 7)	New	Both	Interval arithmetic	Replay; the 94 complex rows by a separate program
Conditions [145, (4.29), (4.34)]	Published; new check	Numerical	Interval arithmetic	Separate programs
Conductor refinement $\frac{1}{108}$ (Section 7.6)	New	Analytic	–	Hand proof
Graded near lemma, response lemma, threshold form (Section 8)	New	Analytic	–	Lean (Comparator, two kernels)
Family, shifted and graded rows (Section 9)	New	Numerical	Interval arithmetic	Lean row checker (16,842 rows); multiple-precision audit
Two-test rows (Section 9.6)	New, after [145]	Both	Interval arithmetic	Hand proof; replay; exact comparison with a separate implementation on every certificate interval (constants: that implementation's interval code only)
Certificate soundness (Theorem 10.7)	New	Analytic	–	Lean (Comparator, two kernels)
Leaf certificates (Definition 10.6)	New	Numerical	Exact integer	Replay; independent checker; Lean checker
Leaf data (Section 11.5)	New	Numerical	Interval arithmetic	Multiple-precision audits; Lean numeric checker (count-type integers: audit only)
Case tree and realization (Section 11)	New	Analytic	–	Hand proof; replay
Tiny exceptional zero (Proposition 12.1); empty rectangle	Published [50]; $W = 0$	Analytic	–	As printed
Small exceptional zero (Theorem 12.5)	New, from lemmas of [51]	Both	Interval arithmetic	Hand proof; exterior replay; margins recomputed independently
Large first zero (Theorem 12.6)	New	Numerical	Exact integer	Replay's exact checker; independent checker; Lean certificate (native and kernel), numeric and row checkers
Assembly (Section 13)	New	Analytic	–	Hand proof

The interval arithmetic is described at the beginning of this section. Floating point is used only to propose parameters, dual multipliers and sieve heights, which need no justification, and to recompute published tables.

TABLE 12. Status of the components. Published inputs are used as printed, and those used in the Lean theorems enter them as hypotheses; “Both” means an analytic argument with numerical margins; “hand proof” means a conventional proof in this paper. The checks are described in Section 14.

File	Bytes
inside.jsonl.gz	298,536,973
539362a03e7b330dad03b33d6b5037871cfcbbf7b0973e7976fccf12509fe63c	
nodes.jsonl.gz	159,624,716
b715ac735b6976da17a2671667ecfdf07c3f19b8ab74ae5f0533dbcf6b2a6cfc	
outside.jsonl.gz	8,626,435
a53d4d5793f71c92647a7f22022110bdc8135cbda1f2575d1dc0c1f487f5ac1d	

The first holds the 2768 inside roots with their 2913 other leaves, the second the 233 supplementary leaves, and the third the 1685 outside roots with their 1642 leaves. A leaf record contains the parameters of its near rows and its certificate tree; it does *not* contain the leaf program itself, which the verifier regenerates from the case tree. Some leaves are split into subcases (Section 11.6), each with its own certificate tree, so the 4788 leaves carry 5591 certificate trees in all: 3949 inside (including the supplementary leaves) and 1642 outside.

14.2. Exact verification. All numbers entering a leaf program are enclosed with outward-rounded interval arithmetic and stored as integers scaled by $S = 10^{16}$, rounded in the safe direction (Section 10). The verification of a certificate uses only integer arithmetic. The following checks have been run.

- (1) *The replay.* A verifier traverses the case tree from its roots. It re-checks every split, location node and exclusion (regenerating the location rows, except that the table of complex polynomial rows and the side conditions [145, (4.29), (4.34)] are checked by separate programs). At every leaf it regenerates the leaf program from the specification and the tree, independently of the stored certificate, and checks the certificate exactly. It passes on all 4453 roots: 4788 leaves, 4,196,879 boxes, and largest certified value $0.9999998223 \dots < 1$ (inside) and $0.9999942478 \dots$ (outside).
- (2) *An independent checker.* A second verifier was written from Section 10 alone, without reading or using the checking code of the first, and it uses only exact integer and rational arithmetic. It takes the leaf data from the same generator as the replay (the leaf data are checked separately, in (4) and in Section 14.3). It accepts every certificate of the corpus (29,397,336 relaxation cases in all) and that of Theorem 12.6, and for every certificate tree its value and number of boxes are exactly those of the first verifier. In a negative test on ten leaves, including the tightest, it rejected all 308 certificates made invalid by deliberate forgeries, among them duals lowered, and coefficients raised, by the smallest amount that breaks (10.1); it accepted all 82 controls perturbed by one unit less. It replaces an earlier independent checker, whose code was not preserved and which had also accepted every certificate of the corpus.
- (3) *A mutation test.* On the 19 leaves of three of the tightest inside roots, 133 perturbations of the regenerated leaf data (far budget, objective, first-family charge, correlation term, features), of certificate multipliers and of the allowance 5η , each in the unfavourable direction, were all rejected by the replay.
- (4) *Audits of the constants.* For 38 leaves of all kinds, the 19,616 objective and 19,616 far coefficients, the tail, hidden and second-family columns, F , first and all 120 family, shifted and graded near rows were recomputed from their definitions in multiple-precision arithmetic by a program sharing no code with the builders; all lie on the safe side of their true values. (The program of this audit was not preserved; the numeric and row checkers of Section 14.3 now check the same quantities for every leaf, except the two-test rows.) Separately, for 117 roots (100 of them chosen at random), all 530,942 column and budget integers of their 154 leaves (objective, far, count, hidden-count and second-family

coefficients, F , first and final) were recomputed from their defining formulas in 50-digit arithmetic and agree exactly with the stored ones.

- (5) *The two-test rows.* A separate program replays the 37 roots whose leaves use the row of [Section 9.6](#) (31 with a single test, 2 with a mixture, 4 paired). It compares the row as the certificates use it with a separate implementation of the same inequality, with its own builder and its own certificate checker. The comparison runs on every threshold interval that the 37 certificates use (1079 intervals, from 23,341 boxes) and on 11,211 seeded random intervals, in the first-order relaxation and in both tangent relaxations (for these, the tangent rule applied to the exact inequality of that implementation). The exact costs and budgets are those of that inequality divided by D , and the integers used are rounded in the safe direction. The row's integers are regenerated by the interval routines of that implementation and agree with the leaves; every one-unit perturbation of them is detected. These constants rest on those interval routines alone.

The exterior regimes. The margins of [Theorem 12.5](#) were regenerated from their definitions and checked in interval arithmetic by a separate replay of the exterior components, and they have been recomputed independently in multiple-precision interval arithmetic. The program of [Theorem 12.6](#) is a leaf program in the sense of [Definition 10.1](#). Its objective and far integers are regenerated and compared with the stored ones, and its near row is built with the normalization of [Section 9.2](#). The exact checking code of the replay (1) and the independent checker (2) accept its certificate, and so do the verified checkers of [Section 14.3](#), which also accept its numeric data and its near row.

14.3. Formal verification. Several parts of the argument have been formalized in the Lean 4 proof assistant [29] with the Mathlib library [125]; see [2] for the role of such formalizations. Earlier formalizations in analytic number theory include the prime number theorem [3, 47, 72] and Dirichlet's theorem [121]. The formalization uses Lean v4.35.0-rc3 with the corresponding Mathlib release, and contains about 24,000 lines and about 800 theorems and lemmas. The 18 statements of the near lemma and the certificate checker were checked against their proofs with the Comparator tool [75], which replays the proofs in the Lean kernel and in the independent kernel nanoda [5]; the leaf-level and near-row theorems were checked by the Lean kernel. All proofs are complete and use only the three standard axioms (propositional extensionality, the axiom of choice, and the soundness of quotients). The published inputs of [Section 3](#) enter as explicit hypotheses, each stated no more strongly than its printed source. The following results are formally verified.

- (a) The graded near-density lemma ([Theorem 8.4](#)), from the local explicit formulas, Burgess's estimate and Graham's estimate as hypotheses. This includes the response lemma, the threshold form, and the entry bounds of [Section 9](#).
- (b) The soundness of the certificate checker of [Section 10](#): an accepted certificate proves that every feasible point of the leaf program has value below 1.
- (c) The passage from a certified leaf with valid data to primes. From Xylouris's criterion and his far-density lemma, stated as printed, and the localized envelope of [Proposition 5.5](#), stated as a hypothesis, a leaf whose certificate and data pass the checks yields a prime $q^A < p < q^L$ in every reduced class, for every modulus whose zeros satisfy the leaf's zero-level hypotheses. The count, hidden-count and second-family integers and final enter as they are (they were recomputed in the audit of (4) above).
- (d) Interval arithmetic with proved inclusion, including the exponential function, and enclosures of the functions entering the objective and far data of the leaves. A numeric checker is proved to imply that each leaf's objective and far coefficients, far budget and first-family charge are valid.

- (e) Heath-Brown’s Conditions 1 and 2 for the parabolic test functions of [Section 9](#).
- (f) The near rows: each family, shifted and graded near row is a concrete instance of [Theorem 8.4](#). A row checker is proved sound, including the three family terms that keep a second zero and the constant C_Z of [Section 9](#). A checked row is a valid row of the leaf program under zero-level hypotheses only, and this is composed with (c). (The formal version of [Corollary 8.10](#) assumes that the numerators $D(\delta_j) - d_1 + e_j$ themselves are positive, which is stronger than what the proof in [Section 8](#) needs; the row checker verifies this for every checked entry.)

The certificate checker has a Mathlib-free copy that is proved to compute the same value; the numeric checker and the row checker are Mathlib-free themselves and are proved sound directly. Run as compiled programs, the certificate and numeric checkers accept every certificate tree and the objective, far and first-family data of every leaf in the corpus: 5591 trees and 4,196,879 boxes. The compiled row checker accepts every near row of the corpus except the 37 two-test rows of [Section 9.6](#), which it does not cover: 16,842 rows with 11,135,323 entries. The same three checkers accept the leaf program of [Theorem 12.6](#): its certificate (36 boxes), its objective and far data, and its near row (150 entries). So (c) and (f) apply to it as to the leaves of the case tree. In addition, seven certificates are checked by evaluation in the Lean kernel itself: six of the case tree (122 boxes) and that of [Theorem 12.6](#) (36 boxes).

The published inputs ([Sections 3, 7 and 12](#)) enter the formal statements as hypotheses. The following parts have conventional proofs in this paper but are *not* formally verified. Their numerical checks are as described here and in the relevant sections; in particular, a multiple-precision audit is distinct from a formally verified enclosure.

- the per-character costs of [Section 5](#): the localized envelope, which enters as a hypothesis, and the first-family bounds (the enclosures of J_{old} and J_{new} are verified, but not the inequalities they bound);
- the conductor coefficient $\frac{1}{4}$ for real characters and the refinement of [Section 7.6](#);
- the zero-location arguments and the case tree ([Sections 7 and 11](#)), that is, the statement that every configuration of zeros falls into a leaf whose zero-level hypotheses it satisfies, including the count, hidden-count and second-family constraints;
- the exterior regimes ([Section 12](#)): the range of [Proposition 12.1](#) and [Theorem 12.5](#), and, for [Theorem 12.6](#), the statement that every configuration with $\lambda_1 \geq 1.5$ satisfies the zero-level hypotheses of its program;
- the two-test row of [Section 9.6](#), which is checked by the replay and compared exactly with a separate implementation in (5), and whose constants were computed only by the interval routines of that implementation;
- the count-type integers of the leaf programs (checked in the audit of (4));
- the assembly of [Section 13](#).

A compiled program is trusted to compute what it is proved to compute; its correctness then rests on the Lean compiler and runtime and on the unverified routines that read the data files. The Lean proofs were written by AI agents, like the rest of this work (see the note after the abstract). The Lean kernel checks the proofs, but whether each formal statement says what the corresponding statement of this paper says has so far been reviewed only by AI agents.

14.4. Data availability. The development repository contains the certificates, the programs that generate and check them, the verification reports, and the Lean formalization:

<https://github.com/enaslund/linniks-constant>.

At the date of this revision the repository is private. A public, versioned archive of these materials is needed before submission so that readers can reproduce the computations and identify the

precise data used in the proof. This paper and the Lean formalization are publicly available in the repository prepared for the Palomar registry of formalized mathematics:

<https://github.com/enaslund/linniks-constant-3.99>.

The repository includes instructions for rerunning the replay, mutation tests, numerical audits, and formal checks. On the machine used, the replay took 1.7 hours of wall-clock time with parallel workers, and the compiled Lean checks of all certificates, leaf data and near rows took about 5 hours of wall-clock time. One caveat concerns exact reproducibility: the sieve heights of the graded rows (Section 9.2) are regenerated in binary64 floating point and are not stored in the records. Any nonnegative heights are admissible, so this does not affect soundness, but on a platform whose elementary functions round differently the regenerated rows, and hence the certificates that use them, could differ and would have to be recomputed. Thus the reproducibility requirement concerns the particular stored certificates; admissibility of the analytic sieve weight alone does not guarantee that a different regenerated row will pass those certificates.

15. CONCLUDING REMARKS

The difficult configurations. The tightest certified cases have a nonreal first character with $\lambda_1 \in [0.64, 0.80]$ and a nearby second family. The numerical experiments of Section 1.7 suggest that stronger location bounds in this range would be useful. One limitation of a single near row can be seen directly: if N entries have the same positive feature v and diagonal D , then its quadratic inequality permits $N \leq D/(v^2 - d)$ when $v^2 > d$ (Example 8.9). This is a restriction of that relaxation, not a lower bound on what other tests or arguments can achieve.

Improved far-density constants may also help. Xylouris estimates [145, p. 87] that halving the constant in his (5.19) would lower the exponent obtained by his method to about 4.6. A corresponding gain for the present method would require new certificates for the complete cover.

Possible improvements. Heath-Brown's list of possible improvements [51, §16] contains several that we have not used: optimal test functions for the combined inequalities (his item 1), positivity beyond the support of the test functions via Brun–Titchmarsh bounds (item 3; see [129, 86, 60, 81]), and modified sieve coefficients in the density argument (item 6; compare [35]). His item 7, a continuous weight in the density argument, is what Xylouris's far-density lemma implements, and we use it. For cube-free moduli Burgess's bounds hold for every k , and even the Weyl bound is known [100]; Heath-Brown already noted that all the arguments improve in that case. Recent large-value estimates for Dirichlet polynomials [44, 17, 122] and new density theorems [105] improve zero-density estimates away from $\sigma = 1$; it is not clear whether they help in the range $\sigma = 1 - O(1/\log q)$ that matters here.

Other approaches. For moduli with bounded cubic part the exponent 4.5 has been claimed [83], and for moduli all of whose prime factors are small much better exponents are known [16]. A different recent approach, through the exceptional set of Goldbach's problem (compare [104]), claims the exponent 5 in a preprint [148].

The role of the computation. The graded lemma gives quadratic constraints for different test functions and anchors. Their threshold forms can be combined with the far-density budget in linear programs, whose bounds are certified by exact duality. This separates the search for effective test parameters from the verification of the resulting bound. Any improvement of the global exponent must still include the zero-location arguments, a complete case cover, and the exterior regimes.

APPENDIX A. THE COMPLEX LOCATION ROWS

The 94 rows below make the implications of Propositions 7.12 and 7.13 explicit. In each row, the hypothesis is that χ_1 is nonreal and $\lambda_1 \in [a, b]$. The number h is the resulting strict lower

bound for λ_2 or λ' , as specified in the caption. The parameters γ and t select the test function (7.1) and the polynomial P_t .

Each margin is the interval-certified amount, divided by $f(0)$, by which the necessary inequality fails at error tolerance 0; the displayed margins are rounded for readability. Their positivity allows a common sufficiently large q . For the λ_2 rows, the real-second-character case uses the test parameter γ_r and $t' = 0.85$; the column labelled “alias” gives the margin in (7.3). The λ' rows use the parent lower bound described after Proposition 7.13 and assume $\text{ord } \chi_1 \geq 5$. For orders 3 and 4, the stronger published bound quoted there applies.

Table 13: Rows of Proposition 7.12: $\lambda_1 \in [a, b]$ implies $\lambda_2 > h$.

$[a, b]$	h	γ	t	γ_r	generic	real	alias
[.66, .6625]	.763	1.3	.95	1.15	0.00146	0.0323	0.1637
[.6625, .665]	.762	1.3	.95	1.15	0.00084	0.0314	0.1639
[.665, .6675]	.76	1.3	.9	1.15	0.00114	0.0311	0.1644
[.6675, .67]	.758	1.3	.9	1.15	0.00148	0.0308	0.1647
[.67, .6725]	.757	1.3	.9	1.15	0.00091	0.0299	0.1649
[.6725, .675]	.755	1.3	.9	1.15	0.00126	0.0296	0.1651
[.675, .6775]	.753	1.3	.9	1.15	0.00162	0.0293	0.1654
[.6775, .68]	.752	1.3	.9	1.15	0.00106	0.0284	0.1655
[.68, .6825]	.75	1.3	.9	1.15	0.00143	0.0281	0.1658
[.6825, .685]	.749	1.3	.9	1.15	0.00088	0.0272	0.1658
[.685, .6875]	.747	1.3	.9	1.15	0.00126	0.0269	0.1658
[.6875, .69]	.745	1.3	.9	1.15	0.00164	0.0266	0.1658
[.69, .6925]	.744	1.3	.9	1.15	0.00110	0.0258	0.1658
[.6925, .695]	.742	1.3	.9	1.15	0.00150	0.0254	0.1658
[.695, .6975]	.741	1.3	.9	1.15	0.00095	0.0246	0.1658
[.6975, .7]	.739	1.3	.9	1.1	0.00136	0.0243	0.1657
[.7, .7025]	.738	1.3	.9	1.1	0.00082	0.0236	0.1657
[.7025, .705]	.736	1.3	.9	1.1	0.00124	0.0234	0.1657
[.705, .7075]	.734	1.3	.9	1.1	0.00167	0.0232	0.1657
[.7075, .71]	.733	1.3	.9	1.1	0.00114	0.0225	0.1657
[.71, .7125]	.731	1.3	.9	1.1	0.00158	0.0223	0.1657
[.7125, .715]	.73	1.3	.9	1.1	0.00106	0.0216	0.1657
[.715, .7175]	.728	1.25	.95	1.1	0.00146	0.0214	0.1657
[.7175, .72]	.727	1.25	.95	1.1	0.00105	0.0207	0.1657
[.72, .7225]	.725	1.25	.95	1.1	0.00155	0.0206	0.1657

Table 14: Rows of Proposition 7.13: $\lambda_1 \in [a, b]$ implies $\lambda' > h$ (for χ_1 of order at least 5).

$[a, b]$	h	γ	t	margin	$[a, b]$	h	γ	t	margin
[.68, .6825]	.975	1.15	.95	0.00179	[.77, .7725]	.908	1.1	.95	0.00169
[.6825, .685]	.973	1.15	.95	0.00175	[.7725, .775]	.906	1.1	.95	0.00187
[.685, .6875]	.971	1.15	.95	0.00171	[.775, .7775]	.904	1.1	.95	0.00205
[.6875, .69]	.968	1.15	.95	0.00228	[.7775, .78]	.903	1.1	.95	0.00161
[.69, .6925]	.966	1.15	.95	0.00225	[.78, .7825]	.901	1.1	.95	0.00180
[.6925, .695]	.964	1.15	.95	0.00222	[.7825, .785]	.899	1.1	.95	0.00199
[.695, .6975]	.962	1.15	.95	0.00220	[.785, .7875]	.897	1.1	.95	0.00219
[.7, .7025]	.958	1.15	.95	0.00216	[.7875, .79]	.896	1.1	.95	0.00175
[.7025, .705]	.956	1.15	.95	0.00215	[.79, .7925]	.894	1.1	.95	0.00196
[.705, .7075]	.954	1.15	.95	0.00214	[.7925, .795]	.892	1.1	.95	0.00217
[.7075, .71]	.952	1.15	.95	0.00213	[.795, .7975]	.891	1.1	.95	0.00173
[.71, .7125]	.95	1.15	.9	0.00214	[.7975, .8]	.889	1.1	.95	0.00195
[.7125, .715]	.948	1.15	.9	0.00218	[.8, .8025]	.887	1.1	.95	0.00217
[.715, .7175]	.946	1.15	.9	0.00222	[.8025, .805]	.886	1.1	.95	0.00174

$[a, b]$	h	γ	t	margin	$[a, b]$	h	γ	t	margin
[.7175, .72]	.945	1.1	.95	0.00170	[.805, .8075]	.884	1.1	.9	0.00197
[.72, .7225]	.943	1.1	.95	0.00181	[.8075, .81]	.883	1.1	.9	0.00157
[.7225, .725]	.941	1.1	.95	0.00193	[.81, .8125]	.881	1.1	.9	0.00184
[.725, .7275]	.939	1.1	.95	0.00205	[.8125, .815]	.879	1.1	.9	0.00212
[.7275, .73]	.937	1.1	.95	0.00217	[.815, .8175]	.878	1.1	.9	0.00173
[.73, .7325]	.936	1.1	.95	0.00170	[.8175, .82]	.876	1.1	.9	0.00202
[.7325, .735]	.934	1.1	.95	0.00183	[.82, .8225]	.875	1.1	.9	0.00163
[.735, .7375]	.932	1.1	.95	0.00196	[.8225, .825]	.873	1.1	.9	0.00192
[.7375, .74]	.93	1.1	.95	0.00210	[.825, .8275]	.872	1.1	.9	0.00153
[.74, .7425]	.928	1.1	.95	0.00223	[.8275, .83]	.87	1.1	.9	0.00183
[.7425, .745]	.927	1.1	.95	0.00177	[.83, .8325]	.868	1.1	.9	0.00214
[.745, .7475]	.925	1.1	.95	0.00192	[.8325, .835]	.867	1.1	.9	0.00175
[.7475, .75]	.923	1.1	.95	0.00207	[.835, .8375]	.865	1.1	.9	0.00207
[.75, .7525]	.921	1.1	.95	0.00222	[.8375, .84]	.864	1.1	.9	0.00169
[.7525, .755]	.92	1.1	.95	0.00176	[.84, .8425]	.862	1.1	.9	0.00200
[.755, .7575]	.918	1.1	.95	0.00192	[.8425, .845]	.861	1.1	.9	0.00163
[.7575, .76]	.916	1.1	.95	0.00208	[.845, .8475]	.859	1.1	.9	0.00195
[.76, .7625]	.915	1.1	.95	0.00163	[.8475, .85]	.858	1.1	.9	0.00158
[.7625, .765]	.913	1.1	.95	0.00179	[.85, .8525]	.856	1.05	.95	0.00193
[.765, .7675]	.911	1.1	.95	0.00196	[.8525, .855]	.855	1.05	.95	0.00166
[.7675, .77]	.909	1.1	.95	0.00214					

APPENDIX B. THE FIXED CONSTANTS

This appendix collects the constants and parameters of the proof. The error budget that they serve is summarized in [Section 2.7](#).

Table 15: Fixed constants and parameters. The last column gives the stage of the order of choices of [Section 13.1](#) in which the constant is fixed; a dash marks a quantity computed from earlier ones. Apart from the tolerances $\varepsilon = \eta H_0$ and $\varepsilon = \eta J^2$, the constants of stages (S2)–(S6) are not computed; they exist by the arguments cited. Decimals are exact rational numbers unless they end with an ellipsis.

Symbol	Value	Role	Where fixed	Stage
<i>Exponent and kernel</i>				
L	3.99	The exponent	Section 2	(S1)
T	0.416829	Support of ψ is $[0, T]$; h is supported in $[L - 2T, L]$	Section 4.1	(S1)
κ	$T/16 = 0.0260518125$	Width of the 16 steps of ψ	Section 4.1	(S1)
$b_0, \dots, b_{15}$	Table 3	Step heights of ψ	Table 3	(S1)
A	$L - 2T = 3.156342$	Lower end of the support of h ; primes are found in (q^A, q^L) ; $A > 3$ and $A > 2x$	Section 4.1 , Lemma 6.2	–
$\Psi(0)$	$\kappa \sum_i b_i = 0.1936422928 \dots$	Integral of ψ	Section 4.1	–
H_0	$\Psi(0)^2 = 0.0374973375 \dots$	$H(0)$; normalizes the zero sum W	(4.1)	–
<i>Tolerances</i>				
η	10^{-6}	Detection needs $W \leq 1 - 2\eta$; far budget $(1 + \eta)V$; allowance 5η	Section 2 , Proposition 4.4	(S1)
η'	10^{-6}	Tolerance of the near rows: in every feature, diagonal and correlation term	Section 9.2	(S1)

Table 15, continued

Symbol	Value	Role	Where fixed	Stage
ε_1	$\eta/23$	Envelope error $\varepsilon_1 e^{-A\lambda}$ per character; $23 = 19 + 4$, where $19 > K_{\text{far}}$ bounds $\sum_{\chi} e^{-A\lambda_{\chi}}$ over the ordinary characters and 4 counts the first and reserved characters	Section 5.8	(S1)
final/ S	5η	Allowance included in every leaf value	Section 11.5	(S1)
ε in Input 4.2	ηH_0	Tolerance of the positivity criterion	Proposition 4.4	(S5)
ε in Input 6.1	ηJ^2	Tolerance of the far-density lemma; gives (6.1)	Section 6	(S5)
<i>Analytic constants</i>				
λ_{11}	0.05	Least anchor of the envelope; the anchor set $\mathcal{A}$ lies in $[\lambda_{11}, C_P]$	Section 5.3	(S1)
$s_{\max}$	3	Largest anchor of an ordinary entry; K_0 counts zeros with $\lambda \leq s_{\max}$	Section 6.3	(S1)
C_P	$\geq \max(C_0(\eta H_0), 3)$	Size of the prime window R_P ; C_0 is the constant of Input 4.2	(4.2)	(S2)
N	$N(C_P)$	Bound for the zeros of one character in R_P	Lemma 5.3	(S3)
$h_{\mathcal{A}}, \mathcal{A}$	Mesh with $ B_{\phi}(x) - B_{\phi}(y) \leq \varepsilon_1/4$ for $ x - y \leq h_{\mathcal{A}}$	Finite anchor set: the grid $\lambda_{11} + h_{\mathcal{A}}\mathbb{Z}$ and the anchors a, p^* , p° of the leaves	Section 5.3	(S3)
θ (smoothing)	In $(0, 1)$, with (5.1)	L^1 -distance of the smooth step function ψ^{θ} from ψ	Section 5.2	(S3)
c_{out}	$C_P f_p^{\theta}(0) + (f_p^{\theta})'(0) + e^{C_P T} \ (f_p^{\theta})''\ _1$, maximized over the anchors p	Cost of the distinguished terms of an outside first family	Proposition 5.9	(S3)
K_0	Independent of q	Bound for the zeros with $\lambda \leq s_{\max}$ and $ \gamma \leq 2$	Section 6.3	(S4)
M	Large	Height separation for Lemmas 6.3 and 8.7; depends on K_0 and $\eta''_{\min}$	Section 13.1	(S4)
C_B	$> \max(C_P + M, 1.5)$, with $4c_{\text{out}}/(H_0 C_B^2) \leq \eta$	Size of the buffer R_B , which separates inside from outside	Section 2, Section 13.1	(S4)
δ_0	In $(0, \frac{1}{3}]$	Common disc radius of the explicit formulas	Section 5.3	(S5)
q_0	Not computed	Maximum of the finitely many thresholds	Section 13.1	(S6)
<i>Far density</i> (profile I; profile II)				
J, ε_0	10, 10^{-7}	Number of weights α_i ; offset in w_0 (both profiles)	Section 6	(S1)
c_1	0.09035; 0.0821922	Parameter of Input 6.1	Section 6	(S1)
c_2	0.235968; 0.2170903	Parameter of Input 6.1	Section 6	(S1)
θ_{far}	1.28683; 1.4964274	Exponent in $w_0(t)^2$	Section 6	(S1)
$\alpha_1, \dots, \alpha_{10}$	Table 4	Weights of Input 6.1, of sum 1	Table 4	(S1)
x	1.1736846...; 1.1303335...	Upper end $\frac{2}{3} + 3c_1 + c_2$ of the profile; $2x < A$	Lemma 6.2	—

Table 15, continued

Symbol	Value	Role	Where fixed	Stage
V	175.26640331...; 243.32098105...	Right side of Input 6.1 with $\varepsilon = 0$; far budget $(1 + \eta)V$	Section 6	–
$w(0)$	10.7054...; 13.1730...	Far weight at $\lambda = 0$	Section 6	–
K_{far}	$(1 + \eta)V/w(0)$: < 16.38 ; < 18.48	Bound for $\sum_{\mathbf{x}} e^{-A\lambda_{\mathbf{x}}}$	(6.2)	–
<i>Leaf programs and certificates</i>				
S	10^{16}	Scale of the stored coefficients	Section 2 , Section 10.1	(S1)
T_S	10^6	Thresholds are measured in ticks of $1/T_S$	Section 10	(S1)
D_S	10^{12}	Scale of the dual multipliers	Section 10	(S1)
den	One of 200, 400, 500, 800, 2000 per leaf	Ordinary bins between r and $R = \max(3, r)$ on the grid $1/\text{den}$	Section 11.5	(S1)
Reserved grid	$1/400$	Grid of the columns of a reserved second family	Section 11.5	(S1)
<i>Near rows</i>				
ε'	$1/2000$	Sieve levels $\vartheta_k = \frac{1}{2}(t_k - \frac{1}{3}) - \varepsilon'$	Section 9.2	(S1)
Cells	2000 sieve cells; 4000 cells for I_u	Equal cells $t_0 < \dots < t_{2000} = 2\gamma_f$; upper Riemann sum for I_B	Section 9.2	(S1)
h_k, μ_{row}	Rational, denominators at most 10^{12} ; $\mu_{\text{row}} = 10^9$ in family and shifted rows	Sieve heights; they vanish in family and shifted rows	Section 9.2	(S1)
$[\delta]_{1/200}$	Grid $1/200$	Offsets at which the diagonal $D(\delta)$ is evaluated	Section 9.2	(S1)
s_1	$[50a]/50$	Safe anchor of a leaf with first-zero cell $[a, b]$	Section 9.3	(S1)
s (shift)	$\min(1.9, p^*, l_2)$	Anchor of the shifted row and of the two-test row	Section 9.5 , Section 11.5	(S1)
Graded anchors	1.5, 1.6, 1.9, 2.1, 2.3	Reference anchors s of the graded rows used	Section 9.5	(S1)
<i>Exterior regimes</i>				
u_{exc}	$\min(1/\eta_{\text{HB}}(\frac{1}{2}), 0.08)$	Below it, Proposition 12.1 ($P(a, q) \leq q^{3.5}$); not computed	Section 12	(S1)
K, A_s	0.1821; $L - 2K = 3.6258$	Single triangle $h_{L,K}$ of the small branch	Theorem 12.5	(S1)
c_1, c_2, ϕ	0.057, 0.1554, $\frac{1}{3}$	Parameters of Input 12.4	Theorem 12.5	(S1)
a_d, b_d, V_d	$\frac{3724}{1875}, \frac{671}{750}, \frac{184750}{6993}$	Exponents and constant of Input 12.4	Theorem 12.5	–
α	$1.09 < \frac{12}{11}$	Exponent in the bounds $\lambda', \lambda_2 \geq \alpha \log(1/\lambda_1)$ on $[u_{\text{exc}}, 0.08]$	Theorem 12.5	(S1)
m^*	$\alpha \log 12.5 = 2.7530442\dots$	Anchor for the other characters on $[u_{\text{exc}}, 0.08]$	Theorem 12.5	–
m_2, m'_2	2.829, 4.959	Bounds $\lambda_2 \geq 2.83 - \varepsilon$, $\lambda' \geq 4.96 - \varepsilon$ with $\varepsilon = 0.001$, on $[0.08, 0.1]$	Theorem 12.5	(S1)
B_1	$K^2 + \frac{K}{4}$	Bound for $\mathcal{B}_{1/4}$ for the first character	Theorem 12.5	–
Large branch	$s = s_1 = \frac{3}{2}$, $\gamma_f = 1$, $\gamma_g = \frac{3}{2}$, $t_0 = 1$	One near row; bins of width $\frac{1}{100}$ on $[1.5, 3)$ and a tail	Theorem 12.6	(S1)

REFERENCES

1. David L. Applegate, William Cook, Sanjeeb Dash, and Daniel G. Espinoza, *Exact solutions to linear programming problems*, Oper. Res. Lett. **35** (2007), no. 6, 693–699.
2. Jeremy Avigad, *Mathematics and the formal turn*, Bull. Amer. Math. Soc. (N.S.) **61** (2024), no. 2, 225–240.
3. Jeremy Avigad, Kevin Donnelly, David Gray, and Paul Raff, *A formally verified proof of the prime number theorem*, ACM Trans. Comput. Log. **9** (2007), no. 1, Art. 2.
4. Eric Bach and Jonathan Sorenson, *Explicit bounds for primes in residue classes*, Math. Comp. **65** (1996), no. 216, 1717–1735.
5. Chris Bailey, *nanoda.lib*, GitHub repository `ammkrn/nanoda.lib`, 2020, an external type checker for Lean 4.
6. William D. Banks and Igor E. Shparlinski, *Bounds on short character sums and L -functions with characters to a powerful modulus*, J. Anal. Math. **139** (2019), no. 1, 239–263.
7. M. B. Barban, Yu. V. Linnik, and N. G. Chudakov, *On prime numbers in an arithmetic progression with a prime-power difference*, Acta Arith. **9** (1964), no. 4, 375–390.
8. Kübra Benli, Shivani Goel, Henry Twiss, and Asif Zaman, *Explicit Deuring–Heilbronn phenomenon for Dirichlet L -functions*, Proc. Amer. Math. Soc. **154** (2026), no. 2, 509–525.
9. Michael A. Bennett, Greg Martin, Kevin O’Byrant, and Andrew Rechnitzer, *Explicit bounds for primes in arithmetic progressions*, Illinois J. Math. **62** (2018), no. 1–4, 427–532.
10. Enrico Bombieri, *Le grand crible dans la théorie analytique des nombres*, second ed., Astérisque, vol. 18, Société Mathématique de France, Paris, 1987 (French), First edition 1974.
11. D. A. Burgess, *On character sums and L -series*, Proc. London Math. Soc. (3) **12** (1962), 193–206.
12. ———, *On character sums and primitive roots*, Proc. London Math. Soc. (3) **12** (1962), 179–192.
13. ———, *On character sums and L -series. II*, Proc. London Math. Soc. (3) **13** (1963), 524–536.
14. ———, *The character sum estimate with $r = 3$* , J. London Math. Soc. (2) **33** (1986), no. 2, 219–226.
15. Emanuel Carneiro, Micah B. Milinovich, Emily Quesada-Herrera, and Antonio Pedro Ramos, *Fourier optimization, the least quadratic non-residue, and the least prime in an arithmetic progression*, Math. Comp. (2025), to appear; published electronically 2 December 2025, doi:10.1090/mcom/4154.
16. Mei-Chu Chang, *Short character sums for composite moduli*, J. Anal. Math. **123** (2014), no. 1, 1–33.
17. Bin Chen, Vishal Gupta, and Yung Chi Li, *Large value estimates for Dirichlet polynomials with characters and zero density of Dirichlet L -functions*, 2025, arXiv:2507.08296.
18. Jingrun Chen, *On the least prime in an arithmetical progression*, Sci. Sinica **14** (1965), 1868–1871. MR 188172
19. ———, *On the least prime in an arithmetical progression and two theorems concerning the zeros of Dirichlet’s L -functions*, Sci. Sinica **20** (1977), no. 5, 529–562. MR 476668
20. ———, *On the least prime in an arithmetical progression and theorems concerning the zeros of Dirichlet’s L -functions. II*, Sci. Sinica **22** (1979), no. 8, 859–889. MR 549597
21. Jingrun Chen and Jianmin Liu, *On the least prime in an arithmetical progression. III*, Sci. China Ser. A **32** (1989), no. 6, 654–673. MR 1056044
22. ———, *On the least prime in an arithmetical progression. IV*, Sci. China Ser. A **32** (1989), no. 7, 792–807. MR 1058000
23. ———, *On the least prime in an arithmetical progression and theorems concerning the zeros of Dirichlet’s L -functions (V)*, International Symposium in Memory of Hua Loo Keng, Vol. I: Number Theory (Sheng Gong, Qi-Keng Lu, Yuan Wang, and Lo Yang, eds.), Springer, Berlin, 1991, pp. 19–42.
24. S. Chowla, *On the least prime in an arithmetical progression*, J. Indian Math. Soc. (N.S.) **1** (1934), 1–3.
25. Vašek Chvátal, *Linear programming*, A Series of Books in the Mathematical Sciences, W. H. Freeman and Company, New York and San Francisco, 1983.
26. Harold Davenport, *Multiplicative number theory*, third ed., Graduate Texts in Mathematics, vol. 74, Springer-Verlag, New York, 2000, Revised and with a preface by Hugh L. Montgomery.
27. Ch.-J. de la Vallée Poussin, *Recherches analytiques sur la théorie des nombres premiers*, Ann. Soc. Sci. Bruxelles **20** (1896), 183–256, 281–397 (French).
28. ———, *Sur la fonction $\zeta(s)$ de Riemann et le nombre des nombres premiers inférieurs à une limite donnée*, Mém. Couronnés et Autres Mém. Publ. Acad. Roy. Sci. Lett. Beaux-Arts Belg. **59** (1899), 1–74 (French).
29. Leonardo de Moura and Sebastian Ullrich, *The Lean 4 theorem prover and programming language*, Automated Deduction – CADE 28 (André Platzer and Geoff Sutcliffe, eds.), Lecture Notes in Comput. Sci., vol. 12699, Springer, 2021, pp. 625–635.
30. Max Deuring, *Imaginäre quadratische Zahlkörper mit der Klassenzahl 1*, Math. Z. **37** (1933), no. 1, 405–415.
31. P. G. Lejeune Dirichlet, *Beweis des Satzes, dass jede unbegrenzte arithmetische Progression, deren erstes Glied und Differenz ganze Zahlen ohne gemeinschaftlichen Factor sind, unendlich viele Primzahlen enthält*, G.

- Lejeune Dirichlet's Werke, Vol. 1 (L. Kronecker, ed.), Reimer, Berlin, 1889, Reissued by Cambridge University Press, 2012, pp. 313–342.
32. E. Fogels, *On the zeros of L -functions*, Acta Arith. **11** (1965), no. 1, 67–96.
 33. J. B. Friedlander and H. Iwaniec, *Exceptional characters and prime numbers in arithmetic progressions*, Int. Math. Res. Not. **2003** (2003), no. 37, 2033–2050.
 34. ———, *Opera de cribro*, Amer. Math. Soc. Colloq. Publ., vol. 57, Amer. Math. Soc., Providence, RI, 2010.
 35. ———, *Selberg's sieve of irregular density*, Acta Arith. **209** (2023), 385–396.
 36. ———, *Sifting for small primes from an arithmetic progression*, Sci. China Math. **66** (2023), no. 12, 2715–2730.
 37. P. X. Gallagher, *A large sieve density estimate near $\sigma = 1$* , Invent. Math. **11** (1970), no. 4, 329–339.
 38. ———, *Primes in progressions to prime-power modulus*, Invent. Math. **16** (1972), no. 3, 191–201.
 39. S. W. Graham, *Applications of sieve methods*, Ph.D. thesis, University of Michigan, 1977, 187 pp., ProQuest 7804710. MR 2627480
 40. ———, *An asymptotic estimate related to Selberg's sieve*, J. Number Theory **10** (1978), no. 1, 83–94.
 41. ———, *On Linnik's constant*, Acta Arith. **39** (1981), no. 2, 163–179.
 42. Andrew Granville and Carl Pomerance, *On the least prime in certain arithmetic progressions*, J. London Math. Soc. (2) **41** (1990), no. 2, 193–200.
 43. T. H. Gronwall, *Sur les séries de Dirichlet correspondant à des caractères complexes*, Rend. Circ. Mat. Palermo **35** (1913), 145–159.
 44. Larry Guth and James Maynard, *New large value estimates for Dirichlet polynomials*, Ann. of Math. (2) **203** (2026), no. 2, 623–675.
 45. H. Halberstam and H.-E. Richert, *Sieve methods*, London Math. Soc. Monogr., vol. 4, Academic Press, London, 1974.
 46. Thomas Hales, Mark Adams, Gertrud Bauer, Tat Dat Dang, John Harrison, Le Truong Hoang, Cezary Kaliszyk, Victor Magron, Sean McLaughlin, Tat Thang Nguyen, Quang Truong Nguyen, Tobias Nipkow, Steven Obua, Joseph Pleso, Jason Rute, Alexey Solovyev, Thi Hoai An Ta, Nam Trung Tran, Thi Diep Trieu, Josef Urban, Ky Vu, and Roland Zumkeller, *A formal proof of the Kepler conjecture*, Forum Math. Pi **5** (2017), e2.
 47. John Harrison, *Formalizing an analytic proof of the prime number theorem*, J. Automat. Reason. **43** (2009), no. 3, 243–261.
 48. D. R. Heath-Brown, *Hybrid bounds for Dirichlet L -functions*, Invent. Math. **47** (1978), no. 2, 149–170.
 49. ———, *The density of zeros of Dirichlet's L -functions*, Canad. J. Math. **31** (1979), no. 2, 231–240.
 50. ———, *Siegel zeros and the least prime in an arithmetic progression*, Quart. J. Math. Oxford Ser. (2) **41** (1990), no. 4, 405–418.
 51. ———, *Zero-free regions for Dirichlet L -functions, and the least prime in an arithmetic progression*, Proc. London Math. Soc. (3) **64** (1992), no. 2, 265–338.
 52. ———, *Burgess's bounds for character sums*, Number Theory and Related Fields: In Memory of Alf van der Poorten (Jonathan M. Borwein, Igor Shparlinski, and Wadim Zudilin, eds.), Springer Proc. Math. Stat., vol. 43, Springer, New York, 2013, pp. 199–213.
 53. Hans Heilbronn, *On the class-number in imaginary quadratic fields*, Quart. J. Math. Oxford Ser. **5** (1934), no. 1, 150–160.
 54. Harald Andrés Helfgott, *The ternary Goldbach problem*, 2015, arXiv:1501.05438.
 55. Q. Huangfu and J. A. J. Hall, *Parallelizing the dual revised simplex method*, Math. Program. Comput. **10** (2018), no. 1, 119–142.
 56. M. N. Huxley, *On the difference between consecutive primes*, Invent. Math. **15** (1972), no. 2, 164–170.
 57. A. E. Ingham, *The distribution of prime numbers*, Cambridge Tracts in Mathematics and Mathematical Physics, vol. 30, Cambridge University Press, Cambridge, 1932.
 58. ———, *On the estimation of $N(\sigma, T)$* , Quart. J. Math. Oxford Ser. **11** (1940), no. 1, 201–202.
 59. H. Iwaniec, *On zeros of Dirichlet's L series*, Invent. Math. **23** (1974), no. 2, 97–104.
 60. ———, *On the Brun–Titchmarsh theorem*, J. Math. Soc. Japan **34** (1982), no. 1, 95–123.
 61. ———, *Conversations on the exceptional character*, Analytic Number Theory: Lectures given at the C.I.M.E. Summer School held in Cetraro, Italy, July 11–18, 2002, Lecture Notes in Math., vol. 1891, Springer, Berlin, 2006, pp. 97–132.
 62. H. Iwaniec and Emmanuel Kowalski, *Analytic number theory*, Amer. Math. Soc. Colloq. Publ., vol. 53, Amer. Math. Soc., Providence, RI, 2004.
 63. Christian Jansson, *Rigorous lower and upper bounds in linear programming*, SIAM J. Optim. **14** (2004), no. 3, 914–935.

64. Matti Jutila, *A new estimate for Linnik's constant*, Ann. Acad. Sci. Fenn. Ser. A I **471** (1970), 8 pp. MR 271056
65. ———, *On a density theorem of H. L. Montgomery for L -functions*, Ann. Acad. Sci. Fenn. Ser. A I **520** (1972), 13 pp. MR 327681
66. ———, *On Linnik's constant*, Math. Scand. **41** (1977), 45–62.
67. Habiba Kadir, *Une région explicite sans zéros pour la fonction ζ de Riemann*, Acta Arith. **117** (2005), no. 4, 303–339 (French).
68. ———, *Explicit zero-free regions for Dirichlet L -functions*, Mathematika **64** (2018), no. 2, 445–474.
69. Anatolij A. Karatsuba, *Basic analytic number theory*, Springer-Verlag, Berlin, 1993, Translated from the Russian by Melvyn B. Nathanson.
70. Bryce Kerr, *Moments of character sums to composite modulus*, 2019, arXiv:1904.04578.
71. S. Knapowski, *On Linnik's theorem concerning exceptional L -zeros*, Publ. Math. Debrecen **9** (1962), 168–178.
72. Alex Kontorovich, Terence Tao, et al., *Prime number theorem and ...*, GitHub repository AlexKontorovich/PrimeNumberTheoremAnd and blueprint (the PNT+ project), 2024.
73. Youness Lamzouri, Xiannan Li, and Kannan Soundararajan, *Conditional bounds for the least quadratic non-residue and related problems*, Math. Comp. **84** (2015), no. 295, 2391–2412, Corrigendum: Math. Comp. **86** (2017), no. 307, 2551–2554, doi:10.1090/mcom/3261.
74. E. Landau, *Über imaginär-quadratische Zahlkörper mit gleicher Klassenzahl*, Nachr. Ges. Wiss. Göttingen, Math.-Phys. Kl. (1918), 277–284.
75. Lean FRO, *Comparator*, GitHub repository leanprover/comparator, 2025, a trustworthy judge for Lean proofs.
76. Junxian Li, Kyle Pratt, and George Shakan, *A lower bound for the least prime in an arithmetic progression*, Quart. J. Math. **68** (2017), no. 3, 729–758.
77. Yu. V. Linnik, *On the least prime in an arithmetic progression. I. The basic theorem*, Rec. Math. [Mat. Sbornik] N.S. **15(57)** (1944), 139–178. MR 12111
78. ———, *On the least prime in an arithmetic progression. II. The Deuring–Heilbronn phenomenon*, Rec. Math. [Mat. Sbornik] N.S. **15(57)** (1944), 347–368. MR 12112
79. Ming-Chit Liu and Tianze Wang, *A numerical bound for small prime solutions of some ternary linear equations*, Acta Arith. **86** (1998), no. 4, 343–383.
80. Kaisa Matomäki, Jori Merikoski, and Joni Teräväinen, *Primes in arithmetic progressions and short intervals without L -functions*, 2024, arXiv:2401.17570.
81. James Maynard, *On the Brun–Titchmarsh theorem*, Acta Arith. **157** (2013), no. 3, 249–296.
82. Kevin S. McCurley, *Explicit zero-free regions for Dirichlet L -functions*, J. Number Theory **19** (1984), no. 1, 7–32.
83. Zaizhao Meng, *A note on the Linnik's constant*, 2010, arXiv:1010.3544.
84. H. L. Montgomery, *Topics in multiplicative number theory*, Lecture Notes in Math., vol. 227, Springer-Verlag, Berlin, 1971.
85. ———, *Ten lectures on the interface between analytic number theory and harmonic analysis*, CBMS Regional Conference Series in Mathematics, vol. 84, Amer. Math. Soc., Providence, RI, 1994.
86. H. L. Montgomery and R. C. Vaughan, *The large sieve*, Mathematika **20** (1973), no. 2, 119–134.
87. ———, *Multiplicative number theory I. Classical theory*, Cambridge Stud. Adv. Math., vol. 97, Cambridge University Press, Cambridge, 2007.
88. Ramon E. Moore, *Interval analysis*, Prentice-Hall Series in Automatic Computation, Prentice-Hall, Englewood Cliffs, N. J., 1966.
89. Ramon E. Moore, R. Baker Kearfott, and Michael J. Cloud, *Introduction to interval analysis*, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2009.
90. Michael J. Mossinghoff and Timothy S. Trudgian, *Nonnegative trigonometric polynomials and a zero-free region for the Riemann zeta-function*, J. Number Theory **157** (2015), 329–349.
91. Y. Motohashi, *On a density theorem of Linnik*, Proc. Japan Acad. **51** (1975), 815–817.
92. ———, *Primes in arithmetic progressions*, Invent. Math. **44** (1978), no. 2, 163–178.
93. ———, *Lectures on sieve methods and prime number theory*, Tata Inst. Fund. Res. Lectures on Math. and Phys., vol. 72, Tata Institute of Fundamental Research and Springer-Verlag, Berlin, 1983.
94. Władysław Narkiewicz, *The development of prime number theory: From Euclid to Hardy and Littlewood*, Springer Monographs in Mathematics, Springer-Verlag, Berlin, 2000.
95. Arnold Neumaier and Oleg Shcherbina, *Safe bounds in linear and mixed-integer linear programming*, Math. Program. **99** (2004), no. 2, 283–296.

96. Steven Obua and Tobias Nipkow, *Flyspeck II: the basic linear programs*, Ann. Math. Artif. Intell. **56** (2009), no. 3–4, 245–272.
97. A. Page, *On the number of primes in an arithmetic progression*, Proc. London Math. Soc. (2) **39** (1935), no. 1, 116–141.
98. Chengdong Pan, *On the least prime in an arithmetical progression*, Sci. Record (N.S.) **1** (1957), 311–313. MR 105398
99. ———, *On the least prime in an arithmetical progression*, Acta Sci. Natur. Univ. Pekinensis (1958), no. 1, 3–36 (Chinese).
100. Ian Petrow and Matthew P. Young, *The Weyl bound for Dirichlet L -functions of cube-free conductor*, Ann. of Math. (2) **192** (2020), no. 2, 437–486.
101. E. Phragmén and Ernst Lindelöf, *Sur une extension d'un principe classique de l'analyse et sur quelques propriétés des fonctions monogènes dans le voisinage d'un point singulier*, Acta Math. **31** (1908), 381–406.
102. J. Pintz, *Elementary methods in the theory of L -functions, V. The theorems of Landau and Page*, Acta Arith. **32** (1977), no. 2, 163–171.
103. ———, *Elementary methods in the theory of L -functions, VIII. Real zeros of real L -functions*, Acta Arith. **33** (1977), no. 1, 89–98.
104. ———, *A new explicit formula in the additive theory of primes with applications II. The exceptional set in Goldbach's problem*, 2018, arXiv:1804.09084.
105. ———, *Some new density theorems for Dirichlet L -functions*, Banach Center Publ. **118** (2019), 231–244.
106. D. J. Platt, *Numerical computations concerning the GRH*, Math. Comp. **85** (2016), no. 302, 3009–3027.
107. D. J. Platt and Tim Trudgian, *The Riemann hypothesis is true up to $3 \cdot 10^{12}$* , Bull. Lond. Math. Soc. **53** (2021), no. 3, 792–797.
108. G. Pólya, *Über die Verteilung der quadratischen Reste und Nichtreste*, Nachr. Ges. Wiss. Göttingen, Math.-Phys. Kl. (1918), 21–29.
109. Carl Pomerance, *A note on the least prime in an arithmetic progression*, J. Number Theory **12** (1980), no. 2, 218–223.
110. Karl Prachar, *Primzahlverteilung*, Grundlehren der Mathematischen Wissenschaften, vol. 91, Springer-Verlag, Berlin, 1957 (German).
111. K. A. Rodoskii, *On the least prime number in an arithmetic progression*, Mat. Sbornik N.S. **34(76)** (1954), 331–356. MR 62156
112. J. Barkley Rosser and Lowell Schoenfeld, *Approximate formulas for some functions of prime numbers*, Illinois J. Math. **6** (1962), no. 1, 64–94.
113. Siegfried M. Rump, *Verification methods: rigorous results using floating-point arithmetic*, Acta Numer. **19** (2010), 287–449.
114. Stelios Sachpazis, *A pretentious proof of Linnik's estimate for primes in arithmetic progressions*, Mathematika **69** (2023), no. 3, 879–902.
115. Alexander Schrijver, *Theory of linear and integer programming*, Wiley-Interscience Series in Discrete Mathematics, John Wiley & Sons, Chichester, 1986.
116. Atle Selberg, *On an elementary method in the theory of primes*, Norske Vid. Selsk. Forh., Trondhjem **19** (1947), no. 18, 64–67. MR 22871
117. ———, *Lectures on sieves*, Collected Papers, Vol. II, Springer-Verlag, Berlin, 1991, pp. 65–247.
118. C. L. Siegel, *Über die Classenzahl quadratischer Zahlkörper*, Acta Arith. **1** (1935), no. 1, 83–86.
119. Alexey Solovyev and Thomas C. Hales, *Efficient formal verification of bounds of linear programs*, Intelligent Computer Mathematics (Calculemus 2011 and MKM 2011, Bertinoro, Italy) (James H. Davenport, William M. Farmer, Josef Urban, and Florian Rabe, eds.), Lecture Notes in Comput. Sci., vol. 6824, Springer, 2011, pp. 123–132.
120. Kannan Soundararajan and Jesse Thorner, *Weak subconvexity without a Ramanujan hypothesis*, Duke Math. J. **168** (2019), no. 7, 1231–1268.
121. Michael Stoll, *Dirichlet's theorem on primes in arithmetic progression*, Lean 4 Mathlib source file `Mathlib/NumberTheory/LSeries/PrimesInAP.lean`, 2024.
122. Terence Tao, Tim Trudgian, and Andrew Yang, *New exponent pairs, zero density estimates, and zero additive energy estimates: a systematic approach*, Math. Comp. **95** (2026), no. 362, 2941–2990.
123. Tikao Tatuzawa, *On a theorem of Siegel*, Jpn. J. Math. **21** (1951), 163–178. MR 51262
124. Gérald Tenenbaum, *Introduction to analytic and probabilistic number theory*, third ed., Graduate Studies in Mathematics, vol. 163, Amer. Math. Soc., Providence, RI, 2015, Translated from the 2008 French edition by Patrick D. F. Ion. MR 3363366

125. The mathlib Community, *The Lean mathematical library*, Proceedings of the 9th ACM SIGPLAN International Conference on Certified Programs and Proofs (CPP 2020), ACM, 2020, arXiv:1910.09336, pp. 367–381.
126. The mpmath development team, *mpmath: a Python library for arbitrary-precision floating-point arithmetic (version 1.3.0)*, 2023, <https://mpmath.org/>.
127. Jesse Thorner and Asif Zaman, *An explicit version of Bombieri’s log-free density estimate and Sárközy’s theorem for shifted primes*, Forum Math. **36** (2024), no. 4, 1059–1080.
128. ———, *Refinements to the prime number theorem for arithmetic progressions*, Math. Z. **306** (2024), no. 3, Paper No. 54.
129. E. C. Titchmarsh, *A divisor problem*, Rend. Circ. Mat. Palermo **54** (1930), 414–429.
130. ———, *The theory of functions*, second ed., Oxford University Press, Oxford, 1939.
131. ———, *The theory of the Riemann zeta-function*, second ed., Clarendon Press, Oxford, 1986, Revised by D. R. Heath-Brown.
132. Warwick Tucker, *Validated numerics: A short introduction to rigorous computations*, Princeton University Press, Princeton, NJ, 2011.
133. P. Turán, *On a density theorem of Yu. V. Linnik*, Magyar Tud. Akad. Mat. Kutató Int. Közl. **6** (1961), 165–179. MR 146155
134. ———, *On some recent results in the analytical theory of numbers*, 1969 Number Theory Institute (State Univ. New York, Stony Brook, N.Y., 1969), Proc. Sympos. Pure Math., vol. 20, Amer. Math. Soc., Providence, RI, 1971, pp. 359–374. MR 316400
135. ———, *On a new method of analysis and its applications*, John Wiley & Sons, New York, 1984, a Wiley-Interscience Publication.
136. I. M. Vinogradov, *Sur la distribution des résidus et des nonrésidus des puissances*, J. Soc. Phys.-Math. Univ. Perm **1** (1918), 94–98.
137. Pauli Virtanen, Ralf Gommers, Travis E. Oliphant, Matt Haberland, Tyler Reddy, David Cournapeau, Evgeni Burovski, Pearu Peterson, Warren Weckesser, Jonathan Bright, et al., *SciPy 1.0: fundamental algorithms for scientific computing in Python*, Nature Methods **17** (2020), no. 3, 261–272.
138. Samuel S. Wagstaff, Jr., *Greatest of the least primes in arithmetic progressions having a given modulus*, Math. Comp. **33** (1979), no. 147, 1073–1080.
139. Wei Wang, *On the least prime in an arithmetic progression*, Acta Math. Sinica **29** (1986), no. 6, 826–836. MR 888852
140. ———, *On the least prime in an arithmetic progression*, Acta Math. Sinica (N.S.) **7** (1991), no. 3, 279–289.
141. André Weil, *Sur les “formules explicites” de la théorie des nombres premiers*, Comm. Sém. Math. Univ. Lund [Medd. Lunds Univ. Mat. Sem.] (1952), no. Tome supplémentaire, 252–265 (French). MR 53152
142. David Vernon Widder, *The Laplace transform*, Princeton Mathematical Series, vol. 6, Princeton University Press, Princeton, N. J., 1941.
143. Triantafyllos Xylouris, *On Linnik’s constant*, 2009, arXiv:0906.2749; Diplomarbeit, Universität Bonn (in German; title page: Über die Linniksche Konstante).
144. ———, *On the least prime in an arithmetic progression and estimates for the zeros of Dirichlet L-functions*, Acta Arith. **150** (2011), no. 1, 65–91.
145. ———, *Über die Nullstellen der Dirichletschen L-Funktionen und die kleinste Primzahl in einer arithmetischen Progression*, Dissertation, Universität Bonn, 2011, Bonner Mathematische Schriften 404.
146. ———, *Linniks Konstante ist kleiner als 5*, Chebyshevskii Sb. **19** (2018), no. 3, 80–94 (German).
147. Asif Zaman, *On the least prime ideal and Siegel zeros*, Int. J. Number Theory **12** (2016), no. 8, 2201–2229.
148. Genheng Zhao, *The exceptional set of Goldbach problem and Linnik’s constant*, 2025, arXiv:2511.05631.

Email address: `naslund.math@gmail.com`