Annals of Mathematics **181** (2015), 361–382
http://dx.doi.org/10.4007/annals.2015.181.1.6

Solution of the minimum modulus problem for covering systems

By Bob Hough

Abstract

We answer a question of Erdős by showing that the least modulus of a distinct covering system is at most 10^{16}.

1. Introduction

In 1934 Romanoff proved that the numbers of form a prime plus a power of two have positive lower density. Writing to Erdős, he asked whether there exists an arithmetic progression of odd numbers none of whose members is of this form. Erdős's positive answer to this question introduced the notion of a *distinct covering system of congruences*, which is a finite collection of congruences

$$a_i \bmod m_i, \qquad 1 < m_1 < m_2 < \cdots < m_k$$

such that every integer satisfies at least one of them. His paper [4] gives the example

$$0 \bmod 2, \quad 0 \bmod 3, \quad 1 \bmod 4, \quad 3 \bmod 8, \quad 7 \bmod 12, \quad 23 \bmod 24.$$

Erdős posed a number of problems concerning covering systems, of which two in particular are well known. From [4], the minimum modulus problem asks whether there exist distinct covering systems for which the least modulus is arbitrarily large. With Selfridge, Erdős asked if there exists a distinct covering system with all moduli odd. These two questions appear frequently in Erdős' collections of open problems [5], [6], [7], [8], [9]. See also [13].

Following Erdős' paper, a number of covering systems have been exhibited with increasing minimum modulus [3], [14], [2], [15], [12], with the current record of 40 due to Nielsen [16]. In [16], Nielsen suggests for the first time that

This research was begun while the author was a postdoctoral research fellow at Department of Pure Maths and Math Stats, Cambridge, and completed while a postdoctoral research fellow at the Mathematical Institute, Oxford. He is grateful for financial support from ERC Research Grant 279438, Approximate Algebraic Structure and Applications.

the answer to the minimum modulus problem may be negative. We confirm this conjecture.

THEOREM 1. *The least modulus of a distinct covering system is at most* 10^{16}.

To obtain the bound of 10^{16} we use some simple numerical calculations performed in Pari/GP [19], together with a standard explicit estimate for the counting function of primes. For the reader interested only in the qualitative statement that the minimum modulus has a uniform upper bound, our presentation is self-contained.

In the spirit of the odd modulus problem, Theorem 1 immediately implies that any covering system contains a modulus divisible by one of an initial segment of primes. We may return to give a stronger quantitative statement of this type at a later time.

Prior to our work, the main theoretical progress on the minimum modulus problem was made recently by Filaseta, Ford, Konyagin, Pomerance and Yu [11], who showed, among other results, a lower bound for the sum of the reciprocals of the moduli of a covering system that grows with the minimum modulus. We build upon their work. In particular, we use an inductive scheme in which we filter the moduli of the congruences according to the size of their prime factors, so that we first consider the subset of congruences all of whose prime factors are below an initial threshold, and we then increase the threshold in stages. The paper [11] roughly makes the first stage of this argument.

A detailed overview of our argument is given in the next section, but we mention here that our proof follows the probabilistic method in the sense that we give a positive lower bound for the density of integers left uncovered by any distinct system of congruences for which the minimum modulus is sufficiently large. The Lovász Local Lemma plays a crucial rôle. The suitability of the Local Lemma for estimating the density of the uncovered set at each stage of the argument relies upon a certain regularity of the uncovered set from the previous stage, and this regularity we are able to guarantee by applying the Local Lemma a second time, in a relative form.

Notation. Throughout we denote $\omega(n)$ the number of distinct prime factors of natural number n.

Acknowledgements. The author is grateful to Ben Green, who read an early version of this paper and made a number of suggestions that dramatically improved the structure and readability. The author is also grateful to Pace Nielsen, Kevin Ford and Michael Filaseta for detailed comments, and to K. Soundararajan and Persi Diaconis, from whom he learned many of the methods applied here. An anonymous referee pointed out a numerical improvement to the parameters that lowered the final bound.

2. **Overview**

We begin by giving a reasonably detailed overview of the argument. In this summary we will consider only congruence systems all of whose moduli are square free. Treating the case of general moduli involves a minor complication, which we address in the next section.

Let $M > 1$, and let

$$\mathscr{M} \subset \{m \in \mathbb{N} : m \text{ square free}, \ m > M\}$$

be a finite set of moduli. We assume that for each $m \in \mathscr{M}$, a residue class $a_m \bmod m$ has been given. For M sufficiently large, we argue that for any $\mathscr{M}$, and for any assignment of the a_m, we can give a positive lower bound for the density of solutions to the system of (non)congruences

$$R = \{z \in \mathbb{Z} : \ \forall m \in \mathscr{M}, \ z \not\equiv a_m \bmod m\}.$$

The bound will, of course, depend upon $\mathscr{M}$.

We estimate the density of R in stages, so we introduce a sequence of thresholds $1 = P_{-1} < P_0 < P_1 < \cdots$ with $P_i \to \infty$. For the purpose of this summary we assume that P_0 is sufficiently small so that $\prod_{p \le P_0} p < M$, although to get a better bound for M, we will in practice choose P_0 to be somewhat larger. Let $1 = Q_{-1}, Q_0, Q_1, \ldots$ be such that

$$Q_i = \prod_{p \le P_i} p, \qquad i \ge 0.$$

We say that a number n is P_i-smooth if $n | Q_i$. Let $\mathscr{M}_0, \mathscr{M}_1, \ldots$ be given by

$$\mathscr{M}_i = \{m \in \mathscr{M} : m | Q_i\}, \qquad i \ge 0;$$

that is, $\mathscr{M}_i$ is the set of P_i-smooth moduli in $\mathscr{M}$. In particular, by our assumption on P_0 we have that $\mathscr{M}_0$ is empty. For this reason we set $R_0 = R_{-1} = \mathbb{Z}$ and consider the sequence of unsifted sets $R_0 \supset R_1 \supset R_2 \supset \cdots$

$$R_i = \bigcap_{m \in \mathscr{M}_i} \{z \in \mathbb{Z} : z \not\equiv a_m \bmod m\}, \qquad i \ge 1.$$

Since the sets $\mathscr{M}_i$ grow to exhaust $\mathscr{M}$, we eventually have $R = R_i$, and so it will suffice to prove that the density of R_i is nonzero for each i. This lower bound we will give in a uniform way for all congruence systems with minimum modulus greater than M.

We may view R_i as a subset of $\mathbb{Z}/Q_i\mathbb{Z}$. Thinking of $\mathbb{Z}/Q_{i+1}\mathbb{Z}$ as fibred over $\mathbb{Z}/Q_i\mathbb{Z}$, we then have that R_{i+1} is contained in fibres over R_i, and we may estimate the density of R_{i+1} by estimating its density in individual fibres over R_i. In fact, we only consider some 'good' fibres over a 'well-distributed' subset of R_i. Thus we do not actually estimate the density of R_{i+1}, but rather that of a somewhat smaller set. Also, rather than explicitly estimate the density of the smaller set, we will check that the smaller set is nonempty and then estimate some statistics related to it.

Let $i \geq 0$, and let $r \in R_i \bmod Q_i$. By definition, r has survived sieving by all of the congruences to moduli dividing Q_i, so that the fraction of the fibre $(r \bmod Q_i)$ that survives into R_{i+1} is determined by congruence conditions to moduli in $\mathscr{M}_{i+1} \setminus \mathscr{M}_i$. Each such modulus m has a unique factorization as $m = m_0 n$ with $m_0 | Q_i$ and n composed of primes in the interval $(P_i, P_{i+1}]$. We call the collection of such n the set of 'new factors'

$$\forall i \geq 0, \qquad \mathscr{N}_{i+1} = \{n \in \mathbb{N} : n > 1, n \text{ square free}, p|n \Rightarrow p \in (P_i, P_{i+1}]\}.$$

This set will play a very important rôle in what follows.

Given $r \in R_i \bmod Q_i$, $a_{m_0 n} \bmod m_0 n$ intersects $(r \bmod Q_i)$ if and only if $a_{m_0 n} \equiv r \bmod m_0$. If this condition is met, the effect within the fibre is determined only by $a_{m_0 n} \bmod n$. For this reason, we group together the congruence conditions according to common r and n: for each $r \in \mathbb{Z}/Q_i\mathbb{Z}$ and each $n \in \mathscr{N}_{i+1}$, we set

$$A_{n,r} = (r \bmod Q_i) \cap \bigcup_{m_0 | Q_i, m_0 n \in \mathscr{M}} (a_{m_0 n} \bmod m_0 n).$$

We then have

$$\forall i \geq 0, \qquad (r \bmod Q_i) \cap R_{i+1} = (r \bmod Q_i) \cap \bigcap_{n \in \mathscr{N}_{i+1}} A_{n,r}^c,$$

with the interpretation that R_{i+1} within $(r \bmod Q_i)$ results from sieving $(r \bmod Q_i)$ by sets of residues to moduli in $\mathscr{N}_{i+1}$.

When $n_1, n_2 \in \mathscr{N}_{i+1}$ are coprime, sieving by the sets $A_{n_1,r}$ and $A_{n_2,r}$ are independent events, by the Chinese Remainder Theorem. If all of the sets $\{A_{n,r}\}_{n \in \mathscr{N}_{i+1}}$ were jointly independent, then the density of the fibre $r \bmod Q_i$ surviving into R_{i+1} would be

$$\prod_{n \in \mathscr{N}_{i+1}} \left(1 - \frac{|A_{n,r} \bmod nQ_i|}{n}\right) \doteq \exp\left(- \sum_{n \in \mathscr{N}_{i+1}} \frac{|A_{n,r} \bmod nQ_i|}{n}\right).$$

For a given n, we can bound the average size of $|A_{n,r} \bmod nQ_i|$ averaged over $r \bmod Q_i$:

$$\begin{aligned}
\frac{1}{Q_i} \sum_{r \bmod Q_i} |A_{n,r} \bmod nQ_i| &\leq \frac{1}{Q_i} \sum_{r \bmod Q_i} \sum_{m_0 | Q_i} \mathbf{1}\{a_{m_0 n} \equiv r \bmod m_0\} \\
&= \frac{1}{Q_i} \sum_{m_0 | Q_i} \sum_{r \bmod Q_i} \mathbf{1}\{r \equiv a_{m_0 n} \bmod m_0\} \\
&= \frac{1}{Q_i} \sum_{m_0 | Q_i} \frac{Q_i}{m_0} = \prod_{p | Q_i} \left(1 + \frac{1}{p}\right) = (\log P_i)^{1+o(1)}.
\end{aligned}$$

With the belief that the typical set $A_{n,r}$ has size $\approx \log P_i$, then since

$$\sum_{n\in\mathscr{N}_{i+1}} \frac{1}{n} = -1 + \prod_{P_i<p\le P_{i+1}} \left(1+\frac{1}{p}\right) \approx \frac{\log P_{i+1}}{\log P_i},$$

we might hope that the typical fibre above R_i has density $P_{i+1}^{-O(1)}$. Thus far our reasoning in the case $i = 0$ roughly follows the treatment of [11], but now we diverge.

One difficulty with this heuristic account is that for generic $n_1, n_2 \in \mathscr{N}_{i+1}$ it is not generally true that $(n_1, n_2) = 1$, so that the congruences in $A_{n_1,r}$ and $A_{n_2,r}$ are not independent. To clarify the situation, we may imagine the numbers in the set $\mathscr{N}_{i+1}$ as being split into two types. Within the collection of numbers that are composed of 'few' prime factors, it is generally true that most pairs of numbers in the set are co-prime. Meanwhile, the numbers composed of many prime factors are large and sparse, and thus they may be expected to not contribute significantly to the sieve. This reasoning makes it plausible that the Lovász Local Lemma can be used to handle the mild dependence that results from sieving by the moduli in $\mathscr{N}_{i+1}$. In practice, rather than split the moduli into two groups, in applying the Local Lemma we are naturally led to make a smoother decomposition, which assigns to each modulus a weight according to its number of prime factors.

Unfortunately, it will not generally be true that the Local Lemma applies to estimate the density of a given fibre, but rather only that it applies on a certain subset $R_i^* \subset R_i$ of 'good' fibres on which the distribution of the sizes $\{|A_{n,r} \bmod nQ_i|\}_{n\in\mathscr{N}_{i+1}}$ is under control. Roughly what is needed for a fibre to be good is that a bound in dilations should hold at each prime $p \in (P_i, P_{i+1}]$,

$$\sum_{n\in\mathscr{N}_{i+1}, p|n} \frac{|A_{n,r} \bmod nQ_i|}{n} \ll 1. \tag{1}$$

Such a bound controls the dependence among the sets $\{A_{n,r}\}_{n\in\mathscr{N}_{i+1}}$. We give a more precise definition of good fibres in the next section.

In order to demonstrate that a reasonable number of fibres are good we wish to understand the distribution of values of $|A_{n,r} \bmod nQ_i|$ for varying r and n. Recall that we gained a heuristic understanding of the typical behavior of $|A_{n,r} \bmod nQ_i|$ by taking the average over $\mathbb{Z}/Q_i\mathbb{Z}$. Similarly, we control the distribution of $|A_{n,r} \bmod nQ_i|$ as r varies in subsets S_i of R_i by bounding the moments

$$\frac{1}{|S_i \bmod Q_i|} \sum_{r\in S_i \bmod Q_i} |A_{n,r} \bmod nQ_i|^k, \qquad k = 1, 2, 3, \dots ,$$

and making a truncation argument. In practice we use only the third moment of the sizes $|A_{n,r} \bmod nQ_i|$, although other choices would work as well with appropriately modified parameters.

It transpires that the moments are controlled by statistics

$$\sum_{m|Q_i} \ell_k(m) \max_{b \bmod m} \frac{|S_i \cap (b \bmod m) \bmod Q_i|}{|S_i \bmod Q_i|}, \qquad k = 1, 2, 3, \dots$$

that measure the bias in the set S_i. Here $\ell_k(m)$ is a weight, equal to $(2^k - 1)^{\omega(m)}$ in the case that m is square free. When $i = 0$, it will not be necessary to consider subsets of $R_0 = \mathbb{Z}/Q_0\mathbb{Z}$, since the statistics taken over R_0 are unbiased, equal to

$$\sum_{m|Q_0} \frac{\ell_k(m)}{m} = \prod_{p<P_0} \left(1 + \frac{2^k - 1}{p}\right) \approx (\log P_0)^{2^k-1}, \tag{2}$$

a rate of growth that will be acceptable for us. When $i > 0$, however, the set R_i will typically be small and irregular as compared to $\mathbb{Z}/Q_i\mathbb{Z}$, so that our argument requires searching for good fibres R_i^* only within a subset $S_i \subset R_i$ chosen to have statistics that approximate (2).

The above discussion suggests that there is a second convenient notion of a good fibre, which is that $(r \bmod Q_i)$ is 'well distributed' if for each $n \in \mathscr{N}_{i+1}$,

$$\max_{b \bmod n} |R_{i+1} \cap (b \bmod n) \cap (r \bmod Q_i) \bmod Q_{i+1}| \approx \frac{1}{n} |R_{i+1} \cap (r \bmod Q_i) \bmod Q_{i+1}|. \tag{3}$$

Thus in a well-distributed fibre $(r \bmod Q_i)$, for each modulus $n \in \mathscr{N}_{i+1}$, any residue class modulo n is allowed to hold at most slightly more than its share of the set R_{i+1}. A pleasant feature of our argument is that a relative form of the Lovász Local Lemma guarantees that good fibres in the sense of (1) are automatically well distributed in the sense of (3), so that with respect to the moduli in $\mathscr{N}_{i+1}$ composed of large prime factors, a reasonable choice for the set S_{i+1} is the union of good fibres from the previous stage, $S_{i+1} = R_i^* \cap R_{i+1}$.

The choice of $S_{i+1} = R_i^* \cap R_{i+1}$ ensures that S_{i+1} is well distributed to the moduli in $\mathscr{N}_{i+1}$ that have only large prime factors, but $R_i^* \cap R_{i+1} \subset S_i$ may have become poorly distributed as compared to S_i with respect to moduli having smaller prime factors as a result of variable sieving in the fibres above R_i^*. We balance this effect by reweighting $R_i^* \cap R_{i+1}$ with a measure μ_{i+1} on $\mathbb{Z}/Q_{i+1}\mathbb{Z}$, with respect to which each fibre over R_i^* has equal weight. Thus at stage $i + 1 \geq 1$ we will in fact consider the bias statistics

$$\beta_k^k(i+1) = \sum_{m|Q_{i+1}} \ell_k(m) \max_{b \bmod m} \frac{\mu_{i+1}(R_i^* \cap R_{i+1} \cap (b \bmod m))}{\mu_{i+1}(R_i^* \cap R_{i+1})}.$$

In general we will be able to show that these statistics approximate the unbiased statistics (2) to within an error determined only in terms of the quality of well-distribution (3) and the proportions of fibres that are good from previous stages.

To summarize, at stage 0 we do no sieving so that, with a uniform measure, the bias statistics are under control. This allows us to say that many fibres over $R_0 = \mathbb{Z}/Q_0\mathbb{Z}$ are good, and thus, that the bias statistics at stage 1 do not grow too rapidly. The argument then iterates, with the possibility of continuing iteration for arbitrarily large values of the parameters P_i depending upon growth of the statistics $\beta(i)$ as compared with growth of the P_i. The proof is completed by making this comparison for an explicit choice of parameters.

3. The complete argument

We turn to the technical details of the argument. As we now treat congruences to general moduli, we briefly recall some notions from the previous section, pointing out the minor variation from the square free case.

As above, $M > 0$ is our upper bound for the minimum modulus of a covering system, and

$$\mathscr{M} \subset \{m \in \mathbb{Z},\ m > M\}$$

is a finite collection of moduli. For each $m \in \mathscr{M}$, we assume that a congruence class $a_m \bmod m$ is given. The uncovered set is

$$R = \bigcap_{m \in \mathscr{M}} (a_m \bmod m)^c,$$

which we show has a nonzero density. In the general case it is convenient to let

$$Q = \mathrm{LCM}(m : m \in \mathscr{M}),$$

so that R is a set defined modulo Q.

We take a sequence of thresholds $1 = P_{-1} < P_0 < P_1 < \cdots$ with $P_0 \geq 2$ and $P_i \to \infty$. Setting $v = v_p = v_p(Q)$ for the multiplicity with which p divides Q, we let

$$Q_{-1} = 1, \qquad \forall i \geq 0,\ Q_i = \prod_{p \leq P_i} p^v.$$

Then $\mathscr{M}_i = \{m \in \mathscr{M} : m | Q_i\}$ is the collection of P_i-smooth moduli in $\mathscr{M}$. The set R is filtered in stages $R_{-1} \supset R_0 \supset R_1 \supset \cdots$ by letting $R_{-1} = \mathbb{Z}$, and, for $i \geq 0$,

$$R_i = \bigcap_{m \in \mathscr{M}_i} (a_m \bmod m)^c.$$

Although Q_i now depends in an essential way on the collection of moduli $\mathscr{M}$, our argument will, for a given i, treat the properties of R_i uniformly for all distinct congruence systems having minimum modulus greater than M.

3.1. *The initial stage.* We are no longer able to assume that $Q_0 < M$ so that $\mathscr{M}_0 = \emptyset$, but we will assume that M is sufficiently large so that $\mathscr{M}_0$ is quite sparse. Specifically, we let $0 < \delta < 1$ be a parameter. We may estimate the density of the set

$$R_0 = \bigcap_{m \in \mathscr{M}_0} (a_m \bmod m)^c$$

by applying the union bound

$$\begin{aligned} |R_0 \bmod Q_0| &\leq Q_0 - \sum_{m \in \mathscr{M}_0} |(a_m \bmod m) \bmod Q_0| \\ &= Q_0 \left(1 - \sum_{m \in \mathscr{M}_0} \frac{1}{m}\right) \leq Q_0 \left(1 - \sum_{\substack{m > M \\ p|m \Rightarrow p \leq P_0}} \frac{1}{m}\right), \end{aligned}$$

and we make the condition that

$$\text{(C0)} \qquad \sum_{\substack{m > M \\ p|m \Rightarrow p \leq P_0}} \frac{1}{m} < \delta.$$

This implies a bound for some bias statistics of R_0 as follows.

Let $\ell_k(m)$ be the number of k-tuples of natural numbers having LCM m. This is a multiplicative function (that is, $\ell_k(mn) = \ell_k(m)\ell_k(n)$ when m and n are co-prime), and it is given at prime powers by

$$\ell_k(p^j) = (j+1)^k - j^k.$$

We define the kth bias statistic at stage 0 to be

$$\beta_k^k(0) = \sum_{m | Q_0} \ell_k(m) \max_{b \bmod m} \frac{|R_0 \cap (b \bmod m) \bmod Q_0|}{|R_0 \bmod Q_0|}.$$

Putting in the trivial bound $|R_0 \cap (b \bmod m) \bmod Q_0| \leq \frac{Q_0}{m}$, we find

$$\beta_k^k(0) \leq \frac{1}{1-\delta} \sum_{m|Q_0} \frac{\ell_k(m)}{m} < \frac{1}{1-\delta} \prod_{p \leq P_0} \left(\sum_{j=0}^{\infty} \frac{(j+1)^k - j^k}{p^j} \right).$$

We now leave the initial stage. We will return to choose δ and P_0 at the end of the argument.

3.2. *The inductive loop.* In sieving stage $i+1$, $i \geq 0$, we view $\mathbb{Z}/Q_{i+1}\mathbb{Z}$ as fibred over $\mathbb{Z}/Q_i\mathbb{Z}$, and we consider the set R_{i+1} within individual fibres over R_i.

Introduce the set of 'new moduli'

$$\mathscr{N}_{i+1} = \{n : n | Q_{i+1},\ n > 1,\ p|n \Rightarrow P_i < p \leq P_{i+1}\},$$

and notice that each $n \in \mathscr{N}_{i+1}$ is coprime to Q_i. Thus each modulus $m \in \mathscr{M}_{i+1} \setminus \mathscr{M}_i$ has a unique factorization as $m = m_0 n$ with $m_0 | Q_i$ and $n \in \mathscr{N}_{i+1}$. Given $r \in R_i$ and $n \in \mathscr{N}_{i+1}$, we set

$$A_{n,r} = (r \bmod Q_i) \cap \bigcup_{m_0 | Q_i, m_0 n \in \mathscr{M}_{i+1}} (a_{m_0 n} \bmod m_0 n).$$

Then

$$(r \bmod Q_i) \cap R_{i+1} = (r \bmod Q_i) \cap \bigcap_{n \in \mathscr{N}_{i+1}} A^c_{n,r}.$$

We wish to consider R_{i+1} only in good fibres $(r \bmod Q_i)$ where the sieve is well behaved. A set of properties that we would like good fibres to have is the following.

Definition. Let $i \geq 0$, and let $\lambda \geq 0$ be a parameter. We say that $r \in \mathbb{Z}/Q_i\mathbb{Z}$ is λ-*well distributed* if $R_{i+1} \cap (r \bmod Q_i)$ is nonempty, and if the fibre satisfies the uniformity property that for each $n \in \mathscr{N}_{i+1}$,

$$\max_{b \bmod n} \frac{|R_{i+1} \cap (b \bmod n) \cap (r \bmod Q_i) \bmod Q_{i+1}|}{|R_{i+1} \cap (r \bmod Q_i) \bmod Q_{i+1}|} \leq \frac{e^{\lambda\omega(n)}}{n}. \tag{4}$$

An alternative, more technical characterization of good fibres is as follows.

Definition. Let $i \geq 0$, and let $\lambda \geq 0$ be a real parameter. We say that the fibre $r \in R_i \bmod Q_i$ is λ-*good* if, for each $p \in (P_i, P_{i+1}]$,

$$\sum_{n \in \mathscr{N}_{i+1}, p|n} \frac{|A_{n,r} \bmod nQ_i| e^{\lambda\omega(n)}}{n} \leq 1 - e^{-\lambda}. \tag{5}$$

If each fibre in a set $S \subset R_i$ is λ-good, then we say that the set S is λ-good as well, similarly λ-well distributed.

A basic observation of our proof is that a λ-good fibre is automatically λ-well distributed.

Proposition 1. *Let $i \geq 0$, $\lambda \geq 0$, and let $r \in \mathbb{Z}/Q_i\mathbb{Z}$ be λ-good. Then r is λ-well distributed.*

The proof of this fact uses a relative form of the Lovász Local Lemma.

Lemma (Lovász Local Lemma, relative form). *Let $\{A_u\}_{u \in V}$ be a finite collection of events in a probability space. Let $D = (V, E)$ be a directed graph, such that, for each $u \in V$, event A_u is independent of the sigma-algebra generated by the events $\{A_v : (u, v) \notin E\}$. Suppose that there exist real numbers $\{x_u\}_{u \in V}$, satisfying $0 \leq x_u < 1$, and for each $u \in V$,*

$$\mathbf{P}(A_u) \leq x_u \prod_{(u,v) \in E} (1 - x_v).$$

Then for any $\emptyset \neq U \subset V$,

$$\mathbf{P}\left(\bigcap_{u\in V} A_u^c\right) \geq \mathbf{P}\left(\bigcap_{u\in U} A_u^c\right) \cdot \prod_{v\in V\setminus U}(1-x_v). \tag{6}$$

In particular, taking U *to be a singleton,*

$$\mathbf{P}\left(\bigcap_{u\in V} A_u^c\right) \geq \prod_{u\in V}(1-x_u). \tag{7}$$

Remark. The conclusion (7) is the standard one; see [1]. The stronger conclusion (6) follows directly from the proof. For completeness, we show the argument in Appendix B; see also [18].

The application of the Local Lemma to prove Proposition 1 is as follows. Write F_r for the fibre $(r \bmod Q_i) \subset \mathbb{Z}/Q_{i+1}\mathbb{Z}$, and make it a probability space with the uniform measure $\mathbf{P}_r$. The events are the collection $\{A_{n,r}\}_{n\in\mathscr{N}_{i+1}}$. Since F_r contains $\frac{Q_{i+1}}{Q_i}$ elements, and since $A_{n,r}$ is a set defined modulo nQ_i,

$$\mathbf{P}_r(A_{n,r}) = \frac{|A_{n,r} \bmod nQ_i|}{n}.$$

By first translating by $-r$ and then dilating by $\frac{1}{Q_i}$, we map F_r onto $\mathbb{Z}/\frac{Q_{i+1}}{Q_i}\mathbb{Z}$. For $n \in \mathscr{N}_{i+1}$, this map gives a bijection between progressions modulo nQ_i constrained to $(r \bmod Q_i)$, and unconstrained progressions modulo n in $\mathbb{Z}/\frac{Q_{i+1}}{Q_i}\mathbb{Z}$. Applying this map, and then the Chinese Remainder Theorem, makes it clear that $A_{n,r}$ is jointly independent of the σ-algebra generated by the events

$$\{(b \bmod n') \cap (r \bmod Q_i) : n' \in \mathscr{N}_{i+1}, (n,n') = 1\}.$$

In particular, a valid dependency graph with which to apply the Local Lemma has edges between $n_1, n_2 \in \mathscr{N}_{i+1}$ if and only if $n_1 \neq n_2$ and $(n_1, n_2) > 1$.

Proof of Proposition 1. We first check that

$$\forall n \in \mathscr{N}_{i+1}, \qquad x_n = e^{\lambda\omega(n)}\frac{|A_{n,r} \bmod nQ_i|}{n}$$

is an admissible set of weights with which to apply the Local Lemma.

Since the fibre r is λ-good, the bound in dilations condition (5) gives that for all $p \in (P_i, P_{i+1}]$,

$$\sum_{n\in\mathscr{N}_{i+1}:p|n} \frac{|A_{n,r} \bmod nQ_i|e^{\lambda\omega(n)}}{n} \leq 1 - e^{-\lambda}.$$

Dropping all but one term in the sum, we see that for each $n \in \mathscr{N}_{i+1}$, $1 - x_n \geq e^{-\lambda}$. Thus, by convexity,

$$1 - x_n \geq \exp\left(\frac{-\lambda}{1-e^{-\lambda}}x_n\right).$$

Therefore, for a given $n \in \mathscr{N}_{i+1}$,

$$\begin{aligned}\prod_{n'\in\mathscr{N}_{i+1}:(n,n')>1}(1-x_{n'}) &\geq \prod_{p|n}\prod_{n'\in\mathscr{N}_{i+1}:p|n'}(1-x_{n'})\\ &\geq \exp\left(\frac{-\lambda}{1-e^{-\lambda}}\sum_{p|n}\sum_{n'\in\mathscr{N}_{i+1}:p|n'}\frac{e^{\lambda\omega(n')}|A_{n',r} \bmod n'Q_i|}{n'}\right)\\ &\geq \exp\left(-\lambda\omega(n)\right).\end{aligned}$$

It follows that

$$x_n \prod_{\substack{n'\in\mathscr{N}_{i+1}:(n,n')>1\\ n'\neq n}}(1-x_{n'}) \geq x_n \prod_{n'\in\mathscr{N}_{i+1}:(n,n')>1}(1-x_{n'}) \geq \frac{|A_{n,r} \bmod nQ_i|}{n}$$

so that the Lovász criterion is satisfied. It is then immediate that the fibre itself is nonempty, since the product in the conclusion (7) of the Local Lemma is nonzero.

For the uniformity property (4), let $n \in \mathscr{N}_{i+1}$ and let $b \bmod n$ maximize

$$\frac{|R_{i+1}\cap(r \bmod Q_i)\cap(b \bmod n) \bmod Q_{i+1}|}{|R_{i+1}\cap(r \bmod Q_i) \bmod Q_{i+1}|} = \frac{\mathbf{P}_r\left(\left(\bigcap_{n'\in\mathscr{N}_{i+1}}A^c_{n',r}\right)\cap(b \bmod n)\right)}{\mathbf{P}_r\left(\bigcap_{n'\in\mathscr{N}_{i+1}}A^c_{n',r}\right)}.$$

Dropping part of the intersection, the numerator is bounded above by

$$\mathbf{P}_r\left(\left(\bigcap_{n'\in\mathscr{N}_{i+1},(n',n)=1}A^c_{n',r}\right)\cap(b \bmod n)\right) = \frac{1}{n}\mathbf{P}_r\left(\bigcap_{n'\in\mathscr{N}_{i+1},(n',n)=1}A^c_{n',r}\right).$$

Now by the stronger conclusion (6) of the Local Lemma,

$$\mathbf{P}_r\left(\bigcap_{n'\in\mathscr{N}_{i+1}}A^c_{n',r}\right) \geq \mathbf{P}_r\left(\bigcap_{n'\in\mathscr{N}_{i+1},(n',n)=1}A^c_{n',r}\right)\prod_{n'\in\mathscr{N}_{i+1},(n',n)>1}(1-x_{n'}).$$

Since we checked above that

$$\prod_{n'\in\mathscr{N}_{i+1},(n',n)>1}(1-x_{n'}) \geq e^{-\lambda\omega(n)},$$

it follows that

$$\begin{aligned}\frac{|R_{i+1}\cap(b \bmod n)\cap(r \bmod Q_i) \bmod Q_{i+1}|}{|R_{i+1}\cap(r \bmod Q_i) \bmod Q_{i+1}|} &\leq \frac{1}{n}\prod_{n'\in\mathscr{N}_{i+1},(n',n)>1}(1-x_{n'})^{-1}\\ &\leq \frac{e^{\lambda\omega(n)}}{n},\end{aligned}$$

which is the condition of uniformity.

□

Let $R^*_{-1} = \mathbb{Z}$, and for $i \geq 0$, let R^*_i be the λ-good fibres within $R^*_{i-1} \cap R_i$. It remains to describe how we may find good fibres above a large well-distributed set.

It will be convenient to reweight $\mathbb{Z}/Q_i\mathbb{Z}$ at each stage with a measure μ_i, supported on the set $R^*_{i-1} \cap R_i$. The advantage of using this measure is that it will balance the effect of the variation in size of the various good fibres from previous stages, so that at stage $i+1$ we isolate the effects of sieving by moduli in $\mathscr{N}_{i+1}$. We define μ_i iteratively by setting

$$\mu_0(r) = \begin{cases} \frac{1}{|R_0 \bmod Q_0|} & r \in R_0 \bmod Q_0, \\ 0 & r \notin R_0 \bmod Q_0. \end{cases}$$

For $i \geq 0$ and for $r \in R^*_i \cap R_{i+1} \bmod Q_{i+1}$, we reduce $r \bmod Q_i$ to determine $\mu_i(r)$, and we set

$$\mu_{i+1}(r) = \begin{cases} \frac{\mu_i(r \bmod Q_i)}{|R_{i+1} \cap (r \bmod Q_i) \bmod Q_{i+1}|} & r \in R^*_i \cap R_{i+1} \bmod Q_{i+1}, \\ 0 & r \notin R^*_i \cap R_{i+1} \bmod Q_{i+1}. \end{cases} \tag{8}$$

Along with the measures μ_i, we track a collection of bias statistics.

Definition. Let $i \geq 0$ and $k \geq 1$. The kth *bias statistic* of set $R^*_{i-1} \cap R_i \subset \mathbb{Z}/Q_i\mathbb{Z}$ is defined by

$$\beta^k_k(i) = \sum_{m|Q_i} \ell_k(m) \max_{b \bmod m} \frac{\mu_i(R^*_{i-1} \cap R_i \cap (b \bmod m))}{\mu_i(R^*_{i-1} \cap R_i)}.$$

Since we require $R^*_{-1} = \mathbb{Z}$ and since μ_0 is uniform on R_0, this agrees with our definition of the bias statistics for R_0 given in the initial stage. These bias statistics will be the main tool used to produce good fibres, a discussion that we briefly postpone.

The primary virtue of the measure μ_i is that it allows us to bound the iterative growth of the bias statistics only in terms of the size of the well-distributed set R^*_i and its parameter of well-distribution, λ. Before demonstrating this, we record the notation

$$\pi^{\text{good}}_i = \frac{\mu_i(R^*_i)}{\mu_i(R^*_{i-1} \cap R_i)}$$

for the proportion relative to μ_i of good fibres in $R^*_{i-1} \cap R_i$, and we record the following simple lemma.

Lemma 2. *Let $i \geq 0$. For a fixed $r \in R^*_i \bmod Q_i$, the measure μ_{i+1} is constant on $R_{i+1} \cap (r \bmod Q_i)$. The total mass of μ_{i+1} is given by*

$$\mu_{i+1}(R^*_i \cap R_{i+1}) = \pi^{\text{good}}_i \mu_i(R^*_{i-1} \cap R_i).$$

Proof. The first observation is immediate from the definition.

The total mass is given by

$$\begin{aligned}\mu_{i+1}(R_i^* \cap R_{i+1}) &= \sum_{r \in R_i^* \cap R_{i+1} \bmod Q_{i+1}} \mu_{i+1}(r) \\ &= \sum_{r_0 \in R_i^* \bmod Q_i} \mu_i(r_0) \\ &\quad \times \sum_{r \in R_{i+1} \cap (r_0 \bmod Q_i) \bmod Q_{i+1}} \frac{1}{|R_{i+1} \cap (r_0 \bmod Q_i) \bmod Q_{i+1}|} \\ &= \sum_{r_0 \in R_i^* \bmod Q_i} \mu_i(r_0) \\ &= \pi_i^{\mathrm{good}} \mu_i(R_{i-1}^* \cap R_i). \end{aligned}$$

□

The main proposition regarding the measures μ_i now is as follows.

Proposition 3. *Let $i \geq 0$ and $k \geq 1$, and suppose that R_i^* is λ-good. We have*

$$\beta_k^k(i+1) \leq \frac{\beta_k^k(i)}{\pi_i^{\mathrm{good}}} \prod_{P_i < p \leq P_{i+1}} \left(1 + e^{\lambda} \sum_{j=1}^{v_p} \frac{(j+1)^k - j^k}{p^j}\right).$$

Proof. Recall that

$$\beta_k^k(i+1) = \sum_{m | Q_{i+1}} \ell_k(m) \max_{b \bmod m} \frac{\mu_{i+1}(R_i^* \cap R_{i+1} \cap (b \bmod m))}{\mu_{i+1}(R_i^* \cap R_{i+1})}. \tag{9}$$

Given $m | Q_{i+1}$, factor $m = m_0 n$ with $m_0 | Q_i$ and $n \in \{1\} \cup \mathcal{N}_{i+1}$. Let $b \bmod m$ maximize $\mu_{i+1}(R_i^* \cap R_{i+1} \cap (b \bmod m))$. Fibring over $\mathbb{Z}/Q_i\mathbb{Z}$, we have

$$\mu_{i+1}(R_i^* \cap R_{i+1} \cap (b \bmod m)) = \sum_{\substack{r_0 \in R_i^* \bmod Q_i \\ r_0 \equiv b \bmod m_0}} \mu_{i+1}((r_0 \bmod Q_i) \cap (b \bmod n))$$

$$= \sum_{\substack{r_0 \in R_i^* \bmod Q_i \\ r_0 \equiv b \bmod m_0}} \mu_i(r_0) \frac{|R_{i+1} \cap (b \bmod n) \cap (r_0 \bmod Q_i) \bmod Q_{i+1}|}{|R_{i+1} \cap (r_0 \bmod Q_i) \bmod Q_{i+1}|}.$$

Since the good set R_i^* is λ-well distributed, the last sum is bounded by

$$\frac{e^{\lambda \omega(n)}}{n} \sum_{\substack{r_0 \in R_i^* \bmod Q_i \\ r_0 \equiv b \bmod m_0}} \mu_i(r_0).$$

Therefore, using the multiplicativity of $\ell_k(m)$, we find

$$\beta_k^k(i+1) \leq \sum_{n \in \{1\} \cup \mathcal{N}_{i+1}} \frac{\ell_k(n) e^{\lambda \omega(n)}}{n} \sum_{m_0 | Q_i} \ell_k(m_0) \max_{b \bmod m_0} \frac{\mu_i(R_i^* \cap (b \bmod m_0))}{\mu_{i+1}(R_i^* \cap R_{i+1})}.$$

Since $\{1\} \cup \mathscr{N}_{i+1}$ has the structure of a direct product, the sum over n factors as the product of the proposition. Meanwhile, using $R_i^* \subset R_{i-1}^* \cap R_i$ and $\mu_{i+1}(R_i^* \cap R_{i+1}) = \pi_i^{\mathrm{good}} \mu_i(R_{i-1}^* \cap R_i)$, we bound the sum over m_0 by

$$\sum_{m_0 | Q_i} \ell_k(m_0) \max_{b \bmod m_0} \frac{\mu_i(R_i^* \cap (b \bmod m_0))}{\mu_{i+1}(R_i^* \cap R_{i+1})}$$
$$\leq \frac{1}{\pi_i^{\mathrm{good}}} \sum_{m_0 | Q_i} \ell_k(m_0) \max_{b \bmod m_0} \frac{\mu_i(R_{i-1}^* \cap R_i \cap (b \bmod m_0))}{\mu_i(R_{i-1}^* \cap R_i)} = \frac{\beta_k^k(i)}{\pi_i^{\mathrm{good}}}. \quad \square$$

It remains to demonstrate the utility of the bias statistics for generating good fibres. For $n \in \mathscr{N}_{i+1}$, $k \geq 1$ and $R_{i-1}^* \cap R_i$ defined modulo Q_i, define the kth moment of $|A_{n,r} \bmod nQ_i|$ to be

$$M_k^k(i,n) = \frac{1}{\mu_i(R_{i-1}^* \cap R_i)} \sum_{r \in R_{i-1}^* \cap R_i \bmod Q_i} \mu_i(r) |A_{n,r} \bmod nQ_i|^k.$$

The bias statistics control these moments.

LEMMA 4. *Let $i \geq 0$ and let $n \in \mathscr{N}_{i+1}$. We have $M_k(i,n) \leq \beta_k(i)$.*

Proof. Recall that

$$A_{n,r} = (r \bmod Q_i) \cap \left(\bigcup_{m_0 | Q_i, m_0 n \in \mathscr{M}} (a_{m_0 n} \bmod m_0 n) \right).$$

A given congruence $(a_{m_0 n} \bmod m_0 n)$ intersects $r \bmod Q_i$ if and only if $r \equiv a_{m_0 n} \bmod m_0$. If it does intersect, it does so in a single residue class modulo nQ_i. Thus, the union bound gives

$$|A_{n,r} \bmod nQ_i| \leq \sum_{m_0 | Q_i} \mathbf{1}\{r \equiv a_{m_0 n} \bmod m_0\}.$$

It follows that, considering $R_{i-1}^* \cap R_i$ as a subset of $\mathbb{Z}/Q_i\mathbb{Z}$,

$$\begin{aligned} M_k^k(i,n) \leq & \frac{1}{\mu_i(R_{i-1}^* \cap R_i)} \sum_{r \in R_{i-1}^* \cap R_i} \mu_i(r) \\ & \times \sum_{m_1, \ldots, m_k | Q_i} \mathbf{1}\{\forall 1 \leq j \leq k,\ r \equiv a_{m_j n} \bmod m_j\} \\ = & \frac{1}{\mu_i(R_{i-1}^* \cap R_i)} \sum_{m_1, \ldots, m_k | Q_i} \\ & \times \sum_{r \in R_{i-1}^* \cap R_i} \mu_i(r) \mathbf{1}\{\forall 1 \leq j \leq k,\ r \equiv a_{m_j n} \bmod m_j\}. \end{aligned}$$

The inner condition restricts r to at most one class modulo the LCM of $m_1, \ldots, m_k$. Grouping $m_1, \ldots, m_k$ according to their LCM, and writing $\ell_k(m)$

for the number of ways in which m is the LCM of a k-tuple of natural numbers, we find

$$M_k^k(i,n) \le \frac{1}{\mu_i(R_{i-1}^* \cap R_i)} \times \sum_{m|Q_i} \ell_k(m) \max_{b \bmod m} \mu_i(R_{i-1}^* \cap R_i \cap (b \bmod m)) = \beta_k^k(i). \qquad \square$$

Since the above estimate is uniform in n, we have convexity-type control over mixtures of the sizes $\{|A_{n,r} \bmod nQ_i|\}_{n\in\mathcal{N}_{i+1}}$.

LEMMA 5. *Let $i \ge 0$ and $k \ge 1$. Let $\{w_n\}_{n\in\mathcal{N}_{i+1}}$ be a set of nonnegative weights, not all zero. Then for all $B > 0$ and any $k \ge 1$,*

$$\frac{1}{\mu_i(R_{i-1}^* \cap R_i)} \mu_i\left(r \in R_{i-1}^* \cap R_i : \sum_{n\in\mathcal{N}_{i+1}} w_n |A_{n,r} \bmod nQ_i| > B \right) \le \frac{\beta_k^k(i)}{B^k} \left(\sum_{n\in\mathcal{N}_{i+1}} w_n \right)^k.$$

Proof. Set $w_n' = \frac{w_n}{\sum_{\tilde{n}} w_{\tilde{n}}}$, which is a probability measure on $\mathcal{N}_{i+1}$. Convexity gives

$$\left(\sum_{n\in\mathcal{N}_{i+1}} w_n' |A_{n,r} \bmod nQ_i| \right)^k \le \sum_{n\in\mathcal{N}_{i+1}} w_n' |A_{n,r} \bmod nQ_i|^k,$$

so that

$$\frac{1}{\mu_i(R_{i-1}^* \cap R_i)} \sum_{r\in R_{i-1}^* \cap R_i} \mu_i(r) \left(\sum_{n\in\mathcal{N}_{i+1}} w_n' |A_{n,r} \bmod nQ_i| \right)^k \le \sum_{n\in\mathcal{N}_{i+1}} w_n' M_k^k(i,n) \le \beta_k^k(i).$$

The result now follows from Markov's inequality. $\square$

We now complete our argument by using the bias statistics to guarantee the existence of good fibres.

For a given $p \in (P_i, P_{i+1}]$, the dilation condition of good fibres (5) at p is the statement that

$$\sum_{n\in\mathcal{N}_{i+1}, p|n} \frac{|A_{n,r} \bmod nQ_i| e^{\lambda\omega(n)}}{n} \le 1 - e^{-\lambda}.$$

By applying the convexity lemma, Lemma 5, with weights

$$w_n = \mathbf{1}_{p|n} \frac{e^{\lambda\omega(n)}}{n},$$

we find that the relative proportion of fibres failing this condition is bounded by

$$\min_k \frac{\beta_k^k(i)}{(1-e^{-\lambda})^k}\left(\sum_{n\in\mathcal{N}_{i+1}:p|n}\frac{e^{\lambda\omega(n)}}{n}\right)^k.$$

Since

$$\sum_{n\in\mathcal{N}_{i+1},p|n}\frac{e^{\lambda\omega(n)}}{n}\le\frac{e^\lambda}{p-1}\sum_{n\in\{1\}\cup\mathcal{N}_{i+1}}\frac{e^{\lambda\omega(n)}}{n}\le\frac{e^\lambda}{p-1}\prod_{P_i<p'\le P_{i+1}}\left(1+\frac{e^\lambda}{p'-1}\right),$$

making a union bound, we find that the total relative proportion of fibres failing some dilation condition is bounded by

$$\min_k \beta_k^k(i)\frac{e^{k\lambda}}{(1-e^{-\lambda})^k}\left(\prod_{P_i<p\le P_{i+1}}\left(1+\frac{e^\lambda}{p-1}\right)\right)^k\sum_{P_i<p\le P_{i+1}}\frac{1}{(p-1)^k}.$$

For a value $0<\pi^{\mathrm{good}}<1$, we make the constraint that this quantity is bounded by $1-\pi^{\mathrm{good}}$; that is,

(C1)

$$\frac{e^\lambda}{1-e^{-\lambda}}\prod_{P_i<p\le P_{i+1}}\left(1+\frac{e^\lambda}{p-1}\right)\le\max_k\frac{(1-\pi^{\mathrm{good}})^{\frac1k}}{\beta_k(i)}\left(\sum_{P_i<p\le P_{i+1}}\frac{1}{(p-1)^k}\right)^{-\frac1k},$$

which guarantees that, with respect to μ_i, the proportion of good fibres in $R_{i-1}^*\cap R_i$ is at least π^{good}.

3.3. *Proof of Theorem* 1. The iterative stage of our argument is summarized in the following technical theorem.

THEOREM 2. *Let $i\ge 0$, and let $0<\pi^{\mathrm{good}}<1$. Let the set $R_{i-1}^*\subset\mathbb{Z}/Q_{i-1}\mathbb{Z}$ be such that $R_{i-1}^*\cap R_i$ is nonempty, let μ_i be a measure on $\mathbb{Z}/Q_i\mathbb{Z}$ with support in $R_{i-1}^*\cap R_i$, and denote the bias statistics of μ_i by $\beta_k(i)$, $k=1,2,3,\dots$. Suppose that $\lambda>0$ and $P_{i+1}>P_i$ satisfy the constraint*

(C1)

$$\prod_{P_i<p\le P_{i+1}}\left(1+\frac{e^\lambda}{p-1}\right)\le\frac{1-e^{-\lambda}}{e^\lambda}\max_k\frac{(1-\pi^{\mathrm{good}})^{\frac1k}}{\beta_k(i)}\left(\sum_{P_i<p\le P_{i+1}}\frac{1}{(p-1)^k}\right)^{-\frac1k}.$$

Then there exists $R_i^*\subset R_{i-1}^*\cap R_i$ defined modulo Q_i with $\frac{\mu_i(R_i^*)}{\mu_i(R_{i-1}^*\cap R_i)}\ge\pi^{\mathrm{good}}$, such that the density of R_{i+1} in each fibre above R_i^* is positive, and such that the associated bias statistics $\beta_k(i+1)$ of $R_i^*\cap R_{i+1}$ with respect to μ_{i+1} defined by (8) *satisfy*

$$\beta_k^k(i+1)\le\frac{\beta_k^k(i)}{\pi^{\mathrm{good}}}\prod_{P_i<p\le P_{i+1}}\left(1+e^\lambda\sum_{j=1}^{v_p}\frac{(j+1)^k-j^k}{p^j}\right),\qquad k=1,2,\dots\ .$$

We now make specific choices for our parameters and prove Theorem 1.

Proof of Theorem 1. Set $M = 10^{16}$ as in Theorem 1. For $i \geq 0$, let $P_i = e^{11+i}$. Set $e^\lambda = 2$, $\pi^{\text{good}} = \frac{1}{2}$. It will suffice to check that the density of the set R_0 is positive and that the constraint (C1) of Theorem 2 is met for every $i \geq 0$.

By Rankin's trick, for any $\sigma > 0$,

$$\sum_{\substack{m>M \\ p|m \Rightarrow p \leq P_0}} \frac{1}{m} \leq M^{-\sigma} \sum_{m: p|m \Rightarrow p \leq P_0} \frac{1}{m^{1-\sigma}} = M^{-\sigma} \prod_{p \leq P_0} \left(1 - \frac{1}{p^{1-\sigma}}\right)^{-1}.$$

Choosing $\sigma = 0.19$, we verify in Pari-GP [19] that the right-hand side is less than 0.859, so that R_0 is nonempty and, in particular, $\delta = 0.86$ in the initial stage is permissible.

We will argue throughout with the third bias statistic. We calculate

$$\beta_3(0) \leq \left((1-\delta)^{-1} \prod_{p \leq P_0} \left(\sum_{j=0}^{\infty} \frac{3j^2+3j+1}{p^j}\right)\right)^{\frac{1}{3}} < 731.8.$$

We use the following explicit estimates, which are verified in Appendix A. For all $n \geq 11$,

$$\prod_{e^n < p \leq e^{n+1}} \left(1 + \frac{2}{p-1}\right) < 1.2,$$

$$\prod_{e^n < p \leq e^{n+1}} \left(1 + 2\sum_{j=1}^{\infty} \frac{(j+1)^3 - j^3}{p^j}\right) < 3.4,$$

$$\left(\sum_{e^n < p \leq e^{n+1}} \frac{1}{(p-1)^3}\right)^{-\frac{1}{3}} > (2ne^{2n})^{\frac{1}{3}}.$$

Thus the constraint (C1) is satisfied at $i = 0$ since

$$\prod_{e^{11} < p \leq e^{12}} \left(1 + \frac{2}{p-1}\right) < 1.2 < \frac{(1-0.5)^{\frac{1}{3}}}{4} \frac{1}{731.8} \left(\sum_{e^{11} < p \leq e^{12}} \frac{1}{(p-1)^3}\right)^{\frac{-1}{3}}.$$

The constraint holds for all i since the growth of the bias statistics guarantees that for $i \geq 0$,

$$\frac{\beta_3(i+1)}{\beta_3(i)} < \left(\frac{3.4}{0.5}\right)^{\frac{1}{3}} < e^{\frac{2}{3}},$$

which is less than the growth of $\left((22+2i)e^{22+2i}\right)^{\frac{1}{3}}$ from i to $i+1$. $\square$

Appendix A. Explicit estimates with primes

A standard reference for explicit prime sum estimates is [17]. Slightly stronger estimates are now known (see, e.g., [10]), but the following will suffice for our purpose.

THEOREM 6 ([17, Cor. 2]). *Let* $\theta(x) = \sum_{p \le x} \log p$. *For* $x \ge 678407$, *we have*

$$|\theta(x) - x| < \frac{x}{40 \log x}. \tag{10}$$

We now check the explicit estimates used in the proof of Theorem 1.

LEMMA 7. *For any* $n \ge 11$,

$$\prod_{e^n < p \le e^{n+1}} \left(1 + \frac{2}{p-1}\right) < 1.2,$$

$$\prod_{e^n < p \le e^{n+1}} \left(1 + 2\sum_{j=1}^{\infty} \frac{(j+1)^3 - j^3}{p^j}\right) < 3.4,$$

$$\sum_{e^n < p \le e^{n+1}} \frac{1}{(p-1)^3} < \frac{1}{2ne^{2n}}.$$

Proof. Using Pari-GP [19] we verified these estimates numerically for $n = 11, 12, 13$. For $n > 13$, they follow by partial summation against (10). For the first,

$$\log \prod_{e^n < p \le e^{n+1}} \left(1 + \frac{2}{p-1}\right) \le 2 \sum_{e^n < p \le e^{n+1}} \frac{1}{p-1} \le \frac{2}{1 - e^{-n}} \int_{e^n}^{e^{n+1}} \frac{d\theta(x)}{x \log x}.$$

Write $d\theta(x) = dx + d(\theta(x) - x)$. Integrating the second term by parts, we obtain

$$\begin{aligned}
\int_{e^n}^{e^{n+1}} \frac{d\theta(x)}{x \log x} \le \log \frac{n+1}{n} &+ \frac{|\theta(e^{n+1}) - e^{n+1}|}{(n+1)e^{n+1}} + \frac{|\theta(e^n) - e^n|}{ne^n} \\
&+ \int_{e^n}^{e^{n+1}} \frac{|\theta(x) - x|}{x^2} \left(\frac{1}{\log x} + \frac{1}{(\log x)^2}\right) dx \\
\le \log \frac{15}{14} &+ \frac{1}{40 \cdot 15^2} + \frac{1}{40 \cdot 14^2} + \frac{2}{40 \cdot 14} \log \frac{15}{14} < 0.0695
\end{aligned}$$

so that

$$\frac{2}{1 - e^{-14}} \int_{e^n}^{e^{n+1}} \frac{d\theta(x)}{x \log x} < 0.14 < \log 1.2.$$

For the second,

$$\begin{aligned}\log \prod_{e^n<p\le e^{n+1}} \left(1+2\sum_{j=1}^{\infty}\frac{(j+1)^3-j^3}{p^j}\right) &\le 2\sum_{e^n<p\le e^{n+1}}\sum_{j=1}^{\infty}\frac{(j+1)^3-j^3}{p^j}\\ &\le 14\sum_{e^n<p\le e^{n+1}}\frac{1}{p-3}\\ &\le \frac{14}{1-3e^{-14}}\sum_{e^n<p\le e^{n+1}}\frac{1}{p}\\ &< \frac{14}{1-3e^{-14}}\cdot 0.07 < 1 < \log(3.4).\end{aligned}$$

For the third, proceed as for the first,

$$\begin{aligned}&\sum_{e^n<p\le e^{n+1}}\frac{1}{(p-1)^3} \le \frac{1}{n(1-e^{-n})^3}\left(\int_{e^n}^{e^{n+1}}\frac{dx}{x^3}+\int_{e^n}^{e^{n+1}}\frac{d(\theta(x)-x)}{x^3}\right)\\ &\le \frac{1}{(1-e^{-n})^3}\left[\frac{1-e^{-2}}{2ne^{2n}}+\frac{1}{40n^2e^{2n}}+\frac{1}{40n(n+1)e^{2(n+1)}}+\frac{3}{40n^2}\int_{e^n}^{e^{n+1}}\frac{dx}{x^3}\right]\\ &\le \frac{1}{2ne^{2n}}\frac{1}{(1-e^{-14})^3}\left[1-e^{-2}+\frac{1}{20\cdot 14}+\frac{1}{20e^2\cdot 15}+\frac{3}{40\cdot 14}\right]\\ &< \frac{0.88}{2ne^{2n}}.\end{aligned}$$

□

Appendix B. **The relative Lovász Local Lemma**

For completeness, and for the reader's convenience, we record a proof of the relative form of the Lovász Local Lemma used in our argument. We emphasize that the proof is the standard one (see, e.g., [1, pp. 54–55]), although the conclusion that we need is not typically recorded.

Recall the statement of the lemma.

Lemma (Lovász Local Lemma, relative form). *Let $\{A_u\}_{u\in V}$ be a finite collection of events in a probability space. Let $D=(V,E)$ be a directed graph, such that, for each $u\in V$, event A_u is independent of the sigma-algebra generated by the events $\{A_v : (u,v)\notin E\}$. Suppose that there exist real numbers $\{x_u\}_{u\in V}$, satisfying $0\le x_u<1$, and for each $u\in V$,*

$$\mathbf{P}(A_u) \le x_u \prod_{(u,v)\in E}(1-x_v).$$

Then for any $\emptyset \ne U\subset V$,

$$\mathbf{P}\left(\bigcap_{u\in V}A_u^c\right) \ge \mathbf{P}\left(\bigcap_{u\in U}A_u^c\right)\cdot \prod_{v\in V\setminus U}(1-x_v). \tag{11}$$

In particular, taking U to be a singleton,

$$\mathbf{P}\left(\bigcap_{u\in V} A_u^c\right) \geq \prod_{u\in V}(1-x_u). \tag{12}$$

Proof. By assigning an ordering to V, identify it with the set $\{1,2,\ldots,n\}$ for some n. Assume that in this ordering U is identified with $\{1,2,\ldots,m\}$ for some m. The following is to be shown by induction. For $k=1,2,\ldots,n$,

(1) For any $S\subset\{1,\ldots,n\}$, $|S|=k-1$, and for any $1\leq i\leq n$, $i\notin S$, we have

$$\mathbf{P}\left(A_i \,\middle|\, \bigcap_{j\in S} A_j^c\right)\leq x_i.$$

(2) For any $S\subset\{1,\ldots,n\}$, $|S|=k$ we have

$$\mathbf{P}\left(\bigcap_{j\in S} A_j^c\right)\geq \prod_{j\in S}(1-x_j).$$

Obviously (12) is the second item when $k=n$. The conclusion (11) is also easily deduced:

$$\mathbf{P}\left(\bigcap_{i=1}^n A_i^c\right)=\mathbf{P}\left(\bigcap_{i=1}^m A_i^c\right)\cdot\prod_{j=m+1}^n \mathbf{P}\left(A_j^c\,\middle|\,\bigcap_{i=1}^{j-1}A_i^c\right)\geq \mathbf{P}\left(\bigcap_{i=1}^m A_i^c\right)\cdot\prod_{j=m+1}^n(1-x_j).$$

When $k=1$, the conditional statement is to be interpreted as if there is no conditioning, and both statements are then obvious.

To induce, let $1<k\leq n$ and assume the truth of both statements for any $1\leq k'<k$. We first prove statement (1) in case k. Note that by the case $k-1$ of statement (2), the conditional probability in (1) is well defined. Let $S_1=\{j\in S:(i,j)\in E\}$, and let $S_2=S\setminus S_1$. We may obviously assume that $S_1=\{j_1<j_2<\cdots<j_r\}$ is nonempty, since otherwise the result is immediate by independence. We have

$$\mathbf{P}\left(A_i\,\middle|\,\bigcap_{j\in S}A_j^c\right)=\frac{\mathbf{P}\left(A_i\cap\bigcap_{j\in S_1}A_j^c\,\middle|\,\bigcap_{j\in S_2}A_j^c\right)}{\mathbf{P}\left(\bigcap_{j\in S_1}A_j^c\,\middle|\,\bigcap_{j\in S_2}A_j^c\right)}.$$

For the denominator, we have the lower bound

$$\mathbf{P}\left(A_{j_1}^c\,\middle|\,\bigcap_{j\in S_2}A_j^c\right)\cdot\mathbf{P}\left(A_{j_2}^c\,\middle|\,A_{j_1}^c\cap\bigcap_{j\in S_2}A_j^c\right)$$
$$\cdots\cdot\mathbf{P}\left(A_{j_r}^c\,\middle|\,\bigcap_{\ell=1}^{r-1}A_{j_\ell}^c\cap\bigcap_{j\in S_2}A_j^c\right)\geq\prod_{\ell=1}^r(1-x_{j_\ell})$$

by applying (1) of the inductive assumption in cases $k'<k$.

For the numerator, we have the upper bound

$$\begin{aligned}\mathbf{P}\left(A_i \cap \bigcap_{j\in S_1} A_j^c \,\middle|\, \bigcap_{j\in S_2} A_j^c\right) &\le \mathbf{P}\left(A_i \,\middle|\, \bigcap_{j\in S_2} A_j^c\right)\\ &= \mathbf{P}(A_i) \le x_i \prod_{j:(i,j)\in E}(1-x_j).\end{aligned}$$

Combined, these two bounds prove (1) in case k.

To prove (2) in case k, let $S = \{j_1 < j_2 < \cdots < j_r\}$ and observe

$$\mathbf{P}\left(\bigcap_{j\in S} A_j^c\right) = \prod_{\ell=1}^{r} \mathbf{P}\left(A_\ell^c \,\middle|\, \bigcap_{1\le m<\ell} A_m^c\right) \ge \prod_{\ell=1}^{r}(1-x_\ell),$$

which uses (1) in case k. □

References

[1] N. Alon and J. H. Spencer, *The Probabilistic Method*, *Wiley-Intersci. Ser. Discrete Math. Optim.*, John Wiley & Sons, New York, 1992. MR 1140703. Zbl 0767.05001.

[2] S. L. G. Choi, Covering the set of integers by congruence classes of distinct moduli, *Math. Comp.* **25** (1971), 885–895. MR 0297692. Zbl 0231.10004. http://dx.doi.org/10.2307/2004353.

[3] R. F. Churchhouse, Covering sets and systems of congruences, in *Computers in Mathematical Research*, North-Holland, Amsterdam, 1968, pp. 20–36. MR 0240045. Zbl 0212.39703.

[4] P. Erdös, On integers of the form $2^k + p$ and some related problems, *Summa Brasil. Math.* **2** (1950), 113–123. MR 0044558. Zbl 0041.36808.

[5] P. Erdős, Some unsolved problems, *Michigan Math. J.* **4** (1957), 291–300. MR 0098702. Zbl 0081.00102. http://dx.doi.org/10.1307/mmj/1028997963.

[6] P. Erdős, Quelques problèmes de théorie des nombres, in *Monographies de L'Enseignement Mathématique*, *No.* 6, Université de Genève, Geneva, 1963, pp. 81–135. MR 0158847. Zbl 0117.02901.

[7] P. Erdős, Some problems in number theory, in *Computers in Number Theory* (Proc. Science Research Council Atlas Sympos. No. 2, Oxford (August 1823, 1969)) (A. O. L. Atkin and B. J. Birch, eds.), Academic Press, New York, 1971, p. 405414. MR 0314733. Zbl 0217.03101.

[8] P. Erdős, Résultats et problèmes en théorie des nombres, in *Séminaire Delange-Pisot-Poitou* (14*e année*: 1972/73), *Théorie des nombres*, *Fasc.* 2, *Exp. No.* 24, Secrétariat Mathématique, Paris, 1973, p. 7. MR 0396376. Zbl 0319.10002.

[9] P. Erdős and R. L. Graham, *Old and New Problems and Results in Combinatorial Number Theory*, *Mongr. Enseign. Math.* **28**, Université de Genève, Geneva, 1980. MR 0592420. Zbl 0434.10001.

[10] L. Faber and H. Kadiri, New bounds for $\psi(x)$, to appear in *Math. Comp.* arXiv 1310.6374v1.

[11] M. Filaseta, K. Ford, S. Konyagin, C. Pomerance, and G. Yu, Sieving by large integers and covering systems of congruences, *J. Amer. Math. Soc.* **20** (2007), 495–517. MR 2276778. Zbl 1210.11020. http://dx.doi.org/10.1090/S0894-0347-06-00549-2.

[12] D. J. Gibson, A covering system with least modulus 25, *Math. Comp.* **78** (2009), 1127–1146. MR 2476575. Zbl 1208.11019. http://dx.doi.org/10.1090/S0025-5718-08-02154-6.

[13] R. K. Guy, *Unsolved Problems in Number Theory*, second ed., *Problem Books in Mathematics*, Springer-Verlag, New York, 1994. MR 1299330. Zbl 0805.11001. http://dx.doi.org/10.1007/978-1-4899-3585-4.

[14] C. E. Kruckenberg, Covering sets of the integers, 1971, Ph.D. thesis, Univ. Illinois at Urbana-Champaigne.

[15] R. Morikawa, Some examples of covering sets, *Bull. Fac. Liberal Arts Nagasaki Univ.* **21** (1981), 1–4. MR 0639635. Zbl 0462.10003.

[16] P. P. Nielsen, A covering system whose smallest modulus is 40, *J. Number Theory* **129** (2009), 640–666. MR 2488595. Zbl 1234.11011. http://dx.doi.org/10.1016/j.jnt.2008.09.016.

[17] J. B. Rosser and L. Schoenfeld, Sharper bounds for the Chebyshev functions $\theta(x)$ and $\psi(x)$, *Math. Comp.* **29** (1975), 243–269. MR 0457373. Zbl 0295.10036. http://dx.doi.org/10.2307/2005479.

[18] T. Tao and V. Vu, *Additive Combinatorics*, *Cambridge Stud. Adv. Math.* no. 105, Cambridge Univ. Press, Cambridge, 2006. MR 2289012. Zbl 1127.11002. http://dx.doi.org/10.1017/CBO9780511755149.

[19] The PARI Group, Pari/gp, version 2.5.5, 2013, Bordeaux. Available at http://pari.math.u-bordeaux.fr.

(Received: December 22, 2013)

Mathematical Institute, University of Oxford, Oxford, UK
E-mail: hough@maths.ox.ac.uk