THE STRUCTURE AND NUMBER OF ERDŐS COVERING SYSTEMS

PAUL BALISTER, BÉLA BOLLOBÁS, ROBERT MORRIS,
JULIAN SAHASRABUDHE, AND MARIUS TIBA

ABSTRACT. Introduced by Erdős in 1950, a *covering system* of the integers is a finite collection of arithmetic progressions whose union is the set $\mathbb{Z}$. Many beautiful questions and conjectures about covering systems have been posed over the past several decades, but until recently little was known about their properties. Most famously, the so-called minimum modulus problem of Erdős was resolved in 2015 by Hough, who proved that in every covering system with distinct moduli, the minimum modulus is at most 10^{16}.

In this paper we answer another question of Erdős, asked in 1952, on the *number* of minimal covering systems. More precisely, we show that the number of minimal covering systems with exactly n elements is

$$\exp\left(\left(\frac{4\sqrt{\tau}}{3} + o(1)\right)\frac{n^{3/2}}{(\log n)^{1/2}}\right)$$

as $n \to \infty$, where

$$\tau = \sum_{t=1}^{\infty}\left(\log\frac{t+1}{t}\right)^2.$$

En route to this counting result, we obtain a structural description of all covering systems that are close to optimal in an appropriate sense.

1. INTRODUCTION

A *covering system* is a finite collection of arithmetic progressions that covers[1] the integers. Erdős [5] initiated the study of covering systems in 1950, and since then numerous beautiful questions have been asked about their properties (see, for example, [5–13, 20, 21]). Until recently little progress had been made on these problems, but following groundbreaking work of Filaseta, Ford, Konyagin, Pomerance and Yu [13] in 2007, a fundamental result was obtained by Hough [16], who resolved a problem from the original paper of Erdős [5] by proving that there do not exist covering systems with distinct moduli and arbitrarily large minimum modulus. Building on his work, the authors of this paper [1, 2] recently made further progress on several related open problems.

In this paper we will study another problem on covering systems, whose study was initiated by Erdős [6] in 1952:

<div align="center">How many minimal covering systems of size n are there?</div>

[1]We emphasize that we do *not* require the progressions to be disjoint. For related work on covering systems with this additional property (sometimes called *exactly covering systems*), see for example [14, 15, 19, 23].

Erdős [6] gave a simple proof that there are only finitely many minimal[2] covering systems of size n, but the bound he obtained on their number was doubly exponential. A more reasonable upper bound follows from a result of Simpson [22], who proved in 1985 (see Section 2) that the largest modulus in a minimal covering system of size n is at most 2^{n-1}. Note that this bound is best possible, since $\mathcal{A} = \left\{ 2^{i-1} \pmod{2^i} : i \in [n-1] \right\} \cup \left\{ 0 \pmod{2^{n-1}} \right\}$ is a minimal covering system, and that it easily implies that there are at most $2^{O(n^2)}$ minimal covering systems of size n. We will show that there are in fact rather fewer such systems, and we will moreover determine asymptotically the logarithm of their number. The main aim of this paper is to prove the following theorem.

Theorem 1.1. *The number of minimal covering systems of $\mathbb{Z}$ of size n is*

$$\exp\left(\left(\frac{4\sqrt{\tau}}{3} + o(1) \right) \frac{n^{3/2}}{(\log n)^{1/2}} \right) \tag{1}$$

as $n \to \infty$, where

$$\tau = \sum_{t=1}^{\infty} \left(\log \frac{t+1}{t} \right)^2.$$

We remark that proving a weaker upper bound, with a different constant in the exponent, is significantly easier, and we will give a short proof of such a bound in Section 6. Let us also note here that we will prove the lower bound under the additional restriction that the moduli are distinct, and so the conclusion of Theorem 1.1 also holds for such systems.

In order to motivate the form of the formula (1), let us begin by describing a simple construction that gives a slightly weaker lower bound. Let $p_1 < \ldots < p_k$ be the first k primes, and for each $i \in [k]$, choose $p_i - 1$ arithmetic progressions $A_1^{(i)}, \ldots, A_{p_i-1}^{(i)}$ with the following properties: for each $j \in [p_i - 1]$, the modulus of $A_j^{(i)}$ is divisible by p_i and divides $Q_i := p_1 \cdots p_i$, and $A_j^{(i)}$ contains $j \cdot Q_{i-1}$. It is not difficult to show that, for each such choice, by adding the progression $\left\{ 0 \pmod{Q_k} \right\}$ we obtain a distinct minimal covering system of size $n = \sum_{i=1}^{k}(p_i - 1) + 1 \approx k^2 \log k$. Since we have 2^{i-1} choices for the progression $A_j^{(i)}$ for each $i \in [k]$ and $j \in [p_i - 1]$, this implies that there are at least

$$\prod_{i=1}^{k} 2^{(i-1)(p_i-1)} = \exp\left(\Omega\left(k^3 \log k\right) \right) = \exp\left(\frac{\Omega\left(n^{3/2}\right)}{(\log n)^{1/2}} \right)$$

minimal covering systems of $\mathbb{Z}$ of size n. In Section 5 we will describe a somewhat more complicated construction that proves the lower bound in Theorem 1.1.

We will refer to collections of progressions as in the construction above as "frames" (see Section 2 for a precise definition). The second main result of this paper, and the key step in the proof of Theorem 1.1, will be a structural description of all "efficient" covering systems; roughly speaking, we will show that every such covering system contains a large "approximate frame". The purpose of the next section is to state this structural theorem.

---

[2]A covering system $\mathcal{A}$ is *minimal* if no proper subset of it covers $\mathbb{Z}$. Without this restriction there are infinitely many covering systems of size 2, since we can take $\mathcal{A} = \{\mathbb{Z}, A\}$ for any arithmetic progression A.

2. The structure of efficient coverings

In this section we will state our main structural theorem. In order to do so, it will be convenient to shift our attention to the following (slightly more general) *geometric* setting. Let $S_1, \ldots, S_k$ be finite sets with at least two elements and set $S_I := \prod_{i \in I} S_i$ for each $I \subseteq [k]$. If $H = H_1 \times \cdots \times H_k \subseteq S_{[k]}$ with each H_i either equal to S_i or a singleton element of S_i, then we say that H is a *hyperplane*. We write $F(H) := \{i \in [k] : |H_i| = 1\}$ for the *fixed coordinates* of H, and $F(\mathcal{A}) := \bigcup_{H \in \mathcal{A}} F(H)$ if $\mathcal{A}$ is a collection of hyperplanes. We will also write $H = [x_1, \ldots, x_k]$, where $x_i \in S_i \cup \{*\}$ for each $i \in [k]$, and $*$ indicates that $H_i = S_i$.

Definition 2.1. A *simple frame* centred at an element $(s_1, \ldots, s_k) \in S_{[k]}$ (which we call the *axis*), is a sequence $(\mathcal{F}_1, \ldots, \mathcal{F}_k)$, where $\mathcal{F}_i$ is a collection of $|S_i| - 1$ hyperplanes of the form

$$[x_1, \ldots, x_{i-1}, a, *, \cdots, *], \tag{2}$$

one for each $a \in S_i \setminus \{s_i\}$, with $x_j \in \{s_j, *\}$ for each $j \in [i-1]$.

A *frame* is obtained from a simple frame by permuting the order of the sets $S_1, \ldots, S_k$.

Observe that if $(\mathcal{F}_1, \ldots, \mathcal{F}_k)$ is a frame centred at $(s_1, \ldots, s_k)$, then the collection

$$\mathcal{A} := \mathcal{F}_1 \cup \cdots \cup \mathcal{F}_k \cup \big\{[s_1, \ldots, s_k]\big\}$$

is a minimal cover of $S_{[k]}$. Indeed, if we remove the hyperplane $[x_1, \ldots, x_{i-1}, a, *, \cdots, *]$ from $\mathcal{A}$, then the element $(s_1, \ldots, s_{i-1}, a, s_{i+1}, \ldots, s_k)$ will be uncovered by the remaining hyperplanes. Note that if we set $S_i = \{0, \ldots, p_i - 1\}$ for each $i \in [k]$, then the construction given in the introduction is equivalent to a frame centred at $(0, \ldots, 0)$. When we (for now informally, but later on precisely) discuss frames in $\mathbb{Z}$, we will always mean that each set $S_i = \{0, \ldots, p - 1\}$ for some prime p (these primes will not generally be distinct), and we will map $S_1 \times \cdots \times S_k$ into $\mathbb{Z}_N$, where $N = \prod_{i=1}^{k} |S_i|$, using the Chinese Remainder Theorem to identify $\mathbb{Z}_N$ with the product of groups $\mathbb{Z}_{p^\gamma}$, and then expanding base p.[3] Note that every arithmetic progression in $\mathbb{Z}$ corresponds to a hyperplane, but not every hyperplane corresponds to an arithmetic progression if primes are repeated (see Sections 5 and 7).

The key idea behind the proof of Theorem 1.1 is the following (imprecise) conjecture:

"Almost every minimal covering system of $\mathbb{Z}$ of size n is close to a frame."

We will not prove a result of this form; instead, we will use a slightly weaker notion, which we call a *δ-generalized frame*. These objects differ from frames in two key ways: the fixed elements "to the left" of i in a hyperplane $H \in \mathcal{F}_i$ are allowed to vary with i, and instead of insisting that "all coordinates to the right are free" (as in (2)), we allow a few "small" coordinates to be fixed (with the product of their sizes bounded by $1/\delta$).

The next definition is both important and somewhat technical, and we will need some additional notation. Given a hyperplane H, we write H_i for its ith coordinate, and for any $I \subseteq [k]$ we will write $H_I = \prod_{i \in I} H_i$ for the hyperplane in S_I obtained by restricting H to the coordinates of I, and define $\mu_I(H) := |H_I| \cdot |S_I|^{-1}$ when $I \neq \emptyset$, and $\mu_\emptyset(H) := 1$.

---

[3]For example, $(a_0, a_1, \ldots, a_{\gamma-1}) \in \mathbb{Z}_p \times \cdots \times \mathbb{Z}_p$ corresponds to the element $\sum_{i=0}^{\gamma-1} a_i p^i \in \mathbb{Z}_{p^\gamma}$.

Definition 2.2 (δ-generalized frames). Let $\delta > 0$, and let $S_1, \ldots, S_k$ be finite sets with at least two elements. A *simple δ-generalized frame* in $S_{[k]}$ is a sequence $(\mathcal{F}_1, \ldots, \mathcal{F}_k)$, where $\mathcal{F}_i$ is a collection of at most $|S_i| - 1$ hyperplanes, satisfying the following conditions. For each $i \in [k]$, there exists a set $I(i) \supseteq \{i+1, \ldots, k\}$, and for each $j \notin I(i) \cup \{i\}$, there exists an element $s_j(i) \in S_j$, such that, for each $H \in \mathcal{F}_i$,

$$i \in F(H), \qquad \mu_{I(i)}(H) > \delta \qquad \text{and} \qquad H_j \in \{s_j(i), S_j\}.$$

Moreover, if $\min\{|S_i|, |S_j|\} \geqslant \delta^{-1}$ and $i \neq j$, then $\mathcal{F}_i$ and $\mathcal{F}_j$ are disjoint. A *δ-generalized frame* is obtained from a simple δ-generalized frame by permuting the sets $S_1, \ldots, S_k$.

We are now ready to state our main structural theorem for covering systems that contain roughly (up to a constant factor) the same number of elements as a frame.

Theorem 2.3. *For every $C, \varepsilon > 0$ there exists $\delta = \delta(C, \varepsilon) > 0$ so that for every collection of finite sets $S_1, \ldots, S_k$ with at least two elements, the following holds. If $\mathcal{A}$ is a minimal cover of $S_{[k]}$ with hyperplanes such that $F(\mathcal{A}) = [k]$ and*

$$|\mathcal{A}| \leqslant C \sum_{i=1}^{k} \big(|S_i| - 1\big), \tag{3}$$

then $\mathcal{A}$ contains a δ-generalized frame $(\mathcal{F}_1, \ldots, \mathcal{F}_k)$, with

$$\sum_{i=1}^{k} |\mathcal{F}_i| \geqslant (1 - \varepsilon) \sum_{i=1}^{k} \big(|S_i| - 1\big). \tag{4}$$

The theorem above can be thought of as an inverse theorem for the following extremal result of Simpson [22]. If $\mathcal{A}$ is a collection of arithmetic progressions, then we write $\mathrm{lcm}(\mathcal{A})$ for the least common multiple of the moduli of the progressions in $\mathcal{A}$.

Theorem 2.4 (Simpson's theorem). *If $\mathcal{A}$ is a minimal cover of $S_{[k]}$ with hyperplanes such that $F(\mathcal{A}) = [k]$, then*

$$|\mathcal{A}| \geqslant \sum_{i=1}^{k} \big(|S_i| - 1\big) + 1.$$

In particular, if $\mathcal{A}$ is a minimal covering system of $\mathbb{Z}$ with $\mathrm{lcm}(\mathcal{A}) = p_1^{\gamma_1} \cdots p_m^{\gamma_m}$, then

$$|\mathcal{A}| \geqslant \sum_{i=1}^{m} \gamma_i \big(p_i - 1\big) + 1.$$

In the appendix, we will provide (for the reader's convenience) a proof of Simpson's theorem. Let us also remark here that, while the form of the function $\delta(C, \varepsilon)$ will not matter for our purposes, we will prove that Theorem 2.3 holds with $\delta = (\varepsilon/C)^{O(\log(1/\varepsilon))}$.

In order to deduce Theorem 1.1 from Theorem 2.3, we will need to count δ-generalized frames quite precisely, and show that there are relatively few choices for the remaining elements; we will also need to show that there are few minimal covering systems that fail to satisfy (3). These calculations are carried out in Sections 6 and 7.

2.1. **An outline of the proof of the structural theorem.** The proof of Theorem 2.3 requires a few somewhat technical definitions, and to prepare the reader for these we will begin by giving an outline of the argument. The idea is to construct a tree that encodes the structure of the covering system by 'exploring' it coordinate by coordinate. To be more precise, given a minimal cover $\mathcal{A}$ of $S_{[k]}$, let us choose a coordinate $i \in [k]$ to explore, and observe (see Section 3.2 for the details) that for each $s \in S_i$ we obtain a covering system of

$$S_1 \times \cdots \times S_{i-1} \times \{s\} \times S_{i+1} \times \cdots \times S_k,$$

which we identify with $S_{[k]\setminus\{i\}}$. (Here the hyperplanes $H \in \mathcal{A}$ with $H_i = S_i$ appear in each of the $|S_i|$ covering systems corresponding to coordinate i.) These covering systems may not be minimal, but for each $s \in S_i$ we can take a minimal sub-covering $\mathcal{A}_s$.

Now, some of the systems $\mathcal{A}_s$ may be trivial (i.e., may consist of a single hyperplane), and when this occurs we are happy, because such hyperplanes can be used in the frame that we are trying to construct. For the remaining elements $s \in S_i$, we consider the set of fixed coordinates $F(\mathcal{A}_s)$ of $\mathcal{A}_s$, and observe (see Lemma 3.7) that every coordinate (except i) is in $F(\mathcal{A}_s)$ for some $s \in S_i$. We may now choose, for each s such that $F(\mathcal{A}_s)$ is non-empty, a coordinate $j \in F(\mathcal{A}_s)$, and repeat the above construction, exploring the minimal covering system $\mathcal{A}_s$, starting with the coordinate j. Iterating this process produces a rooted tree (which we call an 'index tree', see Definition 3.1), each of whose vertices is labelled with a set $I \subseteq [k]$ and a coordinate $i \in I$, which are the fixed coordinates of the corresponding minimal covering system, and the coordinate 'explored' at that vertex, respectively.

So far, we have not said anything about how to construct the sets $\mathcal{F}_i$, or how to choose the coordinate i that we explore in a given step. For simplicity, let us explain this only for the first step (the choice for later steps is similar). First, if there exists $i \in [k]$ such that there are at least $(1 - \varepsilon)(|S_i| - 1)$ hyperplanes $H \in \mathcal{A}$ with $i \in F(H)$ and $\mu_{[k]\setminus\{i\}}(H) > \delta$, then we choose such a coordinate i to explore, and associate this collection of 'frame-like' hyperplanes with the current vertex (in this case, the root of the tree). One of the key ideas of the paper is that, if such a collection of hyperplanes does not exist for any $i \in [k]$, then we may use the Lovász Local Lemma to deduce (see Lemma 3.5) that there exists a coordinate j (which we will choose to explore), and a 'large' collection $\mathcal{G}$ of hyperplanes in the current collection, such that j is a fixed coordinate of each. This collection of 'garbage' hyperplanes will later be used, together with (3), to show that this case does not occur too often.

The plan described above is carried out in Section 3, the main result being Lemma 3.3, which states that if δ is sufficiently small, then there exists a suitable 'exploration tree' $\mathbb{T}$ of $\mathcal{A}$ (see Definition 3.2). This exploration tree can be very large, however, and to extract our δ-generalized frame from it we will need to choose a suitable sub-tree T. To do so, we choose k 'special' vertices of $\mathbb{T}$, one for each coordinate, and take the union of the paths from these vertices to the root. If almost all of these special vertices are 'good' (that is, we found a large collection of frame-like hyperplanes when exploring them), then we obtain a sufficiently large δ-generalized frame. On the other hand, if a positive proportion of them are 'bad', then we use the 'garbage' hyperplanes to show that inequality (3) cannot hold. In order to carry out this argument, we need to choose the special vertices carefully; it turns out that it is sufficient to choose them via a depth-first search, see Section 4 for the details.

3. Exploring the cover

In this section we will take the first step towards Theorem 2.3 by describing a much larger object that is somewhat easier to construct, the *exploration tree*. To define these, we first need to introduce the following simpler objects, which we call *index trees*. Let us fix, for the rest of the proof of Theorem 2.3, a collection of finite sets $S_1, \ldots, S_k$ with at least two elements, and let us write $N(u)$ for the set of out-neighbours of u in a rooted tree, where we orient the edges away from the root.

Definition 3.1. An *index tree* $\mathbb{T}$ of $[k]$ is a rooted tree, equipped with a labelling of its vertices $u \mapsto (I_u, i_u)$, where $I_u \subseteq [k]$ and $i_u \in I_u$, that satisfies the following conditions:

    (i) the root of $\mathbb{T}$ has label $([k], i)$ for some $i \in [k]$;

    (ii) for each vertex $v \in V(\mathbb{T})$,

$$\bigcup_{u \in N(v)} I_u = I_v \setminus \{i_v\}. \tag{5}$$

We can now define the exploration tree of a collection of hyperplanes in $S_{[k]}$. Given a rooted tree T and vertices $u, v \in V(T)$, let us write $u \prec_T v$ to indicate that u lies on the path from v to the root (so, in particular, $v \prec_T v$).

Definition 3.2. Let $\lambda, \varepsilon, \delta > 0$, and let $\mathcal{A}$ be a collection of hyperplanes in $S_{[k]}$. An $(\lambda, \varepsilon, \delta)$-*exploration tree* of $\mathcal{A}$ is an index tree $\mathbb{T}$ of $[k]$ such that, for each vertex $u \in V(\mathbb{T})$:

    (a) there exists a collection $\mathcal{A}_u \subseteq \mathcal{A}$ such that $\mathcal{A}'_u := \{H_{I_u} : H \in \mathcal{A}_u\}$ is a minimal cover of S_{I_u} with $F(\mathcal{A}'_u) = I_u$, and if $u \in N(v)$, then:

        (i) $\mathcal{A}_u \subseteq \mathcal{A}_v$;

        (ii) $F(\mathcal{A}_u) \subseteq \{i_w : w \prec_{\mathbb{T}} v\} \cup I_u$;

        (iii) there exists an element $s_u \in S_{i_v}$ such that $H_{i_v} \in \{s_u, S_{i_v}\}$ for each $H \in \mathcal{A}_u$.

Moreover, for each vertex $u \in V(\mathbb{T})$, one of the following holds:

    (b) u is *good*, which means that there exists a collection of hyperplanes $\mathcal{F}_u \subseteq \mathcal{A}_u$, with

$$|\mathcal{F}_u| \geqslant (1 - \varepsilon)(|S_{i_u}| - 1), \tag{6}$$

    such that $i_u \in F(H)$ and $\mu_{I_u \setminus \{i_u\}}(H) > \delta$ for each $H \in \mathcal{F}_u$.

    (c) u is *bad*, which means that there exists a collection of hyperplanes $\mathcal{G}_u \subseteq \mathcal{A}_u$, with

$$\sum_{H \in \mathcal{G}_u} 2^{-|F(H) \cap I_u|/4} \geqslant \frac{|S_{i_u}|}{\lambda}, \tag{7}$$

    such that $i_u \in F(H)$ and $|F(H) \cap I_u| \geqslant 2$ for each $H \in \mathcal{G}_u$.

We think of the elements of $\mathcal{F}_u$ (when u is good) and $\mathcal{G}_u$ (when u is bad) as hyperplanes that (respectively) do and do not look like parts of a frame from the perspective of the vertex u. We will show (see Lemma 3.3, below) that exploration trees always exist, as long as we choose δ to be sufficiently small, depending on λ and ε. We will then, in Section 4,

carefully choose a subtree T of our exploration tree $\mathbb{T}$, and one 'special' vertex for each coordinate $i \in [k]$, with the following three properties: the frames corresponding to good special vertices are disjoint (unless one of the corresponding sets S_i is very small); if 'many' special vertices of T are bad, then $\mathcal{A}$ fails to satisfy (3); and if 'almost all' of the special vertices of T are good, then there exists a sufficiently large δ-generalized frame in $\mathcal{A}$.

The main aim of this section is to prove the following lemma.

Lemma 3.3. *Let $\lambda, \varepsilon \in (0, 1)$, and let $\mathcal{A}$ be a minimal cover of $S_{[k]}$ with hyperplanes such that $F(\mathcal{A}) = [k]$. If*

$$0 < \delta < 2^{-9} \lambda^2 \varepsilon^{2 \log_2(1/\lambda\varepsilon) + 11}, \tag{8}$$

then there exists an $(\lambda, \varepsilon, \delta)$-exploration tree of $\mathcal{A}$.

Let us fix, for the rest of this section, constants $0 < \lambda, \varepsilon < 1$ and $\delta > 0$ satisfying (8). We will prove Lemma 3.3 by iteratively extending a 'partial' exploration tree by applying the following lemma to a leaf of the current tree.

Lemma 3.4. *Let $\emptyset \neq I \subseteq [k]$, and let $\mathcal{A}$ be a minimal cover of S_I with $F(\mathcal{A}) = I$.*

(a) For each $i \in I$, there exists a map $J \colon S_i \to \mathcal{P}(I \setminus \{i\})$, with

$$\bigcup_{s \in S_i} J(s) = I \setminus \{i\}, \tag{9}$$

and for each $s \in S_i$ there exists $\mathcal{A}_s \subseteq \mathcal{A}$ such that $\mathcal{A}'_s := \{ H_{J(s)} : H \in \mathcal{A}_s \}$ is a minimal cover of $S_{J(s)}$, $F(\mathcal{A}_s) \setminus \{i\} = J(s)$, and $H_i \in \{s, S_i\}$ for each $H \in \mathcal{A}_s$.

Moreover, there exists $i \in I$ such that one of the following holds:

(b) there exists a collection $\mathcal{F} \subseteq \mathcal{A}$, with

$$|\mathcal{F}| \geqslant (1 - \varepsilon)\big(|S_i| - 1\big),$$

such that $i \in F(H)$ and $\mu_{I \setminus \{i\}}(H) > \delta$ for each $H \in \mathcal{F}$.

(c) there exists a collection $\mathcal{G} \subseteq \mathcal{A}$, with

$$\sum_{H \in \mathcal{G}} 2^{-|F(H)|/4} \geqslant \frac{|S_i|}{\lambda}, \tag{10}$$

such that $i \in F(H)$ and $|F(H)| \geqslant 2$ for each $H \in \mathcal{G}$.

Let us fix a set $\emptyset \neq I \subseteq [k]$ until the end of the proof of Lemma 3.4. This section is organised as follows: in Section 3.1 we will prove two (straightforward) technical lemmas, in Section 3.2, we will introduce the operation that we will use to construct the map J and the families $\mathcal{A}_s$, and in Section 3.3, we will prove Lemma 3.4, and deduce Lemma 3.3.

3.1. **Two technical lemmas.** Our first technical lemma (Lemma 3.5, below) follows from a straightforward application of the Lovász Local Lemma. We will apply it, in the case that there does not exist a collection $\mathcal{F} \subseteq \mathcal{A}$ as in Lemma 3.4(b) for any $i \in I$, to a certain subset $\prod_{i \in I} R_i \subseteq S_I$, in order to find an index such that (c) holds. Our second technical lemma

(Lemma 3.6, below) will allow us to deduce the bound (10) from the condition given by the local lemma. Let us say that a hyperplane H in R_I is *non-trivial* if $H \neq R_I$.

Lemma 3.5. *Let $0 < \eta < 1/5$ and let $\{R_i : i \in I\}$ be a collection of finite sets, each with at least two elements. Let $\mathcal{A}$ be a collection of non-trivial hyperplanes in R_I, and let $\tilde{\mu}$ denote the uniform measure on R_I. If*

$$\sum_{H \in \mathcal{A} : i \in F(H)} e^{\eta|F(H)|} \tilde{\mu}(H) < \frac{\eta}{2} \tag{11}$$

for every $i \in [k]$, then R_I is not covered by the hyperplanes in $\mathcal{A}$.

Proof. We choose a point $y \in R_I$ uniformly at random and apply the local lemma. For each hyperplane $H \in \mathcal{A}$ we define E_H to be the ("bad") event that $y \in H$. Observe that $\mathbb{P}(E_H) = \tilde{\mu}(H)$, and define a dependency graph G on the events $\{E_H\}_{H \in \mathcal{A}}$ by setting $E_H \sim E_{H'}$ if $F(H) \cap F(H') \neq \emptyset$. Observe that if $F(H) \cap \big(F(H^{(1)}) \cup \cdots \cup F(H^{(t)})\big) = \emptyset$, then E_H is independent of the collection $\{E_{H^{(1)}}, \ldots, E_{H^{(t)}}\}$, so G is a valid dependency graph.

Next, we define weights

$$x(H) = e^{\eta|F(H)|} \tilde{\mu}(H)$$

for each $H \in \mathcal{A}$. To apply the local lemma we need to show that

$$\mathbb{P}(E_H) \leqslant x(H) \prod_{E_H \sim E_{H'}} \big(1 - x(H')\big).$$

To do so, we first claim that $1 - x(H) \geqslant e^{-2x(H)}$ for every $H \in \mathcal{A}$. This holds because

$$x(H) = e^{\eta|F(H)|} \tilde{\mu}(H) \leqslant \big(e^\eta/2\big)^{|F(H)|} \leqslant \big(1 - e^{-1}\big)^{|F(H)|} \leqslant 1 - e^{-1},$$

where the first inequality is $\tilde{\mu}(H) \leqslant 2^{-|F(H)|}$, which holds because each set S_i has at least two elements, the second follows since $\eta < 1/5$, and the third since the hyperplanes in $\mathcal{A}$ are non-trivial, so $|F(H)| \geqslant 1$. Therefore, for each $H \in \mathcal{A}$, we have

$$\prod_{E_H \sim E_{H'}} \big(1 - x(H')\big) \geqslant \exp\bigg(- 2 \sum_{E_H \sim E_{H'}} x(H') \bigg) \geqslant \exp\bigg(- 2 \sum_{i \in F(H)} \sum_{H' \in \mathcal{A} : i \in F(H')} x(H') \bigg)$$

$$= \exp\bigg(- 2 \sum_{i \in F(H)} \sum_{H' \in \mathcal{A} : i \in F(H')} e^{\eta|F(H')|} \tilde{\mu}(H') \bigg) \geqslant \exp\big(- \eta|F(H)| \big),$$

where the last inequality follows from (11). This implies that

$$x(H) \prod_{E_H \sim E_{H'}} \big(1 - x(H')\big) \geqslant x(H) e^{-\eta|F(H)|} = \tilde{\mu}(H) = \mathbb{P}(E_H),$$

as required. By the local lemma, it follows that the probability that none of the events E_H holds is non-zero, and hence there exists a point $y \in R_I$ that is not covered by $\mathcal{A}$. $\square$

The second technical lemma is even more straightforward. Recall that $S_1, \ldots, S_k$ are fixed finite sets with at least two elements, and that the (non-empty) set $I \subseteq [k]$ and positive constants λ and ε were fixed above.

8

Lemma 3.6. *For each $j \in I$, let $R_j \subseteq S_j$ be such that $|R_j| \geqslant \varepsilon\big(|S_j| - 1\big) + 1$, and let $\tilde{\mu}$ denote the uniform measure on R_I. Let H be a hyperplane in S_I, and let $i \in F(H)$. If*

$$\mu_{I \setminus \{i\}}(H) \leqslant 2^{-9} \lambda^2 \varepsilon^{2 \log_2(1/\lambda\varepsilon)+11},$$

then

$$\tilde{\mu}\big(H \cap R_I\big) \leqslant \frac{\lambda}{2^{|F(H)|/2+4}|S_i|}. \tag{12}$$

Proof. Set $\ell := |F(H)|$ and $\delta_0 := 2^{-9} \lambda^2 \varepsilon^{2 \log_2(1/\lambda\varepsilon)+11}$, and observe that

$$\tilde{\mu}\big(H \cap R_I\big) \cdot |R_i| \leqslant \mu_{I \setminus \{i\}}(H) \prod_{j \in F(H),\, i \neq j} \frac{|S_j|}{|R_j|} \leqslant \delta_0 \varepsilon^{-(\ell-1)}.$$

Now, note that $|R_j| \geqslant 2$ for every $j \in I$, and suppose that (12) does not hold. Then

$$2^{-(\ell-1)} \geqslant \prod_{j \in F(H),\, j \neq i} \frac{1}{|R_j|} = \tilde{\mu}\big(H \cap R_I\big) \cdot |R_i| \geqslant \frac{\lambda}{2^{\ell/2+4}} \cdot \frac{|R_i|}{|S_i|} \geqslant \frac{\lambda\varepsilon}{2^{\ell/2+4}},$$

and hence $\ell \leqslant 2\log_2(1/\lambda\varepsilon) + 10$. It follows that

$$2^{\ell/2+4} \cdot \tilde{\mu}\big(H \cap R_I\big) \leqslant \frac{2^{9/2}\delta_0}{|R_i|} \left(\frac{\sqrt{2}}{\varepsilon}\right)^{\ell-1} \leqslant \frac{2^{9/2}\delta_0}{|R_i|} \left(\frac{\sqrt{2}}{\varepsilon}\right)^{2\log_2(1/\lambda\varepsilon)+9} = \frac{\lambda\varepsilon}{|R_i|} \leqslant \frac{\lambda}{|S_i|},$$

as required. $\qquad\square$

3.2. An operation on a covering system.

We next introduce a simple operation that, given a minimal cover of S_I, produces a map J and a collection $\{\mathcal{A}_s : s \in S_i\}$ as required by Lemma 3.4(*a*). This operation is the basic tool we will use in the construction of our exploration trees. Recall that the (non-empty) set $I \subseteq [k]$ was fixed above, and let $\mathcal{A}$ be a minimal cover of S_I with $F(\mathcal{A}) = I$. For each $i \in I$ and $s \in S_i$, set

$$\mathcal{H}(i,s) := \big\{H \in \mathcal{A} : H_i \in \{s, S_i\}\big\},$$

and observe that the collection $\mathcal{H}'(i,s) := \big\{H_{I \setminus \{i\}} : H \in \mathcal{H}(i,s)\big\}$ is a cover of $S_{I \setminus \{i\}}$. Note that moreover, since $\mathcal{A}$ is minimal, there is a bijection between $\mathcal{H}(i,s)$ and $\mathcal{H}'(i,s)$. Let $\mathcal{A}_s^* \subseteq \mathcal{H}'(i,s)$ be an arbitrary minimal subcover of $S_{I \setminus \{i\}}$, and define

$$J(s) := F(\mathcal{A}_s^*) \qquad \text{and} \qquad \mathcal{A}_s := \big\{H \in \mathcal{H}(i,s) : H_{I \setminus \{i\}} \in \mathcal{A}_s^*\big\}.$$

Note that $\mathcal{A}_s' = \big\{H_{J(s)} : H \in \mathcal{A}_s\big\}$ is a minimal cover of $S_{J(s)}$, that $F(\mathcal{A}_s) \setminus \{i\} = J(s)$, and that $H_i \in \{s, S_i\}$ for each $H \in \mathcal{A}_s$. To verify that J and $\{\mathcal{A}_s : s \in S_i\}$ satisfy Lemma 3.4(*a*), it therefore only remains to check that (9) holds.

Lemma 3.7. *Let $\mathcal{A}$ be a minimal cover of S_I with hyperplanes. If $F(\mathcal{A}) = I$, then*

$$\bigcup_{s \in S_i} J(s) = I \setminus \{i\}$$

for each $i \in I$.

Proof. We will in fact show that for every $H \in \mathcal{A}$, there exists $s \in S_i$ with $H_{I \setminus \{i\}} \in \mathcal{A}_s^*$. Since $J(s) = F(\mathcal{A}_s^*) \subseteq I \setminus \{i\}$ and $F(\mathcal{A}) = I$, this will be enough to prove the lemma.

To prove the claim, let $x \in S_I$ be an element that is only covered by H (recall that $\mathcal{A}$ is minimal), and set $s := x_i$. We claim that $H_{I \setminus \{i\}} \in \mathcal{A}_s^*$. Indeed, since $\mathcal{A}_s^*$ is a cover of $S_{I \setminus \{i\}}$, it must cover the vector x' obtained from x by ignoring the ith coordinate, and $H_{I \setminus \{i\}}$ is the only potential element of $\mathcal{A}_s^*$ that can do so. $\qquad\square$

3.3. Construction of the exploration tree.

Having completed our preparations, we are now ready to prove Lemma 3.4, and deduce Lemma 3.3.

Proof of Lemma 3.4. Let $\emptyset \neq I \subseteq [k]$, and let $\mathcal{A}$ be a minimal cover of S_I such that $F(\mathcal{A}) = I$. To prove part (a), for each $i \in I$ we apply the construction defined in Section 3.2 to obtain a map $J: S_i \to \mathcal{P}(I \setminus \{i\})$ and a collection $\{\mathcal{A}_s : s \in S_i\}$, where $\mathcal{A}_s \subseteq \mathcal{A}$, as required. In particular, $\mathcal{A}_s' = \{H_{J(s)} : H \in \mathcal{A}_s\}$ is a minimal cover of $S_{J(s)}$, $F(\mathcal{A}_s) \setminus \{i\} = J(s)$, $H_i \in \{s, S_i\}$ for each $H \in \mathcal{A}_s$, and the map J satisfies (9) by Lemma 3.7.

To prove that there exists $i \in I$ such that either (b) or (c) holds, let us define an element $s \in S_i$ to be *special for i* if there exists a hyperplane $H \in \mathcal{A}$ such that

$$H_i = s \qquad \text{and} \qquad \mu_{I \setminus \{i\}}(H) > \delta.$$

If this holds, then we say that the hyperplane H is a *witness* for the pair (s, i). Now, for each $i \in I$ define S_i^* to be the set of elements $s \in S_i$ that are special for i, and set $R_i := S_i \setminus S_i^*$. We consider two cases, corresponding to conditions (b) and (c) of Definition 3.2, respectively.

Case 1: There exists $i \in I$ with $|R_i| < \varepsilon(|S_i| - 1) + 1$.

In this case we define

$$\mathcal{F} := \bigcup_{s \in S_i^*} \{H \in \mathcal{A} : H \text{ is a witness for } (s, i)\}.$$

Since a hyperplane H cannot witness (s, i) for more than one element $s \in S_i^*$, we have

$$|\mathcal{F}| \geqslant |S_i^*| \geqslant |S_i| - \varepsilon(|S_i| - 1) - 1 = (1 - \varepsilon)(|S_i| - 1),$$

and by definition $i \in F(H)$ and $\mu_{I \setminus \{i\}}(H) > \delta$ for each $H \in \mathcal{F}$.

Case 2: $|R_i| \geqslant \varepsilon(|S_i| - 1) + 1$ for every $i \in I$.

In this case we shall apply Lemma 3.5 to the set $R_I := \prod_{i \in I} R_i$ with $\eta = 1/6$. Define $\mathcal{A}' \subseteq \mathcal{A}$ by removing all hyperplanes that are witness for (s, i) for some $i \in I$ and $s \in S_i$. Observe that none of the witness hyperplanes intersects R_I, so $\mathcal{A}'' := \{H \cap R_I : H \in \mathcal{A}'\}$ is a cover of R_I. We claim that there exists a coordinate $i \in I$ such that

$$\sum_{H \in \mathcal{A}' : i \in F(H)} e^{\eta |F(H)|} \tilde{\mu}(H \cap R_I) \geqslant \frac{\eta}{2}, \tag{13}$$

where $\tilde{\mu}$ denotes the uniform measure on R_I. Since $\mathcal{A}''$ is a cover of R_I, and noting that $|R_i| \geqslant 2$ for every $i \in I$ (by assumption, and since $|S_i| \geqslant 2$), this will follow from Lemma 3.5 if we show that $\mathcal{A}''$ is a collection of non-trivial hyperplanes in R_I. To do so, suppose for a contradiction that $R_I \subseteq H$ for some $H \in \mathcal{A}'$, and observe that therefore $F(H) \cap I = \emptyset$, and hence also $S_I \subseteq H$. However, since $\mathcal{A}$ is a minimal cover of S_I, this implies that

$\mathcal{A} = \{H\}$, and hence $F(\mathcal{A}) = \emptyset$. This contradicts our assumption that $F(\mathcal{A}) = I$, and thus each hyperplane in $\mathcal{A}''$ is indeed non-trivial. As observed above, it follows by Lemma 3.5 that (13) holds for some $i \in I$, as claimed.

Fix such an $i \in I$, and define $\mathcal{G} := \mathcal{G}^{(1)} \cup \cdots \cup \mathcal{G}^{(k)}$, where

$$\mathcal{G}^{(\ell)} := \{H \in \mathcal{A}' : i \in F(H) \text{ and } |F(H)| = \ell\}$$

for each $\ell \in [k]$. Observe that $\mathcal{G}^{(1)} = \emptyset$, since if $H \in \mathcal{A}$ and $F(H) = \{i\}$, then $\mu_{I \setminus \{i\}}(H) = 1$, and so H would have been removed when we formed $\mathcal{A}'$. Similarly, for each $H \in \mathcal{A}'$ with $i \in F(H)$ we have

$$\mu_{I \setminus \{i\}}(H) \leqslant \delta < 2^{-9}\lambda^2 \varepsilon^{2\log_2(1/\lambda\varepsilon)+11},$$

since otherwise H would witness (s, i) for some $s \in S_i$, and so would been removed when we formed $\mathcal{A}'$. It follows, by Lemma 3.6, that

$$\tilde{\mu}(H \cap R_I) \leqslant \frac{\lambda}{2^{|F(H)|/2+4}|S_i|}$$

for every $H \in \mathcal{A}'$ with $i \in F(H)$. Finally, combining this with (13) gives

$$\sum_{\ell \geqslant 2} |\mathcal{G}^{(\ell)}| \frac{e^{\eta\ell}}{2^{\ell/2}} = \sum_{H \in \mathcal{A}' : i \in F(H)} \frac{e^{\eta|F(H)|}}{2^{|F(H)|/2}} \geqslant \sum_{H \in \mathcal{A}' : i \in F(H)} e^{\eta|F(H)|}\tilde{\mu}(H \cap R_I)\frac{16|S_i|}{\lambda} \geqslant \frac{8\eta|S_i|}{\lambda},$$

and hence

$$\sum_{H \in \mathcal{G}} 2^{-|F(H)|/4} = \sum_{\ell \geqslant 2} \frac{|\mathcal{G}^{(\ell)}|}{2^{\ell/4}} = \sum_{\ell \geqslant 2} |\mathcal{G}^{(\ell)}| \frac{e^{\eta\ell}}{2^{\ell/2}} \geqslant \frac{8\eta|S_i|}{\lambda} \geqslant \frac{|S_i|}{\lambda},$$

as required by (10). $\square$

The deduction of Lemma 3.3 is now straightforward. Let $\partial(T)$ denote the set of vertices of a rooted tree T with no out-neighbours, and call $\partial(T)$ the *boundary* of T.

Proof of Lemma 3.3. We construct $\mathbb{T}$, our exploration tree, inductively, with Lemma 3.4 providing the induction step. We begin our induction by defining T_0 to be a single vertex v, and setting $I_v := [k]$, and $\mathcal{A}_v := \mathcal{A}$. For the induction step, suppose that we have constructed a rooted tree T_t (with root v), a set $\emptyset \neq I_u \subseteq [k]$ and a collection $\mathcal{A}_u \subseteq \mathcal{A}$ for each vertex $u \in V(T_t)$, and an index $i_u \in I_u$ for each non-boundary vertex $u \in V(T_t) \setminus \partial(T_t)$, such that condition (a) of Definition 3.2 holds for every vertex $u \in V(T_t)$, and condition (ii) of Definition 3.1, and either condition (b) or (c) of Definition 3.2, hold for every non-boundary vertex $u \in V(T_t) \setminus \partial(T_t)$. Observe that, since $\mathcal{A}$ is a minimal cover of $S_{[k]}$ with hyperplanes such that $F(\mathcal{A}) = [k]$, these conditions are satisfied in the base case $t = 0$.

To construct T_{t+1}, choose a vertex $u \in \partial(T)$ such that $|I_u| \geqslant 2$, if one exists (we will deal with the other case below), and apply Lemma 3.4 to the set I_u and minimal cover $\mathcal{A}'_u = \{H_{I_u} : H \in \mathcal{A}_u\}$ of S_{I_u} (noting that $F(\mathcal{A}'_u) = I_u$, by the induction hypothesis). We obtain an index $i \in I_u$, a map $J \colon S_i \to \mathcal{P}(I \setminus \{i\})$ and a collection $\{\mathcal{A}_s : s \in S_i\}$ as in part (a) of the lemma, and either a collection $\mathcal{F} \subseteq \mathcal{A}'_u$ as in part (b), or a collection $\mathcal{G} \subseteq \mathcal{A}'_u$ as in part (c). In either case, we set $i_u := i$, add an out-neighbour of u for each element $s \in S_i$ such that $J(s) \neq \emptyset$, and to the new vertex $w(s)$ corresponding to s we assign the set

$I_{w(s)} = J(s)$, the element $s_{w(s)} := s$, and the collection of hyperplanes

$$\mathcal{A}_{w(s)} := \big\{ H \in \mathcal{A}_u : H_{I_u} \in \mathcal{A}_s \big\}.$$

Observe first that condition (a) of Definition 3.2 holds for each vertex $w(s)$. Indeed, we have $\mathcal{A}_{w(s)} \subseteq \mathcal{A}_u \subseteq \mathcal{A}$, and it follows from Lemma 3.4(a) that the family

$$\mathcal{A}'_{w(s)} = \{ H_{J(s)} : H \in \mathcal{A}_{w(s)} \} = \{ H_{J(s)} : H \in \mathcal{A}_s \}$$

is a minimal cover of $S_{J(s)}$ with $F(\mathcal{A}'_{w(s)}) = F(\mathcal{A}_s) \setminus \{i\} = J(s)$, that $H_{i_u} \in \{s, S_{i_u}\}$ for each $H \in \mathcal{A}_{w(s)}$, and that

$$F(\mathcal{A}_{w(s)}) \subseteq \big(F(\mathcal{A}_u) \setminus I_u \big) \cup F(\mathcal{A}_s) \subseteq \{ i_w : w \prec_{\mathbb{T}} u \} \cup J(s).$$

Note also that, by (9), u satisfies condition (ii) of Definition 3.1.

To complete the induction step, it remains to observe that u is either good or bad, i.e., satisfies either condition (b) or (c) of Definition 3.2. Indeed, if Lemma 3.4(b) holds then set

$$\mathcal{F}_u := \big\{ H \in \mathcal{A}_u : H_{I_u} \in \mathcal{F} \big\},$$

and if Lemma 3.4(c) holds then set

$$\mathcal{G}_u := \big\{ H \in \mathcal{A}_u : H_{I_u} \in \mathcal{G} \big\}.$$

In each case, the properties guaranteed by the lemma are exactly those that we require. Since the properties required of all vertices of T_{t+1} other than u and its out-neighbours continue to hold, it follows that T_{t+1} satisfies the same properties that we assumed for T_t.

Finally, observe that in passing from T_t to T_{t+1} we replace a vertex of the boundary by a finite number of boundary elements, each associated with with strictly smaller sets. This process must therefore eventually end, and when it does, it follows that $|I_u| = 1$ for every boundary vertex $u \in \partial(T_t)$. When this happens, we simply set i_u equal to the unique member of I_u for each $u \in \partial(T_t)$, and claim that u is good. Indeed, by the induction hypothesis, the collection $\mathcal{A}'_u$ forms a minimal cover of S_{i_u} with $i_u \in F(\mathcal{A}'_u)$. It follows that $\mathcal{A}'_u$ consists of exactly $|S_{i_u}|$ singleton hyperplanes, and so (6) holds with $\mathcal{F}_u := \mathcal{A}_u$. Since condition (ii) of Definition 3.1 holds automatically (with both sides equal to the empty set), it follows that the tree $\mathbb{T}$ that we have constructed is an $(\lambda, \varepsilon, \delta)$-exploration tree of $\mathcal{A}$, as required. $\qquad \square$

4. Extraction of the frame, and the proof of Theorem 2.3

In order to prove Theorem 2.3, we will use the exploration tree $\mathbb{T}$ constructed in the previous section, together with the bound (3), to find a δ-generalized frame for $\mathcal{A}$. Roughly speaking, we would like to do this by choosing k vertices $\beta(1), \ldots, \beta(k)$ such that the label of $\beta(i)$ in $\mathbb{T}$ is i, define a tree T to be the union of the paths (in $\mathbb{T}$) from $\beta(i)$ to the root, and define the elements $s_j(i)$ using the elements s_u. For each good vertex $\beta(i)$ we have a collection $\mathcal{F}_i$ of hyperplanes more or less as required, and for each bad vertex we obtain a large collection of 'garbage' hyperplanes. We might therefore hope to use (3) to show that there are few bad vertices, and thus to deduce the bound (4).

There are two main problems with the strategy described above: the frame elements obtained for good vertices might not be disjoint, and each hyperplane might be included in

the garbage set $\mathcal{G}_u$ for a very large number of bad vertices. We overcome both obstacles in the same way: by choosing the vertices $\beta(i)$ via a depth-first search algorithm. We do not expect the reader to be able to immediately see why this choice should help in either case, but it turns out that proving that it does is (in both cases) surprisingly simple.

In Section 4.1 we will state precisely the object we will construct, and show that its existence implies the existence of a δ-generalized frame. In Section 4.2 we will describe how we choose the sub-tree $T \subseteq \mathbb{T}$, the frame elements $(\mathcal{F}_1, \ldots, \mathcal{F}_k)$, and the 'garbage' sets $(\mathcal{G}_1, \ldots, \mathcal{G}_k)$; in Section 4.3, we will prove two lemmas on the disjointness of the frame elements and garbage sets; and in Section 4.4 we will complete the proof of Theorem 2.3.

4.1. **Tree-frames.** The purpose of this section is to introduce the following somewhat complicated objects, which also provide significantly more information (though we will not need this) about the covering system. We will use these objects to construct our frames.

Definition 4.1. Let T be a rooted tree equipped with maps

$$\alpha \colon V(T) \to [k], \qquad \beta \colon [k] \to V(T) \qquad \text{and} \qquad \gamma \colon E(T) \to S_1 \cup \cdots \cup S_k$$

such that

(a) $\alpha(u) \neq \alpha(v)$ if $u \prec_T v$ and $u \neq v$;
(b) $\alpha\big(\beta(i)\big) = i$ for each $i \in [k]$;
(c) if $e \in E(T)$ and v is the endpoint of e that is closer to the root, then $\gamma(e) \in S_{\alpha(v)}$;
(d) there exists a permutation π of $[k]$ such that, if for each $i \in [k]$ we set

$$J(i) := \big\{ \alpha(u) : u \prec_T \beta(i) \big\} \qquad \text{and} \qquad I(i) := [k] \setminus J(i),$$

then $J(\pi(i)) \subseteq \{\pi(1), \ldots, \pi(i)\}$ for each $i \in [k]$.

Now, for each $\delta > 0$, a δ-*generalized tree-frame* centred at T is a sequence $(\mathcal{F}_1, \ldots, \mathcal{F}_k)$, where $\mathcal{F}_i$ is a collection of at most $|S_i| - 1$ hyperplanes, satisfying

(i) $i \in F(H)$ for each $H \in \mathcal{F}_i$.
(ii) $\mu_{I(i)}(H) > \delta$ for each $H \in \mathcal{F}_i$.
(iii) $H_j \in \{\gamma(e), S_j\}$ for each $H \in \mathcal{F}_i$ and each $j \in J(i) \setminus \{i\}$, where $e \in E(T)$ is the edge leaving the unique vertex $v \prec_T \beta(i)$ with $\alpha(v) = j$ in the direction of $\beta(i)$;
(iv) If $\min\big\{|S_i|, |S_j|\big\} \geqslant \delta^{-1}$ and $i \neq j$, then $\mathcal{F}_i$ and $\mathcal{F}_j$ are disjoint.

In Sections 4.2–4.4 we will construct, for any $\mathcal{A}$ satisfying (3), a δ-generalized tree-frame satisfying (4). The next lemma shows that this will be sufficient to prove Theorem 2.3.

Lemma 4.2. *If $(\mathcal{F}_1, \ldots, \mathcal{F}_k)$ is a δ-generalized tree-frame centred at a rooted tree T, then $(\mathcal{F}_1, \ldots, \mathcal{F}_k)$ is a δ-generalized frame.*

Proof. Let T be a rooted tree equipped with maps α, β and γ satisfying conditions (a)–(d) of Definition 4.1. In particular, let π be the permutation given by condition (d), and (to simplify the notation) let us permute the sets $S_1, \ldots, S_k$ so that π is the identity, and therefore $J(i) \subseteq \{1, \ldots, i\}$ (and hence $I(i) \supseteq \{i + 1, \ldots, k\}$) for each $i \in [k]$. Now, for each $i \in [k]$ and $j \in J(i) \setminus \{i\}$, set $s_j(i) := \gamma(e) \in S_j$, where $e \in E(T)$ is the edge leaving the unique vertex $v \prec_T \beta(i)$ with $\alpha(v) = j$ in the direction of $\beta(i)$.

We claim that, for each $H \in \mathcal{F}_i$,

$$i \in F(H), \qquad \mu_{I(i)}(H) > \delta \qquad \text{and} \qquad H_j \in \{s_j(i), S_j\}.$$

Indeed, these follow directly from properties (i), (ii) and (iii) of Definition 4.1. Finally, observe that if $\min \{|S_i|, |S_j|\} \geqslant \delta^{-1}$ then $\mathcal{F}_i$ and $\mathcal{F}_j$ are disjoint, by property (iv). $\qquad\square$

4.2. **Constructing the frame.** In this section we will construct the δ-generalized tree-frame $(\mathcal{F}_1, \ldots, \mathcal{F}_k)$, along with the rooted tree T, and a collection $(\mathcal{G}_1, \ldots, \mathcal{G}_k)$ of 'garbage' sets. Let $C > 0$ and $\varepsilon > 0$ be arbitrary, as in the statement of Theorem 2.3, and set

$$\lambda := \frac{\varepsilon}{2^4 C} \qquad \text{and} \qquad \delta := 2^{-23} \lambda^4 \varepsilon^{2 \log_2(1/\lambda\varepsilon)+15}. \tag{14}$$

Recall that the sets $S_1, \ldots, S_k$ were fixed earlier, and let us fix, for the rest of this section, a minimal cover $\mathcal{A}$ of $S_{[k]}$ with hyperplanes such that $F(\mathcal{A}) = [k]$. By Lemma 3.3, there exists an $(\lambda, \varepsilon/2, \delta)$-exploration tree of $\mathcal{A}$; let us also fix such a tree $\mathbb{T}$.

The first step is to observe that every $i \in [k]$ occurs as the label i_u of some vertex $u \in V(\mathbb{T})$. This is an immediate consequence of the following simple observation about index trees.

Observation 4.3. *Let $\mathbb{T}$ be an index tree, let $u \in V(\mathbb{T})$, and let $j \in I_u$. Then there exists $v \in V(\mathbb{T})$, with $u \prec_{\mathbb{T}} v$, such that $i_v = j$.*

Proof. This follows easily from (5): if $j \in I_v$ and $i_v \neq j$, then $j \in I_w$ for some $w \in N(u)$, and if $j \in I_v$ and $v \in \partial(\mathbb{T})$ then $I_v = \{j\}$, so $i_v = j$. $\qquad\square$

To extract our δ-generalized tree-frame from the exploration tree $\mathbb{T}$, we will also need the notion of a *depth-first search ordering* $\prec$ on the vertices of a rooted tree T. This is defined by placing an arbitrary linear order on the out-neighbours of each vertex of T, and then setting $u \prec v$ if either $u \prec_T v$, or if the branch leading to u precedes the branch leading to v in the ordering of the neighbours of the last common ancestor of u and v.

Definition 4.4. Let $\prec$ be a depth-first search ordering on the vertices of $\mathbb{T}$. We define a rooted tree T and a δ-generalized tree-frame centred at T as follows:

1. For each $i \in [k]$, define $\beta(i)$ to be the $\prec$-minimal vertex u of $\mathbb{T}$ such that $i_u = i$.
2. Define T to be the union of the paths in $\mathbb{T}$ from $\beta(1), \ldots, \beta(k)$ to the root.
3. For each $u \in V(T)$, define $\alpha(u) := i_u$.
4. For each edge $uv \in E(T)$, where $u \in N(v)$, define $\gamma(uv) := s_u \in S_{i_v}$.
5. (a) Set $\mathcal{F}_i := \mathcal{F}_{\beta(i)}$ and $\mathcal{G}_i := \emptyset$ for each $i \in [k]$ such that $\beta(i)$ is a good vertex of $\mathbb{T}$.
   (b) Set $\mathcal{G}_i := \mathcal{G}_{\beta(i)}$ and $\mathcal{F}_i := \emptyset$ for each $i \in [k]$ such that $\beta(i)$ is a bad vertex of $\mathbb{T}$.

Properties (a)–(d) of Definition 4.1 follow easily from this construction. Indeed, we have $\alpha(\beta(i)) = i$ for each $i \in [k]$ by our choice of α and β, and $\gamma(uv) \in S_{i_v} = S_{\alpha(v)}$ for all $uv \in E(T)$ with $u \in N(v)$. To see that $\alpha(u) \neq \alpha(v)$ if $u \prec_T v$ and $u \neq v$, recall $\mathbb{T}$ is an index tree, and therefore satisfies (5), so $\alpha(u) = i_u$ is not included in any of the sets associated with the descendants of u. For (d), let π be the permutation of $[k]$ given by the ordering $\prec$ restricted to $\{\beta(1), \ldots, \beta(k)\}$, and observe that if $u \prec_T \beta(\pi(i))$ then $u \prec \beta(\pi(i))$, and hence $\beta(\alpha(u)) \prec \beta(\pi(i))$, by our choice of β. It follows that $\alpha(u) \in \{\pi(1), \ldots, \pi(i)\}$, and therefore that $J(\pi(i)) = \{\alpha(u) : u \prec_T \beta(\pi(i))\} \subseteq \{\pi(1), \ldots, \pi(i)\}$ for each $i \in [k]$, as required.

14

The following lemma shows that properties (i), (ii) and (iii) of Definition 4.1 also hold.

Lemma 4.5. *Let $i \in [k]$ and $H \in \mathcal{F}_i$, let $j \in J(i) \setminus \{i\}$, and let $e_j \in E(T)$ be the edge leaving the unique vertex $v \prec_T \beta(i)$ with $\alpha(v) = j$ in the direction of the vertex $\beta(i)$. Then*

$$i \in F(H), \qquad \mu_{I(i)}(H) > \delta \qquad and \qquad H_j \in \{\gamma(e_j), S_j\}.$$

Proof. Note that $\mathcal{F}_i \neq \emptyset$ implies that the vertex $\beta(i)$ is good. By Definition 3.2(b), it follows that $i \in F(H)$, and also that $\mu_{I_{\beta(i)} \setminus \{i\}}(H) > \delta$. Moreover $\mathcal{F}_i = \mathcal{F}_{\beta(i)} \subseteq \mathcal{A}_{\beta(i)}$ and $F(\mathcal{A}_{\beta(i)}) \subseteq J(i) \cup I_{\beta(i)}$, by Definition 3.2, and therefore $H_j = S_j$ for every $j \in I(i) \setminus I_{\beta(i)}$. Since $i \in J(i)$, and therefore $i \notin I(i)$, it follows that $\mu_{I(i)}(H) > \delta$.

Now suppose that $e_j = uv$, with $u \in N(v)$, and observe that, by Definition 3.2(a), we have $H'_j \in \{s_u, S_j\} = \{\gamma(e_j), S_j\}$ for every hyperplane $H' \in \mathcal{A}_u$. Noting that $\mathcal{F}_i \subseteq \mathcal{A}_{\beta(i)} \subseteq \mathcal{A}_u$ (by Definition 3.2, and since $u \prec_T \beta(i)$), it follows that $H_j \in \{\gamma(e_j), S_j\}$, as claimed. □

It therefore only remains to show that property (iv) of Definition 4.1 and the inequality (4) hold. Both of these properties will follow from our choice of $\beta(1), \dots, \beta(k)$.

4.3. **Lemmas on disjointness.** In this section we will prove two straightforward but crucial lemmas; the first verifies condition (iv) of the definition of a δ-generalized tree-frame.

Lemma 4.6. *If $\min\{|S_i|, |S_j|\} \geqslant \delta^{-1}$ and $i \neq j$, then $\mathcal{F}_i$ and $\mathcal{F}_j$ are disjoint.*

Proof. Suppose that $H \in \mathcal{F}_{\beta(i)} \cap \mathcal{F}_{\beta(j)}$, and suppose that $\beta(i) \prec \beta(j)$ in the depth-first search ordering. Recall that $i, j \in F(H)$, by Definition 3.2(b), and $\mu_{I(i)}(H) > \delta$, by Lemma 4.5. Since $|S_j| \geqslant \delta^{-1}$, it follows that $j \notin I(i)$, and hence $j \in J(i)$, i.e., there exists $u \prec_T \beta(i)$ with $\alpha(u) = j$. However, this is a contradiction, since $u \prec \beta(i) \prec \beta(j)$ in the depth-first search ordering, and $\beta(j)$ was chosen to be the $\prec$-minimal vertex u of $\mathbb{T}$ such that $\alpha(u) = j$. □

The final lemma we need shows that each garbage set only appears on a single path through T. Since the number of fixed coordinates of H_{I_u} decreases along the path (and decreases strictly whenever $\alpha(u) \in F(H)$), this will imply that each hyperplane contributes only $O(1)$ to the sum of the left-hand side of (7) over vertices $u \in \{\beta(1), \dots, \beta(k)\}$.

Lemma 4.7. *Let $H \in \mathcal{G}_i \cap \mathcal{G}_j$. Then either $\beta(i) \prec_{\mathbb{T}} \beta(j)$, or $\beta(j) \prec_{\mathbb{T}} \beta(i)$.*

Proof. Suppose (without loss of generality) that $\beta(i) \prec \beta(j)$ in the depth-first search ordering, and suppose that $\beta(i) \nprec_{\mathbb{T}} \beta(j)$, which implies that $u \prec \beta(j)$ for every $u \in V(\mathbb{T})$ with $\beta(i) \prec_{\mathbb{T}} u$. Note that $j \in F(H) \subseteq I_{\beta(i)} \cup J(i)$, since $H \in \mathcal{G}_i \cap \mathcal{G}_j$ and by Definition 3.2.

If $j \in J(i)$, then $i_v = j$ for some $v \in V(\mathbb{T})$ with $v \prec_{\mathbb{T}} \beta(i)$, and hence $v \prec \beta(i) \prec \beta(j)$. On the other hand, if $j \in I_{\beta(i)}$, then by Observation 4.3 we have $i_v = j$ for some $v \in V(\mathbb{T})$ with $\beta(i) \prec_{\mathbb{T}} v$, and hence (by the observation above) $v \prec \beta(j)$. In either case, this contradicts our choice of $\beta(j)$ as the $\prec$-minimal vertex v of $\mathbb{T}$ such that $i_v = j$. □

Let us record here the following simple consequence of Lemma 4.7.

Lemma 4.8. *If $H \in \mathcal{G}_i \cap \mathcal{G}_j$ and $i \neq j$, then $|F(H) \cap I_{\beta(i)}| \neq |F(H) \cap I_{\beta(j)}|$.*

Proof. By Lemma 4.7, we have (without loss of generality) $\beta(i) \prec_{\mathbb{T}} \beta(j)$, which implies, by (5) and since $\beta(i) \neq \beta(j)$, that $I_{\beta(j)} \subseteq I_{\beta(i)} \setminus \{i\}$. Since $i \in F(H) \cap I_{\beta(i)}$, it follows that $F(H) \cap I_{\beta(j)} \subsetneq F(H) \cap I_{\beta(i)}$, as required. □

15

4.4. The proof of Theorem 2.3. We are now ready to prove our main structural result, Theorem 2.3. It only remains to show that the inequality (4) follows from (3). We will use the following easy consequence of Lemma 4.8. Set $B := \{i \in [k] : \beta(i) \text{ is bad}\}$.

Lemma 4.9.
$$\left| \bigcup_{i \in B} \mathcal{G}_i \right| \geqslant \frac{1}{5\lambda} \sum_{i \in B} |S_i|.$$

Proof. Summing (7) over $i \in B$, we obtain
$$\sum_{i \in B} \sum_{H \in \mathcal{G}_i} 2^{-|F(H) \cap I_{\beta(i)}|/4} \geqslant \frac{1}{\lambda} \sum_{i \in B} |S_i|.$$

Now, by Lemma 4.8, for each H and $\ell \geqslant 2$ there is at most one value of $i \in B$ such that $H \in \mathcal{G}_i$ and $|F(H) \cap I_{\beta(i)}| = \ell$, so for each $H \in \bigcup_{i \in B} \mathcal{G}_i$ we have
$$\sum_{i \in B \,:\, H \in \mathcal{G}_i} 2^{-|F(H) \cap I_{\beta(i)}|/4} \leqslant \sum_{\ell=2}^{\infty} 2^{-\ell/4} = \left(2^{1/2} - 2^{1/4}\right)^{-1} < 5,$$

as required. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\quad\square$

Theorem 2.3 now follows easily from the lemmas above.

Proof of Theorem 2.3. We claim that the sequence $(\mathcal{F}_1, \ldots, \mathcal{F}_k)$ constructed in Definition 4.4 is a δ-generalized tree-frame centred at T, and satisfies (4). By Lemma 4.2, it will follow that $(\mathcal{F}_1, \ldots, \mathcal{F}_k)$ is also a δ-generalized frame, so this will be sufficient to prove the theorem. Note that properties (a)–(d) and (i)–(iv) of Definition 4.1 follow from the comments after Definition 4.4, and by Lemmas 4.5 and 4.6. Moreover, by discarding excess hyperplanes if necessary, we may assume that $|\mathcal{F}_i| \leqslant |S_i| - 1$ for each $i \in [k]$. It therefore only remains to show that (4) holds.

To do so, recall that $\mathbb{T}$ is an $(\lambda, \varepsilon/2, \delta)$-exploration tree of $\mathcal{A}$, and hence
$$|\mathcal{F}_i| > \left(1 - \varepsilon/2\right)\left(|S_i| - 1\right)$$

for each i such that $\beta(i)$ is a good vertex, i.e., for each $i \in [k] \setminus B$. Now, by Lemma 4.9 and the condition (3), we have
$$\frac{1}{5\lambda} \sum_{i \in B} |S_i| \leqslant |\mathcal{A}| \leqslant C \sum_{i=1}^{k} \left(|S_i| - 1\right),$$

and hence, recalling from (14) that $\lambda = \varepsilon/(2^4 C)$, we obtain
$$\sum_{i=1}^{k} |\mathcal{F}_i| \geqslant \sum_{i \in [k] \setminus B} \left(1 - \varepsilon/2\right)\left(|S_i| - 1\right) \geqslant \left(1 - \varepsilon/2\right) \sum_{i=1}^{k} \left(|S_i| - 1\right) - \sum_{i \in B} |S_i|$$
$$\geqslant \left(1 - \varepsilon/2 - 5C\lambda\right) \sum_{i=1}^{k} \left(|S_i| - 1\right) \geqslant \left(1 - \varepsilon\right) \sum_{i=1}^{k} \left(|S_i| - 1\right),$$

as required. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\quad\square$

5. Arithmetic frames and the proof of the lower bound

In order to deduce Theorem 1.1 from Theorem 2.3, we will need to bound the number of δ-generalized frames in the integers. In this section we will warm up for the calculation ahead by counting a simpler set of objects, which we call 'arithmetic frames', and thereby deducing a lower bound on the number of minimal covering systems of $\mathbb{Z}$. Recall that[4]

$$\tau = \sum_{t=1}^{\infty} \left(\log \frac{t+1}{t} \right)^2 \approx 0.977.$$

The following proposition provides the lower bound in Theorem 1.1.

Proposition 5.1. *The number of minimal covering systems of $\mathbb{Z}$ of size n is at least*

$$\exp \left(\left(\frac{4\sqrt{\tau}}{3} + o(1) \right) \frac{n^{3/2}}{(\log n)^{1/2}} \right)$$

as $n \to \infty$.

We shall first prove Proposition 5.1 for an infinite sequence of values of n (see (18)); since this sequence will be sufficiently dense, it will then be easy to deduce the bound for the remaining values of n. For each n in our sequence, we will choose a single value of N, and count only covering systems $\mathcal{A}$ of size n with $\mathrm{lcm}(\mathcal{A}) = N$. We will moreover count only covering systems that correspond to simple frames of a certain family of sets (see below), with a specific (carefully chosen) order, see Definition 5.4. We remark that when N is not square-free, this is not quite as straightforward as counting the simple frames, since there will exist hyperplanes that do not correspond to arithmetic progressions in $\mathbb{Z}_N$. In order to characterise the hyperplanes that do, we need to introduce a little notation.

Given $N = p_1^{\gamma_1} \cdots p_m^{\gamma_m} > 1$, we define

$$\langle N \rangle := \bigcup_{i=1}^{m} \big\{ (p_i, j) : j \in [\gamma_i] \big\}, \tag{15}$$

and set $S_{(p,e)} = \{0, \ldots, p-1\}$ for each $(p, e) \in \langle N \rangle$. Now define a map

$$\varphi_N \colon \mathbb{Z}_N \to S_{\langle N \rangle} = \prod_{(p,e) \in \langle N \rangle} S_{(p,e)}$$

as follows: if $x \in \mathbb{Z}_N$, then $y = \varphi_N(x) \in S_{\langle N \rangle}$ is the vector such that, for each $(p, e) \in \langle N \rangle$, $y_{(p,e)} \in S_{(p,e)}$ is the coefficient of p^{e-1} in the p-ary expansion of x modulo p^e. Observe that φ_N is a bijection, by the Chinese Remainder Theorem.

We say that a hyperplane H in $S_{\langle N \rangle}$ is *arithmetic* if $\varphi_N^{-1}(H)$ is an arithmetic progression in $\mathbb{Z}_N$. The following observation provides a simple characterization of arithmetic hyperplanes.

Observation 5.2. *A hyperplane H in $S_{\langle N \rangle}$ is arithmetic if and only if, for each prime p, the set*

$$\big\{ (p, e) \in \langle N \rangle : (p, e) \in F(H) \big\}$$

forms a (possibly empty) initial segment of the sequence $(p, 1), (p, 2), (p, 3), \ldots$

---

[4]We remark that the constant τ also appears in the study of the iterated divisor function, see [3, 4].

Proof. Suppose first that H is arithmetic, so $A := \varphi_N^{-1}(H)$ is an arithmetic progression in $\mathbb{Z}_N$. Let d be the modulus of A, and observe that $(p, e) \in F(H)$ if and only if p^e divides d, by the definition of φ_N. On the other hand, if $\{(p, 1), \ldots, (p, e(p))\} \subset F(H)$ and $(p, e') \notin F(H)$ for all $e' > e(p)$, then every pair of points of $\varphi_N^{-1}(H)$ differs by a multiple of $p^{e(p)}$, and therefore $\varphi_N^{-1}(H)$ is contained in an arithmetic progression with modulus $d = \prod_p p^{e(p)}$. Since $|H| = N/d$, it follows that $\varphi_N^{-1}(H)$ is in fact the entire arithmetic progression, as claimed. $\square$

Let us now say that a total ordering $\prec$ on the elements of $\langle N \rangle$ is *arithmetic* if

$$(p_i, 1) \prec (p_i, 2) \prec \cdots \prec (p_i, \gamma_i) \tag{16}$$

for all $i \in [m]$. Note that (16) does not impose any constraint on $\prec$ for different primes, and in particular we may have $(p, i) \prec (q, j) \prec (p, i+1)$. We say that a simple frame of $S_{\langle N \rangle}$ is 'arithmetic' if the order of the sets is arithmetic, and if moreover each of the hyperplanes of the frame is arithmetic. We can now prove the following lower bound on the number of minimal covering systems of $\mathbb{Z}$ of size n.

Lemma 5.3. *Let $N = p_1^{\gamma_1} \cdots p_m^{\gamma_m} > 1$, and let $\prec$ be an arithmetic ordering of $\langle N \rangle$. There are at least*

$$\exp\left(\sum_{(p,e) \in \langle N \rangle} (p-1) \sum_{\substack{(q,f) \prec (p,e) \\ q \neq p}} \log\left(\frac{f+1}{f}\right) \right)$$

minimal covering systems of $\mathbb{Z}$ of size $n := \sum_{i=1}^m \gamma_i(p_i - 1) + 1$.

Proof. To prove the lemma we count arithmetic frames of $S_{\langle N \rangle}$ centred at $(0, \ldots, 0)$, where the sets $S_{(p,e)}$ are listed in the order $\prec$. Recall from Definition 2.1 that, for each $(p, e) \in \langle N \rangle$ and each $a \in \{1, \ldots, p-1\}$, we need to choose an arithmetic hyperplane of the form

$$[x_1, \ldots, x_{i-1}, a, *, \cdots, *],$$

with $x_j \in \{0, *\}$ for each $j \in [i-1]$, where (p, e) is the ith element in the ordering $\prec$. To do so, we will choose, for each prime $q \neq p$, an initial segment (in the order $\prec$) of the set

$$J_q(p, e) := \{(q, f) \in \langle N \rangle : (q, f) \prec (p, e)\}$$

set $x_j = 0$ for the corresponding coordinates, and set $x_j = *$ for all other elements of $J_q(p, e)$. If we also set $x_j = 0$ for every $j \in J_p(p, e)$ then, by Observation 5.2, every hyperplane obtained in this way will be arithmetic, and therefore the frame we construct will be arithmetic. We claim that each such choice gives a different minimal covering system of $\mathbb{Z}$ of size n.

To see this, note that the frame consists of $n - 1$ arithmetic hyperplanes, each of which corresponds (via the bijection φ_N^{-1}) to an arithmetic progression in $\mathbb{Z}_N$. Moreover, the only element of $\mathbb{Z}_N$ not covered by these arithmetic progressions is 0, so adding this progression gives a covering system of $\mathbb{Z}$ of size n, and (as observed after Definition 2.1) if we remove the hyperplane $H = [x_1, \ldots, x_{i-1}, a, *, \cdots, *]$, then the element $(0, \ldots, 0, a, 0, \ldots, 0)$ will be uncovered by the remaining hyperplanes, so the covering system we have constructed is minimal. Finally, each hyperplane in the frame has a unique entry $a \notin \{0, *\}$, and therefore each choice leads to a distinct covering system, as claimed.

Finally, since we have exactly

$$\prod_{q \neq p} \left(|J_q(p, e)| + 1 \right) = \exp \left(\sum_{\substack{(q,f) \prec (p,e) \\ q \neq p}} \log \left(\frac{f+1}{f} \right) \right)$$

choices for each hyperplane corresponding to (p, e), the lemma follows. $\qquad\square$

Now, for each arithmetic ordering $\prec$ of $\langle N \rangle$, let us define

$$Q(N, \prec) := \sum_{(p,e) \in \langle N \rangle} (p - 1) \sum_{\substack{(q,f) \prec (p,e) \\ q \neq p}} \log \left(\frac{f+1}{f} \right). \qquad (17)$$

We will use the following particular arithmetic ordering $<$ to prove Proposition 5.1.

Definition 5.4. For each prime p and integer $e \in \mathbb{N}$, set $y_{p,e} := (p - 1)\left(\log \frac{e+1}{e} \right)^{-1}$. Now, given primes p and q, and integers $e, f \in \mathbb{N}$, define

$$(q, f) < (p, e) \qquad \Leftrightarrow \qquad y_{q,f} < y_{p,e}.$$

Moreover, if $x \in \mathbb{R}$ then we write $(p, e) < x$ if and only if $y_{p,e} < x$, and define

$$n(x) := 1 + \sum_{(p,e)<x} (p - 1) \qquad \text{and} \qquad N(x) := \prod_{(p,e)<x} p. \qquad (18)$$

Note that $n(x), N(x) < \infty$ for every $x \in \mathbb{R}$, and that for any $N \in \mathbb{N}$, the ordering $<$ on $\langle N \rangle$ is arithmetic. Our next lemma, combined with Lemma 5.3, implies Proposition 5.1.

Lemma 5.5. *Let $x > 0$, and set $N = N(x)$ and $n = n(x)$. Then*

$$Q(N, <) = \left(\frac{4\sqrt{\tau}}{3} + o(1) \right) \frac{n^{3/2}}{(\log n)^{1/2}}$$

as $x \to \infty$.

Proof. Recalling the definition of $Q(N, <)$, observe first that, for each $(p, e) \in \langle N \rangle$,

$$\sum_{\substack{(q,f)<(p,e) \\ q \neq p}} \log \left(\frac{f+1}{f} \right) = \sum_{f \geqslant 1} \left| \left\{ q \neq p : q - 1 < y_{p,e} \log \left(\frac{f+1}{f} \right) \right\} \right| \cdot \log \left(\frac{f+1}{f} \right). \qquad (19)$$

Now, by the prime number theorem, for each fixed $f \in \mathbb{N}$ and as $y_{p,e} \to \infty$,

$$\left| \left\{ q \neq p : q - 1 < y_{p,e} \log \left(\frac{f+1}{f} \right) \right\} \right| = (1 + o(1)) \frac{y_{p,e} \log \left(\frac{f+1}{f} \right)}{\log \left(y_{p,e} \log \left(\frac{f+1}{f} \right) \right)}.$$

Moreover, the sum in (19) of the terms with $f \geqslant f_0$ is $o\left(y_{p,e} / \log y_{p,e} \right)$ as $f_0 \to \infty$, so

$$\sum_{\substack{(q,f)<(p,e) \\ q \neq p}} \log \left(\frac{f+1}{f} \right) = \frac{(1 + o(1)) y_{p,e}}{\log y_{p,e}} \sum_{f \geqslant 1} \left(\log \frac{f+1}{f} \right)^2 = (\tau + o(1)) \frac{y_{p,e}}{\log y_{p,e}}$$

19

as $y_{p,e} \to \infty$. We next fix $e \in \mathbb{N}$, and sum over primes p. We obtain

$$\sum_{p \,:\, y_{p,e} < x} (p-1) \sum_{\substack{(q,f) < (p,e) \\ q \neq p}} \log\left(\frac{f+1}{f}\right) = \sum_{p-1 < x \log \frac{e+1}{e}} (\tau + o(1)) \frac{(p-1)^2}{\log\left(\frac{e+1}{e}\right) \cdot \log y_{p,e}}$$

$$= (\tau + o(1)) \frac{x^3 \left(\log \frac{e+1}{e}\right)^2}{3(\log x)^2} \tag{20}$$

as $x \to \infty$, again using the prime number theorem.[5] Thus, summing over e, and noting that the left-hand side of (20) is uniformly bounded from above by an absolute constant times the right-hand side (without the $o(1)$ term), we obtain

$$Q(N, <) = (\tau + o(1)) \frac{x^3}{3(\log x)^2} \sum_{e \geqslant 1} \left(\log \frac{e+1}{e}\right)^2 = (\tau^2 + o(1)) \frac{x^3}{3(\log x)^2}$$

as $x \to \infty$. Finally, using the prime number theorem a third time, we obtain

$$n(x) = 1 + \sum_{e \geqslant 1} \sum_{p-1 < x \log \frac{e+1}{e}} (p-1) = (1 + o(1)) \sum_{e \geqslant 1} \frac{\left(x \log \frac{e+1}{e}\right)^2}{2 \log\left(x \log \frac{e+1}{e}\right)}$$

$$= (1 + o(1)) \frac{x^2}{2 \log x} \sum_{e \geqslant 1} \left(\log \frac{e+1}{e}\right)^2 = (\tau + o(1)) \frac{x^2}{2 \log x}, \tag{21}$$

and hence $Q(N, <) \cdot n^{-3/2} \sqrt{\log n} \to \sqrt{2}(\tau^2/3)(\tau/2)^{-3/2} = 4\sqrt{\tau}/3$ as $x \to \infty$, as claimed. □

We can now easily deduce the lower bound in Theorem 1.1, the only remaining difficulty being to deal with those $n \in \mathbb{N}$ that are not of the form $n = n(x)$ for some $x \in \mathbb{R}$.

Proof of Proposition 5.1. It follows immediately from Lemmas 5.3 and 5.5 that the number of minimal covering systems of $\mathbb{Z}$ of size $n(x)$ is at least

$$e^{Q(N(x), <)} = \exp\left(\left(\frac{4\sqrt{\tau}}{3} + o(1)\right) \frac{n(x)^{3/2}}{(\log n(x))^{1/2}}\right)$$

as $x \to \infty$. Let $x > 0$ be maximal such that $n(x) \leqslant n$, and set $t := n - n(x)$. Observe that $t < x = o(n)$, by (21), and that, by removing the hyperplane $[0, \ldots, 0]$ (i.e., the progression $\{0 \pmod{N}\}$) from the construction given in the proof of Lemma 5.3, we obtain a family of minimal covers of $\mathbb{Z}_N \setminus \{0\}$ of size $n(x) - 1$. We complete each to a minimal cover of $\mathbb{Z}$ of size n by adding the progressions $\{2^{\ell-1} N \pmod{2^\ell N}\}$, for each $\ell \in [t]$, and $\{0 \pmod{2^t N}\}$. We obtain a family of

$$\exp\left(\left(\frac{4\sqrt{\tau}}{3} + o(1)\right) \frac{n^{3/2}}{(\log n)^{1/2}}\right)$$

minimal covering systems of $\mathbb{Z}$ of size $n(x) + t = n$, as required. □

---

[5]Indeed, the prime number theorem implies that $\sum_{p < z} p^2 = (1/3 + o(1)) z^3 / \log z$ as $z \to \infty$.

6. Counting coverings that are far from frames

In this section we will begin the deduction of Theorem 1.1 from Theorem 2.3 by bounding the number of minimal covers that fail to satisfy (3). In the process, we will obtain a short proof of weaker version of Theorem 1.1, bounding the number of minimal covering systems of $\mathbb{Z}$ of size n up to a constant factor in the exponent.

Proposition 6.1. *Let $C > 0$ be a constant, and let $n \in \mathbb{N}$ and $N = p_1^{\gamma_1} \cdots p_m^{\gamma_m}$ satisfy*

$$ n > C \sum_{i=1}^{m} \gamma_i (p_i - 1). $$

Then the number of minimal covering systems $\mathcal{A}$ of $\mathbb{Z}$ of size n with $\mathrm{lcm}(\mathcal{A}) = N$ is at most

$$ \exp\left(\left(\frac{2\sqrt{\tau}}{\sqrt{C}} + o(1) \right) \frac{n^{3/2}}{(\log n)^{1/2}} \right) $$

as $n \to \infty$.

In order to bound the number of covering systems, we will need to bound the number of choices for the modulus d and shift a of each arithmetic progression in $\mathcal{A}$. The following simple but important lemma, which we will use again later, shows that, given the moduli, we have relatively few choices for the shifts.

Lemma 6.2. *Let $d_1, \ldots, d_n \in \mathbb{N}$. There are at most $(n!)^2$ minimal covering systems $\mathcal{A} = \{A_1, \ldots, A_n\}$ of $\mathbb{Z}$ of size n such that, for each $i \in [n]$, the modulus of A_i is d_i.*

Proof. Let $\mathcal{A} = \{A_1, \ldots, A_n\}$ be a minimal covering system of $\mathbb{Z}$, and observe that we may reorder the elements of $\mathcal{A}$ so that, for each $i \in [n]$, the arithmetic progression A_i covers at least a $1/i$ proportion of the set

$$ R_i := \mathbb{Z} \setminus \bigcup_{j > i} A_j. $$

Indeed, to see that this is possible we simply choose the sets one by one (in reverse order), letting A_i be the (remaining) progression in $\mathcal{A}$ whose intersection with R_i has largest density, observing that R_i is non-empty (since $\mathcal{A}$ is minimal) and recalling that $\mathcal{A}$ covers $\mathbb{Z}$. The total number of choices for $\mathcal{A}$ is therefore at most the sum over permutations of $(d_1, \ldots, d_n)$ of the number of sequences $(A_1, \ldots, A_n)$ with this additional property.

Now let $i \in [n]$, and suppose that we have already chosen progressions $(A_{i+1}, \ldots, A_n)$. We claim that we have at most i choices for the arithmetic progression A_i. Indeed, since the progressions $\{a \pmod{d_i}\}$ (for $a \in \{0, \ldots, d_i - 1\}$) are disjoint, there are at most i progressions with modulus d_i that cover at least a $1/i$ proportion of R_i. It follows that the number of choices for $\mathcal{A}$ is at most $(n!)^2$, as claimed. $\qquad\square$

It therefore only remains to bound the number of choices for the moduli. Note that if $\mathrm{lcm}(\mathcal{A}) = p_1^{\gamma_1} \cdots p_m^{\gamma_m}$ then we have at most $\prod_i (\gamma_i + 1)$ choices for each modulus. The following lemma, which we will use again later, provides a sharp bound on this product.

Lemma 6.3. *Let $(p_1, \ldots, p_m)$ be a sequence of distinct primes, let $(\gamma_1, \ldots, \gamma_m)$ be a sequence of positive integers, and let $M \geqslant \sum_{i=1}^{m} \gamma_i(p_i - 1)$. Then*

$$\sum_{i=1}^{m} \log\left(\gamma_i + 1\right) \leqslant \left(2\sqrt{\tau} + o(1)\right) \left(\frac{M}{\log M}\right)^{1/2} \tag{22}$$

as $M \to \infty$.

Proof. We may assume that $p_1 < \cdots < p_m$, and reorder the γ_i so that $\gamma_1 \geqslant \cdots \geqslant \gamma_m$, noting that this does not change the left-hand side of (22), and that the inequality $M \geqslant \sum_{i=1}^{m} \gamma_i(p_i - 1)$ still holds under the new ordering. Set $x_t := \max\left\{i : \gamma_i \geqslant t\right\}$, and observe that maximizing the left-hand side of (22) is equivalent to maximizing

$$X := \sum_{t \geqslant 1} x_t \cdot \log\left(\frac{t+1}{t}\right)$$

subject to the constraint

$$M \geqslant \sum_{i \geqslant 1} \gamma_i(p_i - 1) = \sum_{t \geqslant 1} \sum_{i=1}^{x_t} (p_i - 1) \geqslant \sum_{t \geqslant 1} \frac{x_t^2}{2} \max\left\{\log x_t - 3, \, 1\right\}, \tag{23}$$

where in the final step we used the following bound of Massias and Robin [18],

$$\sum_{i=1}^{x} p_i \geqslant \frac{x^2}{2}\left(\log x + \log\log x - \frac{3}{2} - \frac{3.568}{\log x}\right) \geqslant \frac{x^2}{2}\left(\log x - 2\right),$$

which holds for every $x \geqslant e^3$. Note that X is increasing in x_t for each $t \geqslant 1$, and so (by allowing $0 \leqslant x_t \in \mathbb{R}$) we may assume that M is equal to the right-hand side of (23).

Applying the method of Lagrange multipliers, it follows that there exists $\lambda \in \mathbb{R}$ such that, for each $t \geqslant 1$, either

$$\lambda \log \frac{t+1}{t} = x_t\left(\log x_t - \frac{5}{2}\right), \tag{24}$$

or $x_t \leqslant e^4$. We will first show that the contribution to X of those values of t such that $x_t = O(1)$ is small. To do so, note that $x_t = 0$ for all $t > M$, by (23), and observe that therefore

$$\sum_{t \geqslant 1} x_t \cdot \log\left(\frac{t+1}{t}\right) \mathbb{1}\left[x_t \leqslant \log M\right] \leqslant \left(\log(M+1)\right)^2. \tag{25}$$

We may therefore restrict our attention to those values of t for which $x_t > \log M > e^4$, so that, in particular, (24) holds. Let $T = \max\left\{t : x_t > \log M\right\}$ and observe that, by (23), (24) and (25), we have

$$X \leqslant \lambda \sum_{t=1}^{T} \frac{\left(\log \frac{t+1}{t}\right)^2}{\log x_t - 5/2} + O\left(\log M\right)^2 \qquad \text{and} \qquad M \geqslant \frac{\lambda^2}{2} \sum_{t=1}^{T} \frac{\left(\log \frac{t+1}{t}\right)^2}{\log x_t}.$$

To bound these sums, observe that $\lambda \to \infty$ as $M \to \infty$ (by (24) and since $x_1 > \log M$), and that therefore, uniformly in $1 \leqslant t \leqslant \log \lambda$, we have

$$\log x_t = \log \lambda + \log\log \frac{t+1}{t} - \log\left(\log x_t - 5/2\right) = \left(1 + o(1)\right)\log \lambda$$

22

as $M \to \infty$. Moreover, if $\log \lambda \leqslant t \leqslant T$, then $\log \frac{t+1}{t} \leqslant 1/t$ and $\log x_t \geqslant \log \log M$. It follows that, for each fixed $c \in \mathbb{R}$, we have

$$\sum_{t=1}^{T} \frac{\left(\log \frac{t+1}{t}\right)^2}{\log x_t - c} = \frac{1 + o(1)}{\log \lambda} \sum_{t \geqslant 1} \left(\log \frac{t+1}{t}\right)^2 = \frac{\tau + o(1)}{\log \lambda},$$

as $M \to \infty$, and hence $M \geqslant (\tau/2 + o(1)) \lambda^2 / \log \lambda$. Finally, we deduce that

$$X \leqslant (\tau + o(1)) \frac{\lambda}{\log \lambda} \leqslant (2\sqrt{\tau} + o(1)) \left(\frac{M}{\log M}\right)^{1/2}$$

as $M \to \infty$, as required. $\qquad\square$

We can now easily deduce Proposition 6.1.

Proof of Proposition 6.1. We first choose the moduli of the progressions in $\mathcal{A} = \{A_1, \ldots, A_n\}$, and then the shifts. Since $\mathrm{lcm}(\mathcal{A}) = N = p_1^{\gamma_1} \cdots p_m^{\gamma_m}$, for each $j \in [n]$ we have at most

$$\prod_{i=1}^{m} (\gamma_i + 1) \leqslant \exp\left((2\sqrt{\tau} + o(1)) \left(\frac{n}{C \log(n/C)}\right)^{1/2}\right)$$

choices for the modulus of the arithmetic progression A_j, where the inequality follows by applying Lemma 6.3 with $M = n/C$, and using our bound on n. By Lemma 6.2, it follows that the number of choices for $\mathcal{A}$ is at most

$$(n!)^2 \cdot \exp\left(\left(\frac{2\sqrt{\tau}}{\sqrt{C}} + o(1)\right) \frac{n^{3/2}}{\left(\log(n/C)\right)^{1/2}}\right) = \exp\left(\left(\frac{2\sqrt{\tau}}{\sqrt{C}} + o(1)\right) \frac{n^{3/2}}{(\log n)^{1/2}}\right)$$

as $n \to \infty$, as required. $\qquad\square$

Using Simpson's theorem (Theorem 2.4), we can now easily deduce an upper bound on the number of minimal covering systems that is sharp up to a constant factor in the exponent.

Corollary 6.4. *The number of minimal covering systems of $\mathbb{Z}$ of size n is*

$$\exp\left(\frac{\Theta\left(n^{3/2}\right)}{(\log n)^{1/2}}\right).$$

Proof. The lower bound follows by Proposition 5.1 (or by the simpler construction in the introduction). For the upper bound, recall that, by Simpson's theorem, we have

$$|\mathcal{A}| \geqslant \sum_{i=1}^{m} \gamma_i (p_i - 1) + 1 \geqslant \sum_{i=1}^{m} \gamma_i \log_2 p_i$$

for any minimal covering system $\mathcal{A}$ of $\mathbb{Z}$ with $\mathrm{lcm}(\mathcal{A}) = N = p_1^{\gamma_1} \cdots p_m^{\gamma_m}$, and hence $N \leqslant 2^n$. Thus, applying Proposition 6.1 with $C = 1$ (and summing over $N \leqslant 2^n$), there are at most

$$\exp\left((2\sqrt{\tau} + o(1)) \frac{n^{3/2}}{(\log n)^{1/2}}\right)$$

minimal covering systems of $\mathbb{Z}$ of size n, as required. $\qquad\square$

7. Proof of Theorem 1.1

In this section we will complete the proof of Theorem 1.1; we begin by giving an overview of the remaining part of the argument. Let $\mathcal{A}$ be a minimal covering system of $\mathbb{Z}$ of size n, let $N = \text{lcm}(\mathcal{A})$ and, recalling (15), set $S_{(p,e)} = \{0, \ldots, p-1\}$ for each $(p,e) \in \langle N \rangle$. We map $\mathbb{Z}_N$ into $S_{\langle N \rangle} = \prod_{(p,e) \in \langle N \rangle} S_{(p,e)}$ as described in Section 5; that is, we associate $x \in \mathbb{Z}_N$ with the vector $y = \varphi_N(x) \in S_{\langle N \rangle}$, where $y_{(p,e)}$ is the coefficient of p^{e-1} in the p-ary expansion of x modulo p^e. Note that the image of each progression in $A \in \mathcal{A}$ is a hyperplane in $S_{\langle N \rangle}$. Moreover, by Observation 5.2, if $H = \varphi_N(A)$ then, for each prime p, the set

$$\big\{ (p,e) \in \langle N \rangle : (p,e) \in F(H) \big\}$$

forms a (possibly empty) initial segment of the sequence $(p,1), (p,2), \ldots$. Recall that we call hyperplanes that satisfy this condition 'arithmetic'.

We will apply Theorem 2.3 to $\mathcal{A}$ (with $C = 4$ and $\varepsilon > 0$ an arbitrarily small constant), and deduce that either (3) fails to hold, *or* $\mathcal{A}$ contains an almost optimal δ-generalized frame $\big(\mathcal{F}_{(p,e)} : (p,e) \in \langle N \rangle \big)$. In the former case we are done by Proposition 6.1, so let us assume the latter. We will carefully count the number of choices for the fixed sets of the frame elements $\mathcal{F}_{(p,e)}$ such that $p > \delta^{-1}$. The bound we obtain will be sufficiently strong unless N is primarily composed of primes smaller than δ^{-1}; however, for such N it turns out that the simpler argument used in Section 6 suffices to give a sufficiently strong bound.

Next, we bound the number of choices for the fixed sets of the remaining hyperplanes: those in frame sets $\mathcal{F}_{(p,e)}$ for some prime $p \leqslant \delta^{-1}$, and those not used in the frame. Surprisingly, it turns out that we can again obtain a sufficiently strong bound using the method of Section 6. Roughly speaking, these 'extra' hyperplanes are being used inefficiently, and would be better off (in terms of increasing the number of choices) by contributing to the construction of a larger frame (and thus a different value of N).

Finally, noting that the fixed sets of the hyperplanes in $\mathcal{A}$ correspond to the moduli of the original arithmetic progressions, we will use Lemma 6.2 to bound the number of minimal covering systems of $\mathbb{Z}$ of size n with given moduli.

7.1. Choosing the fixed sets of δ-generalized frames.

Let $N \in \mathbb{N}$ and $\delta > 0$, and suppose that $\big(\mathcal{F}_{(p,e)} : (p,e) \in \langle N \rangle \big)$ is a δ-generalized frame in $S_{\langle N \rangle}$ consisting of arithmetic hyperplanes.[6] Recall that $\mathcal{F}_{(p,e)}$ is a collection of at most $p-1$ hyperplanes, and that there exists an ordering $\prec$ on $\langle N \rangle$, and for each $(p,e) \in \langle N \rangle$ a set

$$I(p,e) \supseteq \big\{ (q,f) \in \langle N \rangle : (p,e) \prec (q,f) \text{ and } (q,f) \neq (p,e) \big\}, \tag{26}$$

such that $\mu_{I(p,e)}(H) > \delta$ for each $(p,e) \in \langle N \rangle$ and $H \in \mathcal{F}_{(p,e)}$. Recall also that $(p,e) \in F(H)$, and that the sets $\mathcal{F}_{(p,e)}$ with $p > \delta^{-1}$ are disjoint. We remark that the ordering $\prec$ might not be arithmetic, but the hyperplanes *are* arithmetic, and this will turn out to be sufficient.

In this subsection we will bound the number of choices for the fixed sets of the hyperplanes in $\big(\mathcal{F}_{(p,e)} : (p,e) \in \langle N \rangle \big)$ corresponding to primes larger than δ^{-1}. While doing so, it will be

---

[6]By Definition 2.2, we may suppose that δ is sufficiently small; in particular, we will assume that $\delta < 1/2$.

convenient to write $\langle N \rangle_\delta := \{(p,e) \in \langle N \rangle : p > \delta^{-1}\}$, and to define

$$\Gamma(N) := \sum_{(p,e) \in \langle N \rangle} (p-1) \quad \text{and} \quad \Gamma_\delta(N) := \sum_{(p,e) \in \langle N \rangle_\delta} (p-1).$$

Note that, by Simpson's theorem, if $\mathrm{lcm}(\mathcal{A}) = N$ then $|\mathcal{A}| \geqslant \Gamma(N)$. Given an ordering $\prec$ on $\langle N \rangle$, for each $(p,e) \in \langle N \rangle$ set

$$M_\prec(p,e) := \prod_{(q,f) \prec (p,e)} q,$$

and given a collection $\mathcal{A}$ of hyperplanes, let us write $\mathcal{D}(\mathcal{A}) := (F(H) : H \in \mathcal{A})$ for the corresponding collection of fixed sets. We begin by observing the following upper bound (cf. Lemma 5.3) on the number of choices for the sequence $(\mathcal{D}(\mathcal{F}_{(p,e)}) : (p,e) \in \langle N \rangle_\delta)$.

Lemma 7.1. *Let $N \in \mathbb{N}$ and $\delta > 0$, and let $\prec$ be an ordering on $\langle N \rangle$. There are at most*

$$\exp\left(\sum_{(p,e) \in \langle N \rangle_\delta} (p-1) \left(\sum_{(q,f) \in \langle M_\prec(p,e) \rangle} \log\left(\frac{f+1}{f}\right) + \frac{1}{\delta^2}\right)\right)$$

sequences $(\mathcal{D}(\mathcal{F}_{(p,e)}) : (p,e) \in \langle N \rangle_\delta)$ such that $(\mathcal{F}_{(p,e)} : (p,e) \in \langle N \rangle)$ is a simple δ-generalized frame in $S_{\langle N \rangle}$ with ordering $\prec$ and consisting only of arithmetic hyperplanes.

Proof. Let $(p,e) \in \langle N \rangle_\delta$, let $H \in \mathcal{F}_{(p,e)}$, and let q be a prime. Recall that, since H is an arithmetic hyperplane, it follows by Observation 5.2 that the set $F(H)$ induces a (possibly empty) initial segment of the set $(q,1), (q,2), (q,3), \ldots$

Suppose first that $q > \delta^{-1}$. We claim in this case that there are at most

$$\left|\{f : (q,f) \in \langle M_\prec(p,e) \rangle\}\right| + 1 = \exp\left(\sum_{f : (q,f) \in \langle M_\prec(p,e) \rangle} \log\left(\frac{f+1}{f}\right)\right)$$

choices for this initial segment. To see this, recall that $\mu_{I(p,e)}(H) > \delta$, and therefore $(q,f) \notin F(H)$ for every $(q,f) \in I(p,e)$. By (26), it follows that $F(H)$ does not contain any element (q,f) with $(p,e) \prec (q,f)$ and $(p,e) \neq (q,f)$, and therefore the elements (q,f) in $F(H)$ form an initial segment (in increasing order of f) of the set $\{(q,f) \in \langle N \rangle : (q,f) \prec (p,e)\}$. Since this set has the same size as the set $\{f : (q,f) \in \langle M_\prec(p,e) \rangle\}$ (which is an initial segment of the positive integers), the claimed bound on the number of choices follows.

Now suppose instead that $q \leqslant \delta^{-1}$. In this case the condition $\mu_{I(p,e)}(H) > \delta$ only implies that $F(H)$ contains at most $\log_2(\delta^{-1})$ elements of $I(p,e)$, and hence, by (26), at most $\log_2(\delta^{-1})$ elements (q,f) such that $(q,f) \not\prec (p,e)$. Repeating the argument from the case $q > \delta^{-1}$, it follows that we have at most

$$\left|\{f : (q,f) \in \langle M_\prec(p,e) \rangle\}\right| + 1 + \log_2(\delta^{-1}) \leqslant \exp\left(\sum_{f : (q,f) \in \langle M_\prec(p,e) \rangle} \log\left(\frac{f+1}{f}\right) + \frac{1}{\delta}\right)$$

choices for the initial segment of the set $(q,1), (q,2), (q,3), \ldots$

Finally, recall that there are at most $p-1$ hyperplanes in $\mathcal{F}_{(p,e)}$ for each $(p,e) \in \langle N \rangle_\delta$, and note that there are at most δ^{-1} primes $q \leqslant \delta^{-1}$. Hence, multiplying the number of choices for all $(p,e) \in \langle N \rangle_\delta$, all $H \in \mathcal{F}_{(p,e)}$, and all primes q that divide N, it follows that we have at most

$$\exp\left(\sum_{(p,e) \in \langle N \rangle_\delta} (p-1) \left(\sum_{(q,f) \in \langle M_\prec(p,e) \rangle} \log\left(\frac{f+1}{f}\right) + \frac{1}{\delta^2} \right) \right)$$

choices for the sequence $\left(\mathcal{D}(\mathcal{F}_{(p,e)}) : (p,e) \in \langle N \rangle_\delta \right)$, as claimed. $\qquad\square$

For each $N \in \mathbb{N}$ and $\delta > 0$, and each ordering $\prec$ on $\langle N \rangle$, let us define

$$Q_\delta(N, \prec) := \sum_{(p,e) \in \langle N \rangle_\delta} (p-1) \sum_{(q,f) \in \langle M_\prec(p,e) \rangle} \log\left(\frac{f+1}{f}\right).$$

The following lemma provides a sufficiently strong upper bound on $Q_\delta(N, \prec)$.

Lemma 7.2. *Let $N \in \mathbb{N}$ and $\delta > 0$, and let $\prec$ be an ordering on $\langle N \rangle$. If $\Gamma_\delta(N) > \delta \cdot \Gamma(N)$, then*

$$Q_\delta(N, \prec) \leqslant \left(\frac{4\sqrt{\tau}}{3} + o(1) \right) \frac{\Gamma_\delta(N)^{3/2}}{\left(\log \Gamma_\delta(N) \right)^{1/2}} \tag{27}$$

as $N \to \infty$.

We first use Lemma 6.3 to obtain the following bound when $\Gamma_\delta(M)$ is large.

Lemma 7.3. *Let $M \in \mathbb{N}$ and $\delta > 0$. Then*

$$\sum_{(q,f) \in \langle M \rangle} \log\left(\frac{f+1}{f}\right) \leqslant \left(2\sqrt{\tau} + o(1) \right) \left(\frac{\Gamma_\delta(M)}{\log \Gamma_\delta(M)} \right)^{1/2} + \frac{2}{\delta} \log \Gamma(M)$$

as $\Gamma_\delta(M) \to \infty$.

Proof. Let $M = p_1^{\gamma_1} \cdots p_m^{\gamma_m}$, and observe that

$$\sum_{(q,f) \in \langle M \rangle} \log\left(\frac{f+1}{f}\right) = \sum_{i=1}^{m} \log\left(\gamma_i + 1\right) = \sum_{i \,:\, p_i > \delta^{-1}} \log\left(\gamma_i + 1\right) + \sum_{i \,:\, p_i \leqslant \delta^{-1}} \log\left(\gamma_i + 1\right)$$

$$\leqslant \left(2\sqrt{\tau} + o(1) \right) \left(\frac{\Gamma_\delta(M)}{\log \Gamma_\delta(M)} \right)^{1/2} + \frac{2}{\delta} \log \Gamma(M)$$

as $\Gamma_\delta(M) \to \infty$, as required, by Lemma 6.3 applied to the sequence $\gamma_i \cdot \mathbb{1}\left[p_i > \delta^{-1} \right]$, and by the bound $\gamma_i \leqslant \Gamma(M)$, which holds for every $i \in [m]$. $\qquad\square$

When $\Gamma_\delta(M)$ is bounded and $\Gamma(M) \to \infty$, we will instead use the bound

$$\sum_{(q,f) \in \langle M \rangle} \log\left(\frac{f+1}{f}\right) \leqslant \frac{3}{\delta} \log \Gamma(M), \tag{28}$$

which follows from the proof above by using the trivial bound $\log(\gamma_i + 1) \leqslant \Gamma_\delta(M)$ for the (bounded number of) large primes p_i, instead of applying Lemma 6.3.

26

We will also need the following easy lemma.

Lemma 7.4. *Let* $2 \leqslant m_0 < m_1 < \cdots < m_\ell \leqslant m$. *Then*

$$\sum_{i=0}^{\ell-1} \left(\frac{m_i}{\log m_i} \right)^{1/2} (m_{i+1} - m_i) \leqslant \left(\frac{2}{3} + o(1) \right) \frac{m^{3/2}}{(\log m)^{1/2}}$$

as $m \to \infty$.

We can now prove Lemma 7.2.

Proof of Lemma 7.2. Observe first that, by Lemma 7.3 applied with $M = M_\prec(p, e)$, we have

$$\sum_{(q,f)\in\langle M_\prec(p,e)\rangle} \log\left(\frac{f+1}{f} \right) \leqslant \left(2\sqrt{\tau} + o(1) \right) \left(\frac{\Gamma_\delta(M_\prec(p,e))}{\log \Gamma_\delta(M_\prec(p,e))} \right)^{1/2} + \frac{2}{\delta} \log \Gamma(M_\prec(p,e))$$

as $\Gamma_\delta(M_\prec(p,e)) \to \infty$, and that when $\Gamma_\delta(M_\prec(p,e))$ is bounded, by (28) we have

$$\sum_{(q,f)\in\langle M_\prec(p,e)\rangle} \log\left(\frac{f+1}{f} \right) \leqslant \frac{3}{\delta} \log \Gamma(N),$$

as $\Gamma(N) \to \infty$, since $\Gamma(M_\prec(p,e)) \leqslant \Gamma(N)$. Thus, summing over $(p,e) \in \langle N \rangle_\delta$, it follows that

$$Q_\delta(N, \prec) \leqslant \left(2\sqrt{\tau} + o(1) \right) \sum_{(p,e)\in\langle N \rangle_\delta} (p-1) \left(\frac{\Gamma_\delta(M_\prec(p,e))}{\log \Gamma_\delta(M_\prec(p,e))} \right)^{1/2} + \frac{3}{\delta} \Gamma(N) \log \Gamma(N)$$

as $N \to \infty$ (which, in particular, implies that $\Gamma(N) \to \infty$).

Next, we apply Lemma 7.4 with $(m_0, \ldots, m_\ell) = \left(\Gamma_\delta(M_\prec(p,e)) \right)_{(p,e)\in\langle N \rangle_\delta}$ and $m = \Gamma_\delta(N)$. Note that if (p', e') immediately follows (p, e) in the ordering $\prec$ restricted to $\langle N \rangle_\delta$, then

$$\Gamma_\delta(M_\prec(p', e')) - \Gamma_\delta(M_\prec(p, e)) = p - 1,$$

since Γ_δ only counts the large primes. It follows that

$$Q_\delta(N, \prec) \leqslant \left(\frac{4\sqrt{\tau}}{3} + o(1) \right) \frac{\Gamma_\delta(N)^{3/2}}{\left(\log \Gamma_\delta(N) \right)^{1/2}} + \frac{3}{\delta} \Gamma(N) \log \Gamma(N),$$

as $N \to \infty$. Since $\Gamma_\delta(N) > \delta \cdot \Gamma(N)$, by assumption, we obtain (27), as required. $\qquad \square$

Before continuing, let us observe that the condition $\Gamma_\delta(N) > \delta \cdot \Gamma(N)$ in the statement of Lemma 7.2 (which in any case could be weakened considerably) is not a serious restriction, since we can easily obtain, using the method of Section 6, a suitable bound on the number of minimal covering systems whose least common multiple has mostly small prime factors.

Lemma 7.5. *Let* $\beta, \delta > 0$ *be constants and let* $n \in \mathbb{N}$ *and* $N \in \mathbb{N}$ *with* $\Gamma_\delta(N) \leqslant \beta \cdot \Gamma(N)$. *The number of minimal covering systems* $\mathcal{A}$ *of* $\mathbb{Z}$ *of size* n *with* $\mathrm{lcm}(\mathcal{A}) = N$ *is at most*

$$\exp\left(\left(2\sqrt{\beta\tau} + o(1) \right) \frac{n^{3/2}}{(\log n)^{1/2}} \right)$$

as $n \to \infty$.

Proof. The proof is essentially the same as that of Proposition 6.1, but we use Lemma 7.3 in place of Lemma 6.3 to count the choices of the moduli. To be more precise, in order to count the minimal covering systems $\mathcal{A} = \{A_1, \dots, A_n\}$ of $\mathbb{Z}$ of size n with $\mathrm{lcm}(\mathcal{A}) = N$, we will first choose the moduli, and then the shifts. Observe first that, for each $j \in [n]$, we have at most

$$\prod_{(p,e) \in \langle N \rangle} \left(\frac{e+1}{e} \right) \leqslant \exp \left((2\sqrt{\tau} + o(1)) \left(\frac{\Gamma_\delta(N)}{\log \Gamma_\delta(N)} \right)^{1/2} + \frac{2}{\delta} \log \Gamma(N) \right)$$

choices for the modulus of the arithmetic progression A_j. Indeed, the left-hand side is simply the number of divisors of N, and the inequality follows from Lemma 7.3.

Now, since (by assumption and by Simpson's theorem) we have $\Gamma_\delta(N) \leqslant \beta \cdot \Gamma(N) \leqslant \beta n$, it follows by Lemma 6.2 that the number of choices for $\mathcal{A}$ is at most

$$(n!)^2 \cdot \exp \left((2\sqrt{\beta\tau} + o(1)) \frac{n^{3/2}}{(\log \beta n)^{1/2}} + \frac{2n \log n}{\delta} \right) = \exp \left((2\sqrt{\beta\tau} + o(1)) \frac{n^{3/2}}{(\log n)^{1/2}} \right)$$

as $n \to \infty$, as claimed. $\qquad\square$

7.2. **Proof of Theorem 1.1.**
We are finally ready to put together the pieces and deduce our main counting result. We will need the following easy bound.

Lemma 7.6. *For every $m \in \mathbb{N}$ and $x \geqslant 0$, we have*

$$\frac{(m+x)^{3/2}}{(\log(m+x))^{1/2}} \geqslant \frac{m^{3/2}}{(\log m)^{1/2}} + \left(\frac{3}{2} + o(1) \right) \left(\frac{m}{\log m} \right)^{1/2} x,$$

as $m \to \infty$.

We can now deduce Theorem 1.1 from Theorem 2.3, Propositions 5.1 and 6.1, Lemma 6.2, Simpson's theorem, and the results of this section.

Proof of Theorem 1.1. The lower bound follows immediately from Proposition 5.1, so we will prove the upper bound. Observe first (cf. Section 6) that, by Simpson's theorem, if $\mathcal{A}$ is a minimal covering system of $\mathbb{Z}$ of size n, then $\mathrm{lcm}(\mathcal{A}) \leqslant 2^n$. We may therefore fix $N \leqslant 2^n$, and consider only covering systems $\mathcal{A}$ such that $\mathrm{lcm}(\mathcal{A}) = N$. We associate each progression $A \in \mathcal{A}$ with an arithmetic hyperplane in $S_{\langle N \rangle}$ using the bijection φ_N, as described above.

Let $\varepsilon > 0$ be an arbitrarily small constant, set $C = 4$, and let $\delta = \delta(C, \varepsilon) > 0$ be the constant given by Theorem 2.3. Suppose first that either $n > 4\Gamma(N)$ or $\Gamma(N) > 4\Gamma_\delta(N)$. Then, by Proposition 6.1 and Lemma 7.5, there are at most

$$\exp \left((\sqrt{\tau} + o(1)) \frac{n^{3/2}}{(\log n)^{1/2}} \right)$$

minimal covering systems $\mathcal{A}$ of $\mathbb{Z}$ of size n with $\mathrm{lcm}(\mathcal{A}) = N$, as required. By Simpson's theorem, let us therefore assume from now on that $\Gamma(N) \leqslant n \leqslant 4\Gamma(N) \leqslant 2^4 \Gamma_\delta(N)$.

By Theorem 2.3 (and our choice of δ), every minimal covering system $\mathcal{A}$ of $\mathbb{Z}$ of size $n \leqslant 4\Gamma(N)$ with $\operatorname{lcm}(\mathcal{A}) = N$ contains a δ-generalized frame $\big(\mathcal{F}_{(p,e)} : (p,e) \in \langle N\rangle\big)$, with

$$\sum_{(p,e)\in\langle N\rangle} |\mathcal{F}_{(p,e)}| \geqslant (1-\varepsilon)\Gamma(N). \tag{29}$$

Since $\Gamma(N) \leqslant 4\Gamma_\delta(N)$, it follows by Lemmas 7.1 and 7.2 that the number of sequences $\big(\mathcal{D}(\mathcal{F}_{(p,e)}) : (p,e) \in \langle N\rangle_\delta\big)$ such that $\big(\mathcal{F}_{(p,e)} : (p,e) \in \langle N\rangle\big)$ is a δ-generalized frame in $S_{\langle N\rangle}$ consisting only of arithmetic hyperplanes is at most

$$\Gamma(N)! \cdot \exp\left(\left(\frac{4\sqrt{\tau}}{3} + o(1) \right) \frac{\Gamma_\delta(N)^{3/2}}{\big(\log \Gamma_\delta(N)\big)^{1/2}} + \frac{\Gamma_\delta(N)}{\delta^2} \right), \tag{30}$$

where the factor of $\Gamma(N)!$ bounds (noting that $\Gamma(N) \geqslant |\langle N\rangle|$) the number of choices for the ordering $\prec$ on $\langle N\rangle$ associated with the δ-generalized frame.

We next need to count the choices of the moduli for the remaining arithmetic progressions in $\mathcal{A}$, that is, those corresponding to hyperplanes that are not included in $\mathcal{F}_{(p,e)}$ for any $(p,e) \in \langle N\rangle_\delta$. Recall (from Definition 2.2) that the sets $\mathcal{F}_{(p,e)}$ with $(p,e) \in \langle N\rangle_\delta$ are pairwise disjoint, so there are exactly

$$x := n - \sum_{(p,e)\in\langle N\rangle_\delta} |\mathcal{F}_{(p,e)}|,$$

such arithmetic progressions in $\mathcal{A}$. We bound the number of choices for the fixed sets of these remaining hyperplanes in $\mathcal{A}$ using Lemma 7.3, which implies that we have at most

$$\prod_{(p,e)\in\langle N\rangle} \left(\frac{e+1}{e} \right) \leqslant \exp\left((2\sqrt{\tau} + o(1)) \left(\frac{\Gamma_\delta(N)}{\log \Gamma_\delta(N)} \right)^{1/2} + \frac{2}{\delta}\log\Gamma(N) \right)$$

choices for each. Combining this bound with (30), and recalling that $\Gamma(N) \leqslant n \leqslant 2^4\Gamma_\delta(N)$, it follows that we have at most

$$\exp\left(\left(\frac{4\sqrt{\tau}}{3} + o(1) \right) \frac{\Gamma_\delta(N)^{3/2}}{\big(\log \Gamma_\delta(N)\big)^{1/2}} + (2\sqrt{\tau} + o(1)) \left(\frac{\Gamma_\delta(N)}{\log \Gamma_\delta(N)} \right)^{1/2} x \right) \tag{31}$$

choices for the moduli of the arithmetic progressions in $\mathcal{A}$, given N and x.

In order to bound (31), we apply Lemma 7.6 with $m = \Gamma_\delta(N)$. Since $0 \leqslant x \leqslant n$ and $2^{-4}n \leqslant m \leqslant n$, we obtain

$$\frac{\Gamma_\delta(N)^{3/2}}{\big(\log \Gamma_\delta(N)\big)^{1/2}} + \frac{3}{2}\left(\frac{\Gamma_\delta(N)}{\log \Gamma_\delta(N)} \right)^{1/2} x \leqslant \frac{\big(\Gamma_\delta(N) + x\big)^{3/2}}{\big(\log(\Gamma_\delta(N) + x)\big)^{1/2}} + \frac{o(n^{3/2})}{(\log n)^{1/2}} \tag{32}$$

as $n \to \infty$. Now, since $|\mathcal{F}_{(p,e)}| \leqslant p - 1$ for each $(p,e) \in \langle N\rangle$, it follows from (29) that

$$n - x = \sum_{(p,e)\in\langle N\rangle_\delta} |\mathcal{F}_{(p,e)}| \geqslant \Gamma_\delta(N) - \varepsilon\Gamma(N) \geqslant \Gamma_\delta(N) - \varepsilon n,$$

so $\Gamma_\delta(N) + x \leqslant (1 + \varepsilon)n$, and hence

$$\frac{\big(\Gamma_\delta(N) + x\big)^{3/2}}{\big(\log(\Gamma_\delta(N) + x)\big)^{1/2}} \leqslant (1 + \varepsilon)^{3/2} \frac{n^{3/2}}{(\log n)^{1/2}}.$$

Combining this with (31) and (32), it follows that we have at most

$$\exp\left(\left(\frac{4\sqrt{\tau}}{3} + o(1) \right) (1 + \varepsilon)^{3/2} \frac{n^{3/2}}{(\log n)^{1/2}} \right)$$

choices for the moduli of the progressions in $\mathcal{A}$.

Finally, by Lemma 6.2, it follows that there are at most

$$\exp\left(\left(\frac{4\sqrt{\tau}}{3} + O(\varepsilon) \right) \frac{n^{3/2}}{(\log n)^{1/2}} \right)$$

minimal covering systems of $\mathbb{Z}$ of size n with $\mathrm{lcm}(\mathcal{A}) = N$. Since $\varepsilon > 0$ was arbitrarily small, this completes the proof of Theorem 1.1. $\qquad\square$

Acknowledgements

References

[1] P. Balister, B. Bollobás, R. Morris, J. Sahasrabudhe and M. Tiba, On the Erdős Covering Problem: the density of the uncovered set, *Invent. Math.*, **228** (2022), 377–414.

[2] P. Balister, B. Bollobás, R. Morris, J. Sahasrabudhe and M. Tiba, The Erdős–Selfridge problem with square-free moduli, *Algebra & Number Theory*, **15** (2021), 609–626.

[3] Y. Buttkewitz, C. Elsholtz, K. Ford and J.-C. Schlage-Puchta, A Problem of Ramanujan, Erdős, and Kátai on the Iterated Divisor Function, *Int. Math. Res. Notices*, **17** (2012), 4051–4061.

[4] C. Elsholtz, M. Technau and N. Technau, The maximal order of iterated multiplicative functions, *Mathematika*, **65** (2019), 990–1009.

[5] P. Erdős, On integers of the form $2^k + p$ and some related problems, *Summa Brasil. Math.*, **2** (1950), 113–123.

[6] P. Erdős, Egy kongruenciarendszerekről szóló problémáról (On a problem concerning congruence-systems, in Hungarian), *Mat. Lapok*, **4** (1952), 122–128.

[7] P. Erdős, Problems and results in combinatorial number theory, A survey of combinatorial theory, J. N. Srivastava, ed., North-Holland, Amsterdam, 1973, 117–138.

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

[9] P. Erdős, Problems and results on combinatorial number theory III, in *Number Theory Day* (Proc. Conf., Rockefeller Univ., New York, 1976), pp. 43–72. Lecture Notes in Math., Vol. 626, Springer, Berlin, 1977.

[10] P. Erdős, A survey of problems in combinatorial number theory, Combinatorial mathematics, optimal designs and their applications, *Ann. Discrete Math.*, **6** (1980), 89–115.

[11] P. Erdős, Some of my favourite problems in number theory, combinatorics, and geometry, *Resenhas*, **2,** (1995), 165–186.

[12] P. Erdős and R. L. Graham, Old and new problems and results in combinatorial number theory, Monographies de L'Enseignement Mathématique, No. 28, 1980.

[13] 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.

[14] J. Friedlander, On exact coverings of the integers, *Israel J. Math.,* **12** (1972), 299–305.

[15] I. P. Goulden, A. Granville, L. B. Richmond and J. Shallit, Natural exact covering systems and the reversion of the Möbius series, *Ramanujan J.,* **50** (2019), 211–235.

[16] R. Hough, Solution of the minimum modulus problem for covering systems, *Ann. Math.,* **181** (2015), 361–382.

[17] R. Hough and P. Nielsen, Covering systems with restricted divisibility, *Duke Math. J.,* **168** (2019), 3261–3295.

[18] J.-P. Massias and G. Robin, Bornes effectives pour certaines fonctions concernant les nombres premiers, *Journal Th. Nombres de Bordeaux,* **8** (1996), 213–238.

[19] M. Newman, Roots of unity and covering sets, *Math. Ann.,* **191** (1971), 279–282.

[20] S. Porubský and J. Schönheim, Covering systems of Paul Erdős: Past, present and future, in Paul Erdős and his mathematics, vol. I, János Bolyai Math. Soc., Budapest, 2002, pp. 581–627.

[21] A. Schinzel, Reducibility of polynomials and covering systems of congruences, *Acta Arith.,* **13** (1967), 91–101.

[22] R. J. Simpson, Regular coverings of the integers by arithmetic progressions, *Acta Arith.,* **45** (1985), 145–152.

[23] S. Znám, On exactly covering systems of arithmetic sequences, *Math. Ann.,* **180** (1969), 227–232.

Appendix A. Proof of the geometric Simpson's theorem

In this appendix we will provide, for the reader's convenience, a proof of the following slight generalization of Simpson's theorem [22].

Theorem A.1 (Simpson's theorem)**.** *Let $\mathcal{A}$ be a minimal cover of $S_{[k]}$ with hyperplanes, and let $I \subsetneq F(\mathcal{A})$. Then*

$$\left|\left\{H \in \mathcal{A} : F(H) \not\subseteq I\right\}\right| \geqslant \sum_{i \in F(\mathcal{A}) \setminus I} \left(|S_i| - 1\right) + 1.$$

Note that Theorem 2.4 follows from Theorem A.1 by setting $I = \emptyset$.

Proof. The proof is by induction on $|F(\mathcal{A})|$. Set

$$L_s^{(i)} := S_1 \times \cdots S_{i-1} \times \{s\} \times S_{i+1} \times \cdots S_k$$

for each $i \in [k]$ and $s \in S_i$, and note that if $F(\mathcal{A}) = \{i\}$, then $\mathcal{A}$ (being minimal) must consist precisely of the $|S_i|$ hyperplanes $L_s^{(i)}$, one for each $s \in S_i$. Moreover, if $I \subsetneq F(\mathcal{A})$ then $I = \emptyset$, and hence

$$\left|\left\{H \in \mathcal{A} : F(H) \not\subseteq I\right\}\right| = |\mathcal{A}| = |S_i|,$$

31

as required. So suppose that $|F(\mathcal{A})| \geqslant 2$, and (recalling that $I \subsetneq F(\mathcal{A})$) choose an element $i \in F(\mathcal{A}) \setminus I$. For each $s \in S_i$, let $\mathcal{A}_s \subseteq \mathcal{A}$ be a minimal cover of $L_s^{(i)}$, and observe that

$$\mathcal{A} = \bigcup_{s \in S_i} \mathcal{A}_s,$$

since $\mathcal{A}$ is minimal. For convenience, let us assume (without loss) that $S_i = \{1, \ldots, p\}$.

Now, set $F_i(H) := F(H) \setminus \{i\}$ for each $H \in \mathcal{A}_s$ (and similarly for a family of hyperplanes), and define a sequence of sets $(R_0, \ldots, R_p)$ by setting $R_0 := I$, and

$$R_s := R_{s-1} \cup F_i(\mathcal{A}_s)$$

for each $s \in S_i$, so in particular $R_p = F(\mathcal{A}) \setminus \{i\}$. Now set $I_s := R_{s-1} \cap F_i(\mathcal{A}_s)$, and define

$$\mathcal{Q}_s := \big\{ H \in \mathcal{A}_s : F_i(H) \not\subseteq I_s \big\}.$$

We claim that, applying the induction hypothesis to the minimal cover $\mathcal{A}_s$ of $L_s^{(i)}$ (which we naturally identify with $S_1 \times \cdots S_{i-1} \times S_{i+1} \times \cdots S_k$), we have either $R_{s-1} = R_s$, or

$$|\mathcal{Q}_s| \geqslant \sum_{j \in R_s \setminus R_{s-1}} \big(|S_j| - 1 \big) + 1 \tag{33}$$

for each $s \in S_i$. To see this, simply note that $R_s \setminus R_{s-1} = F_i(\mathcal{A}_s) \setminus I_s$, so if $R_{s-1} \neq R_s$ then $I_s \subsetneq F_i(\mathcal{A}_s)$, and that $F_i(\mathcal{A}_s) \subseteq F(\mathcal{A}) \setminus \{i\}$, so (since $i \in F(\mathcal{A})$) we have $|F_i(\mathcal{A}_s)| < |F(\mathcal{A})|$.

Set $J := \big\{ s \in S_i : R_{s-1} \neq R_s \big\}$, and recall that (33) holds for each $s \in J$. We claim that

$$\left| \bigcup_{s \in J} \mathcal{Q}_s \right| = \sum_{s \in J} |\mathcal{Q}_s| \geqslant |J| - \big(|S_i| - 1 \big) + \sum_{j \in F(\mathcal{A}) \setminus I} \big(|S_j| - 1 \big).$$

The inequality follows from summing (33) over $s \in J$, and recalling that $i \in F(\mathcal{A}) \setminus I$, so it remains to show that the sets $\mathcal{Q}_s$ are disjoint. To see this, observe that, if $H \in \mathcal{A}_s$, then

$$F_i(H) \not\subseteq I_s \qquad \Leftrightarrow \qquad F_i(H) \subseteq R_s \quad \text{and} \quad F_i(H) \not\subseteq R_{s-1},$$

and so $H \in A_s$ for at most one element $s \in S_i$.

Finally, we claim that for each $s \in S_i \setminus J$, there exists a hyperplane $H \in \mathcal{A}$ such that $H \subseteq L_s^{(i)}$ and $F(H) \not\subseteq I$. To see this, observe first that $L_s^{(i)}$ is not covered by $\big\{ H \in \mathcal{A} : i \notin F(H) \big\}$, as otherwise $S_{[k]}$ would be covered by $\big\{ H \in \mathcal{A} : i \notin F(H) \big\}$, contradicting the minimality of $\mathcal{A}$ and the fact that $i \in F(\mathcal{A})$. It follows that there exists $H \in \mathcal{A}$ with $H \subseteq L_s^{(i)}$, and we have $F(H) \not\subseteq I$ because $i \in F(H) \setminus I$. Moreover, none of these $|S_i| - |J|$ distinct hyperplanes is included in $\mathcal{Q}_s$ for any $s \in J$, since they do not intersect the set $\bigcup_{s \in J} L_s^{(i)}$.

Hence, noting that $F(H) \not\subseteq I$ for each $H \in \mathcal{Q}_s$ (since $I \subseteq R_{s-1}$ and $F(H) \not\subseteq R_{s-1}$), we obtain

$$\big| \{ H \in \mathcal{A} : F(H) \not\subseteq I \} \big| \geqslant \left| \bigcup_{s \in J} \mathcal{Q}_s \right| + |S_i| - |J| \geqslant \sum_{j \in F(\mathcal{A}) \setminus I} \big(|S_j| - 1 \big) + 1,$$

as required. $\qquad \square$

Mathematical Institute, University of Oxford, Radcliffe Observatory Quarter, Woodstock Road, Oxford, OX2 6GG, UK

Email address: `Paul.Balister@maths.ox.ac.uk`

Department of Pure Mathematics and Mathematical Statistics, Wilberforce Road, Cambridge, CB3 0WA, UK, and Department of Mathematical Sciences, University of Memphis, Memphis, TN 38152, USA

Email address: `b.bollobas@dpmms.cam.ac.uk`

IMPA, Estrada Dona Castorina 110, Jardim Botânico, Rio de Janeiro, 22460-320, Brazil

Email address: `rob@impa.br`

Department of Pure Mathematics and Mathematical Statistics, Wilberforce Road, Cambridge, CB3 0WA, UK

Email address: `jdrs2@cam.ac.uk`

IMPA, Estrada Dona Castorina 110, Jardim Botânico, Rio de Janeiro, 22460-320, Brazil

Email address: `mt576@dpmms.cam.ac.uk`