



# ERDŐS COVERING SYSTEMS

P. BALISTER<sup>1,†</sup>, B. BOLLOBÁS<sup>1,2,†</sup>, R. MORRIS<sup>3,\*</sup>, J. SAHASRABUDHE<sup>4</sup> and  
M. TIBA<sup>2</sup>

<sup>1</sup>Department of Mathematical Sciences, University of Memphis, Memphis, TN 38152, USA  
e-mail: pbalistr@memphis.edu

<sup>2</sup>Department of Pure Mathematics and Mathematical Statistics, Wilberforce Road, Cambridge,  
CB3 0WA, United Kingdom  
e-mails: bb12@cam.ac.uk, mt576@dpmmms.cam.ac.uk

<sup>3</sup>IMPA, Estrada Dona Castorina 110, Jardim Botânico, Rio de Janeiro, 22460-320, Brazil  
e-mail: rob@impa.br

<sup>4</sup>Peterhouse, Trumpington Street, University of Cambridge, CB2 1RD, United Kingdom  
e-mail: jdrs2@cam.ac.uk

(Received January 24, 2020; revised March 20, 2020; accepted March 21, 2020)

*To Endre Szemerédi, a truly extraordinary mathematician,  
a wonderful friend, and a great inspiration, on his 80th birthday*

**Abstract.** A *covering system* is a finite collection of arithmetic progressions whose union is the set of integers. The study of these objects was initiated by Erdős in 1950, and over the following decades he asked many questions about them. Most famously, he asked whether there exist covering systems with distinct moduli whose minimum modulus is arbitrarily large. This problem was resolved in 2015 by Hough, who showed that in any such system the minimum modulus is at most  $10^{16}$ .

The purpose of this note is to give a gentle exposition of a simpler and stronger variant of Hough's method, which was recently used to answer several other questions about covering systems. We hope that this technique, which we call the *distortion method*, will have many further applications in other combinatorial settings.

## 1. Introduction

We say that a finite collection  $\{A_1, \dots, A_k\}$  of arithmetic progressions is a *covering system* if  $\bigcup_{i=1}^k A_i = \mathbb{Z}$ , that is, if their union covers the integers.

---

\* Corresponding author.

† The first two authors were partially supported by NSF grant DMS 11855745.

*Key words and phrases:* covering system, distortion method, hyperplane.

*Mathematics Subject Classification:* 05B40, 05C70, 11A07, 11B25.

The study of covering systems with distinct moduli (common differences) was initiated almost 70 years ago by Erdős [3], who posed many problems about these systems over the following decades. The most famous of these was his so-called ‘minimum modulus problem’, which asked whether there exist such systems with arbitrarily large minimum modulus. This problem was resolved by Hough [5] in 2015, following important work by Filaseta, Ford, Konyagin, Pomerance and Yu [4].

**THEOREM 1.1** (Hough [5]). *In any covering system of the integers with distinct moduli, the minimum modulus is at most  $10^{16}$ .*

Hough’s paper moreover introduced a new method, which we call the *distortion method*. In this method, we reveal the progressions in stages, and define a sequence of probability measures, each of which depends only on the progressions revealed up to that point. These measures concentrate on the set of uncovered points, and allow us to maintain a constant lower bound on the measure of this set (which may be very small in the uniform measure).

The purpose of this note is to give a gentle introduction to a simpler and more powerful variant of Hough’s method, which was introduced by the authors in two recent papers [1,2]. We will illustrate this method by giving a simple proof of Hough’s theorem in the case of square-free moduli.<sup>1</sup> Our aim is to make this method more widely known amongst the combinatorial community, in the hope that further applications will be discovered.

## 2. A geometric setting

For the purposes of exposition, it will be convenient to work in a (slightly more general) geometric setting. Let  $S_1, \dots, S_n$  be finite sets with at least two elements, and set

$$Q = S_1 \times \cdots \times S_n.$$

A *hyperplane* in  $Q$  is a set  $A = Y_1 \times \cdots \times Y_n \subset Q$ , with each  $Y_i$  either equal to  $S_i$  or a singleton element of  $S_i$ , and the set of *fixed coordinates* of  $A$  is

$$F(A) := \{k : Y_k \neq S_k\}.$$

We say that two hyperplanes  $A$  and  $A'$  are *parallel* if  $F(A) = F(A')$ .

The following theorem was proved in [2]; we will give the proof in Sections 3–5, below.

---

<sup>1</sup> We emphasize that this proof can be extended to prove Theorem 1.1 without much difficulty (see [1]), but this requires some tedious calculations that, for the sake of clarity, we wish to avoid.

**THEOREM 2.1.** *For every sequence of finite sets  $(S_k)_{k \geq 1}$  such that  $|S_k| \geq 2$  for each  $k \in \mathbb{N}$  and*

$$(1) \quad \liminf_{k \rightarrow \infty} \frac{|S_k|}{k} > 3,$$

*there exists a constant  $C$  such that the following holds. Let  $\mathcal{A}$  be a collection of hyperplanes that cover  $Q := S_1 \times \cdots \times S_n$  for some  $n \in \mathbb{N}$ . Then either two of the hyperplanes are parallel, or there exists a hyperplane  $A \in \mathcal{A}$  with  $F(A) \subset \{1, \dots, C\}$ .*

Before continuing, let us note that Theorem 2.1 implies Hough's theorem for covering systems with square-free moduli.

**COROLLARY 2.2.** *In any covering system of the integers with distinct square-free moduli, the minimum modulus is bounded by an absolute constant.*

**PROOF.** Simply apply Theorem 2.1 with  $S_k = \{1, \dots, p_k\}$  for each  $k \in \mathbb{N}$ , where  $p_1 < p_2 < \cdots$  are the prime numbers, listed in increasing order. To spell out the details, let  $\mathcal{A}$  be a covering system of the integers with distinct square-free moduli, let  $p_n$  be the largest prime that divides one of the moduli, and set  $Q := S_1 \times \cdots \times S_n$ . Now, by the Chinese Remainder Theorem, each arithmetic progression  $A = a + d\mathbb{Z} \in \mathcal{A}$  corresponds to the hyperplane  $Y_1 \times \cdots \times Y_n \subset Q$ , where  $Y_k = \{a \pmod{p_k}\}$  if  $p_k$  divides  $d$ , and  $Y_k = S_k$  otherwise. We may therefore map  $\mathcal{A}$  into a finite collection  $\mathcal{H}$  of hyperplanes that covers  $Q$ , and since the moduli of  $\mathcal{A}$  are distinct, the hyperplanes in  $\mathcal{H}$  are non-parallel.

Now, by Theorem 2.1, there exists an arithmetic progression  $A = a + d\mathbb{Z} \in \mathcal{A}$  such that the set of fixed coordinates of the corresponding hyperplane is contained in  $\{1, \dots, C\}$  (where  $C$  is the constant given by the theorem). But this means that  $d$  divides (and hence at most)  $p_1 \cdots p_C$ , which is an absolute constant, as required.  $\square$

Note that  $p_k \sim k \log k$ , whereas in Theorem 2.1 we allow the size of the sets  $S_k$  to grow only linearly. We showed in [2] that Theorem 2.1 is close to best possible, since there exists a sequence with  $|S_k| \sim k$  for which the conclusion of the theorem fails.

In the next section we will give an overview of the distortion method, and prove a general lemma regarding covering. In Section 4 we will perform a simple moment calculation, and in Section 5 we will deduce Theorem 2.1.

### 3. The distortion method

In this section we will give an outline of the proof of Theorem 2.1. We will work in the following general setting: let  $S_1, \dots, S_n$  be finite sets with

at least two elements, set

$$Q := S_1 \times \cdots \times S_n,$$

and let  $\mathcal{A}$  be a collection of hyperplanes in  $Q$ . Our task is to show that if  $|S_k|$  grows sufficiently quickly, and  $\mathcal{A}$  does not contain parallel hyperplanes, then  $\mathcal{A}$  cannot cover  $Q$ .

To do so, we will reveal the elements of  $\mathcal{A}$  in  $n$  rounds, corresponding to the  $n$  sets  $S_1, \dots, S_n$ , and define a sequence of probability measures  $\mathbb{P}_0, \dots, \mathbb{P}_n$  on  $Q$  that gradually distort the space. The measure  $\mathbb{P}_k$  will depend on the elements of  $\mathcal{A}$  that were revealed in the first  $k$  rounds, and will be chosen so that the  $\mathbb{P}_k$ -measure of the set covered in the  $k$ th round is small. However, it will be important that we do not change the measure of the set of points that were covered earlier, and we do not increase the measure of any set too much.

In order to define these measures, recall that  $F(A)$  is the set of fixed coordinates of a hyperplane  $A$ , and define

$$\mathcal{A}_k := \{A \in \mathcal{A} : \max(F(A)) = k\}$$

to be the set of hyperplanes that we reveal in round  $k$ , and

$$B_k := \bigcup_{A \in \mathcal{A}_k} A$$

to be the set that is covered by those hyperplanes. Note that, since  $F(A) \subset [k] = \{1, \dots, k\}$  for every  $A \in \mathcal{A}_k$ , we can consider  $B_k$  to be a subset of

$$Q_k := S_1 \times \cdots \times S_k$$

by identifying  $X \subset Q_k$  with  $X \times S_{k+1} \times \cdots \times S_n$ . We call a set of this form  $Q_k$ -measurable.

Let  $\mathbb{P}_0$  be the uniform probability measure on  $Q$ , and let us think of this as being the trivial measure on  $Q_0$ , the empty product. Let  $1 \leq k \leq n$ , and suppose that we have already defined a probability measure  $\mathbb{P}_{k-1}$  on  $Q_{k-1}$  (which we extend uniformly to a measure on  $Q$ ). A natural way (cf. [5]) to define the measure  $\mathbb{P}_k$  on  $Q_k$  would be to set  $\mathbb{P}_k(B_k) = 0$ , and redistribute the removed measure over the remaining elements (taking care not to change the measure of any  $Q_{k-1}$ -measurable set). However, it turns out to be helpful to define the measure  $\mathbb{P}_k$  in the following, slightly more subtle way.

Recall that  $Q_k = Q_{k-1} \times S_k$ , so the elements of  $Q_k$  can be written as pairs  $(x, y)$ , where  $x \in Q_{k-1}$  and  $y \in S_k$ . Now, for each  $x \in Q_{k-1}$ , define

$$(2) \quad \alpha_k(x) := \frac{|\{y \in S_k : (x, y) \in B_k\}|}{|S_k|},$$

that is, the proportion of the ‘fibre’  $F_x := \{(x, y) : y \in S_k\} \subset Q_k$  that is covered in round  $k$ . Now, for some  $\delta \in [0, 1/2]$ , we do one of two things on the fibre  $F_x$ , depending on whether or not  $\alpha_k(x) \leq \delta$ :

- If  $\alpha_k(x) \leq \delta$ , then we set  $\mathbb{P}_k(x, y) = 0$  for every element of  $F_x \cap B_k$ , and increase the measure proportionally on the rest of  $F_x$ ;

- If  $\alpha_k(x) > \delta$ , then we ‘cap’ the distortion by increasing the measure at each point of  $F_x \setminus B_k$  by a factor of  $1/(1 - \delta)$ , and decreasing the measure on points of  $F_x \cap B_k$  by a corresponding factor.

To be precise, the probability measure  $\mathbb{P}_k$  is defined as follows.

DEFINITION 3.1. For each  $(x, y) \in Q_k$ , define

$$\mathbb{P}_k(x, y) := \begin{cases} \max\left\{0, \frac{\alpha_k(x) - \delta}{\alpha_k(x)(1 - \delta)}\right\} \cdot \frac{\mathbb{P}_{k-1}(x)}{|S_k|}, & \text{if } (x, y) \in B_k; \\ \min\left\{\frac{1}{1 - \alpha_k(x)}, \frac{1}{1 - \delta}\right\} \cdot \frac{\mathbb{P}_{k-1}(x)}{|S_k|}, & \text{if } (x, y) \notin B_k. \end{cases}$$

Note that  $\sum_{y \in S_k} \mathbb{P}_k(x, y) = \mathbb{P}_{k-1}(x)$  for every  $x \in Q_{k-1}$ , and hence  $\mathbb{P}_k(X) = \mathbb{P}_{k-1}(X)$  for any  $Q_{k-1}$ -measurable set  $X$ . We can now easily prove the following key lemma, which (despite its simplicity) is the main step in the proof of Theorem 2.1.

LEMMA 3.2. Let  $\mathcal{A}$  be a collection of hyperplanes in  $Q = S_1 \times \cdots \times S_n$ . If

$$(3) \quad \frac{1}{4\delta(1 - \delta)} \sum_{k=1}^n \mathbb{E}_{k-1}[\alpha_k(x)^2] < 1,$$

then  $\mathcal{A}$  does not cover  $Q$ .

PROOF. Recall from (2) that  $|F_x \cap B_k| = \alpha_k(x) \cdot |S_k|$ . By Definition 3.1, it follows that

$$\begin{aligned} \mathbb{P}_k(B_k) &= \sum_{x \in Q_{k-1}} |F_x \cap B_k| \cdot \max\left\{0, \frac{\alpha_k(x) - \delta}{\alpha_k(x)(1 - \delta)}\right\} \cdot \frac{\mathbb{P}_{k-1}(x)}{|S_k|} \\ &= \frac{1}{1 - \delta} \sum_{x \in Q_{k-1}} \max\{0, \alpha_k(x) - \delta\} \cdot \mathbb{P}_{k-1}(x) \\ &\leq \frac{1}{1 - \delta} \sum_{x \in Q_{k-1}} \frac{\alpha_k(x)^2}{4\delta} \cdot \mathbb{P}_{k-1}(x) = \frac{\mathbb{E}_{k-1}[\alpha_k(x)^2]}{4\delta(1 - \delta)}. \end{aligned}$$

Indeed,  $\max\{a - b, 0\} \leq a^2/4b$  follows from  $(a - 2b)^2 \geq 0$ , and holds for all  $a, b > 0$ .

Now, since  $\mathbb{P}_n(B_k) = \mathbb{P}_k(B_k)$  for every  $1 \leq k \leq n$ , it follows that the set  $R \subset Q$  of points not covered by  $\mathcal{A}$  satisfies

$$\mathbb{P}_n(R) \geq 1 - \sum_{k=1}^n \mathbb{P}_n(B_k) \geq 1 - \frac{1}{4\delta(1-\delta)} \sum_{k=1}^n \mathbb{E}_{k-1}[\alpha_k(x)^2] > 0,$$

by (3), and hence  $\mathcal{A}$  does not cover  $Q$ , as claimed.  $\square$

#### 4. Bounding the moments of $\alpha_k(x)$

In order to use Lemma 3.2 to prove Theorem 2.1, we need to bound, for each  $1 \leq k \leq n$ , the second moment of  $\alpha_k(x)$  with respect to the measure  $\mathbb{P}_{k-1}$ . The following lemma provides the bound we need.

LEMMA 4.1. *Let  $\mathcal{A}$  be a collection of hyperplanes in  $Q = S_1 \times \cdots \times S_n$ , no two of which are parallel. Then, for each  $1 \leq k \leq n$ ,*

$$(4) \quad \mathbb{E}_{k-1}[\alpha_k(x)^2] \leq \frac{1}{|S_k|^2} \prod_{j=1}^{k-1} \left(1 + \frac{3}{(1-\delta)|S_j|}\right).$$

The first step in the proof of Lemma 4.1 is the following straightforward bound on the  $\mathbb{P}_k$ -measure of a  $Q_k$ -measurable hyperplane.

LEMMA 4.2. *Let  $A$  be a hyperplane, and let  $0 \leq k \leq n$ . If  $F(A) \subset [k]$ , then*

$$(5) \quad \mathbb{P}_k(A) \leq \prod_{j \in F(A)} \frac{1}{(1-\delta)|S_j|}.$$

In the proof of Lemma 4.2 we will use the following simple properties of the measures  $\mathbb{P}_k$ . Recall from Section 3 that

$$(6) \quad \mathbb{P}_k(X) = \mathbb{P}_{k-1}(X)$$

for any  $Q_{k-1}$ -measurable set  $X$ , and observe that

$$(7) \quad \mathbb{P}_k(X) \leq \frac{1}{1-\delta} \cdot \mathbb{P}_{k-1}(X)$$

for any set  $X \subset Q$ , by Definition 3.1. We will find it useful to define

$$\nu(J) := \prod_{j \in J} \frac{1}{(1-\delta)|S_j|}$$

for each  $J \subset [n]$  and, given a hyperplane  $A = Y_1 \times \cdots \times Y_n$  and a set  $U \subset [n]$ , to define  $A^U := Y_1^U \times \cdots \times Y_n^U$  to be the hyperplane with  $Y_i^U := Y_i$  if  $i \in U$ , and  $Y_i^U := S_i$  otherwise.

PROOF OF LEMMA 4.2. We will prove, by induction on  $k$ , that  $\mathbb{P}_k(A) \leq \nu(J)$  for all  $0 \leq k \leq n$ , every set  $J \subset [k]$ , and every hyperplane  $A$  with  $F(A) = J$ . For  $k = 0$  this follows because  $\nu(\emptyset) = 1$ , so let  $1 \leq k \leq n$ , and assume that the induction hypothesis holds for  $\mathbb{P}_{k-1}$ .

Suppose first that  $k \notin F(A)$ . Then  $A$  is  $Q_{k-1}$ -measurable and  $J \subset [k-1]$ , and it follows by (6) and the induction hypothesis that  $\mathbb{P}_k(A) = \mathbb{P}_{k-1}(A) \leq \nu(J)$ , as required.

On the other hand, if  $k \in F(A)$ , then it follows from (7) that

$$\mathbb{P}_k(A) \leq \frac{1}{1-\delta} \cdot \mathbb{P}_{k-1}(A) = \frac{1}{(1-\delta)|S_k|} \cdot \mathbb{P}_{k-1}(A^{[k-1]}),$$

since the probability measure  $\mathbb{P}_{k-1}$  is extended uniformly on each fibre. Since  $F(A^{[k-1]}) = J \setminus \{k\} \subset [k-1]$ , it follows from the induction hypothesis that

$$\mathbb{P}_{k-1}(A^{[k-1]}) \leq \nu(J \setminus \{k\}).$$

Hence, by the definition of  $\nu$ , we obtain  $\mathbb{P}_k(A) \leq \nu(J)$ , as claimed.  $\square$

Using Lemma 4.2, we can now prove the following bound on the second moment of  $\alpha_k(x)$ .

LEMMA 4.3. *Let  $\mathcal{A}$  be a collection of hyperplanes in  $Q = S_1 \times \cdots \times S_n$ , no two of which are parallel. Then, for each  $1 \leq k \leq n$ ,*

$$\mathbb{E}_{k-1}[\alpha_k(x)^2] \leq \frac{1}{|S_k|^2} \sum_{F_1, F_2 \subset [k-1]} \nu(F_1 \cup F_2).$$

PROOF. Recalling the definitions of  $\alpha_k$  and  $B_k$ , and using the union bound, we obtain

$$\alpha_k(x) = \frac{1}{|S_k|} \sum_{y \in S_k} \mathbb{1}[(x, y) \in B_k] \leq \frac{1}{|S_k|} \sum_{y \in S_k} \sum_{A \in \mathcal{A}_k} \mathbb{1}[(x, y) \in A]$$

for each  $x \in Q_{k-1}$ , and therefore

$$\alpha_k(x) \leq \frac{1}{|S_k|} \sum_{A \in \mathcal{A}_k} \mathbb{1}[x \in A^{[k-1]}],$$

since for each  $x \in Q_{k-1}$  and  $A \in \mathcal{A}_k$ , there exists  $y \in S_k$  with  $(x, y) \in A$  if and only if  $x \in A^{[k-1]}$ , and moreover such a  $y$  (if it exists) is unique, since  $k \in F(A)$ . It follows that

$$\mathbb{E}_{k-1}[\alpha_k(x)^2] \leq \frac{1}{|S_k|^2} \sum_{A_1, A_2 \in \mathcal{A}_k} \mathbb{P}_{k-1}(A_1^{[k-1]} \cap A_2^{[k-1]}).$$

Now, if  $A_1^{[k-1]} \cap A_2^{[k-1]}$  is non-empty, then it is a hyperplane whose set of fixed coordinates is  $F_1 \cup F_2$ , where  $F_1 = F(A_1) \cap [k-1]$  and  $F_2 = F(A_2) \cap [k-1]$ . Moreover, the sets  $F_i$  determine the hyperplanes  $A_i \in \mathcal{A}_k$  uniquely, since no two of the hyperplanes of  $\mathcal{A}$  are parallel. Hence, applying Lemma 4.2 and recalling the definition of  $\nu$ , it follows that

$$\mathbb{E}_{k-1}[\alpha_k(x)^2] \leq \frac{1}{|S_k|^2} \sum_{F_1, F_2 \subset [k-1]} \nu(F_1 \cup F_2),$$

as required.  $\square$

The claimed bound on  $\mathbb{E}_{k-1}[\alpha_k(x)^2]$  now follows easily.

PROOF OF LEMMA 4.1. Observe that

$$\sum_{F_1, F_2 \subset [k-1]} \nu(F_1 \cup F_2) = \sum_{J \subset [k-1]} \sum_{\substack{F_1, F_2 \subset [k-1] \\ F_1 \cup F_2 = J}} \nu(J) = \sum_{J \subset [k-1]} 3^{|J|} \nu(J).$$

Hence, by Lemma 4.3, and recalling again the definition of  $\nu$ , we have

$$\mathbb{E}_{k-1}[\alpha_k(x)^2] \leq \frac{1}{|S_k|^2} \sum_{J \subset [k-1]} 3^{|J|} \nu(J) = \frac{1}{|S_k|^2} \prod_{j=1}^{k-1} \left(1 + \frac{3}{(1-\delta)|S_j|}\right),$$

as required.  $\square$

## 5. The proof of Theorem 2.1

Theorem 2.1 is a straightforward consequence of Lemmas 3.2 and 4.1; we just need to choose  $C$  and  $\delta$  so that if  $F(A) \not\subset \{1, \dots, C\}$  for every  $A \in \mathcal{A}$ , then the bound given by Lemma 4.1 is strong enough to imply that (3) holds.

PROOF OF THEOREM 2.1. Let  $(S_k)_{k \geq 1}$  be a sequence of sets as in the statement of the theorem, so there exist  $N \in \mathbb{N}$  and  $0 < \varepsilon \leq 1$  such that  $|S_k| \geq (3 + \varepsilon)k$  for all  $k \geq N$ , and moreover  $|S_k| \geq 2$  for each  $k \in \mathbb{N}$ . We will show that if  $C = C(N, \varepsilon)$  is sufficiently large, then the conclusion of the theorem holds. Let  $\mathcal{A}$  be a collection of hyperplanes in  $Q = S_1 \times \dots \times S_n$ , no two of

which are parallel, and with  $F(A) \not\subset \{1, \dots, C\}$  for every  $A \in \mathcal{A}$ . To prove the theorem it will suffice to show that  $\mathcal{A}$  does not cover  $Q$ .

Set  $\delta := \varepsilon/6 \in (0, 1/2]$ , and observe that  $\alpha_k(x) = 0$  for every  $1 \leq k \leq C$  and  $x \in Q_{k-1}$ , since  $F(A) \not\subset \{1, \dots, C\}$  for every  $A \in \mathcal{A}$ . Moreover, by Lemma 4.1,

$$\mathbb{E}_{k-1}[\alpha_k(x)^2] \leq \frac{1}{|S_k|^2} \prod_{j=1}^{k-1} \left(1 + \frac{3}{(1-\delta)|S_j|}\right)$$

for each  $C < k \leq n$ . Now, note that  $(1-\delta)|S_j| \geq 1$  for every  $j \in \mathbb{N}$ , and that if  $j \geq N$  then  $(1-\delta)|S_j| \geq (1-\varepsilon/6)(3+\varepsilon) \cdot j$ . Thus

$$\prod_{j=1}^{k-1} \left(1 + \frac{3}{(1-\delta)|S_j|}\right) \leq 4^N \exp\left(\frac{3}{(1-\varepsilon/6)(3+\varepsilon)} \sum_{j=N}^{k-1} \frac{1}{j}\right) \leq 4^N \cdot k^{1-\varepsilon/10},$$

where the final inequality holds since

$$\sum_{j=N}^{k-1} 1/j \leq \log k \quad \text{and} \quad (1-\varepsilon/6)(3+\varepsilon)(1-\varepsilon/10) \geq 3.$$

It follows that

$$\mathbb{E}_{k-1}[\alpha_k(x)^2] \leq \frac{4^N}{|S_k|^2} \cdot k^{1-\varepsilon/10} \leq \frac{4^N}{9 \cdot k^{1+\varepsilon/10}}$$

for every  $C < k \leq n$  (as long as we chose  $C \geq N$ , so that  $|S_k| \geq 3k$ ), and hence

$$\frac{1}{4\delta(1-\delta)} \sum_{k=1}^n \mathbb{E}_{k-1}[\alpha_k(x)^2] \leq \frac{4^N}{\varepsilon} \sum_{k=C}^n \frac{1}{k^{1+\varepsilon/10}} < 1$$

if  $C = C(N, \varepsilon)$  is sufficiently large. By Lemma 3.2 it follows that  $\mathcal{A}$  does not cover  $Q$ , as required.  $\square$

**REMARK 5.1.** When  $|S_k| = p_k$ , the  $k$ th prime, for each  $k \in \mathbb{N}$ , we can choose  $\varepsilon = 1$  and  $N = 31$ , in which case the final inequality in the proof above holds as long as  $C \geq 10^{200}$ . By the proof of Corollary 2.2, this gives a (fairly terrible) bound of roughly  $\exp(10^{200})$  for the minimum modulus in a covering system with distinct square-free moduli. However, it is clear that one could do rather better with a little more effort, and in [1] we used a variant of the proof above to reduce the bound in Hough's theorem to less than  $10^6$ .

**Acknowledgement.** The authors would like to thank Noga Alon for an interesting conversation that motivated us to write this expository note.

## References

- [1] P. Balister, B. Bollobás, R. Morris, J. Sahasrabudhe and M. Tiba, On the Erdős covering problem: the density of the uncovered set (submitted), arXiv:1811.03547.
- [2] P. Balister, B. Bollobás, R. Morris, J. Sahasrabudhe and M. Tiba, The Erdős–Selfridge problem with square-free moduli (submitted), arXiv:1901.11465.
- [3] P. Erdős, On integers of the form  $2^k + p$  and some related problems, *Summa Brasil. Math.*, **2** (1950), 113–123.
- [4] 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.
- [5] R. Hough, Solution of the minimum modulus problem for covering systems, *Ann. Math.*, **181** (2015), 361–382.