



# Equal sums in random sets and the concentration of divisors

Kevin Ford<sup>1</sup> · Ben Green<sup>2</sup> ·  
Dimitris Koukoulopoulos<sup>3</sup>

Received: 25 November 2019 / Accepted: 9 December 2022  
© The Author(s) 2023

**Abstract** We study the extent to which divisors of a typical integer  $n$  are concentrated. In particular, defining  $\Delta(n) := \max_t \#\{d|n, \log d \in [t, t+1]\}$ , we show that  $\Delta(n) \geq (\log \log n)^{0.35332277\dots}$  for almost all  $n$ , a bound we believe to be sharp. This disproves a conjecture of Maier and Tenenbaum. We also prove analogs for the concentration of divisors of a random permutation and of a random polynomial over a finite field. Most of the paper is devoted to a study of the following much more combinatorial problem of independent interest. Pick a random set  $\mathbf{A} \subset \mathbb{N}$  by selecting  $i$  to lie in  $\mathbf{A}$  with probability  $1/i$ . What is the supremum of all exponents  $\beta_k$  such that, almost surely as  $D \rightarrow \infty$ , some integer is the sum of elements of  $\mathbf{A} \cap [D^{\beta_k}, D]$  in  $k$  different ways? We characterise  $\beta_k$  as the solution to a certain optimisation problem over measures on the discrete cube  $\{0, 1\}^k$ , and obtain lower bounds for  $\beta_k$  which we believe to be asymptotically sharp.

---

✉ Ben Green  
ben.green@maths.ox.ac.uk

Kevin Ford  
ford126@illinois.edu

Dimitris Koukoulopoulos  
dimitris.koukoulopoulos@umontreal.ca

- <sup>1</sup> Department of Mathematics, University of Illinois at Urbana–Champaign, Urbana, IL 61801, USA
- <sup>2</sup> Mathematical Institute, Andrew Wiles Building, Radcliffe Observatory Quarter, Woodstock Road, Oxford OX2 6GG, UK
- <sup>3</sup> Département de Mathématiques et de Statistique, Université de Montréal, CP 6128 succ. Centre-Ville, Montréal, QC H3C 3J7, Canada

**Mathematics Subject Classification** Primary 11N25 · Secondary 05A05 ,  
11S05

## Contents

Part I. Main results and overview of the paper	.....
1 Introduction	.....
2 Application to random integers, random permutations and random polynomials	.....
3 Overview of the paper	.....
Part II. Equal sums and the optimisation problem	.....
4 The upper bound $\beta_k \leq \gamma_k$	.....
5 The lower bound $\beta_k \geq \tilde{\gamma}_k$	.....
6 An argument of Maier and Tenenbaum	.....
Part III. The optimisation problem	.....
7 The optimisation problem—basic features	.....
8 The strict entropy condition	.....
Part IV. Binary systems	.....
9 Binary systems and a lower bound for $\beta_k$	.....
10 Binary systems: proofs of the basic properties	.....
11 The limit of the $\rho_i$	.....
12 Calculating the $\rho_i$ and $\rho$	.....
Appendices	.....
Appendix A. Some probabilistic lemmas	.....
Appendix B. Basic properties of entropy	.....
Appendix C. Maier–Tenenbaum flags	.....
References	.....

## Part I. Main results and overview of the paper

### 1 Introduction

#### 1.1 The concentration of divisors

Given an integer  $n$ , we define the Delta function

$$\Delta(n) := \max_t \#\{d|n, \log d \in [t, t + 1]\},$$

that is to say the maximum number of divisors  $n$  has in any interval of logarithmic length 1. Its normal order (almost sure behaviour) has proven quite mysterious, and indeed it was a celebrated achievement of Maier and Tenen-

baum [20], answering a question of Erdős from 1948 [9], to show that  $\Delta(n) > 1$  for almost all<sup>1</sup>  $n$ .

Work on the distribution of  $\Delta$  began in the 1970s with Erdős and Nicolas [7, 8]. However, it was not until the work of Hooley [16] that the Delta function received proper attention. Among other things, Hooley showed how bounds on the average size of  $\Delta$  can be used to count points on certain algebraic varieties. Further work on the normal and average behavior of  $\Delta$  can be found in the papers of Tenenbaum [23, 24], Hall and Tenenbaum [12–14], and of Maier and Tenenbaum [20–22]. See also [15, Ch. 5, 6, 7]. Finally, Tenenbaum’s survey paper [26, p. 652–658] includes a history of the Delta function and description of many applications in number theory.

The best bounds for  $\Delta(n)$  for “normal”  $n$  currently known were obtained in a more recent paper of Maier and Tenenbaum [22].

**Theorem MT (Maier–Tenenbaum [22])** *Let  $\varepsilon > 0$  be fixed. Then*

$$(\log \log n)^{c_1 - \varepsilon} \leq \Delta(n) \leq (\log \log n)^{\log 2 + \varepsilon},$$

for almost all  $n$ , where

$$c_1 = \frac{\log 2}{\log \left( \frac{1 - 1/\log 27}{1 - 1/\log 3} \right)} \approx 0.33827.$$

It is conjectured in [22] that the lower bound is optimal.

One of the main results of this paper is a disproof of this conjecture.

**Theorem 1** *Let  $\varepsilon > 0$  be fixed. Then*

$$\Delta(n) \geq (\log \log n)^{\eta - \varepsilon}$$

for almost all  $n$ , where  $\eta = 0.35332277270132346711 \dots$

The constant  $\eta$ , which we believe to be sharp, is described in relation (1.3) below, just after the statement of Theorem 2.

## 1.2 Packing divisors

Let us briefly attempt to explain, without details, why it was natural for Maier and Tenenbaum to make their conjecture, and what it is that allows us to find even more tightly packed divisors.

---

<sup>1</sup> A property of natural numbers is said to occur for *almost all*  $n$  if the number of exceptions below  $x$  is  $o(x)$  as  $x \rightarrow \infty$ .

We start with a simple observation. Let  $n$  be an integer, and suppose we can find pairs of divisors  $d_i, d'_i$  of  $n$ ,  $i = 1, \dots, k$ , such that

- $1 < d_i/d'_i \leq 2^{1/k}$ ;
- The sets of primes dividing  $d_i d'_i$  are disjoint, as  $i$  varies in  $\{1, \dots, k\}$ .

Then we can find  $2^k$  different divisors of  $n$  in a dyadic interval, namely all products  $a_1 \dots a_k$  where  $a_i$  is either  $d_i$  or  $d'_i$ .

In [22], Maier and Tenenbaum showed how to find many such pairs of divisors  $d_i, d'_i$ . To begin with, they look only at the large prime factors of  $n$ . They first find one pair  $d_1, d'_1$  using the technique of [20]. Then, using a modification of the argument, they locate a further pair  $d_2$  and  $d'_2$ , but with these divisors not having any primes in common with  $d_1, d'_1$ . They continue in this fashion to find  $d_3, d'_3, d_4, d'_4$ , etc., until essentially all the large prime divisors of  $n$  have been used. After this, they move on to a smaller range of prime factors of  $n$ , and so on.

By contrast, we eschew an iterative approach and select  $2^k$  close divisors from amongst the large prime divisors of  $n$  in one go, in a manner that is combinatorially quite different to that of Maier and Tenenbaum. We then apply a similar technique to a smaller range of prime factors of  $n$ , and so on. This turns out to be a more efficient way of locating proximal divisors.

In fact, we provide a general framework that encapsulates all possible combinatorial constructions one might use to pack many divisors close to each other. To work in this generality it is necessary to use a probabilistic formalism. One effect of this is that, even though our work contains that of Maier and Tenenbaum as a special case, the arguments here will look totally different.

### 1.3 Random sets and equal sums

For most of the paper we do not talk about integers and divisors, but rather about the following model setting. Throughout the paper,  $\mathbf{A}$  will denote a random set of positive integers in which  $i$  is included in  $\mathbf{A}$  with probability  $1/i$ , these choices being independent for different  $i$ s. We refer to  $\mathbf{A}$  as a *logarithmic random set*.

A large proportion of our paper will be devoted to understanding conditions under which there is an integer which can be represented as a sum of elements of  $\mathbf{A}$  in (at least)  $k$  different ways. In particular, we wish to obtain bounds on the quantities  $\beta_k$  defined in the following problem.

**Problem 1** Let  $k \geq 2$  be an integer. Determine  $\beta_k$ , the supremum of all exponents  $c < 1$  for which the following is true: with probability tending to 1 as  $D \rightarrow \infty$ , there are distinct sets  $A_1, \dots, A_k \subset \mathbf{A} \cap [D^c, D]$  with equal sums, i.e.,  $\sum_{a \in A_1} a = \dots = \sum_{a \in A_k} a$ .

The motivation for the random set  $\mathbf{A}$  comes from our knowledge of the anatomy of integers, permutations and polynomials. For a random integer  $m \leq x$ , with  $x$  large, let  $U_k$  be the event that  $m$  has a prime factor in the interval  $(e^k, e^{k+1}]$ . For a random permutation  $\sigma \in S_n$ , let  $V_k$  be the event that  $\sigma$  has a cycle of size  $k$ , and for a random monic polynomial  $f$  of degree  $n$  over  $\mathbb{F}_q$ , with  $n$  large, let  $W_k$  be the event that  $f$  has an irreducible factor of degree  $k$ . Then it is known (see e.g., [2,3,15]) that  $U_k, V_k$  and  $W_k$  each occur with probability close to  $1/k$ , and also that the  $U_k$  are close to independent for  $k = o(\log x)$ , the  $V_k$  are close to independent for  $k = o(n)$ , and the  $W_k$  are close to independent for  $k$  large and  $k = o(n)$ . Thus, the model set  $\mathbf{A}$  captures the factorization structure of random integers, random permutations and random polynomials over a finite field. It is then relatively straightforward to transfer results about subset sums of  $\mathbf{A}$  to divisors of integers, permutations and polynomials. Section 2 below contains details of the transference principle.

The main result of this paper is an asymptotic lower bound on  $\beta_k$ .

**Theorem 2** *We have  $\liminf_{r \rightarrow \infty} (\beta_{2^r})^{1/r} \geq \rho/2$ , where*

$$\rho = 0.28121134969637466015 \dots$$

*is a specific constant defined as the unique solution in  $[0, 1/3]$  of*

$$\frac{1}{1 - \rho/2} = \lim_{j \rightarrow \infty} \frac{\log a_j}{2^{j-2}}, \tag{1.1}$$

*where the sequence  $a_j$  is defined by*

$$a_1 = 2, \quad a_2 = 2 + 2^\rho, \quad a_j = a_{j-1}^2 + a_{j-1}^\rho - a_{j-2}^{2\rho} \quad (j \geq 3).$$

The proof of Theorem 2 will occupy the bulk of this paper, and has three basic parts:

- (a) Showing that for every  $r \geq 1$ ,  $\beta_{2^r} \geq \theta_r$  for a certain explicitly defined constant  $\theta_r$ ;
- (b) Showing that  $\lim_{r \rightarrow \infty} \theta_r^{1/r}$  exists;
- (c) Showing that (1.1) has a unique solution  $\rho \in [0, 1/3]$  and that

$$\rho = 2 \lim_{r \rightarrow \infty} \theta_r^{1/r}.$$

In the sequel we shall refer to “Theorem 2 (a)”, “Theorem 2 (b)” and “Theorem 2 (c)”. Parts (a), (b) and (c) are quite independent of one another, with the proof of (a) (given in Sect. 9.2) being by far the longest of the three. The definition of  $\theta_r$ , while somewhat complicated, is fairly self-contained: see Definition 9.6. Parts (b) and (c) are then problems of an analytic and combinatorial

flavour which can be addressed largely independently of the main arguments of the paper. The formula (1.1) allows for a quick computation of  $\rho$  to many decimal places, as the limit on the right side converges extremely rapidly. See Sect. 12 for details.

Let us now state an important corollary of Theorem 2.

**Corollary 1** *Define*

$$\zeta_+ = \limsup_{k \rightarrow \infty} \frac{\log k}{\log(1/\beta_k)} \quad \text{and} \quad \zeta_- = \liminf_{k \rightarrow \infty} \frac{\log k}{\log(1/\beta_k)}. \quad (1.2)$$

*Then*

$$\zeta_+ \geq \zeta_- \geq \eta := \frac{\log 2}{\log(2/\rho)} = 0.3533227\dots \quad (1.3)$$

*Proof* Evidently,  $\zeta_+ \geq \zeta_-$ . In addition, observe the trivial bound  $\beta_k \leq \beta_{k+1}$ . Hence,

$$\zeta_+ = \limsup_{r \rightarrow \infty} \frac{r \log 2}{\log(1/\beta_{2r})} \quad \text{and} \quad \zeta_- = \liminf_{r \rightarrow \infty} \frac{r \log 2}{\log(1/\beta_{2r})}. \quad (1.4)$$

We then use Theorem 2 to find that  $\zeta_- \geq \eta$ . □

We conjecture that our lower bounds on  $\beta_k$  are asymptotically sharp, so that the following holds:

**Conjecture 1** *We have  $\zeta_+ = \zeta_- = \eta$ .*

We will address the exact values of  $\beta_k$  in a future paper; in particular, we will show that

$$\beta_3 = \frac{\log 3 - 1}{\log 3 + \frac{1}{\xi}} = 0.02616218797316965133\dots$$

and

$$\beta_4 = \frac{\log 3 - 1}{\log 3 + \frac{1}{\xi} + \frac{1}{\xi\lambda}} = 0.01295186091360511918\dots$$

where

$$\xi = \frac{\log 2 - \log(e-1)}{\log(3/2)}, \quad \lambda = \frac{\log 2 - \log(e-1)}{1 + \log 2 - \log(e-1) - \log(1 + 2^{1-\xi})}.$$

### 1.4 Application to divisors of integers, permutations and polynomials

The link between Problem 1 and the concentration of divisors is given by the following Theorems. The proofs are relatively straightforward and given in the next section. Recall from (1.2) the definition of  $\zeta_+$ .

**Theorem 3** *For any  $\varepsilon > 0$ , we have*

$$\Delta(n) \geq (\log \log n)^{\zeta_+ - \varepsilon}$$

*for almost every  $n$ .*

*Remark* In principle, the proof of Theorem 3 yields an explicit bound on the size of the set of integers  $n$  with  $\Delta(n) \leq (\log \log n)^{\zeta_+ - \varepsilon}$ . However, incorporating such an improvement is a very complicated task. In addition, the obtained bound will presumably be rather weak without a better understanding of the theoretical tools we develop (cf. Sect. 3).

The same probabilistic setup allows us to quickly make similar conclusions about the distribution of divisors (product of cycles) of permutations and of polynomials over finite fields.

**Theorem 4** *For a permutation  $\sigma$  on  $S_n$ , denote by*

$$\Delta(\sigma) := \max_r \#\{d|\sigma : \text{length}(d) = r\},$$

*where  $d$  denotes a generic divisor of  $\sigma$ ; that is,  $d$  is the product of a subset of the cycles of  $\sigma$ .*

*Let  $\varepsilon > 0$  be fixed. If  $n$  is sufficiently large in terms of  $\varepsilon$ , then for at least  $(1 - \varepsilon)(n!)$  of the permutations  $\sigma \in S_n$ , we have*

$$\Delta(\sigma) \geq (\log n)^{\zeta_+ - \varepsilon}.$$

**Theorem 5** *Let  $q$  be any prime power. For a polynomial  $f \in \mathbb{F}_q[t]$ , let*

$$\Delta(f) = \max_r \#\{g|f : \text{deg}(g) = r\}.$$

*Let  $\varepsilon > 0$  be fixed. If  $n$  is sufficiently large in terms of  $\varepsilon$ , then at least  $(1 - \varepsilon)q^n$  monic polynomials of degree  $n$  satisfy*

$$\Delta(f) \geq (\log n)^{\zeta_+ - \varepsilon}.$$

**Conjecture 2** *The lower bounds given in Theorems 3, 4 and 5 are sharp. That is, corresponding upper bounds with exponent  $\zeta_+ + \varepsilon$  hold.*

If both Conjectures 1 and 2 hold, then we deduce that the optimal exponent in the above theorems is equal to  $\eta$ .

*Remark* The exponent  $\zeta_+ - \varepsilon$  in Theorems 3, 4 and 5 depends only on accurate asymptotics for  $\beta_k$  as  $k \rightarrow \infty$  or, even more weakly, for  $\beta_{2^r}$  as  $r \rightarrow \infty$  (cf. (1.4)). In this work, however, we develop a framework for determining  $\beta_k$  exactly for each  $k$ .

The quantity  $\beta_k$  is also closely related to the densest packing of  $k$  divisors of a typical integer. To be specific, we define  $\alpha_k$  be the supremum of all real numbers  $\alpha$  such that for almost every  $n \in \mathbb{N}$ ,  $n$  has  $k$  divisors  $d_1 < \dots < d_k$  with  $d_k \leq d_1(1 + (\log n)^{-\alpha})$ . In 1964, Erdős [10] conjectured that  $\alpha_2 = \log 3 - 1$ , and this was confirmed by Erdős and Hall [6] (upper bound) and Maier and Tenenbaum [20] (lower bound). The best bounds on  $\alpha_k$  for  $k \geq 3$  are given by Maier and Tenenbaum [22], who showed that

$$\alpha_k \leq \frac{\log 2}{k+1} \quad (k \geq 3)$$

and (this is not stated explicitly in [22])

$$\alpha_k \geq \frac{(\log 3 - 1)^m 3^{m-1}}{(3 \log 3 - 1)^{m-1}} \quad (2^{m-1} < k \leq 2^m, m \in \mathbb{N}). \quad (1.5)$$

See also [26, p. 655–656].<sup>2</sup> In particular, it is not known if  $\alpha_3 > \alpha_4$ , although Tenenbaum [26] conjectures that the sequence  $(\alpha_k)_{k \geq 2}$  is strictly decreasing.

We can quickly deduce a lower bound for  $\alpha_k$  in terms of  $\beta_k$ .

**Theorem 6** *For all  $k \geq 2$  we have  $\alpha_k \geq \beta_k / (1 - \beta_k)$ .*

In particular,

$$\alpha_3 \geq \frac{\beta_3}{1 - \beta_3} = 0.0268650\dots,$$

which is substantially larger than the bound from (1.5), which is  $\alpha_3 \geq 0.0127069\dots$

Combining Theorem 6 with the bounds on  $\beta_k$  given in Theorem 2, we have improved the lower bounds (1.5) for large  $k$ .

The upper bound on  $\alpha_k$  is more delicate, and a subject which we will return to in a future paper. For now, we record our belief that the lower bound in Theorem 6 is sharp.

**Conjecture 3** *For all  $k \geq 2$  we have  $\alpha_k = \beta_k / (1 - \beta_k)$ .*

<sup>2</sup> The factor  $3^{m-1}$  is missing in the stated lower bounds for  $\alpha_k$  in [26].



## 2 Application to random integers, random permutations and random polynomials

In this section we assume the validity of Theorem 2 and use it to prove Theorems 3, 4, 5 and 6. The two main ingredients in this deduction are a simple combinatorial device (Lemma 2.1), of a type often known as a “tensor power trick”, used for building a large collection of equal subset sums, and transference results (Lemmas 2.2, 2.3 and 2.4) giving a correspondence between the random set  $\mathbf{A}$  and prime factors of a random integer, the cycle structure of a random permutation and the factorization of a random polynomial over a finite field. In the integer setting, this is a well-known principle following, e.g. from the Kubilius model of the integers (Kubilius, Elliott [4,5], Tenenbaum [25]). We give a self-contained (modulo using the sieve) proof below.

Throughout this section,  $\mathbf{A}$  denotes a logarithmic random set.

### 2.1 A “tensor power” argument

In this section we give a simple combinatorial argument, first used in a related context in the work of Maier–Tenenbaum [20], which shows how to use equal subsums in multiple intervals  $((D')^c, D']$  to create many more common subsums in  $\mathcal{A}$ .

**Lemma 2.1** *Let  $k \in \mathbb{Z}_{\geq 2}$  and  $\varepsilon > 0$  be fixed. Let  $D_1, D_2$  be parameters depending on  $D$  with  $3 \leq D_1 < D_2 \leq D$ ,  $\log \log D_1 = o(\log \log D)$  and  $\log \log D_2 = (1 - o(1)) \log \log D$  as  $D \rightarrow \infty$ . Then, with probability  $\rightarrow 1$  as  $D \rightarrow \infty$ , there are distinct  $A_1, \dots, A_M \subset \mathbf{A} \cap [D_1, D_2]$  with  $\sum_{a \in A_1} a = \dots = \sum_{a \in A_M} a$  and  $M \geq (\log D)^{(\log k) / \log(1/\beta_k) - \varepsilon}$ .*

*Remark* In particular, the result applies when  $D_1 = 3$  and  $D_2 = D$ , in which case it has independent combinatorial interest, giving a (probably tight) lower bound on the growth of the representation function for a random set.

*Proof* Since increasing the value of  $D_1$  only makes the proposition stronger, we may assume that  $D_1 \rightarrow \infty$  as  $D \rightarrow \infty$ . Let  $0 < \delta < \beta_k$ , and set  $\alpha := \beta_k - \delta$ . Set

$$m := \left\lfloor \frac{\log \log D_2 - \log \log D_1}{-\log(\beta_k - \delta)} \right\rfloor$$

and consider the intervals  $[D_2^{\alpha^{i+1}}, D_2^{\alpha^i}), i = 0, 1, \dots, m - 1$ . Due to the choice of  $m$ , these all lie in  $[D_1, D_2]$ .

Let  $E_i, i = 0, 1, 2, \dots$  be the event that there are distinct  $A_1^{(i)}, \dots, A_k^{(i)} \subset [D_2^{\alpha^{i+1}}, D_2^{\alpha^i})$  with  $\sum_{a \in A_1^{(i)}} a = \dots = \sum_{a \in A_k^{(i)}} a$ . Then, by the definition of

$\beta_k$  and the fact that  $D_1 \rightarrow \infty$ , we have  $\mathbb{P}(E_i) = 1 - o(1)$ , uniformly in  $i = 0, 1, \dots, m - 1$ . Here and throughout the proof,  $o(1)$  means a function tending to zero as  $D \rightarrow \infty$ , at a rate which may depend on  $k, \delta$ . These events  $E_i$  are all independent. The Law of Large Numbers then implies that, with probability  $1 - o(1)$ , at least  $(1 - o(1))m$  of them occur, let us say for  $i \in I$ ,  $|I| = (1 - o(1))m$ .

From the above discussion, we have found  $M := k^{|I|} = k^{(1-o(1))m}$  distinct sets  $B_j = \bigcup_{i \in I} A_{j_i}^{(i)}$ ,  $j \in [k]^I$ , such that all of the sums  $\sum_{a \in B_j} a$  are the same. Note that

$$M = k^{(1+O_k(\delta)+o(1)) \log \log D / \log(1/\beta_k)}.$$

Taking  $\delta$  small enough and  $D$  large enough, the result follows. □

### 2.2 Modeling prime factors with a logarithmic random set

Let  $X$  be a large parameter, suppose that

$$1 \leq K \leq (\log X)^{1/2}, \tag{2.1}$$

and let  $I = [i_1, i_2] \cap \mathbb{N}$ , where

$$i_1 = \lfloor K(\log \log X)^3 \rfloor, \quad i_2 = \left\lfloor \frac{K \log X}{2 \log \log \log X} \right\rfloor. \tag{2.2}$$

For a uniformly random positive integer  $\mathbf{n} \leq X$ , let  $\mathbf{n} = \prod_p p^{v_p}$  be the prime factorization of  $\mathbf{n}$ , where the product is over all primes. Let  $\mathcal{P}_i$  be the set of primes in  $(e^{i/K}, e^{(i+1)/K}]$ , and define the random set

$$\mathbf{I} = \{i \in I : \exists p \in \mathcal{P}_i \text{ such that } p|\mathbf{n}\}. \tag{2.3}$$

that is, the set of  $i$  for which  $\mathbf{n}$  has a prime factor in  $\mathcal{P}_i$ . By the sieve, it is known that the random variables  $v_p$  are nearly independent for  $p = X^{o(1)}$ , and thus the probability that  $b_i \geq 1$  is roughly

$$R_i := \sum_{p \in \mathcal{P}_i} \frac{1}{p} \approx \frac{1}{i}.$$

The next lemma makes this precise.

Recall the notion of *total variation distance*  $d_{TV}(X, Y)$  between two discrete real random vectors  $X, Y$  defined on the same probability space  $(\Omega, \mathcal{F}, \mathbb{P})$ :

$$d_{TV}(X, Y) = \max_{A \in \mathcal{F}} |\mathbb{P}(X \in A) - \mathbb{P}(Y \in A)|.$$

We have

$$d_{\text{TV}}((X_1, \dots, X_k), (Y_1, \dots, Y_k)) \leq \sum_{j=1}^k d_{\text{TV}}(X_j, Y_j), \tag{2.4}$$

provided that the random variables  $X_j, Y_j$  live on the same probability space for each  $j$ , that  $X_1, \dots, X_k$  are independent, and  $Y_1, \dots, Y_k$  are also independent. Although we believe this is a standard inequality, we could not find a good reference for it and give a proof of (2.4) in Lemma A.8. In addition, recall the identity

$$d_{\text{TV}}(X, Y) = \frac{1}{2} \sum_{t \in \Omega} |\mathbb{P}(X = t) - \mathbb{P}(Y = t)| \tag{2.5}$$

when  $X$  and  $Y$  take values in a probability space  $(\Omega, \mathcal{F}, \mathbb{P})$  with  $\Omega$  countable and  $\mathcal{F}$  being the power set of  $\Omega$ . See, e.g. [19, Proposition 4.2].

**Lemma 2.2** *Uniformly for any collection  $\mathcal{I}$  of subsets of  $I$ , we have*

$$\mathbb{P}(\mathbf{A} \cap I \in \mathcal{I}) = \mathbb{P}(\mathbf{I} \in \mathcal{I}) + O(1/\log \log X).$$

*Proof* For  $i_1 \leq i \leq i_2$ , let  $\omega_i$  be the indicator function of the event that  $\mathbf{n}$  has a prime factor from  $\mathcal{P}_i$ , let  $Q_i$  be a Poisson random variable with parameter  $R_i$ , with the different  $Q_i$  independent, and let  $Z_i = 1_{Q_i \geq 1}$ .<sup>3</sup> Also, let  $Y_i$  be a Bernoulli random variable with  $\mathbb{P}(Y_i = 1) = 1/i$ , again with the  $Y_i$  independent. Let  $\boldsymbol{\omega}, \mathbf{Z}, \mathbf{Y}$  denote the vectors of the variables  $\omega_i, Z_i, Y_i$ , respectively. By assumption, each  $\mathcal{P}_i \subset [\log X, X^{1/3 \log \log \log X}]$ . Hence, Theorem 1 of [11] implies that

$$d_{\text{TV}}(\boldsymbol{\omega}, \mathbf{Z}) \ll \frac{1}{\log \log X}.$$

In addition, note that  $d_{\text{TV}}(Z_i, Y_i) \ll 1/i^2$  for all  $i$ , something that can be easily proven using (2.5). Combining this estimate with (2.4), we find that

$$d_{\text{TV}}(\mathbf{Z}, \mathbf{Y}) \leq \sum_{i=i_1}^{i_2} d_{\text{TV}}(Z_i, Y_i) \ll \sum_{i=i_1}^{i_2} \frac{1}{i^2} \ll \frac{1}{\log \log X}.$$

The triangle inequality then implies that  $d_{\text{TV}}(\boldsymbol{\omega}, \mathbf{Y}) \ll 1/\log \log X$ , as desired. □

---

<sup>3</sup> We use  $1_E$  for the indicator function of a statement  $E$ ; that is,  $1_E = 1$  if  $E$  is true and  $1_E = 0$  if  $E$  is false.

### 2.3 The concentration of divisors of integers

In this section we prove Theorems 3 and 6. Recall from (1.2) the definition of  $\zeta_+$ .

*Proof of Theorem 3* Fix  $\varepsilon > 0$  and let  $X$  be large enough in terms of  $\varepsilon$ , and let  $\mathbf{n} \leq X$  be a uniformly sampled random integer. Generate a logarithmic random set  $\mathbf{A}$ . Set  $K = 10 \log \log X$ ,  $D_1 = i_1$ ,  $D = D_2 = i_2$ , where  $i_1$  and  $i_2$  are defined by (2.2). With our choice of parameters, the hypotheses of Lemma 2.1 hold and hence, with probability  $1 - o(1)$  as  $X \rightarrow \infty$ , there are distinct sets  $A_1, \dots, A_M \subset \mathbf{A} \cap [D_1, D_2]$  with  $\sum_{a \in A_1} a = \dots = \sum_{a \in A_M} a$  and  $M := \lceil (\log \log X)^{\zeta_+ - \varepsilon} \rceil$ . By Lemma A.2, with probability  $1 - o(1)$ , we have

$$|\mathbf{A} \cap [D_1, D_2]| \leq 2 \log D_2 \leq 2 \log \log X + 2 \log K.$$

Write  $F$  for the event that both of these happen.

Let  $\mathbf{n}$  be a random integer chosen uniformly in  $[1, X]$ , and let  $\mathbf{I}$  be the random set associated to  $\mathbf{n}$  via (2.3). By Lemma 2.2, the corresponding event  $F'$  for  $\mathbf{I}$  also holds with probability  $1 - o(1)$ ; that is,  $F'$  is the event that  $|\mathbf{I} \cap [D_1, D_2]| \leq 2 \log D_2$  and that there are distinct subsets  $I_1, \dots, I_M$  with equal sums. Assume we are in the event  $F'$ . For each  $i \in \mathbf{I}$ ,  $\mathbf{n}$  is divisible by some prime  $p_i \in \mathcal{P}_i$ . In addition, for each  $r, s \in \{1, 2, \dots, M\}$ , we have

$$\begin{aligned} \left| \sum_{i \in I_r} \log p_i - \sum_{i \in I_s} \log p_i \right| &\leq \frac{|I_r| + |I_s|}{K} + \frac{1}{K} \left| \sum_{i \in I_r} i - \sum_{i \in I_s} i \right| \\ &\leq \frac{4 \log \log X + 4 \log K}{K} < \frac{1}{2}. \end{aligned}$$

Writing  $d_r := \prod_{i \in I_r} p_i$  for each  $i$ , we thus see that the  $d_r$ 's are all divisors of  $\mathbf{n}$  and their logarithms all lie in an interval of length 1. It follows that  $\mathbb{P}(\Delta(\mathbf{n}) \geq M) = 1 - o(1)$  when  $\mathbf{n}$  is a uniformly sampled random integer from  $[1, X]$ , as required for Theorem 3.  $\square$

*Proof of Theorem 6* Fix  $0 < c < \beta_k / (1 - \beta_k)$ , let  $X$  be large and set  $K = (\log X)^c$ . Define  $i_1, i_2$  by (2.2), let  $D = i_2$  and define  $c'$  by  $D^{c'} = i_1$ . Let  $\mathbf{n}$  be a random integer chosen uniformly in  $[1, X]$ . We have

$$c' = \frac{c}{c + 1} + o(1) \quad (X \rightarrow \infty),$$

and therefore  $c' \leq \beta_k - \delta$  for some  $\delta > 0$ , which depends only on  $c$ . By the definition of  $\beta_k$  and Lemma 2.2, it follows that with probability  $1 - o(1)$ ,

the set  $\mathbf{I}$  defined in (2.3) has  $k$  distinct subsets  $I_1, \dots, I_k$  with equal sums, and moreover (cf. the proof of Theorem 3 above),  $|\mathbf{I}| \leq 2 \log i_2$ , so that  $|I_j| \leq 2 \log i_2$  for each  $j$ . Thus, with probability  $1 - o(1)$ , there are primes  $p_i \in \mathcal{P}_i$  ( $i \in \mathbf{I}$ ) such that for any  $r, s \in \{1, \dots, k\}$  we have

$$\left| \sum_{i \in I_r} \log p_i - \sum_{i \in I_s} \log p_i \right| \leq \frac{|I_r| + |I_s|}{K} \leq \frac{4 \log \log X}{(\log X)^c}.$$

Thus, setting  $d_r = \prod_{i \in I_r} p_i$ , we see that  $d_r \leq d_s \exp \left\{ O\left(\frac{\log \log X}{(\log X)^c}\right) \right\}$  for any  $r, s \in \{1, \dots, k\}$ . Since  $c$  is arbitrary subject to  $c < \beta_k / (1 - \beta_k)$ , we conclude that  $\alpha_k \geq \beta_k / (1 - \beta_k)$ . □

### 2.4 Permutations and polynomials over finite fields

The connection between random logarithmic sets, random permutations and random polynomials is more straightforward, owing to the well-known approximations of these objects by a vector of Poisson random variables.

For each  $j$ , let  $Z_j$  be a Poisson random variable with parameter  $1/j$ , and such that  $Z_1, Z_2, \dots$ , are independent. The next proposition states that, apart from the very longest cycles, the cycle lengths of a random permutation have a joint Poisson distribution.

**Lemma 2.3** *For a random permutation  $\sigma \in S_n$ , let  $C_j(\sigma)$  denote the number of cycles in  $\sigma$  of length  $j$ . Then for  $r = o(n)$  as  $n \rightarrow \infty$  we have*

$$d_{\text{TV}}\left(\left(C_1(\sigma), \dots, C_r(\sigma)\right), \left(Z_1, \dots, Z_r\right)\right) = o(1).$$

*Proof* In fact there is a bound  $\ll e^{-n/r}$  uniformly in  $n$  and  $r$ ; see [3]. □

The next proposition states a similar phenomenon for the degrees of the irreducible factors of a random polynomial over  $\mathbb{F}_q$ , except that now one must also exclude the very smallest degrees as well.

**Lemma 2.4** *Let  $q$  be a prime power. Let  $f$  be a random, monic polynomial in  $\mathbb{F}_q[t]$  of degree  $n$ . Let  $Y_d(f)$  denote the number of monic, irreducible factors of  $f$  which have degree  $d$ . Suppose that  $10 \log n \leq r \leq s \leq \frac{n}{10 \log n}$ . Then*

$$d_{\text{TV}}\left(\left(Y_r(f), \dots, Y_s(f)\right), \left(Z_r, \dots, Z_s\right)\right) = o(1)$$

as  $n \rightarrow \infty$ .

*Proof* For  $r \leq i \leq s$ , let  $\hat{Z}_i$  be a negative binomial random variable<sup>4</sup>  $\text{NB}(\frac{1}{i} \sum_{j|i} \mu(i/j)q^j, q^{-i})$ . Corollary 3.3 in [2] implies that

$$d_{\text{TV}}\left((Y_r(f), \dots, Y_s(f)), (\hat{Z}_r, \dots, \hat{Z}_s)\right) \ll 1/n \tag{2.6}$$

uniformly in  $q, n, r, s$  as in the statement of the lemma. Note that

$$\frac{1}{i} \sum_{j|i} \mu(i/j)q^j = \frac{1}{i}q^i(1 + O(q^{-i/2})) = \frac{1}{i}q^i(1 + O(1/n))$$

for  $i \geq r \geq 10 \log n$ . A routine if slightly lengthy calculation with (2.5) gives

$$d_{\text{TV}}(Z_i, \hat{Z}_i) \ll 1/n.$$

Combining this with (2.4), we arrive at

$$d_{\text{TV}}((Z_r, \dots, Z_s), (\hat{Z}_r, \dots, \hat{Z}_s)) \ll s/n = o(1).$$

The conclusion follows from this, (2.6) and the triangle inequality. □

*Proof of Theorem 4* Fix  $\varepsilon > 0$ , let  $n$  be large enough in terms of  $\varepsilon$ , let  $u = \log n$  and  $v = n/\log n$ . For a random permutation  $\sigma \in S_n$ , let  $\mathbf{C} = \{j : C_j(\sigma) \geq 1\}$ , and define the random set  $\tilde{\mathbf{A}} = \{j : Z_j \geq 1\}$ . As in the proof of Lemma 2.2, (2.4) and (2.5) imply that

$$d_{\text{TV}}(\mathbf{A} \cap (u, v], \tilde{\mathbf{A}} \cap (u, v]) \ll \sum_{u < i \leq v} \frac{1}{i^2} \ll \frac{1}{u}.$$

Lemma 2.3 implies that

$$d_{\text{TV}}(\tilde{\mathbf{A}} \cap (u, v], \mathbf{C} \cap (u, v]) = o(1) \quad (n \rightarrow \infty).$$

Hence,

$$\begin{aligned} d_{\text{TV}}(\mathbf{A} \cap (u, v], \mathbf{C} \cap (u, v]) &\leq d_{\text{TV}}(\mathbf{A} \cap (u, v], \tilde{\mathbf{A}} \cap (u, v]) \\ &\quad + d_{\text{TV}}(\tilde{\mathbf{A}} \cap (u, v], \mathbf{C} \cap (u, v]) \\ &= o(1) \end{aligned}$$

---

<sup>4</sup> We say that the random variable  $X$  has the distribution  $\text{NB}(r, p)$  with  $r \in \mathbb{N}$  and  $p \in (0, 1]$  if  $X$  takes values in  $\mathbb{Z}_{\geq 0}$  with the following frequency:  $\mathbb{P}(X = k) = \binom{k+r-1}{r-1}(1-p)^k p^r$  for each  $k \in \mathbb{Z}_{\geq 0}$ .

as  $n \rightarrow \infty$ . By Lemma 2.1, with probability  $\rightarrow 1$  as  $n \rightarrow \infty$ ,  $\mathbf{A} \cap (u, v]$  has  $M$  distinct subsets  $A_1, \dots, A_M$  with equal sums, where  $M = \lceil (\log n)^{\zeta + \varepsilon} \rceil$ . Hence,  $\mathbf{C}$  has distinct subsets  $S_1, \dots, S_M$  with equal sums with probability  $\rightarrow 1$  as  $n \rightarrow \infty$ . Each subset  $S_j$  corresponds to a distinct divisor of  $\sigma$ , the size of the divisor being the sum of elements of  $S_j$ .  $\square$

*Proof of Theorem 5* The proof is essentially the same as that of Theorem 4, except now we take  $u = 10 \log n$ ,  $v = \frac{n}{10 \log n}$ ,  $\mathbf{C} = \{j : Y_j(f) \geq 1\}$  and use Lemma 2.4 in place of Lemma 2.3.  $\square$

### 3 Overview of the paper

The purpose of this section is to explain the main ideas that go into the proof of Theorem 2 in broad strokes, as well as to outline the structure of the rest of the paper. The remainder of the paper splits into three parts, and we devote a subsection to each of these. Finally, in Sect. 3.4, we make some brief comments about the relationship of our work to previous work of Maier and Tenenbaum [20, 22]. Further comments on this connection are made in Appendix C.

#### 3.1 Part II: equal sums and the optimization problem

Part II provides a very close link between the key quantity  $\beta_k$  (which is defined in Problem 1 and appears in all four of Theorems 2, 3, 4 and 5) and a quantity  $\gamma_k$ , which on the face of it appears to be of a completely different nature, being the solution to a certain optimization problem (Problem 3.7 below) involving the manner in which linear subspaces of  $\mathbb{Q}^k$  intersect the cube  $\{0, 1\}^k$ .

At the heart of this connection is a fairly simple way of associating a *flag* to  $k$  distinct sets  $A_1, \dots, A_k \subset A$ , where  $A$  is a given set of integers (that we typically generate logarithmically).

**Definition 3.1 (Flags)** Let  $k \in \mathbb{N}$ . By an  $r$ -step *flag* we mean a nested sequence

$$\mathcal{V} : \langle \mathbf{1} \rangle = V_0 \leq V_1 \leq V_2 \leq \dots \leq V_r \leq \mathbb{Q}^k$$

of vector spaces.<sup>5</sup> Here  $\mathbf{1} = (1, 1, \dots, 1) \in \mathbb{Q}^k$ . A flag is *complete* if  $\dim V_{i+1} = \dim V_i + 1$  for  $i = 0, 1, \dots, r - 1$ .

To each choice of distinct sets  $A_1, \dots, A_k \subset A$ , we associate a flag as follows. The Venn diagram of the subsets  $A_1, \dots, A_k$  produces a natural

---

<sup>5</sup> In the literature, the term “flag” means that the inclusions are proper, i.e.,  $\dim(V_{i+1}) > \dim V_i$  for all  $i$ . In this paper, we will use the term more broadly to refer to an arbitrary nested sequence of subspaces.

partition of  $A$  into  $2^k$  subsets, which we denote by  $B_\omega$  for  $\omega \in \{0, 1\}^k$ . Here  $A_i = \sqcup_{\omega:\omega_i=1} B_\omega$ . We iteratively select vectors  $\omega^1, \dots, \omega^r$  to maximize  $\prod_{j=1}^r (\max B_{\omega^j})$  subject to the constraint that  $\mathbf{1}, \omega^1, \dots, \omega^r$  are linearly independent over  $\mathbb{Q}$ . We then define<sup>6</sup>  $V_j = \text{Span}(\mathbf{1}, \omega^1, \dots, \omega^j)$  for  $j = 0, 1, \dots, r$ .

The purpose of making this construction is difficult to describe precisely in a short paragraph. However, the basic idea is that the vectors  $\omega^1, \dots, \omega^r$  and the flag  $\mathcal{V}$  provide a natural frame of reference for studying the equal sums equation

$$\sum_{a \in A_1} a = \dots = \sum_{a \in A_k} a. \tag{3.1}$$

Suppose now that  $A_1, \dots, A_k \subset [D^c, D]$ . Then the construction just described naturally leads, in addition to the flag  $\mathcal{V}$ , to the following further data: thresholds  $c_j$  defined by  $\max B_{\omega^j} \approx D^{c_j}$ , and measures  $\mu_j$  on  $\{0, 1\}^k$ , which capture the relative sizes of the sets  $B_\omega \cap (D^{c_{j+1}}, D^{c_j}]$ ,  $\omega \in \{0, 1\}^k$ . Full details of these constructions are given in Sect. 4.

The above discussion motivates the following definition, which will be an important one in our paper.

**Definition 3.2** (*Systems*) Let  $(\mathcal{V}, \mathbf{c}, \boldsymbol{\mu})$  be a triple such that:

- (a)  $\mathcal{V}$  is an  $r$ -step flag whose members  $V_j$  are distinct and spanned by elements of  $\{0, 1\}^k$ ;
- (b)  $\mathcal{V}$  is *nondegenerate*, which means that  $V_r$  is not contained in any of the subspaces  $\{x \in \mathbb{Q}^k : x_i = x_j\}, i \neq j$ ;
- (c)  $\mathbf{c} = (c_1, \dots, c_r, c_{r+1})$  with  $1 \geq c_1 \geq \dots \geq c_{r+1} \geq 0$ ;
- (d)  $\boldsymbol{\mu} = (\mu_1, \dots, \mu_r)$  is an  $r$ -tuple of probability measures;
- (e)  $\text{Supp}(\mu_i) \subset V_i \cap \{0, 1\}^k$  for all  $i$ .

Then we say that  $(\mathcal{V}, \mathbf{c}, \boldsymbol{\mu})$  is a *system*. We say that a system is complete if its underlying flag is, in the sense of Definition 3.1.

*Remark* The nondegeneracy condition (b) arises naturally from the construction described previously, provided one assumes the sets  $A_1, \dots, A_k$  are distinct.

We have sketched how a system  $(\mathcal{V}, \mathbf{c}, \boldsymbol{\mu})$  may be associated to any  $k$  distinct sets  $A_1, \dots, A_k \subset [D^c, D]$ . Full details are given in Sect. 4.1. There is certainly no canonical way to reverse this and associate sets  $A_i$  to a system  $(\mathcal{V}, \mathbf{c}, \boldsymbol{\mu})$ , even if the numbers  $\mu_j(\omega)$  are all rational. However, given a set

<sup>6</sup> Here and throughout the paper,  $\text{Span}(v_1, \dots)$  denotes the  $\mathbb{Q}$ -span of vectors  $v_1, \dots$



$\mathbf{A} \subset [D^c, D]$  (which, in our paper, will be a logarithmic random set) and a system  $(\mathcal{V}, \mathbf{c}, \boldsymbol{\mu})$ , there is a natural *probabilistic* way to construct subsets  $A_1, \dots, A_k \subset \mathbf{A}$  via their Venn diagram  $(B_\omega)_{\omega \in \{0,1\}^k}$ : if  $a \in \mathbf{A} \cap (D^{c_{j+1}}, D^{c_j}]$  then we put  $a$  in  $B_\omega$  with probability  $\mu_j(\omega)$ , these choices being independent for different  $a$ s.

This will be indeed be roughly our strategy for constructing, given a logarithmic random set  $\mathbf{A} \subset [D^c, D]$ , distinct subsets  $A_1, \dots, A_k \subset \mathbf{A} \cap [D^c, D]$  satisfying the equal sums condition (3.1). Very broadly speaking, we will enact this plan in two stages, described in Sects. 5 and 6 respectively. In Sect. 5, which is by far the deeper part of the argument, we will show that (almost surely in  $\mathbf{A}$ ) the distribution of tuples  $(\sum_{a \in A_i} a)_{i=1}^k$  is dense in a certain box adapted to the flag  $\mathcal{V}$ , as the  $A_i$  range over the random choices just described. Then, in Sect. 6, we will show that (almost surely) one of these tuples can be “corrected” to give the equal sums condition (3.1). This general mode of argument has its genesis in the paper [20] of Maier and Tenenbaum, but the details here will look very different. In addition to the fact that linear algebra and entropy play no role in Maier and Tenenbaum’s work, they use a second moment argument which does not work in our setting. Instead we use an  $l^p$  estimate with  $p \approx 1$ , building on ideas in [17, 18].

In analysing the distribution of tuples  $(\sum_{a \in A_i} a)_{i=1}^k$ , the notion of entropy comes to the fore.

**Definition 3.3** (*Entropy of a subspace*) Suppose that  $\nu$  is a finitely supported probability measure on  $\mathbb{Q}^k$  and that  $W \leq \mathbb{Q}^k$  is a vector subspace. Then we define

$$\mathbb{H}_\nu(W) := - \sum_x \nu(x) \log \nu(W + x).$$

*Remark* This the (Shannon) entropy of the distribution on cosets  $W+x$  induced by  $\nu$ . Entropy will play a key role in our paper, and basic definitions and properties of it are collected in Appendix B.

More important than the entropy itself will be a certain quantity  $e(\mathcal{V}', \mathbf{c}, \boldsymbol{\mu})$ , assigned to *subflags* of  $\mathcal{V}$ . We give the relevant definitions now.

**Definition 3.4** (*Subflags*) Suppose that

$$\mathcal{V}' : \langle \mathbf{1} \rangle = V_0 \leq V_1 \leq V_2 \leq \dots \leq V_r \leq \mathbb{Q}^k$$

is a flag. Then another flag

$$\mathcal{V}' : \langle \mathbf{1} \rangle = V'_0 \leq V'_1 \leq V'_2 \leq \dots \leq V'_r \leq \mathbb{Q}^k$$

is said to be a *subflag* of  $\mathcal{V}$  if  $V'_i \leq V_i$  for all  $i$ . In this case we write  $\mathcal{V}' \leq \mathcal{V}$ . It is a *proper subflag* if it is not equal to  $\mathcal{V}$ .

**Definition 3.5** (*e-value*) Let  $(\mathcal{V}, \mathbf{c}, \boldsymbol{\mu})$  be a system, and let  $\mathcal{V}' \leq \mathcal{V}$  be a subflag. Then we define the *e-value*

$$e(\mathcal{V}', \mathbf{c}, \boldsymbol{\mu}) := \sum_{j=1}^r (c_j - c_{j+1}) \mathbb{H}_{\mu_j}(V'_j) + \sum_{j=1}^r c_j \dim(V'_j/V'_{j-1}). \quad (3.2)$$

*Remark* Note that

$$e(\mathcal{V}, \mathbf{c}, \boldsymbol{\mu}) = \sum_{j=1}^r c_j \dim(V_j/V_{j-1}), \quad (3.3)$$

since condition (e) of Definition 3.2 implies that  $\mathbb{H}_{\mu_j}(V_j) = 0$  for  $1 \leq j \leq r$ .

**Definition 3.6** (*Entropy condition*) Let  $(\mathcal{V}, \mathbf{c}, \boldsymbol{\mu})$  be a system. We say that this system satisfies the *entropy condition* if

$$e(\mathcal{V}', \mathbf{c}, \boldsymbol{\mu}) \geq e(\mathcal{V}, \mathbf{c}, \boldsymbol{\mu}) \quad \text{for all subflags } \mathcal{V}' \text{ of } \mathcal{V}, \quad (3.4)$$

and the *strict entropy condition* if

$$e(\mathcal{V}', \mathbf{c}, \boldsymbol{\mu}) > e(\mathcal{V}, \mathbf{c}, \boldsymbol{\mu}) \quad \text{for all proper subflags } \mathcal{V}' \text{ of } \mathcal{V}. \quad (3.5)$$

We cannot give a meaningful discussion of exactly why these definitions are the right ones to make in this overview. Indeed, it took the authors over a year of working on the problem to arrive at them. Let us merely say that

- If a random logarithmic set  $\mathbf{A} \cap [D^c, D]$  almost surely admits distinct subsets  $A_1, \dots, A_k$  satisfying the equal sums condition (3.1), then some associated system  $(\mathcal{V}, \mathbf{c}, \boldsymbol{\mu})$  satisfies the entropy condition (3.4). For detailed statements and proofs, see Sect. 4.
- If a system  $(\mathcal{V}, \mathbf{c}, \boldsymbol{\mu})$  satisfies the strict entropy condition (3.5) then the details of the construction of sets  $A_1, \dots, A_k$  satisfying the equal sums condition, outlined above, can be made to work. For detailed statements and proofs, see Sects. 5 and 6.

With the above definitions and discussion in place, we are finally ready to introduce the key optimization problem, the study of which will occupy a large part of our paper.

**Problem 3.7** (*The optimisation problem*) Determine the value of  $\gamma_k$ , defined to be the supremum of all constants  $c$  for which there is a system  $(\mathcal{V}, \mathbf{c}, \boldsymbol{\mu})$  such that  $c_{r+1} = c$  and the entropy condition (3.4) holds.

Similarly, determine  $\tilde{\gamma}_k$ , defined to be the supremum of all constants  $c$  for which there is a system  $(\mathcal{V}, \mathbf{c}, \boldsymbol{\mu})$  such that  $c_{r+1} = c$  and the strict entropy condition (3.5) holds.

The precise content of the two bullet points above, and the main result of Part II of the paper, is then the following theorem.

**Theorem 7** *For every  $k \geq 2$ , we have*

$$\tilde{\gamma}_k \leq \beta_k \leq \gamma_k.$$

*Remark 3.1* (a) Presumably  $\gamma_k = \beta_k = \tilde{\gamma}_k$ . Indeed, it is natural to think that any system satisfying (3.4) can be perturbed an arbitrarily small amount to satisfy (3.5). However, we have not been able to show that this is possible in general.

(b) It is not a priori clear that  $\gamma_k$  and  $\tilde{\gamma}_k$  exist and are positive. This will follow, e.g., from our work on “binary systems” in part IV of the paper, although there is an easier way to see this using the original Maier–Tenenbaum argument, adapted to our setting; see Appendix C for a sketch of the details.

### 3.2 Part III: the optimization problem

Part III of the paper is devoted to the study of Problem 3.7 in as much generality as we can manage. Unfortunately we have not yet been able to completely resolve this problem, and indeed numerical experiments suggest that a complete solution, for all  $k$ , could be very complicated.

The main achievement of Part III is to provide a solution of sorts when the flag  $\mathcal{V}$  is fixed, but one is free to choose  $\mathbf{c}$  and  $\boldsymbol{\mu}$ . Write  $\gamma_k(\mathcal{V})$  (or  $\tilde{\gamma}_k(\mathcal{V})$ ) for the solution to this problem, that is, the supremum of values  $c = c_{r+1} \geq 0$  for which a system  $(\mathcal{V}, \mathbf{c}, \boldsymbol{\mu})$  exists satisfying (3.4) (or (3.5)).

Our solution applies only to rather special flags  $\mathcal{V}$ , but this is unsurprising: for “generic” flags  $\mathcal{V}$ , one would not expect there to be any choice of  $\mathbf{c}, \boldsymbol{\mu}$ , for which  $c_{r+1} > 0$ , and so  $\gamma_k(\mathcal{V}) = 0$  in these cases. Such flags are of no interest in this paper.

We begin, in Sect. 7, by solving an even more specific problem in which the entropy condition (3.4) is only required to hold for certain very special subflags  $\mathcal{V}'$  of  $\mathcal{V}$ , which we call *basic flags*. These are flags of the form

$$\mathcal{V}'_{\text{basic}(m)} : \langle \mathbf{1} \rangle = V_0 \leq V_1 \leq \dots \leq V_{m-1} \leq V_m = V_m = \dots = V_m.$$

We call this the *restricted entropy condition*; to spell it out, this is the condition that

$$e(\mathcal{V}'_{\text{basic}(m)}, \mathbf{c}, \boldsymbol{\mu}) \geq e(\mathcal{V}, \mathbf{c}, \boldsymbol{\mu}) \quad (3.6)$$

for  $m = 0, 1, \dots, r - 1$  (the case  $m = r$  being vacuous).

We write  $\gamma_k^{\text{res}}(\mathcal{V})$  for the maximum value of  $c_{r+1}$  (over all choices of  $\mathbf{c}$  and  $\boldsymbol{\mu}$  such that  $(\mathcal{V}, \mathbf{c}, \boldsymbol{\mu})$  is a system) subject to this condition. Clearly

$$\gamma_k^{\text{res}}(\mathcal{V}) \geq \gamma_k(\mathcal{V}). \quad (3.7)$$

The main result of Sect. 7 is Proposition 7.7, which states that under certain conditions we have

$$\gamma_k^{\text{res}}(\mathcal{V}) = \frac{\log 3 - 1}{\log 3 + \sum_{i=1}^{r-1} \frac{\dim(V_{i+1}/V_i)}{\rho_1 \cdots \rho_{r-1}}}, \quad (3.8)$$

for certain parameters  $\rho_1, \dots, \rho_{r-1}$  depending on the flag  $\mathcal{V}$ .

To define these, one considers the “tree structure” on  $\{0, 1\}^k \cap V_r$  induced by the flag  $\mathcal{V}$ : the “cells at level  $j$ ” are simply intersections with cosets of  $V_j$ , and we join a cell  $C$  at level  $j$  to a “child” cell  $C'$  at level  $j - 1$  iff  $C' \subset C$ . The  $\rho_i$  are then defined by setting up a certain recursively-defined function on this tree and then solving what we term the  $\rho$ -equations. The details may be found in Sect. 7.2. Proposition 7.7 also describes the measures  $\boldsymbol{\mu}$  and the parameters  $\mathbf{c}$  for which this optimal value is attained.

In Sect. 8, we relate the restricted optimisation problem to the real one, giving fairly general conditions under which we in fact have equality in (3.7), that is to say  $\gamma_k^{\text{res}}(\mathcal{V}) = \gamma_k(\mathcal{V})$ . The basic strategy of this section is to show that for the  $\mathbf{c}$  and  $\boldsymbol{\mu}$  which are optimal for the restricted optimisation problem, the full entropy condition (3.4) is in fact a consequence of the restricted condition (3.6).

The arguments of this section make heavy use of the submodularity inequality for entropy, using this to drive a kind of “symmetrisation” argument. In this way one can show that an arbitrary  $e(\mathcal{V}', \mathbf{c}, \boldsymbol{\mu})$  is greater than or equal to one in which  $\mathcal{V}'$  is *almost* a basic flag; these “semi-basic” flags are then dealt with by hand.

To add an additional layer of complexity, we build a perturbative device into this argument so that our results also apply to  $\tilde{\gamma}_k(\mathcal{V})$ .

### 3.3 Part IV: binary systems

The final part of the paper is devoted to a discussion of a particular type of flag  $\mathcal{V}$ , the *binary flags*, and the associated optimal systems  $(\mathcal{V}, \mathbf{c}, \boldsymbol{\mu})$ , which we call *binary systems*.

**Definition 3.8** (*Binary flag of order  $r$* ) Let  $k = 2^r$  be a power of two. Identify  $\mathbb{Q}^k$  with  $\mathbb{Q}^{\mathcal{P}[r]}$  (where  $\mathcal{P}[r]$  means the power set of  $[r] = \{1, \dots, r\}$ ) and define an  $r$ -step flag  $\mathcal{V}$ ,  $\langle \mathbf{1} \rangle = V_0 \leq V_1 \leq \dots \leq V_r = \mathbb{Q}^{\mathcal{P}[r]}$ , as follows:  $V_i$  is the subspace of all  $(x_S)_{S \subset [r]}$  for which  $x_S = x_{S \cap [i]}$  for all  $S \subset [r]$ .

Whilst the definition is, in hindsight, rather simple and symmetric, it was motivated by extensive numerical experiment. We believe these flags to be asymptotically optimal for Problem 3.7, though we currently lack a proof.

There are two main tasks in Part IV. First, we must verify that the various conditions necessary for the results of Part III hold for the binary flags. This is accomplished in Sect. 10, the main statements being given in Sect. 9. At the end of Sect. 9 we give the proof (and complete statement) of Theorem 2(a), conditional upon the results of Sect. 10. This is the deepest result in the paper.

Following this we turn to Theorem 2(b). There are two tasks here. First, we prove that the parameters  $\rho_i$  for the binary flags (which do not depend on  $r$ ) tend to a limit  $\rho$ . This is not at all straightforward, and is accomplished in Sect. 11.

After that, in Sect. 12, we describe this limit in terms of certain recurrence relations, which also provide a useful means of calculating it numerically. Theorem 2(b) is established at the very end of the paper.

Most of Part IV could, if desired, be read independently of the rest of the paper.

### 3.4 Relation to previous work

Previous lower bounds for the a.s. behaviour of  $\Delta$  are contained in two papers of Maier and Tenenbaum [20,22]. Both of these bounds can be understood within the framework of our paper.

The main result of [20] follows from the fact that

$$\tilde{\gamma}_2 \geq 1 - \frac{1}{\log 3}. \tag{3.9}$$

Indeed by Theorem 7 it then follows that  $\beta_2 \geq 1 - \frac{1}{\log 3}$ , and then from Theorem 3 it follows that for almost every  $n$  we have

$$\Delta(n) \gg (\log \log n)^{-\log 2 / \log(1 - \frac{1}{\log 3}) + o(1)}. \tag{3.10}$$

The exponent appearing here is  $0.28754048957\dots$  and is exactly the one in [20, Theorem 2].

The bound (3.9) is very easy to establish, and a useful exercise in clarifying the notation we have set up. Take  $k = 2, r = 1$  and let  $\mathcal{V}$  be the flag

$\langle \mathbf{1} \rangle = V_0 \leq V_1 = \mathbb{Q}^2$ . Let  $\mathbf{c} = (c_1, c_2)$  with  $c_1 = 1$  and

$$c_2 < 1 - \frac{1}{\log 3}. \tag{3.11}$$

Let  $\mu_1$  be the measure which assigns weight  $\frac{1}{3}$  to the points  $\mathbf{0} = (0, 0)$ ,  $(0, 1)$  and  $(1, 0)$  in  $\{0, 1\}^2$  (this being a pullback of the uniform measure on  $\{0, 1\}^2 / V_0$ ).

There are only two subflags  $\mathcal{V}'$  of  $\mathcal{V}$ , namely  $\mathcal{V}$  itself and the basic flag  $\mathcal{V}'_{\text{basic}(\mathbf{0})} : \langle \mathbf{1} \rangle = V'_0 \leq V'_1$  with  $V'_0 = V'_1 = V_0 = \langle \mathbf{1} \rangle$ . The entire content of the strict entropy condition (3.5) is therefore that

$$e(\mathcal{V}'_{\text{basic}(\mathbf{0})}, \mathbf{c}, \mu) > e(\mathcal{V}, \mathbf{c}, \mu),$$

which translates to

$$(c_1 - c_2)\mathbb{H}_{\mu_1}(V_0) > c_1.$$

We have  $\mathbb{H}_{\mu_1}(V_0) = \log 3$  and  $c_1 = 1$ , and so this translates to precisely condition (3.11).

*Remark* (a) With very little more effort (appealing to Lemma B.2) one can show that  $\gamma_2 = \beta_2 = \tilde{\gamma}_2 = 1 - \frac{1}{\log 3}$ .

(b) This certainly does not provide a shorter proof of Theorem 3.10 than the one Maier and Tenenbaum gave, since our deductions are reliant on the material in Sects. 5 and 6, which constitute a significant elaboration of the ideas from [20].

The main result of [22] (Theorem 1.4 there) follows from the lower bound

$$\tilde{\gamma}_{2r} \geq \left(1 - \frac{1}{\log 3}\right) \left(\frac{1 - 1/\log 3}{1 - 1/\log 27}\right)^{r-1}, \tag{3.12}$$

which of course includes (3.9) as the special case  $r = 1$ . Applying Theorem 7 and Theorem 3, then letting  $r \rightarrow \infty$ , we recover [22, Theorem 1.4] (quoted as Theorem MT in Sect. 1), namely the bound

$$\Delta(n) \geq (\log \log n)^{\frac{\log 2}{\log\left(\frac{1-1/\log 27}{1-1/\log 3}\right)} - o(1)}$$

for almost all  $n$ . The exponent here is 0.33827824168 . . .

To explain how (3.12) may be seen within our framework requires a little more setting up. Since it is not directly relevant to our main arguments, we defer this to Appendix C.

**Part II. Equal sums and the optimisation problem**

**4 The upper bound  $\beta_k \leq \gamma_k$**

In this section we establish the bound in the title. We recall the definitions of  $\beta_k$  (Problem 1) and  $\gamma_k$  (Problem 3.7). We will in fact show a bit more, that if  $c > \gamma_k$  then

$$\mathbb{P}(\text{there are distinct } A_1, \dots, A_k \in [D^c, D] \text{ with equal sums}) \rightarrow 0 \tag{4.1}$$

as  $D \rightarrow \infty$ .

**4.1 Venn diagrams and linear algebra**

Let  $0 < c < 1$  be some fixed quantity, and let  $D$  be a real number, large in terms of  $c$ . Suppose that  $A_1, \dots, A_k \subset [D^c, D]$  are distinct sets. In this section we show that there is a rather natural way to associate a complete system  $(\mathcal{V}, \mathbf{c}, \boldsymbol{\mu})$  (in the sense of Definition 3.2) to these sets. This system encodes the “linear algebra of the Venn diagram of the  $A_i$ ” in a way that turns out to be extremely useful.

The Venn diagram of the  $A_i$  has  $2^k$  cells, indexed by  $\{0, 1\}^k$  in a natural way. Thus for each  $\omega = (\omega_1, \dots, \omega_k) \in \{0, 1\}^k$ , we define

$$B_\omega := \bigcap_{i:\omega_i=1} A_i \bigcap_{i:\omega_i=0} (A_i)^c, \tag{4.2}$$

*The flag  $\mathcal{V}$ .* Set  $\Omega := \{\omega : B_\omega \neq \emptyset\}$ . We may put a total order  $<$  on  $\Omega$  by writing  $\omega' < \omega$  if and only if  $\max B_{\omega'} < \max B_\omega$ . We now select  $r$  special vectors  $\omega^1, \dots, \omega^r \in \Omega$ , with  $r \leq k - 1$ , in the following manner. Let  $\omega^1 = \max_{<}(\Omega \setminus \{\mathbf{0}, \mathbf{1}\})$ . Assuming we have chosen  $\omega^1, \dots, \omega^j$  such that  $\mathbf{1}, \omega^1, \dots, \omega^j$  are linearly independent over  $\mathbb{Q}$ , let  $\omega^{j+1} = \max(\Omega \setminus \text{Span}(\mathbf{1}, \omega^1, \dots, \omega^j))$ , as long as such a vector exists.

Let  $\mathbf{1}, \omega^1, \dots, \omega^r$  be the set of vectors produced when this algorithm terminates. By construction,  $\Omega \subset \text{Span}(\mathbf{1}, \omega^1, \dots, \omega^r)$ , or in other words  $B_\omega = \emptyset$  whenever  $\omega \in \{0, 1\}^k \setminus \text{Span}(\mathbf{1}, \omega^1, \dots, \omega^r)$ .

Now define an  $r$ -step flag  $\mathcal{V} : \langle \mathbf{1} \rangle = V_0 < V_1 < \dots < V_r$  by setting  $V_j := \text{Span}(\mathbf{1}, \omega^1, \dots, \omega^j)$  for  $1 \leq j \leq r$ .

*The parameters  $\mathbf{c}$ .* Now we construct the parameters  $\mathbf{c} : 1 \geq c_1 \geq c_2 \geq \dots \geq c_{r+1}$ . For  $j = 1, \dots, r$ , we define

$$c_j = 1 + \frac{\lceil \log \max B_{\omega^j} - \log D \rceil}{\log D}. \tag{4.3}$$

Thus

$$\max B_{\omega^j} \in \left(\frac{1}{e}D^{c_j}, D^{c_j}\right] \tag{4.4}$$

for  $j = 1, \dots, r$ . Also set  $c_{r+1} = c$ . (The ceiling function  $\lceil \cdot \rceil$  produces a ‘‘coarse’’ or discretised set of possible thresholds  $c_i$ , suitable for use in a union bound later on; see Lemma 4.2 below. The offset of  $-\log D$  is to ensure that  $c_1 \leq 1$ .)

The measures  $\mu$ . Set

$$B'_\omega := \begin{cases} B_\omega \setminus \{\max B_{\omega^j}\} & \text{if } \omega = \omega^j \text{ for some } j, \\ B_\omega & \text{otherwise.} \end{cases} \tag{4.5}$$

Define

$$\mu_j(\omega) := \frac{\#(B'_\omega \cap (D^{c_{j+1}}, D^{c_j}))}{\sum_{\omega'} \#(B'_\omega \cap (D^{c_{j+1}}, D^{c_j}))}, \tag{4.6}$$

with the convention that if the denominator vanishes, then  $\mu_j(\omega) = 1_{\omega=0}$ .

*Remark* It is important that we use the  $B'_\omega$  here, rather than the  $B_\omega$ , for technical reasons that will become apparent in the proof of Proposition 4.4 below.

**Lemma 4.1**  $(\mathcal{V}, \mathbf{c}, \mu)$  is a complete system (in the sense of Definition 3.2).

*Proof* We need to check that  $\text{Supp}(\mu_j) \subset V_j$  for  $j = 1, \dots, r$ . By definition, if  $\mu_j(\omega) > 0$  then  $B_\omega \cap (D^{c_{j+1}}, D^{c_j}] \neq \emptyset$ . This implies that  $\max B_\omega > D^{c_{j+1}}$ . On the other hand, (4.4) implies that  $D^{c_{j+1}} \geq \max B_{\omega^{j+1}}$ , and thus  $\max B_\omega > \max B_{\omega^{j+1}}$ . By the construction of the vectors  $\omega^i$ , we must have  $\omega \in \text{Span}(\mathbf{1}, \omega^1, \dots, \omega^j) = V_j$ .

We also need to check that  $\mathcal{V}$  is nondegenerate, also in the sense of Definition 3.2, that is to say  $V_r$  is not contained in any hyperplane  $\{\omega \in \mathbb{Q}^k : \omega_i = \omega_j\}$ . This follows immediately from the fact that the  $A_i$  are distinct. Since

$$A_i \Delta A_j = \bigcup_{\substack{\omega \in \{0,1\}^k \\ \omega_i \neq \omega_j}} B_\omega,$$

and so there is certainly some  $\omega$  with  $\omega_i \neq \omega_j$  and  $B_\omega \neq \emptyset$ . □

Note that, in addition to the system  $(\mathcal{V}, \mathbf{c}, \mu)$ , the procedure described above outputs a sequence  $\omega^1, \dots, \omega^r$  of elements of  $\{0, 1\}^k$ . We call the ensemble



consisting of the system and the  $\omega^i$  the linear data associated to  $A_1, \dots, A_k$ . We will only consider the event  $\mathbf{A} \in \mathcal{E}$ , where

$$\mathcal{E} := \left\{ A \subseteq [D^c, D] : \left| \#(A \cap (D^\alpha, D^\beta]) - (\beta - \alpha) \log D \right| \leq \log^{3/4} D \quad (c \leq \alpha \leq \beta \leq 1) \right\}. \tag{4.7}$$

By Lemma A.5,  $\mathbb{P}(\mathbf{A} \in \mathcal{E}) = 1 - o(1)$  as  $D \rightarrow \infty$ . In particular, if  $A \in \mathcal{E}$ , we have  $|A \cap [D^c, D]| \leq 2 \log D$  for large enough  $D$ .

**Lemma 4.2** Fix  $k \in \mathbb{Z}_{\geq 2}$  and suppose that  $A \in \mathcal{E}$ . The number of different ensembles of linear data arising from distinct sets  $A_1, \dots, A_k \subset A$  is  $\ll (\log D)^{O(1)}$ .

*Proof* The number of choices for  $\omega^1, \dots, \omega^r$  is  $O(1)$ , and hence the number of  $\mathcal{V}$  is also  $O_k(1)$ . The thresholds  $c_j$  are drawn from a fixed set of size  $\log D$ , and the numerators and denominators of the  $\mu_j(\omega)$  are all integers  $\leq 2 \log D$ .  $\square$

*Remark 4.1* The  $O(1)$  and the  $\ll$  here both depend on  $k$ . However we regard  $k$  as fixed here and do not indicate this dependence explicitly. If one is more careful then one can obtain results that are effective up to about  $k \sim \log \log D$ .

### 4.2 A local-to-global estimate

Our next step towards establishing the bound  $\beta_k \leq \gamma_k$  is to pass from the “local” event that a random logarithmic set  $\mathbf{A}$  possesses a  $k$ -tuple of equal subsums  $(\sum_{a \in A_1} a, \dots, \sum_{a \in A_k} a)$  to the “global” distribution of such subsums (with the subtlety that we must mod out by  $\mathbf{1}$ ). The latter is controlled by the set  $\mathcal{L}^{\mathcal{V}, \mathbf{c}, \boldsymbol{\mu}}(\mathbf{A})$  defined below.

**Definition 4.3** Given a set of integers  $A$  and a system  $(\mathcal{V}, \mathbf{c}, \boldsymbol{\mu})$ , we write  $\mathcal{L}^{\mathcal{V}, \mathbf{c}, \boldsymbol{\mu}}(A)$  for the set of vectors

$$\sum_{\omega \in \{0,1\}^k} \omega \sum_{a \in B_\omega} a \pmod{\mathbf{1}},$$

where  $(B_\omega)_{\omega \in \{0,1\}^k}$  runs over all partitions of  $A$  such that

$$\mu_j(\omega) = \frac{\#(B_\omega \cap (D^{c_{j+1}}, D^{c_j}])}{\#(A \cap (D^{c_{j+1}}, D^{c_j}])} \quad (1 \leq j \leq r, \omega \in \{0,1\}^k). \tag{4.8}$$

**Proposition 4.4** Fix an integer  $k \geq 2$  and a parameter  $0 < c < 1$ . Let  $D$  be large in terms of  $c$  and  $k$ , and let  $\mathbf{A} \subset [D^c, D]$  be a logarithmic random set. Let

$$\begin{aligned} \tilde{\mathcal{E}} = \left\{ A \subseteq [D^c, D] : \left| \#(A \cap (D^\alpha, D^\beta]) - (\beta - \alpha) \log D \right| \right. \\ \left. \leq 2 \log^{3/4} D \quad (c \leq \alpha \leq \beta \leq 1) \right\}. \end{aligned} \tag{4.9}$$

Then we have

$$\begin{aligned} \mathbb{P} \left( \exists \text{ distinct } A_1, \dots, A_k \subseteq \mathbf{A} \text{ such that } \sum_{a \in A_1} a = \dots = \sum_{a \in A_k} a \right) \\ \leq (\log D)^{O(1)} \sup_{(\mathcal{V}, \mathbf{c}, \boldsymbol{\mu})} D^{-(c_1 + \dots + c_r)} \mathbb{E} \mathbf{1}_{\mathbf{A} \in \tilde{\mathcal{E}}} |\mathcal{L}_{\mathcal{V}, \mathbf{c}, \boldsymbol{\mu}}(\mathbf{A})| + \mathbb{P}(\mathcal{E}^c). \end{aligned} \tag{4.10}$$

Here, the supremum is over all complete systems  $(\mathcal{V}, \mathbf{c}, \boldsymbol{\mu})$  with  $c_{r+1} = c$ .

*Proof* Recall the definition of the set  $\mathcal{E}$ , given in Eq. (4.7). We have

$$\begin{aligned} \mathbb{P} \left( \exists \text{ distinct } A_1, \dots, A_k \subseteq \mathbf{A} \text{ such that } \sum_{a \in A_1} a = \dots = \sum_{a \in A_k} a \right) \\ \leq \mathbb{P}(\mathcal{E}^c) + \sum_{\mathcal{V}, \mathbf{c}, \boldsymbol{\mu}, (\omega^i)} \sum_{A \in \mathcal{S}(\mathcal{V}, \mathbf{c}, \boldsymbol{\mu}, (\omega^i))} \mathbb{P}(\mathbf{A} = A), \end{aligned}$$

where, given linear data  $\{(\mathcal{V}, \mathbf{c}, \boldsymbol{\mu}), \omega^1, \dots, \omega^r\}$ , we write  $\mathcal{S}(\mathcal{V}, \mathbf{c}, \boldsymbol{\mu}, (\omega^i))$  to denote the set of all  $A \in \mathcal{E}$  that have  $k$  distinct subsets  $(A_1, \dots, A_k)$  with equal sums-of-elements and associated linear data  $\{(\mathcal{V}, \mathbf{c}, \boldsymbol{\mu}), \omega^1, \dots, \omega^r\}$ . (The set  $\mathbf{A}$  appearing in (4.10) will be constructed below by removing certain elements from the logarithmic set  $\mathbf{A}$  we started with; this new set belongs to  $\tilde{\mathcal{E}}$ , but not necessarily to  $\mathcal{E}$ .)

Let us fix a choice of linear data  $\{(\mathcal{V}, \mathbf{c}, \boldsymbol{\mu}), \omega^1, \dots, \omega^r\}$  and let us abbreviate  $\mathcal{S}$  for the set  $\mathcal{S}(\mathcal{V}, \mathbf{c}, \boldsymbol{\mu}, (\omega^i))$ . An elementary probability calculation gives

$$E(\mathcal{S}) := \sum_{A \in \mathcal{S}} \mathbb{P}(\mathbf{A} = A) = \sum_{A \in \mathcal{S}} \prod_{D^c < a \leq D} \left(1 - \frac{1}{a}\right) \prod_{a \in A} \frac{1}{a-1}. \tag{4.11}$$

For each  $A \in \mathcal{S}$ , fix a choice of  $(A_1, \dots, A_k)$  with equal sums and such that the linear data associated to  $(A_1, \dots, A_k)$  is  $\{(\mathcal{V}, \mathbf{c}, \boldsymbol{\mu}), \omega^1, \dots, \omega^r\}$ . Let  $B_\omega$  be the cells of the Venn diagram corresponding to the  $A_i$ , as in (4.2), and then

define the  $B'_\omega$  as in (4.5). Recall that (4.6) holds, and define  $K_j = \max B_{\omega^j}$  for  $1 \leq j \leq r$ . In particular,  $K_1 > \dots > K_r$ . Let  $A' = A \setminus \{K_1, \dots, K_r\}$ . Note that  $A' \in \tilde{\mathcal{E}}$  if  $D$  is large enough in terms of  $k$ . Moreover, we have

$$\sum_{a \in A_i} a = \sum_{\omega \in \{0,1\}^k} \omega_i \sum_{a \in B_\omega} a.$$

Therefore, the equal sums condition is equivalent to

$$\sum_{\omega \in \{0,1\}^k} \omega \sum_{a \in B_\omega} a = \mathbf{0} \pmod{\mathbf{1}},$$

and hence

$$\sum_{j=1}^r K_j \omega^j = - \sum_{\omega} \omega \sum_{a' \in B'_\omega} a' \pmod{\mathbf{1}}. \tag{4.12}$$

Since  $\mathbf{1}, \omega^1, \dots, \omega^r$  are linearly independent, the value of the right-hand side of (4.12) uniquely determines the numbers  $K_j$ , which themselves uniquely determine  $A$  in terms of the sets  $B'_\omega$ . Therefore, given  $A' \in \tilde{\mathcal{E}}$ , the number of possible sets  $A$  is, by Definition 4.3, at most  $|\mathcal{L}_{\mathcal{V}, \mathbf{c}, \mu}(A')|$ . Moreover by (4.4) we have  $K_j > \frac{1}{e} D^{c_j}$  for every  $j$ , and therefore

$$\prod_{a \in A} \frac{1}{a-1} \ll D^{-(c_1+\dots+c_r)} \prod_{a \in A'} \frac{1}{a-1}. \tag{4.13}$$

We sum over  $A'$ , and reinterpret the product on the right-hand side of (4.13) in terms of  $\mathbb{P}(\mathbf{A} = A')$ . This gives

$$\begin{aligned} E(\mathcal{S}) &\ll D^{-(c_1+\dots+c_r)} \sum_{A' \in \tilde{\mathcal{E}}} |\mathcal{L}_{\mathcal{V}, \mathbf{c}, \mu}(A')| \prod_{D^c < a \leq D} \left(1 - \frac{1}{a}\right) \prod_{a \in A'} \frac{1}{a-1} \\ &= D^{-(c_1+\dots+c_r)} \sum_{A' \in \tilde{\mathcal{E}}} |\mathcal{L}_{\mathcal{V}, \mathbf{c}, \mu}(A')| \cdot \mathbb{P}(\mathbf{A} = A') \\ &= D^{-(c_1+\dots+c_r)} \mathbb{E} \mathbf{1}_{\mathbf{A} \in \tilde{\mathcal{E}}} \cdot |\mathcal{L}_{\mathcal{V}, \mathbf{c}, \mu}(\mathbf{A})|. \end{aligned}$$

By Lemma 4.2 there are  $(\log D)^{O(1)}$  possible choices for the linear data  $\{(\mathcal{V}, \mathbf{c}, \mu), \omega^1, \dots, \omega^r\}$ , and the proof is complete. □

### 4.3 Upper bounds in terms of entropies

Having established Proposition 4.4, we turn to the study of the sets  $\mathcal{L}_{\mathcal{V}, \mathbf{c}, \mu}(A)$ . We will bound their cardinality in terms of the quantities  $e(\mathcal{V}', \mathbf{c}, \mu)$  from Definition 3.2 with  $\mathcal{V}'$  a subflag of  $\mathcal{V}$ .

**Lemma 4.5** *Let  $(\mathcal{V}, \mathbf{c}, \mu)$  be a system and let  $A \in \tilde{\mathcal{E}}$ , where  $\tilde{\mathcal{E}}$  is defined in (4.9). Then, for any subflag  $\mathcal{V}'$  of  $\mathcal{V}$ ,*

$$|\mathcal{L}_{\mathcal{V}, \mathbf{c}, \mu}(A)| \ll_{\mathcal{V}'} e^{O(\log^{3/4} D)} D^{e(\mathcal{V}', \mathbf{c}, \mu)}. \tag{4.14}$$

*Remark* The implied constant in the  $\ll_{\mathcal{V}'}$  could be made explicit if desired (in terms of the quantitative rationality of a basis for the spaces in  $\mathcal{V}'$ ) but we have no need to do this.

*Proof of Lemma 4.5* Given a set  $X \subset [D^c, D]$ , write  $X^{(j)} := X \cap (D^{c_{j+1}}, D^{c_j}]$  for  $j = 1, \dots, r$ . Throughout the proof, we will assume that  $A$  is a set of integers and that  $(B_\omega)_{\omega \in \{0,1\}^k}$  runs over all partitions of  $A$  such that (4.8) is satisfied. In our new notation, this may be rewritten as

$$|B_\omega^{(j)}| = \mu_j(\omega) |A^{(j)}|, \quad j = 1, \dots, r, \quad \omega \in \{0, 1\}^k. \tag{4.15}$$

For each  $j$ ,  $1 \leq j \leq r$ , fix a linear projection  $P_j : V_j \rightarrow V'_j$ , and set  $Q_j := \text{id}_{V_j} - P_j$ , so that  $Q_j$  maps  $V_j$  to itself. Set

$$\mathcal{L}^P(A) := \left\{ \sum_{j=1}^r \sum_{\substack{\omega \in \{0,1\}^k \\ \omega \in V_j}} P_j(\omega) \sum_{a \in B_\omega^{(j)}} a \pmod{\mathbf{1}} : (4.15) \text{ is satisfied} \right\}$$

and

$$\mathcal{L}^Q(A) := \left\{ \sum_{j=1}^r \sum_{\substack{\omega \in \{0,1\}^k \\ \omega \in V_j}} Q_j(\omega) \sum_{a \in B_\omega^{(j)}} a \pmod{\mathbf{1}} : (4.15) \text{ is satisfied} \right\}.$$

Since

$$\sum_{\omega \in \{0,1\}^k} \omega \sum_{a \in B_\omega} a = \sum_{j=1}^r \sum_{\substack{\omega \in \{0,1\}^k \\ \omega \in V_j}} P_j(\omega) \sum_{a \in B_\omega^{(j)}} a + \sum_{j=1}^r \sum_{\substack{\omega \in \{0,1\}^k \\ \omega \in V_j}} Q_j(\omega) \sum_{a \in B_\omega^{(j)}} a,$$

it follows immediately from the definition of  $\mathcal{L}_{\mathcal{V}, \mathbf{c}, \mu}(A)$  (Definition 4.3) that

$$|\mathcal{L}_{\mathcal{V}, \mathbf{c}, \mu}(A)| \leq |\mathcal{L}^P(A)| \cdot |\mathcal{L}^Q(A)|. \tag{4.16}$$

We claim that

$$|\mathcal{L}^P(A)| \ll_{\mathcal{V}'} (\log D)^r D^{\sum_{j=1}^r c_j \dim(V'_j/V'_{j-1})} \tag{4.17}$$

and that

$$|\mathcal{L}^Q(A)| \leq e^{O(\log^{3/4} D)} D^{\sum_{j=1}^r (c_j - c_{j+1}) \mathbb{H}_{\mu_j}(V'_j)}. \tag{4.18}$$

These bounds, substituted into (4.16), immediately imply Lemma 4.5.

It remains to establish (4.17) and (4.18), which are proven in quite different ways. We begin with (4.18), which is a “combinatorial” bound, in that there cannot be too many choices for the data making up the sums in  $\mathcal{L}^Q(A)$ . For this, observe that  $Q_j$  vanishes on  $V'_j$  and hence is constant on cosets of  $V'_j$ . Therefore the elements of  $\mathcal{L}^Q(A)$  are determined by the sets  $\bigcup_{\omega \in v_j + V'_j} B_\omega^{(j)}$ , over all  $v_j \in V_j/V'_j$  and  $1 \leq j \leq r$ . By (4.15),

$$\left| \bigcup_{\omega \in v_j + V'_j} B_\omega^{(j)} \right| = \mu_j(v_j + V'_j) |A^{(j)}|,$$

and by Lemma B.1 the number of ways of partitioning  $A^{(j)}$  into sets of these sizes is bounded above by  $e^{\mathbb{H}(\mathbf{p}^{(j)}) |A^{(j)}|}$ , where  $\mathbf{p}^{(j)} = (\mu_j(v_j + V'_j))_{v_j \in V_j/V'_j}$ . By Definition 3.3,  $\mathbb{H}(\mathbf{p}^{(j)}) = \mathbb{H}_{\mu_j}(V'_j)$ . Taking the product over  $j = 1, \dots, r$  gives

$$|\mathcal{L}^Q(A)| \leq e^{\sum_{j=1}^r \mathbb{H}_{\mu_j}(V'_j) |A^{(j)}|}.$$

From the assumption that  $A \in \tilde{\mathcal{E}}$ , where  $\tilde{\mathcal{E}}$  is defined in (4.9), we have

$$|A^{(j)}| = (c_j - c_{j+1}) \log D + O(\log^{3/4} D).$$

Using this, and the trivial bound  $\mathbb{H}_{\mu_j}(V'_j) \leq \log |\text{Supp}(\mu_j)| \leq \log(2^k)$ , (4.18) follows.

Now we prove (4.17), which is a “metric” bound, the point being that none of the sums in  $\mathcal{L}^P(A)$  can be too large in an appropriate sense. Pick a basis for  $\mathbb{Q}^k$  adapted to  $\mathcal{V}'$ : that is, a basis  $e_1, \dots, e_k$  such that  $V'_j = \text{Span}(e_1, \dots, e_{\dim V'_j})$  for each  $j$ , and  $e_1 = \mathbf{1}$ . There are positive integers  $M, N = O_{\mathcal{V}', \mathcal{Y}}(1)$  such that, in this basis, the  $e_i$ -coordinates of  $P_j(\omega)$  are all rationals with denominator  $M$  and absolute value at most  $N$ .

Now for fixed  $j$  and  $\omega$ , if  $D$  is large then  $\sum_{a \in B_\omega^{(j)}} a \leq D^{c_j} \log D$ , since  $B_\omega^{(j)} \subset (D^{c_{j+1}}, D^{c_j}]$  and by the assumption that  $A \in \tilde{\mathcal{E}}$ . Thus

$$\sum_{\substack{\omega \in \{0,1\}^k \\ \omega \in V_j}} P_j(\omega) \sum_{a \in B_\omega^{(j)}} a \in \left\{ \sum_{1 \leq i \leq \dim(V'_j)} x_i e_i \in \mathbb{Q}^k : Mx_i \in \mathbb{Z}, \right. \\ \left. |x_i| \leq rND^{c_j} \log D \text{ (for all } i) \right\},$$

and so the expression  $\sum_{i=1}^r \sum_{\omega \in \{0,1\}^k \cap V_j} P_j(\omega) \sum_{a \in B_\omega^{(j)}} a$  belongs to the set

$$\left\{ \sum_{1 \leq i \leq k} x_i e_i \in \mathbb{Q}^k : Mx_i \in \mathbb{Z} \text{ and } |x_i| \leq r^2ND^{c_j} \log D \right. \\ \left. \text{for } \dim V'_{j-1} < i \leq \dim V'_j \text{ and } 1 \leq j \leq r \right\}.$$

We must bound the number of different values that the expression  $\sum_{i=1}^k x_i e_i$  can take mod  $\mathbf{1}$  when the coefficients  $x_1, \dots, x_k$  are as above. Since  $e_1 = \mathbf{1}$  and  $x_1 M \in \mathbb{Z}$ , given  $x_2, \dots, x_k$  there are at most  $M$  possibilities for  $x_1$  mod  $\mathbf{1}$ . In addition, there are

$$\ll (r^2MN)^{k-1} (\log D)^r D^{\sum_{j=1}^r c_j \dim(V'_j/V'_{j-1})}$$

possibilities for  $x_2, \dots, x_k$ , thereby concluding the proof of (4.17) and hence of Lemma 4.5. □

A potential problem with applying Lemma 4.5 is that there may be infinitely many subflags  $\mathcal{V}'$  to consider, and the constant implied by the  $\ll$ -symbol depends on  $\mathcal{V}'$ . As we shall see in the next Lemma, however, we may reduce the problem to consideration of a finite number of subflags, a tool which will be used in several parts of this paper.

**Lemma 4.6** *For a given  $k$ , the set of all flags*

$$\mathcal{V}' : (\mathbf{1}) = V'_0 \leq V'_1 \leq V'_2 \leq \dots \leq V'_r \leq \mathbb{Q}^k$$

*may be partitioned into  $O_k(1)$  equivalence classes such that any two flags  $\mathcal{V}', \mathcal{V}''$  in the same equivalence class satisfy  $\dim V'_j = \dim V''_j$  for all  $j$ , and for any thresholds  $\mathbf{c}$  satisfying  $c_1 \geq c_2 \geq \dots \geq c_{r+1}$  and probability measures  $\mu$  supported on  $\{0, 1\}^k$ , we have  $\mathbb{H}_{\mu_j}(V'_j) = \mathbb{H}_{\mu_j}(V''_j)$  for all  $j$  and  $e(\mathcal{V}', \mathbf{c}, \mu) = e(\mathcal{V}'', \mathbf{c}, \mu)$ .*

*Proof* We say that two subflags  $\mathcal{V}', \mathcal{V}''$  are *equivalent* if  $V'_j, V''_j$  have the same intersection with  $\{0, 1\}^k$  and  $\dim V'_j = \dim V''_j$ , for all  $j = 1, \dots, r$ . There are clearly only  $O_k(1)$  equivalence classes, and the desired properties hold for members of the same equivalence class by the definition of  $\mathbb{H}_{\mu_j}(V'_j)$  and  $e(\mathcal{V}', \mathbf{c}, \boldsymbol{\mu})$ .  $\square$

Armed with Lemma 4.6, we immediately obtain from Lemma 4.5, applied to one representative from each class, the following corollary.

**Corollary 4.7** *Let  $(\mathcal{V}, \mathbf{c}, \boldsymbol{\mu})$  be a system and suppose that  $A \in \tilde{\mathcal{E}}$ . Then*

$$|\mathcal{L}_{\mathcal{V}, \mathbf{c}, \boldsymbol{\mu}}(A)| \ll e^{O(\log^{3/4} D)} \min_{\mathcal{V}' \leq \mathcal{V}} D^{e(\mathcal{V}', \mathbf{c}, \boldsymbol{\mu})}.$$

#### 4.4 The upper bound in Theorem 7

We can now establish the upper bound in Theorem 7, that is to say the inequality  $\beta_k \leq \gamma_k$ .

We start by applying Proposition 4.4. Together with Lemma A.5, it implies that

$$\begin{aligned} &\mathbb{P}(\exists \text{ distinct } A_1, \dots, A_k \subseteq \mathbf{A} \cap (D^c, D] \text{ with equal sums}) \\ &\leq (\log D)^{O(1)} \sup_{(\mathcal{V}, \mathbf{c}, \boldsymbol{\mu})} D^{-e(\mathcal{V}, \mathbf{c}, \boldsymbol{\mu})} \mathbb{E} 1_{\mathbf{A} \in \tilde{\mathcal{E}}} |\mathcal{L}_{\mathcal{V}, \mathbf{c}, \boldsymbol{\mu}}(\mathbf{A})| + O(e^{-\frac{1}{4} \log^{1/2} D}). \end{aligned}$$

Here, the supremum is over complete systems  $(\mathcal{V}, \mathbf{c}, \boldsymbol{\mu})$  with  $c_{r+1} = c$ , and we made the observation that for such systems we have

$$e(\mathcal{V}, \mathbf{c}, \boldsymbol{\mu}) = c_1 + \dots + c_r,$$

an immediate consequence of the definition of  $e(\mathcal{V}, \mathbf{c}, \boldsymbol{\mu})$  and the fact that  $\mathbb{H}_{\mu_j}(V_j) = 0$  for all  $j$  and that  $\dim V_j = j + 1$ . Thus we may apply Corollary 4.7, concluding that

$$\begin{aligned} &\mathbb{P}(\exists \text{ distinct } A_1, \dots, A_k \subseteq \mathbf{A} \cap (D^c, D] \text{ with equal sums}) \\ &\leq D^{\theta + o(1)} + O(e^{-\frac{1}{4} \log^{1/2} D}), \end{aligned}$$

where

$$\theta := \sup_{(\mathcal{V}, \mathbf{c}, \boldsymbol{\mu}) : c_{r+1} = c} \min_{\mathcal{V}' \leq \mathcal{V}} (e(\mathcal{V}', \mathbf{c}, \boldsymbol{\mu}) - e(\mathcal{V}, \mathbf{c}, \boldsymbol{\mu})); \tag{4.19}$$

the supremum is over all complete systems  $(\mathcal{V}, \mathbf{c}, \boldsymbol{\mu})$  with  $c_{r+1} = c$ , and the minimum is over all subflags  $\mathcal{V}' \leq \mathcal{V}$ . Note that the minimum exists

by Lemma 4.6, since we may restrict attention to a finite set of subflags  $\mathcal{V}'$ . Moreover, the supremum is realised, meaning there is a system  $(\mathcal{V}, \mathbf{c}, \boldsymbol{\mu})$  for which the right side of (4.19) equals  $\theta$ . Indeed, there are  $O(1)$  choices for  $\mathcal{V}$ , and with  $\mathcal{V}$  fixed the quantities  $\mathbf{c}, \boldsymbol{\mu}$  range over compact subsets of Euclidean space, with the right side of (4.19) continuous in these variables.

Now, if we assume that  $c > \gamma_k$ , then the definition of  $\gamma_k$  in Problem 3.7 implies that there is no system  $(\mathcal{V}, \mathbf{c}, \boldsymbol{\mu})$  with  $c_{r+1} = c$  and that satisfies the entropy condition (3.4). Equivalently, if  $c_{r+1} = c$ , then  $\min_{\mathcal{V}' \leq \mathcal{V}} (e(\mathcal{V}', \mathbf{c}, \boldsymbol{\mu}) - e(\mathcal{V}, \mathbf{c}, \boldsymbol{\mu})) < 0$ . In particular, we have  $\theta < 0$ . We have thus established (4.1), as required.

*Remark* In the above proof,  $(\mathcal{V}, \mathbf{c}, \boldsymbol{\mu})$  is a *complete* system. However, for other aspects of our problem it is not natural to focus on the completeness condition, for which reason we omit it from the definition of  $\gamma_k$ .

## 5 The lower bound $\beta_k \geq \tilde{\gamma}_k$

### 5.1 Introduction and simple reductions

The aim of this section and the next is to establish the lower bound  $\beta_k \geq \tilde{\gamma}_k$ . We begin, in Lemma 5.3 below, by showing that we may restrict our attention to certain systems satisfying some additional regularity conditions.

We isolate a “folklore” lemma from the proof for which it is not easy to find a good reference. The authors thank Carla Groenland for a helpful conversation on this topic.

**Lemma 5.1** *Let  $V$  be a subspace of  $\mathbb{Q}^k$ . Then  $\#(V \cap \{0, 1\}^k) \leq 2^{\dim V}$ .*

*Proof* We outline two quite different short proofs. Let  $d := \dim V$ .

*Proof 1.* We claim that there is a projection from  $\mathbb{Q}^k$  onto some set of  $d$  coordinates which is injective on  $V$ . From this, the result is obvious, since the image of  $\{0, 1\}^k$  under any such projection has size  $2^d$ . To prove the claim, let  $e_1, \dots, e_n$  denote the standard basis on  $\mathbb{Q}^n$ . Note that if  $W \leq \mathbb{Q}^n$  and if none of the quotient maps  $\mathbb{Q}^n \mapsto \mathbb{Q}^n / \langle e_i \rangle$  is injective on  $W$ , then  $W$  must contain a multiple of each  $e_i$ , and therefore  $W = \mathbb{Q}^n$ . Thus if  $W$  is a proper subspace of  $\mathbb{Q}^n$  then there is a projection onto some set of  $(n - 1)$  coordinates which is injective on  $W$ . Repeated use of this fact establishes the claim.

*Proof 2.* Suppose that  $\#(V \cap \{0, 1\}^k)$  contains  $2^d + 1$  points. These are all distinct under the natural ring homomorphism  $\pi : \mathbb{Z}^k \rightarrow \mathbb{F}_2^k$ , and so their images cannot lie in a subspace (over  $\mathbb{F}_2$ ) of dimension  $d$ . Hence there are  $v_1, \dots, v_{d+1} \in V$  such that  $\pi(v_1), \dots, \pi(v_{d+1})$ , are linearly independent over  $\mathbb{F}_2$ . The  $(d + 1) \times k$  matrix formed by these  $\pi(v_i)$  therefore has a  $(d + 1) \times (d + 1)$ -subminor which is nonzero in  $\mathbb{F}_2$ . The corresponding subminor



of the matrix formed by the  $v_i$  is therefore an odd integer, and in particular not zero. This means that  $v_1, \dots, v_{d+1}$  are linearly independent over  $\mathbb{Q}$ , contrary to the assumption that  $\dim(V) = d$ . □

We now record an immediate corollary of Lemma 4.6, which provides a “gap condition” on the e-quantities.

**Lemma 5.2** *If the system  $(\mathcal{V}, \mathbf{c}, \boldsymbol{\mu})$  satisfies (3.5) then there is an  $\varepsilon > 0$  such that for all proper subflags  $\mathcal{V}'$ ,*

$$e(\mathcal{V}', \mathbf{c}, \boldsymbol{\mu}) \geq e(\mathcal{V}, \mathbf{c}, \boldsymbol{\mu}) + \varepsilon. \tag{5.1}$$

For future reference, the next two lemmas record more information about optimal systems for  $\tilde{\gamma}_k$  and for  $\gamma_k$ , respectively.

**Lemma 5.3** *Let  $k \in \mathbb{Z}_{\geq 2}$ . We have that  $\tilde{\gamma}_k$  is the supremum of all  $c > 0$  for which there is a system  $(\mathcal{V}, \mathbf{c}, \boldsymbol{\mu})$  such that  $c_{r+1} = c$ , (3.5) holds and we further have:*

- (a)  $1 = c_1 > c_2 > \dots > c_{r+1} = c$ ;
- (b)  $\mathbb{H}_{\mu_j}(V_{j-1}) > \dim(V_j/V_{j-1})$  for  $1 \leq j \leq r - 1$  and

$$\mathbb{H}_{\mu_r}(V_{r-1}) > \frac{c_r}{c_r - c_{r+1}} \dim(V_r/V_{r-1});$$

- (c)  $\dim(V_1/V_0) = 1$ ;
- (d)  $\text{Supp}(\mu_j) = V_j \cap \{0, 1\}^k$  for  $j = 1, 2, \dots, r$ ;
- (e) for all  $j$  and  $\omega$ ,  $\mu_j(\omega) = \mu_j(\mathbf{1} - \omega)$ .

*Proof* First of all, we show that we may assume that  $c > 0$  and that statement (d) holds. Indeed, if a system  $(\mathcal{V}, \boldsymbol{\mu}, \mathbf{c})$  satisfies (3.5), then Lemma 5.2 implies that (5.1) holds for some  $\varepsilon > 0$ . As the difference between the left and right sides of (5.1) is continuous in the quantities  $c_j$  and  $\mu_j(\omega)$ , we may increase  $c_{r+1}$  (and possibly some of the other  $c_j$ 's) a tiny bit and we may also adjust the measures  $\mu_j$  by a small amount, so that  $c_{r+1} > 0$ , statement (d) holds, and we also have that

$$e(\mathcal{V}', \mathbf{c}, \boldsymbol{\mu}) \geq e(\mathcal{V}, \mathbf{c}, \boldsymbol{\mu}) + \varepsilon/2$$

for every proper subflag  $\mathcal{V}'$ .

Next, we show that we may take  $c_1 = 1$ . Indeed, condition (3.5) implies that  $e(\mathcal{V}', \mathbf{c}, \boldsymbol{\mu}) \geq e(\mathcal{V}, \mathbf{c}, \boldsymbol{\mu}) \geq 0$  for all  $\mathcal{V}' \leq \mathcal{V}$  by (3.3). Now if  $c_1 < 1$  and  $\tilde{c}_j = c_j/c_1$  for each  $j$ , then the perturbed system  $(\mathcal{V}, \tilde{\mathbf{c}}, \boldsymbol{\mu})$  has a larger

value of  $c_{r+1}$ , and moreover also satisfies (3.5), since for any subflag  $\mathcal{V}'$  we have

$$e(\mathcal{V}', \tilde{\mathbf{c}}, \boldsymbol{\mu}) = (1/c_1)e(\mathcal{V}', \mathbf{c}, \boldsymbol{\mu}).$$

Next, consider a system  $(\mathcal{V}, \mathbf{c}, \boldsymbol{\mu})$  satisfying  $c_1 = 1$  and  $c_{r+1} = c > 0$ , and consider the subflag  $\mathcal{V}' : \langle \mathbf{1} \rangle = V'_0 \leq V'_1 \leq \dots \leq V'_r$ , where  $V'_i = V_i$  for  $i \neq j$ , and  $V'_j = V_{j-1}$ ; that is,  $\mathcal{V}'$  has two consecutive copies of  $V_{j-1}$ . By assumption (Definition 3.2), we have  $V_{j-1} \neq V_j$ , and thus  $\mathcal{V}'$  is a proper subflag of  $\mathcal{V}$ . Thus

$$e(\mathcal{V}', \mathbf{c}, \boldsymbol{\mu}) - e(\mathcal{V}, \mathbf{c}, \boldsymbol{\mu}) = \begin{cases} (c_j - c_{j+1})(\mathbb{H}_{\mu_j}(V_{j-1}) - \dim(V_j/V_{j-1})) & \text{if } j \leq r - 1, \\ (c_r - c_{r+1})\mathbb{H}_{\mu_r}(V_{r-1}) - c_r \dim(V_r/V_{r-1}) & \text{if } j = r. \end{cases}$$

Since the left-hand side is positive, we conclude that (a) and (b) hold.

(c) Let  $d = \dim(V_1/V_0)$ . By Lemma 5.1,  $|V_1 \cap \{0, 1\}^k| \leq 2^{\dim V_1} = 2^{d+1}$  and hence  $\mu_1$  is supported on at most  $2^{d+1} - 1$  cosets of  $V_0$  (since  $\mathbf{1} \in V_0$ , the points  $\mathbf{0}$  and  $\mathbf{1}$  lie in the same coset). In particular, by Lemma B.2,  $\mathbb{H}_{\mu_1}(V_0) \leq \log(2^{d+1} - 1)$ . On the other hand,  $\mathbb{H}_{\mu_1}(V_0) > d$  by statement (b). We must thus have  $d = 1$ , which is exactly statement (c).

(e) Assume the system  $(\mathcal{V}, \mathbf{c}, \boldsymbol{\mu})$  satisfies (3.5) and (a). For every  $j$  and  $\omega \in V_j$ , we define

$$\tilde{\mu}_j(\omega) = \frac{\mu_j(\omega) + \mu_j(\mathbf{1} - \omega)}{2}.$$

We then consider the system  $(\mathcal{V}, \mathbf{c}, \tilde{\boldsymbol{\mu}})$ , and must show that it also satisfies (3.5). For this, it is enough to show that

$$\mathbb{H}_{\tilde{\mu}_j}(V'_j) \geq \mathbb{H}_{\mu_j}(V'_j) \tag{5.2}$$

for all  $j$ . Indeed, we then have, for every proper subflag  $\mathcal{V}'$ ,

$$e(\mathcal{V}', \mathbf{c}, \tilde{\boldsymbol{\mu}}) \geq e(\mathcal{V}', \mathbf{c}, \boldsymbol{\mu}) > e(\mathcal{V}, \mathbf{c}, \boldsymbol{\mu}) = e(\mathcal{V}, \mathbf{c}, \tilde{\boldsymbol{\mu}}).$$

To prove (5.2), write

$$\mathbb{H}_{\mu_j}(V'_j) = \sum_C L(\mu_j(C)), \quad \mathbb{H}_{\tilde{\mu}_j}(V'_j) = \sum_C L(\tilde{\mu}_j(C)),$$

where the sum is over all cosets  $C$  of  $V'_j$  and  $L(t) = -t \log t$ . Thus, since  $-C$  runs over all cosets as  $C$  does, we have

$$\mathbb{H}_{\mu_j}(V'_j) = \sum_C \frac{L(\mu_j(C)) + L(\mu_j(-C))}{2}.$$

By the concavity of  $L$ , we have

$$\frac{L(\mu_j(C)) + L(\mu_j(-C))}{2} \leq L\left(\frac{\mu_j(C) + \mu_j(-C)}{2}\right) = L(\tilde{\mu}_j(C)).$$

Claim (5.2) then readily follows. □

**Lemma 5.4** *Let  $k \in \mathbb{Z}_{\geq 2}$  be such that  $\gamma_k > 0$ . Then we have that  $\gamma_k$  is the supremum of all  $c > 0$  for which there is a system  $(\mathcal{V}, \mathbf{c}, \boldsymbol{\mu})$  such that  $c_{r+1} = c$ , (3.4) holds and we further have:*

- (a)  $1 = c_1 > c_2 > \dots > c_{r+1} = c$ ;
- (b)  $\mathbb{H}_{\mu_j}(V_{j-1}) \geq \dim(V_j/V_{j-1})$  for  $1 \leq j \leq r - 1$  and

$$\mathbb{H}_{\mu_r}(V_{r-1}) \geq \frac{c_r}{c_r - c_{r+1}} \dim(V_r/V_{r-1});$$

- (c)  $\dim(V_1/V_0) = 1$ ;
- (d)  $\bigcup_{i=1}^j \text{Supp } \mu_i$  spans  $V_j$  for  $j = 1, 2, \dots, r$ ;
- (e) for all  $j$  and  $\omega$ ,  $\mu_j(\omega) = \mu_j(\mathbf{1} - \omega)$ .

*Remark* As we will see in Part IV, we always have  $\gamma_k > 0$ .

*Proof* The proof that we may take  $c_1 = 1$  is the same as in Lemma 5.3.

Next, consider a system  $(\mathcal{V}, \mathbf{c}, \boldsymbol{\mu})$  satisfying  $c_1 = 1$  and  $c_{r+1} = c > 0$ , and consider the subflag  $\mathcal{V}' : \langle \mathbf{1} \rangle = V'_0 \leq V'_1 \leq \dots \leq V'_r$ , where  $V'_i = V_i$  for  $i \leq r - 1$ , and  $V'_r = V_{r-1}$ . Thus

$$e(\mathcal{V}', \mathbf{c}, \boldsymbol{\mu}) - e(\mathcal{V}, \mathbf{c}, \boldsymbol{\mu}) = (c_r - c_{r+1})\mathbb{H}_{\mu_r}(V_{r-1}) - c_r \dim(V_r/V_{r-1}).$$

Since the left-hand side is  $\geq 0$  and we have assumed that  $c_{r+1} = c > 0$  and that  $V_{r-1} \neq V_r$ , the latter being true from Definition 3.2, we conclude that

$$c_r > c_{r+1} \quad \text{and} \quad \mathbb{H}_{\mu_r}(V_{r-1}) \geq \frac{c_r}{c_r - c_{r+1}} \dim(V_r/V_{r-1}). \tag{5.3}$$

This proves part of statements (a) and (b). We shall now prove them fully.

(a) There are always indices  $1 = i_1 < i_2 < \dots < i_s < i_{s+1} = r + 1$  such that

$$c_{i_j} = \dots = c_{i_{j+1}-1} > c_{i_{j+1}} \quad \text{for } j = 1, \dots, s.$$

Crucially, note that  $i_{s+1} = r + 1$  because  $c_r > c_{r+1}$  by (5.3). Next, we define the system  $(\mathscr{W}, \mathbf{v}, \mathbf{d})$ , where  $\mathscr{W}$  is an  $s$ -step flag and, for all  $j \in \{1, \dots, s\}$ , we have

$$W_j = V_{i_{j+1}-1}, \quad v_j = \mu_{i_{j+1}-1}, \quad \text{and} \quad d_j = c_{i_{j+1}-1}.$$

In particular,  $W_s = V_{i_{s+1}-1} = V_r$  because  $i_{s+1} = r$ , and thus  $\mathscr{W}$  is a non-degenerate flag system as per Definition 3.2 (b). Clearly,  $1 = d_1 > d_2 > \dots > d_s > d_{s+1} = c$ , so in order to prove part (a), all that remains to show is that the system  $(\mathscr{W}, \mathbf{v}, \mathbf{d})$  satisfies the entropy condition (3.4). This follows by a simple computation. Indeed, let  $\mathscr{W}'$  be a subflag of  $\mathscr{W}$ . We then define  $\mathscr{V}' \leq \mathscr{V}$  by letting  $V'_m = W_j$  whenever  $i_j \leq m < i_{j+1}$ . Hence,

$$\begin{aligned} e(\mathscr{V}', \boldsymbol{\mu}, \mathbf{c}) &= \sum_{m=1}^r (c_m - c_{m+1}) \mathbb{H}_{\mu_m}(V'_m) + \sum_{m=1}^r c_m \dim(V'_m / V'_{m-1}) \\ &= \sum_{j=1}^s (c_{i_{j+1}-1} - c_{i_{j+1}}) \mathbb{H}_{\mu_m}(V'_m) + \sum_{j=1}^s c_{i_j} \dim(V'_{i_j} / V'_{i_j-1}) \\ &= e(\mathscr{W}', \mathbf{v}, \mathbf{d}). \end{aligned}$$

Consequently, since the system  $(\mathscr{V}, \boldsymbol{\mu}, \mathbf{c})$  satisfies condition (3.4), so does  $(\mathscr{W}, \mathbf{v}, \mathbf{d})$ . This proves that we may always assume condition (a).

(b) Consider a system  $(\mathscr{V}, \mathbf{c}, \boldsymbol{\mu})$  satisfying (a). We then argue as in Lemma 5.3, by considering the subflag  $\mathscr{V}'$  with  $V'_i = V_i$  for  $i \neq j$ , and  $V'_j = V_{j-1}$ . We then have

$$\begin{aligned} e(\mathscr{V}', \mathbf{c}, \boldsymbol{\mu}) - e(\mathscr{V}, \mathbf{c}, \boldsymbol{\mu}) &= \begin{cases} (c_j - c_{j+1})(\mathbb{H}_{\mu_j}(V_{j-1}) - \dim(V_j / V_{j-1})) & \text{if } j \leq r - 1, \\ (c_r - c_{r+1})\mathbb{H}_{\mu_r}(V_{r-1}) - c_r \dim(V_r / V_{r-1}) & \text{if } j = r. \end{cases} \end{aligned}$$

Since the left-hand side is  $\geq 0$  and  $c_j - c_{j+1} > 0$  for all  $j = 1, \dots, r$ , statement (b) follows.

(c) Assuming statement (b), we may prove statement (c) by arguing as in Lemma 5.3.

(d) Suppose that (a) holds. Consider the flag  $\mathscr{V}' : \langle \mathbf{1} \rangle \leq V'_1 \leq \dots \leq V'_r$ , where

$$V'_j = \text{Span} \left( \bigcup_{i=1}^j \text{Supp}(\mu_i) \right) \quad (1 \leq j \leq r).$$

It is easy to see from the definition of a system (Definition 3.2) that  $\mathcal{V}'$  is a subflag of  $\mathcal{V}$ . We have  $\mathbb{H}_{\mu_j}(V'_j) = 0$  for all  $j$ , and hence

$$\begin{aligned} e(\mathcal{V}', \mathbf{c}, \boldsymbol{\mu}) &= \sum_{i=1}^r c_i \dim(V'_i / V'_{i-1}) \\ &= -c_1 + c_r \dim(V'_r) + \sum_{i=1}^{r-1} (c_i - c_{i+1}) \dim(V'_i) \\ &\geq -c_1 + c_r \dim(V_r) + \sum_{i=1}^{r-1} (c_i - c_{i+1}) \dim(V_i) = e(\mathcal{V}, \mathbf{c}, \boldsymbol{\mu}), \end{aligned}$$

by (3.5). Since  $c_i - c_{j+1} > 0$  for all  $i \leq r - 1$ , and  $c_r > c_{r+1} \geq 0$ , we must have that  $V'_i = V_i$  for all  $i$ , which is precisely statement (d).

(e) This statement is proven as in Lemma 5.3. □

The bound  $\beta_k \geq \tilde{\gamma}_k$  will now follow from the following proposition, as long as we can show that the quantity  $\tilde{\gamma}_k$  is well-defined and positive. The latter will be accomplished in Sect. 9, where we construct a system satisfying the strict entropy condition 3.5. An alternative construction is given in Appendix C.

As usual,  $\mathbf{A}$  is a logarithmic random set.

**Proposition 5.5** *Let  $c > 0$  and suppose that there is a system  $(\mathcal{V}, \mathbf{c}, \boldsymbol{\mu})$  such that:*

- (i)  $1 = c_1 > c_2 > \dots > c_{r+1} = c$ ;
- (ii) *There is some  $\varepsilon > 0$  such that  $e(\mathcal{V}', \mathbf{c}, \boldsymbol{\mu}) \geq e(\mathcal{V}, \mathbf{c}, \boldsymbol{\mu}) + \varepsilon$  for all proper subflags  $\mathcal{V}'$  of  $\mathcal{V}$ .*
- (iii)  $\text{Supp}(\mu_j) = V_j \cap \{0, 1\}^k$  for  $j = 1, 2, \dots, r$ .

*Let  $\delta > 0$ , and assume that  $D$  is large enough in terms of  $\delta, \varepsilon$  and  $(\mathcal{V}, \mathbf{c}, \boldsymbol{\mu})$ . Then the probability that  $\mathbf{A} \cap [D^c, D]$  has  $k$  distinct subsets with equal sums is  $\geq 1 - \delta$ .*

The proof of Proposition 5.5 is perhaps the most difficult part of this paper, and will occupy this and the next section. Throughout the remainder of this section and throughout the next section, we will fix a system  $(\mathcal{V}, \mathbf{c}, \boldsymbol{\mu})$  with  $c_{r+1} = c$  satisfying conditions (i)–(iii) of Proposition 5.5. Constants implied by  $O$ – and  $\ll$ –symbols may depend on this system.

The main result, which we will prove in this section and the next, is Proposition 5.7 below.

**Definition 5.6** (*Nondegenerate maps*) A map  $\psi : X \rightarrow \{0, 1\}^k$  is said to be *nondegenerate* if the image of  $\psi$  is not contained in any of the subspaces  $\{x \in \mathbb{Q}^k : x_i = x_j\}$ .

The map  $\psi$  is a ‘‘Venn diagram selection function’’, that is, the value of  $\psi(b)$  specifies which piece of the Venn diagram of  $k$  subsets  $X_1, \dots, X_k$  of  $X$  that  $b$  belongs to. In the notation (4.6) of the previous section,  $\psi(a) = \omega$  means that  $a \in B_\omega$ . The condition that  $\psi$  is nondegenerate is equivalent to  $X_1, \dots, X_k$  being distinct, and is similar to the property of a flag  $\mathcal{V}$  being nondegenerate.

**Proposition 5.7** *With probability tending to 1 as  $D \rightarrow \infty$ , there exists a nondegenerate map  $\psi : \mathbf{A} \cap (D^c, D] \rightarrow \{0, 1\}^k$  such that  $\sum_{a \in \mathbf{A}} a\psi(a) \in \langle \mathbf{1} \rangle$ .*

The map  $\psi$  will be constructed using the data from the system  $(\mathcal{V}, \mathbf{c}, \boldsymbol{\mu})$ . Before we embark on the proof of this result, we show how to deduce Proposition 5.5 from it.

*Proof of Proposition 5.5, assuming Proposition 5.7* By Proposition 5.7, we know that with probability  $1 - o_{D \rightarrow \infty}(1)$  there is a nondegenerate map  $\psi : \mathbf{A} \cap (D^c, D] \rightarrow \{0, 1\}^k$  such that  $\sum_{a \in \mathbf{A}} a\psi(a)$  lies in  $\langle \mathbf{1} \rangle$ , that is to say, it is a constant vector. We will show that this map induces  $k$  distinct subsets of  $\mathbf{A}$  with equal sums.

Let  $\psi_i : \mathbf{A} \cap (D^c, D] \rightarrow \mathbb{Q}, i = 1, \dots, k$ , denote the projection of  $\psi$  onto the  $i$ -th coordinate of  $\mathbb{Q}^k$ , so that  $\psi = (\psi_1, \dots, \psi_k)$ . Define

$$A_i := \{a \in \mathbf{A} : \psi_i(a) = 1\}.$$

These sets are distinct because if  $A_i = A_j$ , then the image of  $\psi$  would take values in the hyperplane  $\{x \in \mathbb{Q}^k : x_i = x_j\}$ , contrary to the fact that  $\psi$  is nondegenerate. Moreover, for all  $i, j$  we have

$$\sum_{a \in A_i} a - \sum_{a \in A_j} a = \sum_{a \in \mathbf{A}} a\psi_i(a) - \sum_{a \in \mathbf{A}} a\psi_j(a) = 0,$$

and so  $A_1, \dots, A_k$  do indeed have equal sums. □

### 5.2 Many values of $\sum_{a \in \mathbf{A}} a\psi(a)$ , and a moment bound

We turn now to the task of proving Proposition 5.7. We will divide the proof of Proposition 5.7 into two parts. The first and more difficult part, which we prove in this section, states that (with high probability)  $\sum_{a \in \mathbf{A}} a\psi(a)$  takes many different values modulo  $\langle \mathbf{1} \rangle$  as  $\psi$  ranges over all nondegenerate maps  $\psi : \mathbf{A} \cap (D^c, D] \rightarrow \{0, 1\}^k$ . The precise statement is Proposition 5.9 below. The deduction of Proposition 5.7 from Proposition 5.9 will occupy Sect. 6.

Let  $0 < \kappa \leq \min_{1 \leq j \leq r} (c_j - c_{j+1}) - 2/\log D$  be a small quantity, which may depend on  $D$ . Let

$$\begin{aligned} \mathbf{A}^j &= \{a \in \mathbf{A} : D^{c_{j+1}+\kappa} < a \leq D^{c_j}/e\} \quad (1 \leq j \leq r), \\ \mathbf{A}' &:= \bigcup_{j=1}^r \mathbf{A}^j. \end{aligned} \tag{5.4}$$

The purpose of working with  $\mathbf{A}'$  rather than  $\mathbf{A}$  is to ensure that some gaps are left for the subsequent argument in the next section (based on ideas of Maier and Tenenbaum [20]), in which we show that one of the many sums  $\sum_{a \in \mathbf{A}'} a\psi(a)$  guaranteed by Proposition 5.9 may be modified, using the elements of  $\mathbf{A} \cap (D^c, D] \setminus \mathbf{A}'$ , to be in  $\langle \mathbf{1} \rangle$ .

**Definition 5.8** (*Compatible functions*) We say that a map  $\psi : \mathbf{A}' \rightarrow \{0, 1\}^k$  is *compatible* if, for all  $j, a \in \mathbf{A}^j$  implies  $\psi(a) \in V_j$ .

*Remark* Recall that  $\text{Supp}(\mu_j) = V_j \cap \{0, 1\}^k$  for all  $j$  by condition (iii) of Proposition 5.5. Setting  $B_\omega^{(j)} = \{a \in \mathbf{A}^j : \psi(a) = \omega\}$ , we see that  $\psi$  being compatible is equivalent to  $B_\omega^{(j)} \neq \emptyset$  only if  $\mu_j(\omega) > 0$ , and is consistent with earlier notation (4.6).

**Proposition 5.9** *There exist real numbers  $\kappa^* > 0, p > 1$  and  $t > 0$  (which depend on the system  $(\mathcal{V}, \mathbf{c}, \boldsymbol{\mu})$ ) so that the following is true. Let  $\delta > 0$  and suppose that  $D$  is sufficiently large as a function of  $\delta$ . Uniformly for  $0 \leq \kappa \leq \kappa^*$ , we have with probability at least  $1 - \delta$ , that  $\sum_{a \in \mathbf{A}'} a\psi(a)$  takes at least*

$$(t\delta)^{\frac{1}{p-1}} D^{\sum_j c_j \dim(V_j/V_{j-1})}$$

*different values modulo  $\langle \mathbf{1} \rangle$ , as  $\psi$  ranges over all nondegenerate, compatible maps  $\psi$ .*

*Remark* By (5.4), it clearly suffices to prove Proposition 5.9 for  $\kappa = \kappa^*$ .

We will deduce Proposition 5.9 from a moment bound. Firstly, define the representation function  $r_{\mathbf{A}'} : \mathbb{Q}^k / \langle \mathbf{1} \rangle \rightarrow \mathbb{R}$  by

$$r_{\mathbf{A}'}(x) := \sum_{\substack{\psi: \mathbf{A}' \rightarrow \{0,1\}^k \\ \sum_{a \in \mathbf{A}'} a\psi(a) - x \in \langle \mathbf{1} \rangle}} w_{\mathbf{A}'}(\psi),$$

where the summation is over all maps  $\psi : \mathbf{A}' \rightarrow \{0, 1\}^k$ , and where

$$w_{\mathbf{A}'}(\psi) := \prod_{j=1}^r \prod_{a \in \mathbf{A}^j} \mu_j(\psi(a)).$$

This weight function  $w_{\mathbf{A}'}$  is chosen so that it is large only when  $\psi$  is *balanced*, that is, when for all  $j$  and  $\omega$ , the set  $\mathbf{A}^j$  has about  $\mu_j(\omega)|\mathbf{A}^j|$  elements  $a$  with

$\psi(a) = \omega$ . Observe that if  $\psi(a) \notin \text{Supp}(\mu_j)$  for some  $j$  and some  $a \in \mathbf{A}^j$ , then  $w_{\mathbf{A}'}(\psi) = 0$ , and thus only compatible  $\psi$  contribute to the sum  $r_{\mathbf{A}'}(x)$ . However,  $w_{\mathbf{A}'}(\psi)$  might be non-zero for some degenerate maps  $\psi$ , and these will be removed by a separate argument below.

The crucial moment bound for the deduction of Proposition 5.9 is given below.

**Proposition 5.10** *Let*

$$\mathcal{E}^* = \left\{ A \subseteq [D^c, D] : \#(A \cap (y/e, y]) \leq \sqrt{y}/100 \quad (D^c \leq y \leq D) \right\}.$$

There is a  $p > 1$  and  $\kappa^* > 0$  so that uniformly for  $0 \leq \kappa \leq \kappa^*$  and for all  $D \geq e^{100/c}$  we have the moment bound

$$\mathbb{E} \left[ 1_{\mathbf{A}' \in \mathcal{E}^*} \sum_x r_{\mathbf{A}'}(x)^p \right] \ll D^{-(p-1) \sum_j c_j \dim(V_j/V_{j-1})}.$$

*Proof of Proposition 5.9, assuming Proposition 5.10* Define also

$$\tilde{r}_{\mathbf{A}'}(x) := \sum_{\substack{\psi: \mathbf{A}' \rightarrow \{0,1\}^k \\ \psi \text{ is compatible and nondegenerate} \\ \sum_{a \in \mathbf{A}'} a \psi(a) - x \in \mathbf{1}}} w_{\mathbf{A}'}(\psi).$$

We have

$$\sum_x r_{\mathbf{A}'}(x) = \prod_{j=1}^r \left( \sum_{\omega} \mu_j(\omega) \right)^{|\mathbf{A}^j|} = \prod_{j=1}^r 1 = 1$$

for any  $\mathbf{A}'$ . On the other hand, when  $\psi$  is non-compatible, then  $w_{\mathbf{A}'}(\psi) = 0$  because we know that  $\text{Supp}(\mu_j) = V_j \cap \{0, 1\}^k$  for all  $j$  by our assumption of condition (iii) of Proposition 5.5. In addition, if  $\psi$  is degenerate, then its image is contained in  $\{x \in \mathbb{Q}^k : x_i = x_j\} \cap \{0, 1\}^k$  for some  $i \neq j$ . Since  $V_r \not\subset \{x \in \mathbb{Q}^k : x_i = x_j\}$ , there must exist some  $\omega \in V_r \cap \{0, 1\}^k = \text{Supp}(\mu_r)$  that is not in the support of  $\psi$ . Therefore,

$$\sum_x (r_{\mathbf{A}'}(x) - \tilde{r}_{\mathbf{A}'}(x)) \leq \sum_{\omega \in \text{Supp}(\mu_r)} (1 - \mu_r(\omega))^{|\mathbf{A}^r|}.$$

Since  $c_r > c_{r+1}$  by our assumption of condition (i) of Proposition 5.5, Lemma A.5 implies  $|\mathbf{A}^r| \geq \frac{1}{2}(c_r - c_{r+1}) \log D$  with probability  $> 1 - O(e^{-(1/4) \log^{1/2} D})$ , and thus the right side above is  $o(1)$  with this



same probability. The same lemma also implies that  $\mathbf{A}' \in \mathcal{E}^*$  with probability  $> 1 - O(e^{-(1/4)\log^{1/2} D})$ .

Now fix a small  $\delta > 0$ . The above discussion implies that, with probability at least  $1 - \delta/2$  (for  $D$  sufficiently large), we have

$$\sum_x \tilde{r}_{\mathbf{A}'}(x) \geq \frac{1}{2} \quad \text{and} \quad \mathbf{A}' \in \mathcal{E}^*. \tag{5.5}$$

On the other hand, Markov’s inequality and Proposition 5.10 imply that, with probability at least  $1 - \delta/2$ , we have

$$\begin{aligned} 1_{\mathbf{A}' \in \mathcal{E}^*} \sum_x \tilde{r}_{\mathbf{A}'}(x)^p &\leq 1_{\mathbf{A}' \in \mathcal{E}^*} \sum_x r_{\mathbf{A}'}(x)^p \\ &\ll \delta^{-1} D^{-(p-1) \sum_j c_j \dim(V_j/V_{j-1})}. \end{aligned} \tag{5.6}$$

By Hölder’s inequality,

$$1_{\mathbf{A}' \in \mathcal{E}^*} \sum_x \tilde{r}_{\mathbf{A}'}(x) \leq |\text{Supp}(\tilde{r}_{\mathbf{A}'})|^{1-1/p} (1_{\mathbf{A}' \in \mathcal{E}^*} \sum_x \tilde{r}_{\mathbf{A}'}(x)^p)^{1/p}. \tag{5.7}$$

With probability at least  $1 - \delta$ , both (5.5) and (5.6) hold, and in this case (5.7) gives

$$|\text{Supp}(\tilde{r}_{\mathbf{A}'})| \gg_p \delta^{\frac{1}{p-1}} D^{\sum_j c_j \dim(V_j/V_{j-1})}.$$

This completes the proof of Proposition 5.9. □

The rest of the section is devoted to the proof of Proposition 5.10.

### 5.3 An entropy condition for adapted systems

For reasons that will become apparent, in the proof of Proposition 5.10 we will need to apply the entropy gap condition not only with subflags  $\mathcal{V}'$  of  $\mathcal{V}$ , but with a more general type of system.

**Definition 5.11** (*Adapted system*) Given a system  $(\mathcal{V}, \mathbf{c}, \boldsymbol{\mu})$ , the pair  $(\mathcal{W}, \mathbf{b})$  is *adapted* to  $(\mathcal{V}, \mathbf{c}, \boldsymbol{\mu})$  if  $\mathcal{W} : \mathbf{1} = W_0 \leq W_1 \leq \dots \leq W_s$  is a complete flag with  $W_s \leq V_r$ , and  $\mathbf{b} = (b_1, \dots, b_s)$  satisfies  $1 \geq b_1 \geq \dots \geq b_s \geq 0$  and the condition

$$W_i \leq V_j \quad \text{whenever} \quad b_i > c_{j+1}.$$

We say that  $(\mathscr{W}, \mathbf{b})$  is *saturated* if  $s = \dim(V_r) - 1$  and if for all  $j \leq r$ , there are exactly  $\dim V_j - 1$  values of  $i$  with  $b_i > c_{j+1}$ . Otherwise, we call  $(\mathscr{W}, \mathbf{b})$  *unsaturated*.

*Remark* For the definition of complete flag, see Definition 3.1. We make a few comments to motivate the term *saturated*. Let

$$m_j = \#\{i : b_i > c_{j+1}\} \quad (0 \leq j \leq r), \tag{5.8}$$

so that the  $b_i$ 's belonging to the interval  $(c_{j+1}, c_j]$  are precisely  $b_{m_{j-1}+1}, \dots, b_{m_j}$ . Since  $W_i \leq V_j$  whenever  $b_i > c_{j+1}$ , we infer that

$$W_{m_j} \leq V_j \quad (1 \leq j \leq r). \tag{5.9}$$

Since  $\mathscr{W}$  is complete, we have  $\dim(W_i) = i + 1$ , and thus  $m_j \leq \dim(V_j) - 1$ . In particular,  $(\mathscr{W}, \mathbf{b})$  is saturated if, and only if, we have equality in (5.9) for all  $j$ . □

We need some further notation, which reflects that  $\mathbf{A}'$  is supported on intervals with gaps. For  $1 \leq j \leq r$ , let

$$I_j = (c_{j+1} + \kappa, c_j]. \tag{5.10}$$

Recall that we take  $\kappa$  small enough so that each  $I_j$  has length  $\geq 2/\log D$ , that is,  $\kappa \leq \min_j(c_j - c_{j+1}) - 2/\log D$ .

There is a natural analogue of the e-value (cf. Definition 3.5) for adapted systems.

**Definition 5.12** Given an adapted system  $(\mathscr{W}, \mathbf{b})$ , we define

$$e(\mathscr{W}, \mathbf{b}) = e(\mathscr{W}, \mathbf{b}; \mathscr{V}, \mathbf{c}, \boldsymbol{\mu}) := \sum_{i,j} \lambda([b_{i+1}, b_i] \cap I_j) \mathbb{H}_{\mu_j}(W_i) + \sum_i b_i,$$

where  $\lambda$  denotes the Lebesgue measure on  $\mathbb{R}$ .

Finally, we define

$$\delta(\mathbf{b}) = \max_{i,j} \{c_j - b_i : b_i \in I_j\}, \tag{5.11}$$

that is to say  $\delta(\mathbf{b})$  is the smallest non-negative real number with the property that

$$c_j - \delta(\mathbf{b}) \leq b_i \leq c_j \quad (1 \leq j \leq r, i \in I_j).$$

Adapted systems  $(\mathcal{W}, \mathbf{b})$  can, in a certain sense, be interpreted in terms of convex superpositions of pairs  $(\mathcal{V}', \mathbf{c})$ ,  $\mathcal{V}' \leq \mathcal{V}$  a subflag. The next lemma gives us a strict inequality analogous to condition (ii) of Proposition 5.5, unless  $\mathcal{W}$  is saturated and has a small value of  $\delta(\mathbf{b})$ , which corresponds to the convex superposition which gives rise to  $(\mathcal{W}, \mathbf{b})$  having weight  $\approx 1$  on the trivial subflag  $(\mathcal{V}, \mathbf{c})$ .

**Lemma 5.13** *Let  $(\mathcal{V}, \boldsymbol{\mu}, \mathbf{c})$  be a system satisfying conditions (i)–(ii) of Proposition 5.5. Let  $\varepsilon$  be as in condition (ii). Suppose that  $(\mathcal{W}, \mathbf{b})$  is an adapted system to  $(\mathcal{V}, \boldsymbol{\mu}, \mathbf{c})$  such that  $b_i$  lies in some set  $I_j$  for each  $i$ . Suppose, further, that  $\kappa$  is small enough in terms of  $\varepsilon$ , and that  $\kappa \leq \frac{1}{2} \min_j (c_j - c_{j+1})$ .*

- (a) *If  $(\mathcal{W}, \mathbf{b})$  is unsaturated, then  $e(\mathcal{W}, \mathbf{b}) \geq e(\mathcal{V}, \mathbf{c}, \boldsymbol{\mu}) + \varepsilon/2$ .*
- (b) *If  $(\mathcal{W}, \mathbf{b})$  is a saturated, then  $e(\mathcal{W}, \mathbf{b}) \geq e(\mathcal{V}, \mathbf{c}, \boldsymbol{\mu}) + \varepsilon\delta(\mathbf{b})/2$ .*

*Proof* We treat both parts together for most of the proof. Let  $m_j$  be defined by (5.8). In particular,  $m_0 = 0$  because  $c_1 = 1$ . Note that

$$\max_{i \in I_j} (c_j - b_i) = c_j - b_{m_j},$$

and let  $h$  be such that

$$\delta(\mathbf{b}) = c_h - b_{m_h}.$$

Without loss of generality, we may assume that  $b_{m_h} < c_h$ ; the case  $b_{m_h} = c_h$  will then follow by continuity.

Set  $b = b_{m_h}$  and note that

$$e(\mathcal{W}, \mathbf{b}) \geq \min \left\{ e(\mathcal{W}, \mathbf{b}') : \begin{array}{l} b'_i \in [c_{j+1} + \kappa, c_j] \text{ when } i \in (m_{j-1}, m_j] \text{ and } j \neq h, \\ b'_i \in [b, c_h] \text{ when } i \in (m_{h-1}, m_h), \ b'_{m_h} = b, \\ b'_1 \geq b'_2 \geq \dots \geq b'_s \end{array} \right\}.$$

The quantity  $e(\mathcal{W}, \mathbf{b}')$  is linear in each variable  $b'_i$  and the region over which we consider the above minimum is a polytope. As a consequence, the minimum of  $e(\mathcal{W}, \mathbf{b}')$  must occur at one of the vertices of the polytope. In particular, there are indices  $\ell_j \in (m_{j-1}, m_j]$  for  $j = 1, \dots, r$  such that

$$e(\mathcal{W}, \mathbf{b}) \geq e(\mathcal{W}, \mathbf{b}^*), \quad \text{where} \quad b_i^* = \begin{cases} c_j & \text{if } m_{j-1} < i \leq \ell_j, \\ c_{j+1} + \kappa & \text{if } \ell_j < i \leq m_j, \ j \neq h, \\ b & \text{if } \ell_h < i \leq m_h. \end{cases} \quad (5.12)$$

In fact, note that we must have  $\ell_h < m_h$  because  $b_{m_h}^* = b$  and we have assumed that  $b < c_h$ .

Using the linearity of  $e(\mathcal{W}, \cdot)$  once again, we find that

$$e(\mathcal{W}, \mathbf{b}^*) = \frac{c_h - b}{c_h - c_{h+1} - \kappa} e(\mathcal{W}, \mathbf{b}^{(1)}) + \frac{b - c_{h+1} - \kappa}{c_h - c_{h+1} - \kappa} e(\mathcal{W}, \mathbf{b}^{(2)}), \tag{5.13}$$

where  $b_i^{(1)} = b_i^{(2)} = b_i^*$  for  $i \in \{1, \dots, s\} \setminus (\ell_h, m_h]$ ,  $b_i^{(1)} = c_{h+1} + \kappa$  for  $i \in (\ell_h, m_h]$  and  $b_i^{(2)} = c_h$  for  $i \in (\ell_h, m_h]$ .

Fix  $\mathbf{b}' \in \{\mathbf{b}^{(1)}, \mathbf{b}^{(2)}\}$ . In addition, define the indices  $i_1, \dots, i_r$  by letting  $i_j = \ell_j$  when  $j \neq h$  or  $\mathbf{b}' = \mathbf{b}^{(1)}$ , while letting  $i_h = m_h$  when  $\mathbf{b}' = \mathbf{b}^{(2)}$ . We then have

$$b'_i = \begin{cases} c_j & \text{if } m_{j-1} < i \leq i_j, \\ c_{j+1} + \kappa & \text{if } i_j < i \leq m_j. \end{cases}$$

A straightforward calculation implies that

$$e(\mathcal{W}, \mathbf{b}') = e(\mathcal{V}', \mathbf{c}, \boldsymbol{\mu}) + S\kappa + (m_r - i_r)c_{r+1}, \tag{5.14}$$

where  $\mathcal{V}'$  is the subflag of  $\mathcal{V}$  with  $V'_j = W_{i_j}$  and

$$S = \sum_{j=1}^r (m_j - i_j - \mathbb{H}_{\mu_j}(W_{i_j})).$$

(Note that  $\mathcal{V}'$  is indeed a subflag since  $W_{i_j} \leq W_{m_j} \leq V_j$  by (5.9).)

If  $\mathcal{V}' = \mathcal{V}$ , we must have that  $W_{i_j} = V_j$  for all  $j$ . Since  $W_{i_j} \leq W_{m_j} \leq V_j$ , we infer that  $W_{m_j} = V_j$ , as well as that  $i_j = m_j$  for all  $j$ . In particular, the flag  $(\mathcal{W}, \mathbf{b})$  we started with must be saturated and  $S = 0$  (since  $i_j = m_j$  and  $\mathbb{H}_{\mu_j}(W_{i_j}) = \mathbb{H}_{\mu_j}(V_j) = 0$  for all  $j$ ).

We are now ready to complete the proof of both parts of the lemma.

(a) By the above discussion, if  $(\mathcal{W}, \mathbf{b})$  is unsaturated, then  $\mathcal{V}' \neq \mathcal{V}$ . Therefore, by assumption of condition (ii) of Proposition 5.5, we have

$$e(\mathcal{W}, \mathbf{b}') \geq e(\mathcal{V}, \mathbf{c}, \boldsymbol{\mu}) + \varepsilon$$

for  $\mathbf{b}' \in \{\mathbf{b}^{(1)}, \mathbf{b}^{(2)}\}$ . Inserting this inequality into (5.13) implies that

$$e(\mathcal{W}, \mathbf{b}^*) \geq e(\mathcal{V}, \mathbf{c}, \boldsymbol{\mu}) + \varepsilon + O(\kappa).$$

Since  $e(\mathcal{W}, \mathbf{b}) \geq e(\mathcal{W}, \mathbf{b}^*)$ , the proof of part (a) is complete by assuming that  $\kappa$  is small enough in terms of  $\varepsilon$ .

(b) Assume that  $(\mathcal{W}, \mathbf{b})$  is saturated. We can only have  $\mathcal{V}' = \mathcal{V}$  if  $i_h = m_h$ . Since  $\ell_h < m_h$ , this can only happen when  $\mathbf{b}' = \mathbf{b}^{(2)}$ . As a consequence, assuming again that  $\kappa$  is small enough in terms of  $\varepsilon$ , we have that

$$e(\mathcal{W}, \mathbf{b}') \geq e(\mathcal{V}, \mathbf{c}, \boldsymbol{\mu}) + 1_{\mathbf{b}'=\mathbf{b}^{(1)}} \cdot \varepsilon/2.$$

Inserting this into (5.13) yields the inequality

$$e(\mathcal{W}, \mathbf{b}^*) \geq e(\mathcal{V}, \mathbf{c}, \boldsymbol{\mu}) + \frac{c_h - b}{c_h - c_{h+1} - \kappa} \cdot \frac{\varepsilon}{2}.$$

Since  $b = c_h - \delta(\mathbf{b})$ ,  $0 < c_h - c_{h+1} - \kappa \leq 1$ , and  $e(\mathcal{W}, \mathbf{b}) \geq e(\mathcal{W}, \mathbf{b}^*)$ , we find that  $e(\mathcal{W}, \mathbf{b}) \geq e(\mathcal{V}, \mathbf{c}, \boldsymbol{\mu}) + \varepsilon\delta(\mathbf{b})/2$ . This completes the proof of part (b) of the lemma. □

### 5.4 Proof of the moment bound

In this subsection we prove Proposition 5.10. For a vector

$$\mathbf{n} = (n_0, n_1, n_2, \dots, n_r) \text{ with}$$

$$0 = n_0 \leq n_1 \leq \dots \leq n_r,$$

define the event

$$S(\mathbf{n}) = \{\mathbf{A}' : \#\mathbf{A}^j = n_j - n_{j-1} \quad (1 \leq j \leq r)\}.$$

When  $\mathbf{A}'$  lies in  $S(\mathbf{n})$ , we write

$$\mathbf{A}' = \{a_1, a_2, \dots, a_{n_r}\}, \quad a_1 > a_2 > \dots > a_{n_r},$$

so that

$$a_t \in \mathbf{A}^j \quad \text{if and only if} \quad n_{j-1} < t \leq n_j. \tag{5.15}$$

We may define, for any compatible  $\psi$ , the auxiliary function

$$\theta : [n_r] \rightarrow V_r \cap \{0, 1\}^k \quad \text{such that} \quad \theta(t) = \psi(a_t). \tag{5.16}$$

The salient property of  $\theta$  is that it is determined by the ordering of the elements in  $\mathbf{A}^j$  and not by the elements themselves. We denote by  $\Theta_{\mathbf{n}}$  the set of compatible functions  $\theta$ , that is, those functions satisfying

$$\theta(t) \in \text{Supp}(\mu_j) \quad \text{whenever} \quad t \leq n_j, \quad 1 \leq j \leq r. \quad (5.17)$$

In the event  $S(\mathbf{n})$ , if  $\psi$  is an compatible function and  $\theta$  is defined by (5.16), we have

$$w_{\mathbf{A}'}(\psi) = w_{\mathbf{n}}(\theta) := \prod_{j=1}^r \prod_{n_{j-1} < t \leq n_j} \mu_j(\theta(t)), \quad (5.18)$$

where the notation  $w_{\mathbf{n}}$  (in place of  $w_{\mathbf{A}}$ ) reflects the fact that  $w$  only depends on  $\theta$ , and not otherwise on  $\mathbf{A}$ . In this notation,

$$r_{\mathbf{A}'}(x) = \sum_{\substack{\theta \in \Theta_{\mathbf{n}} \\ \sum_t \theta(t) a_t - x \in \langle \mathbf{1} \rangle}} w_{\mathbf{n}}(\theta).$$

Writing  $r_{\mathbf{A}'}^p = r_{\mathbf{A}'}^{p-1} r_{\mathbf{A}'}$  and interchanging the order of summation, it follows that if  $\mathbf{A}'$  lies in  $S(\mathbf{n})$ , then

$$\begin{aligned} \sum_x r_{\mathbf{A}'}(x)^p &= \sum_{\theta \in \Theta_{\mathbf{n}}} \left( r_{\mathbf{A}'} \left( \sum_t a_t \theta(t) \right) \right)^{p-1} w_{\mathbf{n}}(\theta) \\ &= \sum_{\theta \in \Theta_{\mathbf{n}}} \left( \sum_{\substack{\theta' \in \Theta_{\mathbf{n}} \\ (5.20)}} w_{\mathbf{n}}(\theta') \right)^{p-1} w_{\mathbf{n}}(\theta), \end{aligned} \quad (5.19)$$

where the inner summation is over all compatible functions  $\theta'$  satisfying

$$\sum_t a_t (\theta'(t) - \theta(t)) \in \langle \mathbf{1} \rangle. \quad (5.20)$$

Similar to the argument in Sect. 4.2, we find a flag  $\mathscr{W}$  and special values of  $i$  which have the effect of isolating terms in the relation (5.20). With  $\theta, \theta', \mathbf{n}$  fixed, let

$$\Omega = \Omega(\theta, \theta') = \{\theta'(t) - \theta(t) : 1 \leq t \leq n_r\}$$

and

$$s = \dim (\text{Span}(\mathbf{1}, \Omega)) - 1.$$

We now choose a special basis of  $\text{Span}(\mathbf{1}, \Omega)$ . For each  $\omega \in \Omega$ , let

$$K_\omega = \min\{t : \theta'(t) - \theta(t) = \omega\},$$

and place a total ordering on  $\Omega$  by saying that  $\omega < \omega'$  if  $K_\omega < K_{\omega'}$ . Let  $\omega^1$  be the minimum element in  $\Omega \setminus \langle \mathbf{1} \rangle$ ,

$$\omega^2 = \min(\Omega \setminus \text{Span}(\mathbf{1}, \omega^1)), \dots, \omega^s = \min(\Omega \setminus \text{Span}(\mathbf{1}, \omega^1, \dots, \omega^{s-1})),$$

where  $s$  is such that  $\Omega \subset \text{Span}(\mathbf{1}, \omega^1, \dots, \omega^s)$ . Finally, let

$$\begin{aligned} W_j &= \text{Span}(\mathbf{1}, \omega^1, \dots, \omega^j), & \tau_j &= K_{\omega^j} \quad (1 \leq j \leq s), \\ \boldsymbol{\tau}(\theta, \theta', \mathbf{n}) &= (\tau_1, \dots, \tau_s), \end{aligned}$$

and form the flag

$$\mathscr{W} = \mathscr{W}(\theta, \theta', \mathbf{n}) : W_0 \leq W_1 \leq \dots \leq W_s.$$

We note that in the special case  $\theta = \theta'$ , we have  $s = 0$  and  $\mathscr{W}$  is a trivial flag with only one space  $W_0$ .

Now we divide up the sample space of  $\mathbf{A}'$  into events describing the rough size of the critical elements  $a_{\tau_j}$ . By construction,

$$a_{\tau_j} = \max\{a_t \in \mathbf{A}' : \theta'(t) - \theta(t) = \omega^j\}.$$

Similarly to Sect. 4, for  $1 \leq i \leq s$  let

$$b_i = 1 + \frac{\lceil \log a_{\tau_i} - \log D \rceil}{\log D} \quad \text{so that} \quad a_{\tau_i} \in (D^{b_i}/e, D^{b_i}]. \quad (5.21)$$

The definition of  $\mathbf{A}'$  implies that for each  $i$ , there is some  $j$  with  $b_i \in I_j = (c_{j+1} + \kappa, c_j]$ . Moreover, we have the implications

$$b_i > c_{j+1} \implies \tau_i \leq n_j \implies \omega^i = \theta(\tau_i) - \theta'(\tau_i) \in V_j,$$

where we used (5.17) to obtain the second implication. Since  $b_1 \geq b_2 \geq \dots \geq b_i$ , we infer the stronger relation

$$b_i > c_{j+1} \implies W_i \leq V_j. \quad (5.22)$$

Therefore, the pair  $(\mathscr{W}, \mathbf{b})$  is adapted to  $(\mathscr{V}, \mathbf{c}, \boldsymbol{\mu})$ .

Using the inequality  $(x + y)^{p-1} \leq x^{p-1} + y^{p-1}$  repeatedly, we may partition (5.19) according to the values of  $\mathscr{W}(\theta, \theta')$  and  $\tau(\theta, \theta')$ , obtaining (still assuming  $S(\mathbf{n})$ )

$$\sum_x r_{\mathbf{A}'}(x)^p \leq \sum_{\mathscr{W}, \tau, \theta} \left( \sum_{\substack{\theta' \in \Theta_{\mathbf{n}}, \\ \mathscr{W}(\theta, \theta', \mathbf{n}) = \mathscr{W}, \tau(\theta, \theta', \mathbf{n}) = \tau}} w_{\mathbf{n}}(\theta') \right)^{p-1} w_{\mathbf{n}}(\theta).$$

We need to separately consider other elements of  $\mathbf{A}'$  that lie in the intervals  $(D^{b_i}/e, D^{b_i}]$ , and so we define

$$\mathcal{B} = \{b_i : 1 \leq i \leq s\} \quad \text{and} \quad \ell = (\ell_b)_{b \in \mathcal{B}},$$

where  $\ell_b = \#(\mathbf{A}' \cap (D^b/e, D^b])$ .

By assumption,  $\sum_b \ell_b \geq s$ . It may happen that  $b_i = b_{i+1}$  for some  $i$ , in which case  $|\mathcal{B}| < s$ . With  $\mathbf{n}, \tau, \mathbf{b}, \ell$  all fixed, consider the event

$$E(\mathbf{b}, \tau, \mathbf{n}, \ell)$$

defined as the intersection of

- $S(\mathbf{n})$ ;
- $a_{\tau_i} \in (D^{b_i}/e, D^{b_i}]$  for all  $i$ ;
- $|\mathbf{A}' \cap (D^b/e, D^b]| = \ell_b$  for all  $b \in \mathcal{B}$ .

Taking expectations over  $\mathbf{A}'$ , we get

$$\begin{aligned} & \mathbb{E} \left[ \mathbf{1}_{\mathbf{A}' \in S(\mathbf{n}) \cap \mathcal{E}^*} \sum_x r_{\mathbf{A}'}(x)^p \right] \\ & \leq \mathbb{E} \left[ \sum_{\substack{\mathscr{W}, \tau, \mathbf{b}, \theta, \ell \\ \ell_b \leq D^{b_i/2}/100 \ \forall b \in \mathcal{B}}} w_{\mathbf{n}}(\theta) \left( \sum_{\substack{\theta' \in \Theta_{\mathbf{n}}, \\ \mathscr{W}(\theta, \theta', \mathbf{n}) = \mathscr{W}, \\ \tau(\theta, \theta', \mathbf{n}) = \tau}} w_{\mathbf{n}}(\theta') \right)^{p-1} \mathbf{1}_{E(\mathbf{b}, \tau, \mathbf{n}, \ell)} \right], \end{aligned}$$

where the condition that  $\ell_b \leq D^{b_i/2}/100$  comes from the fact that we taking expectations over  $\mathbf{A}' \in \mathcal{E}^*$ . By Hölder’s inequality with exponents  $\frac{1}{p-1}, \frac{1}{2-p}$ , this implies that



$$\begin{aligned}
 & \mathbb{E} \left[ \mathbf{1}_{\mathbf{A}' \in \mathcal{S}(\mathbf{n}) \cap \mathcal{E}^*} \sum_x r_{\mathbf{A}'}(x)^p \right] \\
 & \leq \sum_{\substack{\mathscr{W}, \boldsymbol{\tau}, \mathbf{b}, \theta, \boldsymbol{\ell} \\ \ell_b \leq D^{b/2}/100 \ \forall b \in \mathcal{B}}} w_{\mathbf{n}}(\theta) \mathbb{P}(E(\mathbf{b}, \boldsymbol{\tau}, \mathbf{n}, \boldsymbol{\ell}))^{2-p} \times \\
 & \quad \times \left\{ \sum_{\substack{\theta' \in \Theta_{\mathbf{n}} \\ \mathscr{W}(\theta, \theta', \mathbf{n}) = \mathscr{W} \\ \boldsymbol{\tau}(\theta, \theta', \mathbf{n}) = \boldsymbol{\tau}}} w_{\mathbf{n}}(\theta') \mathbb{P}[E(\mathbf{b}, \boldsymbol{\tau}, \mathbf{n}, \boldsymbol{\ell}) \wedge (5.20)] \right\}^{p-1}. \tag{5.23}
 \end{aligned}$$

*Claim.* Let  $\ell_b \leq D^{b/2}/100$  for all  $b \in \mathcal{B}$ . Then we have

$$\mathbb{P}((5.20) \mid E(\mathbf{b}, \boldsymbol{\tau}, \mathbf{n}, \boldsymbol{\ell})) \ll D^{-(b_1 + \dots + b_s)} e^{\sum_b \ell_b}. \tag{5.24}$$

*Proof of Claim* Let us begin by analyzing the event  $E(\mathbf{b}, \boldsymbol{\tau}, \mathbf{n}, \boldsymbol{\ell})$  we are conditioning on. Consider the set  $\bigcup_j (D^{c_{j+1} + \kappa}, D^{c_j}] \setminus \bigcup_{b \in \mathcal{B}} (D^b/e, D^b]$ . There is a unique way to write it as  $\bigcup_{m=1}^M I_m$ , where the sets  $I_m$  are intervals of the form  $(A, B]$  with their closures  $\bar{I}_m$  mutually disjoint. Now, the event  $E(\mathbf{b}, \boldsymbol{\tau}, \mathbf{n}, \boldsymbol{\ell})$  is equivalent to there being mutually disjoint sets of consecutive integers  $\mathcal{I}_m$  ( $1 \leq m \leq M$ ) and  $\mathcal{J}_b$  ( $b \in \mathcal{B}$ ) such that:

- The sets  $\mathcal{I}_m$  ( $1 \leq m \leq M$ ) and  $\mathcal{J}_b$  ( $b \in \mathcal{B}$ ) together form a partition of the set  $[n_r]$ ;
- For all  $m \in \{1, \dots, M\}$ , we have  $a_n \in I_m$  if and only if  $n \in \mathcal{I}_m$ ;
- For all  $b \in \mathcal{B}$ , we have  $a_n \in (D^b/e, D^b]$  if and only if  $n \in \mathcal{J}_b$ ;
- $\tau_i \in \mathcal{J}_{b_i}$  for all  $i$ ;
- $|\mathcal{J}_b| = \ell_b$  for all  $b \in \mathcal{B}$ .

The above discussion allows us to describe the distribution law of  $\mathbf{A}'$  under the event  $E(\mathbf{b}, \boldsymbol{\tau}, \mathbf{n}, \boldsymbol{\ell})$ : given a choice of the intervals  $\mathcal{I}_m$  and  $\mathcal{J}_b$ , we construct independent logarithmic random sets  $\mathbf{A}_m^*$  on  $I_m$  and  $\tilde{A}_b$  on  $(D^b/e, D^b]$  such that  $\#\mathbf{A}' \cap I_m = \#\tilde{\mathcal{I}}_m$  for all  $m$  and  $\#\tilde{A}_b = \ell_b$  for all  $b$ . Then  $\mathbf{A}'$  is the union of all  $\mathbf{A}_m^*$ 's and all  $\tilde{\mathbf{A}}_b$ 's.

Having explained how the distribution of  $\mathbf{A}'$  looks like under the event  $E(\mathbf{b}, \boldsymbol{\tau}, \mathbf{n}, \boldsymbol{\ell})$ , let us now prove our claim. We argue as in the proof of Proposition (4.4). Relation (5.20) implies

$$\sum_{i=1}^s \omega^i a_{\tau_i} + \sum_{t \notin \{\tau_1, \dots, \tau_s\}} a_t(\theta'(t) - \theta(t)) = a_0 \mathbf{1}$$

for some  $a_0 \in \mathbb{Z}$ . Since  $\mathbf{1}, \omega^1, \dots, \omega^s$  are linearly independent, this uniquely determines their coefficients  $a_0, a_{\tau_1}, \dots, a_{\tau_s}$  in terms of the other  $a_i$ 's. For each  $b \in \mathcal{B}$ , let

$$m_b = \#\{i : b_i = b\} \quad \text{and} \quad N_b = \#(\mathbb{Z} \cap (D^b/e, D^b]) = (1 - 1/e)D^b + O(1).$$

Then, given  $\mathbf{A}_m^*$  for all  $m$  and  $b \in \mathcal{B}$ , there are at most

$$\begin{aligned} \binom{N_b}{\ell_b - m_b} &\leq \frac{N_b^{\ell_b - m_b}}{(\ell_b - m_b)!} \\ &< \ell_b^{m_b} \cdot \frac{((1 - 1/e)D)^{b(\ell_b - m_b)}}{\ell_b!} \ll \frac{D^{b(\ell_b - m_b)}}{\ell_b!} \end{aligned}$$

choices for  $\tilde{\mathbf{A}}_b$  (since  $m_b$  of each element are determined by the remaining  $\ell_b - m_b$  elements and by the elements of the  $\mathbf{A}_m^*$  that we have fixed), where we used that  $\ell_b^{m_b} \leq \ell_b^k \ll (1 - 1/e)^{-\ell_b}$ . In addition, Lemma A.4 implies that the probability of occurrence of a given set  $X_b \subset \mathbb{Z} \cap (D^b/e, D^b]$  as the set  $\tilde{\mathbf{A}}_b$ , conditionally to the event that  $\#\tilde{\mathbf{A}}_b = \ell_b$ , is

$$\begin{aligned} &\ll \frac{\ell_b!}{(\sum_{D^b/e < m \leq D^b} 1/(m - 1))^{\ell_b}} \prod_{x \in X_b} \frac{1}{x} \prod_{D^b/e < m \leq D^b} \left(1 - \frac{1}{m}\right) \\ &\ll \frac{\ell_b!}{(D^b/e)^{\ell_b}}. \end{aligned}$$

Putting the above estimates together, we conclude that

$$\mathbb{P}(\text{(5.20)} \mid E(\mathbf{b}, \boldsymbol{\tau}, \mathbf{n}, \boldsymbol{\ell})) \ll \prod_{b \in \mathcal{B}} \frac{e^{\ell_b}}{D^{b m_b}} = D^{-(b_1 + \dots + b_s)} e^{\sum_b \ell_b},$$

upon noticing that  $\sum_{b \in \mathcal{B}} m_b b = \sum_i b_i$ . This proves our claim that (5.24) holds. □

In the light of (5.24), relation (5.23) becomes

$$\begin{aligned} &\mathbb{E} \left[ \mathbf{1}_{\mathbf{A}' \in \mathcal{S}(\mathbf{n}) \cap \mathcal{E}^*} \sum_x r_{\mathbf{A}'}(x)^p \right] \\ &\ll \sum_{\mathcal{W}, \boldsymbol{\tau}, \mathbf{b}, \boldsymbol{\ell}} D^{-(p-1) \sum_j b_j} e^{\sum_b \ell_b} \\ &\quad \times \mathbb{E} \left[ \sum_{\theta \in \Theta_{\mathbf{n}}} w_{\mathbf{n}}(\theta) \left( \sum_{\substack{\theta' \in \Theta_{\mathbf{n}} \\ \mathcal{W}(\theta, \theta', \mathbf{n}) = \mathcal{W} \\ \boldsymbol{\tau}(\theta, \theta', \mathbf{n}) = \boldsymbol{\tau}}} w_{\mathbf{n}}(\theta') \right)^{p-1} \mathbf{1}_{E(\mathbf{b}, \boldsymbol{\tau}, \mathbf{n}, \boldsymbol{\ell})} \right]. \end{aligned} \tag{5.25}$$

To evaluate the bracketed expression, first recall the definition (5.18) of  $w_{\mathbf{n}}(\theta')$ , and note that the conditions  $\mathscr{W}(\theta, \theta', \mathbf{n}) = \mathscr{W}$ ,  $\tau(\theta, \theta', \mathbf{n}) = \tau$  together imply that

$$\theta'(t) - \theta(t) \in W_i \quad (1 \leq t < \tau_{i+1}, 0 \leq i \leq s),$$

where we have defined  $\tau_0 := 0$  and  $\tau_{s+1} := n_r + 1$ . For brevity, write

$$T_{i,j} = (n_{j-1}, n_j] \cap [\tau_i, \tau_{i+1}) \cap \mathbb{N}, \quad (0 \leq i \leq s, 1 \leq j \leq r).$$

Some of these sets are empty. In any case, we have

$$\sum_{\substack{\theta' \in \Theta_{\mathbf{n}} \\ \mathscr{W}(\theta, \theta', \mathbf{n}) = \mathscr{W} \\ \tau(\theta, \theta', \mathbf{n}) = \tau}} w_{\mathbf{n}}(\theta') \leq \prod_{\substack{0 \leq i \leq s \\ 1 \leq j \leq r}} \prod_{t \in T_{i,j}} \mu_j(\theta(t) + W_i). \quad (5.26)$$

From (5.18), and the fact that the discrete intervals  $T_{i,j}$  are disjoint and cover  $[n_r]$ , we have

$$w_{\mathbf{n}}(\theta) = \prod_{i,j} \prod_{t \in T_{i,j}} \mu_j(\theta(t)).$$

With these observations, we conclude that

$$\begin{aligned} \sum_{\theta \in \Theta_{\mathbf{n}}} w_{\mathbf{n}}(\theta) & \left( \sum_{\substack{\theta' \in \Theta_{\mathbf{n}} \\ \mathscr{W}(\theta, \theta', \mathbf{n}) = \mathscr{W} \\ \tau(\theta, \theta', \mathbf{n}) = \tau}} w_{\mathbf{n}}(\theta') \right)^{p-1} \\ & \leq \sum_{\theta \in \Theta_{\mathbf{n}}} \prod_{i,j} \prod_{t \in T_{i,j}} \mu_j(\theta(t)) \mu_j(W_i + \theta(t))^{p-1} \\ & = \prod_{i,j} \eta(i, j, p, \mathscr{W})^{|T_{i,j}|}, \end{aligned} \quad (5.27)$$

where

$$\eta(i, j, p, \mathscr{W}) := \sum_{\omega \in \text{Supp}(\mu_j)} \mu_j(\omega) \mu_j(W_i + \omega)^{p-1}. \quad (5.28)$$

Substituting into (5.25), and summing over  $\mathbf{n}$ , we get

$$\begin{aligned} & \mathbb{E}\left[1_{\mathbf{A}' \in \mathcal{E}^*} \sum_x r_{\mathbf{A}'}(x)^p\right] \\ & \ll \sum_{\mathscr{W}, \mathbf{b}} D^{-(p-1)\sum_j b_j} \sum_{\boldsymbol{\tau}, \mathbf{n}, \ell} e^{\sum_b \ell_b} \mathbb{E}\left[1_{E(\mathbf{b}, \boldsymbol{\tau}, \mathbf{n}, \ell)} \prod_{i,j} \eta(i, j, p, \mathscr{W})^{|T_{i,j}|}\right]. \end{aligned} \tag{5.29}$$

If  $V_j \leq W_i$ , then  $\mu_j(W_i + \omega) = 1$  for all  $\omega$  and thus  $\eta(i, j, p, \mathscr{W}) = 1$ . For all  $i, j, p, \mathscr{W}$  we have  $\eta(i, j, p, \mathscr{W}) \leq 1$ . Thus, we require lower bounds on  $|T_{i,j}|$  in the case  $V_j \not\leq W_i$ .

*Claim.* Assume that  $E(\mathbf{b}, \boldsymbol{\tau}, \mathbf{n}, \ell)$  holds. Given  $i$  such that  $b_{i+1} < b_i$  and  $j \in \{1, \dots, r\}$ , define

$$M_{i,j} := (D^{c_{j+1}+\kappa}, D^{c_j}/e] \cap (D^{b_{i+1}}, D^{b_i}/e]$$

Then

$$\{t : a_t \in M_{i,j}\} \subset T_{i,j}. \tag{5.30}$$

*Proof of Claim* Let  $t$  be such that  $a_t \in M_{i,j}$ . In particular,

$$D^{b_{i+1}} < a_t \leq D^{b_i}/e.$$

This relation and the definition of  $b_i$  in (5.21) imply that  $a_{\tau_{i+1}} < a_t < a_{\tau_i}$  and hence  $\tau_i < t < \tau_{i+1}$ , where we used that  $a_1 > a_2 > \dots > a_{n_r}$ . In addition, since  $D^{c_{j+1}+\kappa} < a_t \leq D^{c_j}$ , we have that  $a_t \in \mathbf{A}^j$ . Thus,  $n_{j-1} < t \leq n_j$  by (5.15). This completes the proof of the claim.  $\square$

A direct consequence of (5.30) is that

$$|T_{i,j}| \geq |\mathbf{A}' \cap M_{i,j}|.$$

Combining this inequality with (5.29), we get

$$\begin{aligned} \mathbb{E}\left[1_{\mathbf{A}' \in \mathcal{E}^*} \sum_x r_{\mathbf{A}'}(x)^p\right] & \ll \sum_{\mathscr{W}, \mathbf{b}} D^{-(p-1)\sum_j b_j} \\ & \times \sum_{\mathbf{n}, \boldsymbol{\tau}, \ell} e^{\sum_b \ell_b} \mathbb{E}\left[1_{E(\mathbf{b}, \boldsymbol{\tau}, \mathbf{n}, \ell)} \prod_{i,j} \eta(i, j, p, \mathscr{W})^{|\mathbf{A}' \cap M_{i,j}|}\right]. \end{aligned}$$

Fix  $\mathbf{b}$  and  $\mathscr{W}$ , and let  $E'(\mathbf{b}, \ell)$  be the event that  $|\mathbf{A}' \cap (D^b/e, D^b]| = \ell_b$  for all  $b \in \mathcal{B}$ . Given  $\mathbf{A}' \in E'(\mathbf{b}, \ell)$ , we have at most  $\prod_b \ell_b \leq e^{\sum_b \ell_b}$  choices for

$\tau_1, \dots, \tau_s$ . Hence,

$$\begin{aligned} & \sum_{\mathbf{n}, \tau, \ell} e^{\sum_b \ell_b} \mathbb{E} \left[ 1_{E(\mathbf{b}, \tau, \mathbf{n}, \ell)} \eta(i, j, p, \mathscr{W})^{|\mathbf{A}' \cap M_{i,j}|} \right] \\ & \leq \sum_{\mathbf{n}, \ell} e^{2 \sum_b \ell_b} \mathbb{E} \left[ 1_{S(\mathbf{n})} 1_{E'(\mathbf{b}, \ell)} \prod_{i,j} \eta(i, j, p, \mathscr{W})^{|\mathbf{A}' \cap M_{i,j}|} \right]. \end{aligned}$$

Since the events  $S(\mathbf{n})$  are mutually disjoint, we arrive at the inequality

$$\begin{aligned} & \mathbb{E} \left[ 1_{\mathbf{A}' \in \mathcal{E}^*} \sum_x r_{\mathbf{A}'}(x)^p \right] \\ & \leq \sum_{\mathscr{W}, \mathbf{b}} D^{-(p-1) \sum_j b_j} \mathbb{E} \left[ \prod_{b \in \mathcal{B}} e^{2|\tilde{\mathbf{A}}_b|} \prod_{i,j} \eta(i, j, p, \mathscr{W})^{|\mathbf{A}' \cap M_{i,j}|} \right]. \quad (5.31) \end{aligned}$$

Next, we estimate the right hand side of (5.31). The intervals  $M_{i,j}$  and  $(D^b/e, D^b]$  are mutually disjoint by (5.30), hence the quantities  $|\mathbf{A}' \cap M_{i,j}|$  and  $|\tilde{\mathbf{A}}_b|$  are independent. Using Lemma A.3, we obtain

$$\begin{aligned} & \mathbb{E} \left[ \prod_{b \in \mathcal{B}} e^{2|\tilde{\mathbf{A}}_b|} \prod_{i,j} \eta(i, j, p, \mathscr{W})^{|\mathbf{A}' \cap M_{i,j}|} \right] \\ & \leq \exp \left\{ \sum_{b \in \mathcal{B}} \sum_{D^b/e < m \leq D^b} \frac{2e-1}{m} + \sum_{i,j} (\eta(i, j, p, \mathscr{W}) - 1) \sum_{m \in M_{i,j}} \frac{1}{m} \right\} \\ & \ll \exp \left\{ \sum_{i,j} (\eta(i, j, p, \mathscr{W}) - 1) \sum_{m \in M_{i,j}} \frac{1}{m} \right\}. \end{aligned}$$

Recall that  $I_j = (c_{j+1} + \kappa, c_j]$ , define

$$G_i = G_i(\mathbf{b}) = (b_{i+1}, b_i],$$

and recall that  $\lambda$  denotes the Lebesgue measure on  $\mathbb{R}$ . Then, by the definition of  $M_{i,j}$ , we have

$$\sum_{m \in M_{i,j}} \frac{1}{m} = \lambda(I_j \cap G_i) \log D + O(1).$$

Substituting into the definition of  $e()$  (Definition 5.12), this gives

$$\mathbb{E} \left[ 1_{\mathbf{A}' \in \mathcal{E}^*} \sum_x r_{\mathbf{A}'}(x)^p \right] \ll \sum_{\mathscr{W}, \mathbf{b}} D^{-E(p, \mathscr{W}, \mathbf{b})}, \quad (5.32)$$

where

$$\begin{aligned}
 E(p, \mathscr{W}, \mathbf{b}) &:= (p - 1) \sum_j b_j - \sum_{i,j} (\eta(i, j, p, \mathscr{W}) - 1) \lambda(I_j \cap G_i) \\
 &= (p - 1)e(\mathscr{W}, \mathbf{b}) - \sum_{i,j} [\eta(i, j, p, \mathscr{W}) - 1 + (p - 1)\mathbb{H}_{\mu_j}(W_i)] \lambda(I_j \cap G_i).
 \end{aligned}$$

Recall the definition (5.28) of  $\eta(i, j, p, \mathscr{W})$ . If  $W_i \geq V_j$ , then  $\mu_j(W_i + x) = 1$  whenever  $x \in \text{Supp}(\mu_j)$ , and so in this case  $\eta(i, j, p, \mathscr{W}) = 1$ . Since  $\mathbb{H}_{\mu_j}(W_i) = 0$  in this case, we have

$$\eta(i, j, p, \mathscr{W}) - 1 + (p - 1)\mathbb{H}_{\mu_j}(W_i) = 0 \quad (V_j \leq W_i). \tag{5.33}$$

For any fixed  $i, j, \mathscr{W}$ , we have

$$\left. \frac{d}{dp} \eta(i, j, p, \mathscr{W}) \right|_{p=1} = -\mathbb{H}_{\mu_j}(W_i),$$

and so

$$\eta(i, j, p, \mathscr{W}) - 1 + (p - 1)\mathbb{H}_{\mu_j}(W_i) \ll (p - 1)^2 \quad (V_j \not\leq W_i). \tag{5.34}$$

We deduce from (5.32), (5.33) and (5.34) that

$$E(p, \mathscr{W}, \mathbf{b}) = (p - 1)e(\mathscr{W}, \mathbf{b}) - \sum_{i,j: V_j \not\leq W_i} \lambda(I_j \cap G_i) O((p - 1)^2). \tag{5.35}$$

To continue, we separate two cases.

*Case 1.*  $(\mathscr{W}, \mathbf{b})$  is unsaturated.

In the above case, Lemma 5.13(a) implies that  $e(\mathscr{W}, \mathbf{b}) \geq e(\mathscr{V}, \mathbf{c}, \boldsymbol{\mu}) + \varepsilon/2$ . Consequently,

$$\begin{aligned}
 E(p, \mathscr{W}, \mathbf{b}) &\geq (p - 1)e(\mathscr{V}, \mathbf{c}, \boldsymbol{\mu}) + \frac{(p - 1)\varepsilon}{2} + O((p - 1)^2) \\
 &\geq (p - 1)e(\mathscr{V}, \mathbf{c}, \boldsymbol{\mu}) + \frac{(p - 1)\varepsilon}{4},
 \end{aligned}$$

provided that  $p - 1$  is small enough in terms of  $\varepsilon$  (and  $k$ ).

Since there are  $O(1)$  choices for  $\mathscr{W}$  and  $\log^{O(1)} D$  choices for  $\mathbf{b}$ , the contribution of such flags to the right hand side of (5.32) is

$$\sum_{(\mathscr{W}, \mathbf{b}) \text{ unsaturated}} D^{-E(p, \mathscr{W}, \mathbf{b})} \ll D^{-(p-1)e(\mathscr{V}, \mathbf{c}, \boldsymbol{\mu})}. \tag{5.36}$$

*Case 2.*  $(\mathscr{W}, \mathbf{b})$  is saturated. (Recall from Definition 5.11 that  $(\mathscr{W}, \mathbf{b})$  is called saturated when  $s = \dim(V_r) - 1$  and for all  $j \leq r$ , there are exactly  $\dim V_j - 1$  values of  $i$  with  $b_i > c_{j+1}$ .)

Fix for the moment a pair  $(i, j)$  such that

$$V_j \not\leq W_i \quad \text{and} \quad \lambda(I_j \cap G_i) > 0. \tag{5.37}$$

The second condition is equivalent to knowing that

$$b_i > c_{j+1} \quad \text{and} \quad b_{i+1} < c_j.$$

In particular, we have  $W_i \leq V_j$  by (5.22). Note though that we have assumed  $V_j \not\leq W_i$ . Therefore,  $W_i < V_j$ . Since  $\dim(W_i) = i + 1$ , we infer that

$$i \leq \dim(V_j) - 2.$$

Since we have assumed that  $(\mathscr{W}, \mathbf{b})$  is saturated, the above inequality implies that  $b_{i+1} > c_{j+1}$ . Recalling the definition (5.11) of  $\delta(\mathbf{b})$ , we conclude that

$$b_{i+1} \geq c_j - \delta(\mathbf{b}).$$

This implies that  $G_i \cap I_j \subset [c_j - \delta(\mathbf{b}), c_j]$  for any pair  $(i, j)$  satisfying (5.37). As a consequence,

$$\sum_{i: V_j \not\leq W_i} \lambda(I_j \cap G_i) \leq \delta(\mathbf{b}) \quad (1 \leq j \leq r).$$

Since we also have that  $e(\mathscr{W}, \mathbf{b}) \geq e(\mathscr{V}, \mathbf{c}, \boldsymbol{\mu}) + \varepsilon\delta(\mathbf{b})/2$  by Lemma (5.13)(b), it follows that

$$\begin{aligned} E(p, \mathscr{W}, \mathbf{b}) &\geq (p - 1)e(\mathscr{V}, \mathbf{c}, \boldsymbol{\mu}) + \varepsilon\delta(\mathbf{b})/2 + O((p - 1)^2\delta(\mathbf{b})) \\ &\geq (p - 1)e(\mathscr{V}, \mathbf{c}, \boldsymbol{\mu}) + \varepsilon\delta(\mathbf{b})/4, \end{aligned} \tag{5.38}$$

provided that  $p - 1$  is small enough compared to  $\varepsilon$ .

Using (5.38), we see that the contribution of saturated flags to the right hand side of (5.32) is

$$\sum_{(\mathscr{W}, \mathbf{b}) \text{ saturated}} D^{-E(p, \mathscr{W}, \mathbf{b})} \ll D^{-(p-1)\epsilon(\mathscr{V}, \mathbf{c}, \boldsymbol{\mu})} \sum_{s=0}^r \sum_{b_1, \dots, b_s} D^{-(p-1)\epsilon\delta(\mathbf{b})/4},$$

where we used that there are  $O(1)$  choices for  $\mathscr{W}$ . Recall (5.21), which implies that the numbers  $b_i$  are restricted to the set  $\{m/\log D : m \in \mathbb{N}\}$ . Thus the number of  $\mathbf{b}$  with  $\delta(\mathbf{b}) = m/\log D$  is at most  $(m + 1)^s$  and

$$\sum_{s=0}^r \sum_{b_1, \dots, b_s} D^{-(p-1)\epsilon\delta(\mathbf{b})/4} \leq \sum_{s=0}^r \sum_{m \geq 0} (m + 1)^s e^{-(p-1)(\epsilon/4)m} \ll_{\epsilon, p} 1.$$

We thus conclude that

$$\sum_{(\mathscr{W}, \mathbf{b}) \text{ saturated}} D^{-E(p, \mathscr{W}, \mathbf{b})} \ll D^{-(p-1)\epsilon(\mathscr{V}, \mathbf{c}, \boldsymbol{\mu})}.$$

If we combine the above inequality with (5.36) and (5.32), we establish Proposition 5.10. □

### 6 An argument of Maier and Tenenbaum

The aim of this section is to prove Proposition 5.7. The reader may care to recall the statement of that proposition now, as well as the definition of a compatible map (Definition 5.8). As in the previous section, the system  $(\mathscr{V}, \mathbf{c}, \boldsymbol{\mu})$  is fixed, and satisfies conditions (i)–(iii) of Proposition 5.5. We also fix a basis  $\{\mathbf{1}, \omega^1, \dots, \omega^d\}$  of  $V_r$  such that  $V_j = \text{Span}(\mathbf{1}, \omega^1, \dots, \omega^{\dim(V_j)-1})$  for each  $j$  and such that  $\omega^i \in \{0, 1\}^k$  for each  $i$ . Denote  $\Omega = \text{Supp}(\mu_r) = V_r \cap \{0, 1\}^k$ .

We begin with an observation related to the solvability of (4.12), which we recall here for the convenience of the reader:

$$\sum_{j=1}^r K_j \omega^j = - \sum_{\omega} \omega \sum_{a' \in B'_\omega} a' \pmod{\mathbf{1}}. \tag{6.1}$$

Let  $\Lambda$  denote the  $\mathbb{Z}$ -span of  $\mathbf{1}, \omega^1, \dots, \omega^d$  (that is, the lattice generated by  $\mathbf{1}, \omega^1, \dots, \omega^d$ ). Every vector  $\omega \in \Omega$  is a rational combination of the basis elements  $\mathbf{1}, \omega^1, \dots, \omega^d$ . Hence, there is some  $M \in \mathbb{N}$  such that  $M\omega \in \Lambda$  for each  $\omega \in \Omega$ . In particular, note that the right-hand side of (6.1) lies generically in the lattice  $\Lambda/M = \{x/M : x \in \Lambda\}$ . However, we must ensure that (6.1) is solvable with  $K_1, \dots, K_r \in \mathbb{Z}$ . Equivalently, the right-hand side of (6.1) must



lie in  $\Lambda$ , which can be guaranteed when the coefficients of all vectors  $\omega$  in it lie in  $M\mathbb{Z}$ .

In this section, implied constants in  $O()$  and  $\ll$  notations may depend on the system  $(\mathcal{V}, \mathbf{c}, \boldsymbol{\mu})$  and basis  $\omega^1, \dots, \omega^d$ ; in particular, on  $k, d$  and  $M$ .

### 6.1 The sets $\mathcal{L}_i(\mathbf{A})$ and lower bounds for their size

The main statement of this subsection, Proposition 6.2, is a variant of Proposition 5.9, where we stipulate that all elements lie in  $\Lambda$ . This will later ensure that (6.1) is solvable with  $K_1, \dots, K_r \in \mathbb{Z}$ .

Fix  $\kappa > 0$  satisfying  $\kappa \leq \frac{\kappa^*}{2}$ , where  $\kappa^*$  is the constant from Proposition 5.9. In particular,  $\kappa \leq 1/2$ . We introduce the sets

$$I_i(D) := \bigcup_{j=1}^r (D^{c_{j+1}}, D^{c_j(1-\kappa/i)}], \quad i = 1, 2, \dots \tag{6.2}$$

Thus each  $I_i(D)$  is simply a union of  $r$  intervals in  $\Lambda$ , and we have the nesting

$$I_1(D) \subset I_2(D) \subset \dots \subset (D^c, D].$$

For any  $\omega \in V_r$  we denote by  $\bar{\omega}$  the projection of  $\omega$  onto

$$\bar{V}_r := V_r / \langle \mathbf{1} \rangle = \text{Span}\{\omega^1, \dots, \omega^d\}.$$

In addition let  $\bar{\psi}(a) = \overline{\psi(a)}$  for  $a \in \mathbf{A}$ .

The reader may wish to recall the definition of nondegenerate (Definition 5.6) and compatible (Definition 5.8) maps.

**Definition 6.1** Write  $\mathcal{L}_i(\mathbf{A})$  for the set of all  $\sum_{a \in \mathbf{A}} a \bar{\psi}(a)$  that lie in  $\Lambda$ , where  $\psi$  ranges over all nondegenerate, compatible maps supported on  $I_i(D)$ .

**Proposition 6.2** Let  $\delta > 0$  and  $i \in \mathbb{N}$ , and let  $D$  be sufficiently large in terms of  $\delta$ . Then with probability at least  $1 - \delta$  in the choice of  $\mathbf{A} \cap I_i(D)$ ,

$$|\mathcal{L}_i(\mathbf{A})| \gg \delta^\alpha D^{(1-\kappa/i) \sum_j c_j \dim(V_j/V_{j-1})}, \tag{6.3}$$

where  $\alpha$  is a positive constant depending at most on  $(\mathcal{V}, \mathbf{c}, \boldsymbol{\mu})$ .

*Proof* Let

$$\begin{aligned}
 I'_i(D) &= \bigcup_{j=1}^r (D^{(c_{j+1} + \kappa^*)(1-\kappa/i)}, D^{c_j(1-\kappa/i)}) \\
 &\subset \bigcup_{j=1}^r (D^{c_{j+1}(1+\kappa/2)}, D^{c_j(1-\kappa/i)}) \subset I_i(D),
 \end{aligned}$$

where the first inclusion follows by noticing that

$$(c_{j+1} + \kappa^*)(1 - \kappa/i) \geq c_{j+1}(1 + \kappa/2)$$

for  $c_{j+1} \in [0, 1]$ ,  $0 \leq \kappa \leq \kappa^*/2 \leq 1/2$  and  $i \geq 1$ . Write  $\mathcal{L}'_i(\mathbf{A})$  for the set of all  $\sum_{a \in \mathbf{A}} a \bar{\psi}(a)$ , where  $\psi$  ranges over all nondegenerate, compatible maps supported on  $I'_i(D)$ , but without the stipulation that the sum is in  $\Lambda$ . We now apply Proposition 5.9 with  $D$  replaced by  $D^{1-\kappa/i}$  and  $\delta$  replaced by  $\delta/2$  to conclude that

$$|\mathcal{L}'_i(\mathbf{A})| \gg \delta^\alpha D^{(1-\kappa/i) \sum_j c_j \dim(V_j/V_{j-1})}$$

with probability at least  $1 - \delta/2$ , where  $\alpha = 1/(p - 1)$  with  $p$  as in Proposition 5.9.

We now use the elements of  $\mathbf{A} \cap (I_i(D) \setminus I'_i(D))$  to create many sums  $\sum_{a \in \mathbf{A}} \bar{\psi}(a)$  which do lie in  $\Lambda$ . Let  $G := (D^{c_{r+1}(1-\kappa/i)}, \delta^{-1} D^{c_{r+1}(1-\kappa/i)})$ , which is a subset of  $I_i(D) \setminus I'_i(D)$ . Let  $\mathcal{E}$  be the event that  $\mathbf{A} \cap G$  contains at least  $2^k$  elements that are  $\equiv m \pmod{M}$  for each  $m \in \{1, \dots, M\}$ . Lemma A.2 (applied with  $B = \{b \in \mathbb{Z} \cap G : b \equiv m \pmod{M}\}$  and  $\varepsilon = 1/3$ ) implies that if  $\delta$  is sufficiently small then  $\mathbb{P}(\mathcal{E}) \geq 1 - \delta/2$ .

Assume now that we are in the event  $\mathcal{E}$ . Let us fix a set  $\mathcal{K} \subset \mathbf{A} \cap G$  that contains exactly  $2^k$  elements that are  $\equiv m \pmod{M}$  for each  $m \in \{1, \dots, M\}$ . Take any nondegenerate, compatible function  $\psi : \mathbf{A} \rightarrow \{0, 1\}^k$  supported on  $I'_i(D)$ , and write

$$\sum_{a \in I'_i(D)} a \psi(a) = \sum_{\omega \in \Omega} \omega N_\omega.$$

Recall that  $\text{Supp}(\mu_r) = V_r \cap \{0, 1\}^k$  by condition (iii) of Proposition 5.5. Hence, for each  $\omega \in \Omega$ , we may find an element  $a_\omega \in \mathcal{K}$  satisfying  $a_\omega \equiv -N_\omega \pmod{M}$ . Setting  $\psi_0(a_\omega) = \omega$  for each  $\omega$ , and  $\psi_0(a) = \psi(a)$  for  $a \in I'_i(D)$ , and  $\psi_0(a) = \mathbf{0}$  for all other  $a \in I_i(D)$ . We have

$$\sum_{a \in I_i(D)} a \psi_0(a) = \sum_{\omega \in \Omega} (a_\omega + N_\omega) \omega \in \Lambda,$$

since  $M|(a_\omega + N_\omega)$  for all  $\omega$ . Moreover,  $\psi_0$  is nondegenerate and compatible by construction. Consequently,  $\sum_a a \bar{\psi}_0(a) \in \Lambda$  (by removing the coefficient of  $\mathbf{1}$ ). Since there are at most  $2^{|\mathcal{K}|} \leq 2^{M2^k}$  choices for  $\{a_\omega : \omega \in \Omega\}$ , the map from  $\sum_{a \in I'_i(D)} a \bar{\psi}(a)$  to  $\sum_{a \in I_i(D)} a \bar{\psi}_0(a)$  is at most  $2^{M2^k}$ -to-1. We conclude that with probability  $\geq 1 - \delta$ ,

$$|\mathcal{L}_i(\mathbf{A})| \geq 2^{-M2^k} |\mathcal{L}'_i(\mathbf{A})| \gg \delta^\alpha D^{(1-\kappa/i) \sum_j c_j \dim(V_j/V_{j-1})},$$

the implied constant only depending on  $k, M$  and  $\alpha$ , which are all fixed.  $\square$

### 6.2 Putting $\mathcal{L}_i(\mathbf{A})$ in a box

In the last section, we showed that (with high probability)  $\mathcal{L}_i(\mathbf{A})$  is large. In this section we show that with high probability it is contained in a box (in coordinates  $\omega^1, \dots, \omega^d$ ); putting these results together one then sees that  $\mathcal{L}_i(\mathbf{A})$  occupies a positive proportion of lattice points in the box, the bound being independent of  $D$ .

For  $t \in \{1, \dots, d\}$ , write  $j(t)$  for the unique  $j$  such that

$$\dim V_{j-1} < t \leq \dim V_j.$$

In addition, let  $C$  be the largest coordinate in absolute value of any element in  $V_r \cap \{0, 1\}^k$  when written with respect to the base  $\mathbf{1}, \omega^1, \dots, \omega^d$ . We then set

$$N_j^{(i)} := \delta^{-1} \cdot C \cdot D^{(1-\kappa/i)c_j} \quad \text{and} \quad N^{(i)} := \prod_{t=1}^d N_{j(t)}^{(i)}. \tag{6.4}$$

**Lemma 6.3** *Assume  $\delta > 0$  is small enough so that  $re^{-2/\delta} \leq \delta$ . Then, we have*

$$\mathcal{L}_i(\mathbf{A}) \subset \bigoplus_{t=1}^d [-N_{j(t)}^{(i)}, N_{j(t)}^{(i)}] \omega^t \tag{6.5}$$

with probability at least  $1 - \delta$  in the choice of  $\mathbf{A} \cap I_i(D)$ .

*Proof* This follows quickly from the fact that  $\psi$  is compatible and by Lemma A.6, the latter implying that

$$\sum_{a \in \mathbf{A} \cap [2, D^{(1-\kappa/i)c_j}]} a \leq \delta^{-1} D^{(1-\kappa/i)c_j} \quad (1 \leq j \leq r)$$

with probability  $\geq 1 - re^{-2/\delta} \geq 1 - \delta$ .  $\square$

**Proposition 6.4** *Let  $\delta$  and  $\alpha$  be as in Proposition 6.2 and in Lemma 6.3. With probability at least  $1 - 2\delta$  in the choice of  $\mathbf{A} \cap I_i(D)$ ,  $\mathcal{L}_i(\mathbf{A})$  is a subset of the box  $\bigoplus_{t=1}^d [-N_{j(t)}^{(i)}, N_{j(t)}^{(i)}]\omega^t$  of size  $\gg \delta^{d+\alpha} N^{(i)}$ .*

*Proof* This follows immediately upon combining Proposition 6.2 and Lemma 6.3. □

### 6.3 Zero sums with positive probability

**Lemma 6.5** *Let  $\delta$  and  $\alpha$  be as in Proposition 6.2 and Lemma 6.3, and let  $D$  be large enough in terms of  $\delta$  and  $(\mathcal{V}, \mathbf{c}, \boldsymbol{\mu})$ . Let  $i \in \mathbb{Z} \cap [1, (\log D)^{1/3}]$ . In addition, let  $S \subset \bigoplus_{t=1}^d [-N_{j(t)}^{(i)}, N_{j(t)}^{(i)}]\omega^t$  with  $|S| \gg \delta^{d+\alpha} N^{(i)}$  and with  $S \subset \Lambda$ . Then*

$$\mathbb{P}(0 \in \mathcal{L}_{i+1}(\mathbf{A}) \mid \mathcal{L}_i(\mathbf{A}) = S) \gg \delta^{2d(d+\alpha)}.$$

*Proof* We condition on a fixed choice of  $\mathbf{A} \cap I_i(D)$  for which  $\mathcal{L}_i(\mathbf{A}) = S$ . Note that

$$I_{i+1}(D) \setminus I_i(D) = \bigcup_{j=1}^r (D^{(1-\kappa/i)c_j}, D^{(1-\kappa/(i+1))c_j}] \supset \bigcup_{j=1}^r [N_j^{(i)}, 100dN_j^{(i)}]. \tag{6.6}$$

Then it is enough to show that with probability  $\gg \delta^{2d(d+\alpha)}$ , the set  $\mathbf{A}$  contains  $2d$  distinct elements  $a_t$  and  $a'_t$ ,  $1 \leq t \leq d$ , such that

$$\sum_t (a'_t - a_t)\omega^t \in S \quad \text{and} \quad a_t, a'_t \in [N_j^{(i)}, 100dN_j^{(i)}] \quad \text{for } t = 1, \dots, d. \tag{6.7}$$

To see why this is sufficient, let  $s = \sum_t (a'_t - a_t)\omega^t$ , which we know belongs to  $S = \mathcal{L}_i(\mathbf{A})$ . In particular, there is an compatible map  $\psi$  supported on  $I_i(D)$  such that  $\sum_{a \in \mathbf{A}} a\psi(a) = s$ . Now, consider the function

$$\psi' : \mathbf{A} \cap I_{i+1}(D) \rightarrow \{0, 1\}^k$$

with  $\psi'(a) = \psi(a)$  for  $a \in \mathbf{A} \cap I_i(D)$ ,  $\psi'(a'_t) = \mathbf{1} - \omega^t$  and  $\psi'(a_t) = \omega^t$  for  $1 \leq t \leq d$ , and  $\psi'(a) = \mathbf{0}$  for all other values of  $a \in \mathbf{A} \cap I_{i+1}(D)$ . Notice that  $\psi'$  is compatible according to Definition 5.8 by the second part of (6.7). It is now clear that  $0 \in \mathcal{L}_{i+1}(\mathbf{A})$ . Hence, if the conditional probability that (6.7) holds is  $\gg \delta^{2\beta d}$ , so is the probability that  $0 \in \mathcal{L}_{i+1}(\mathbf{A})$ .

To find  $a_t$  and  $a'_t$  satisfying (6.7), let

$$n := \left\lceil d3^{d+1} N^{(i)} / |S| \right\rceil \ll \delta^{-(d+\alpha)}.$$

The number of elements  $\sum_t s_t \omega^t \in S$  with  $n|s_t$  for some  $t$  is

$$\begin{aligned} &\leq \sum_{t=1}^d (2N_{j(t)}^{(i)} / n + 1) \prod_{t' \neq t} (2N_{j(t')}^{(i)} + 1) \leq d3^{d-1} \left( \frac{2N^{(i)}}{n} + \frac{N^{(i)}}{\min_j N_j^{(i)}} \right) \\ &\leq |S|/2 \end{aligned}$$

as long as  $D$  is large enough in terms of  $\delta$  and  $(\mathcal{V}, \mathbf{c}, \boldsymbol{\mu})$ . Thus, there is a subset  $S' \subset S$  of size at least  $|S|/2$  and with  $n \nmid s_t$  for all  $t$ . We will choose the sets  $\{a_t : 1 \leq t \leq d\}$  and  $\{a'_t : 1 \leq t \leq d\}$  independently, by selecting  $a_t \equiv 0 \pmod{n}$  and  $a'_t \not\equiv 0 \pmod{n}$ .

Note that

$$I_{i+1}(D) \setminus I_i(D) = \bigcup_{j=1}^r (D^{(1-\kappa/i)c_j}, D^{(1-\kappa/(i+1))c_j}) \supset \bigcup_{j=1}^r [N_j^{(i)}, 100dN_j^{(i)}]$$

provided that  $i \leq (\log D)^{1/3}$ . For each given  $t, i$  and  $j$ , the probability that the interval  $[4tN_j^{(i)}, (4t+2)N_j^{(i)}]$  contains no element  $a_t \equiv 0 \pmod{n}$  of  $\mathbf{A}$  equals

$$\prod_{\substack{4tN_j^{(i)} \leq a \leq (4t+2)N_j^{(i)} \\ a \equiv 0 \pmod{n}}} (1 - 1/a) \leq 1 - \gamma/n$$

for some small positive constant  $\gamma = \gamma(d)$ . Thus, the probability that, for each  $t = 1, 2, \dots, d$ , the set  $\mathbf{A}$  contains some  $a_t \equiv 0 \pmod{n}$  in the interval  $[4tN_{j(t)}^{(i)}, (4t+2)N_{j(t)}^{(i)}]$  is  $\gg 1/n^d \gg \delta^{d(d+\alpha)}$ .

Fix a choice of  $a_1, \dots, a_d$  as described above, and set

$$X := \{(a_1 + s_1, \dots, a_d + s_d) : s_1 \omega^1 + \dots + s_d \omega^d \in S'\}. \tag{6.8}$$

By construction, every coordinate of  $x \in X$  is  $\not\equiv 0 \pmod{n}$ . Also,

$$X \subset \prod_{t=1}^d [(4t-1)N_{j(t)}^{(i)}, (4t+3)N_{j(t)}^{(i)}]. \tag{6.9}$$

Now the intervals on the right-hand side above are disjoint, and

$$|X| \geq \frac{|S|}{2} \gg \delta^\beta \prod_{t=1}^d N_{j(t)}^{(i)}.$$

Thus, by Lemma A.7, with probability  $\gg (\delta^{d+\alpha})^d$ , there are  $a'_1, \dots, a'_d \in \mathbf{A}$  such that  $(a'_1, \dots, a'_d) \in X$ . The relation (6.7) follows for such  $a_t, a'_t$ , which exist with probability  $\gg \delta^{d(d+\alpha)} \cdot \delta^{d(d+\alpha)}$ .  $\square$

#### 6.4 An iterative argument

To complete the proof of Proposition 5.7, we apply Lemma 6.5 iteratively. Let  $\mathcal{S}$  be the set of sets  $S$  satisfying the assumptions of Lemma 6.5. We say that  $\mathcal{L}_i(\mathbf{A})$  is *large* if it satisfies the conclusions of Proposition 6.4, or equivalently if  $\mathcal{L}_i(\mathbf{A}) = S$  with  $S \in \mathcal{S}$ . Thus Lemma 6.5 implies that

$$\begin{aligned} & \mathbb{P}(0 \in \mathcal{L}_{i+1}(\mathbf{A}) \setminus \mathcal{L}_i(\mathbf{A}), \mathcal{L}_i(\mathbf{A}) \text{ large}) \\ &= \sum_{\substack{S \text{ large} \\ 0 \notin S}} \mathbb{P}(\mathcal{L}_i(\mathbf{A}) = S) \cdot \mathbb{P}(0 \in \mathcal{L}_{i+1}(\mathbf{A}) \mid \mathcal{L}_i(\mathbf{A}) = S) \\ &\gg \delta^{2d\alpha} \mathbb{P}(\mathcal{L}_i(\mathbf{A}) \text{ large}, 0 \notin \mathcal{L}_i(\mathbf{A})). \end{aligned}$$

We conclude there is some  $\varepsilon = \delta^{O(1)}$  such that

$$\mathbb{P}(0 \in \mathcal{L}_{i+1}(\mathbf{A}) \mid \mathcal{L}_i(\mathbf{A}) \text{ large}, 0 \notin \mathcal{L}_i(\mathbf{A})) \geq \varepsilon. \quad (6.10)$$

For brevity, write  $E_i$  for the event that  $0 \notin \mathcal{L}_i(\mathbf{A})$ , and  $F_i$  for the event that  $\mathcal{L}_i(\mathbf{A})$  is large. In this notation, (6.10) becomes

$$\mathbb{P}(E_{i+1}^c \mid E_i \cap F_i) \geq \varepsilon. \quad (6.11)$$

Moreover, Proposition 6.4 implies that

$$\mathbb{P}(F_i) \geq 1 - 2\delta. \quad (6.12)$$

Lastly, note that  $E_1 \supset E_2 \supset \dots$  because  $\mathcal{L}_1(\mathbf{A}) \subset \mathcal{L}_2(\mathbf{A}) \subset \dots$

We claim that  $\mathbb{P}(E_i) < 4\delta$  for some  $i \leq I := \lfloor (\log D)^{1/3} \rfloor$ . Indeed, for each  $i \leq I$ , we have

$$\begin{aligned}
 \mathbb{P}(E_{i+1}) &= \mathbb{P}(E_{i+1}|E_i \cap F_i)\mathbb{P}(E_i \cap F_i) + \mathbb{P}(E_{i+1}|E_i \cap F_i^c)\mathbb{P}(E_i \cap F_i^c) \\
 &\leq (1 - \varepsilon)\mathbb{P}(E_i \cap F_i) + \mathbb{P}(E_i \cap F_i^c) \quad \text{by (6.11)} \\
 &= \mathbb{P}(E_i) - \varepsilon\mathbb{P}(E_i \cap F_i) \\
 &\leq \mathbb{P}(E_i) - \varepsilon(\mathbb{P}(E_i) - 2\delta) \quad \text{by (6.12)}.
 \end{aligned}$$

Thus, if  $\mathbb{P}(E_i) \geq 4\delta$ , then  $\mathbb{P}(E_{i+1}) \leq (1 - \varepsilon/2)\mathbb{P}(E_i)$ . If this holds for all  $i \leq I$ , then  $\mathbb{P}(E_I) \leq (1 - \varepsilon/2)^{I-1} < 4\delta$ , a contradiction. Therefore,  $\mathbb{P}(E_{i^*}) < 4\delta$  for some  $i^* \leq I$ , as long as  $D$  is large enough in terms of  $\delta$  and the (fixed) system  $(\mathcal{V}, \mathbf{c}, \boldsymbol{\mu})$ . This completes the proof of Proposition 5.7.

### Part III. The optimisation problem

#### 7 The optimisation problem—basic features

In this section we consider Problem 3.7, the optimisation problem on the cube, which is a key feature of our paper. We will give some kind of a solution to this for a fixed nondegenerate flag  $\mathcal{V}$ , leaving aside the question of how to choose  $\mathcal{V}$  optimally.

Let us refresh ourselves on the main elements of the setup of Problem 3.7. We have a nondegenerate,  $r$ -step flag

$$\mathcal{V} : \langle \mathbf{1} \rangle = V_0 \leq V_1 \leq V_2 \leq \dots \leq V_r \leq \mathbb{Q}^k$$

of distinct vector spaces. In light of Lemma 5.4, we may restrict our attention to flags such that

$$\dim(V_1/V_0) = 1,$$

which we henceforth assume. With the flag  $\mathcal{V}$  fixed, we wish to find  $\gamma_k(\mathcal{V})$ , the supremum of numbers  $c \geq 0$  such that there are thresholds

$$1 = c_1 \geq c_2 \geq \dots \geq c_{r+1} = c$$

(we may assume that  $c_1 = 1$  by arguing as in Lemmas 5.3 and 5.4) and probability measures  $\mu_1, \dots, \mu_r$  on  $\{0, 1\}^k$  satisfying  $\text{Supp}(\mu_j) \subset V_j$  for each  $j$ , and such that the entropy condition (3.4) holds, that is to say

$$e(\mathcal{V}', \mathbf{c}, \boldsymbol{\mu}) \geq e(\mathcal{V}, \mathbf{c}, \boldsymbol{\mu}) \tag{7.1}$$

for all subflags  $\mathcal{V}' \leq \mathcal{V}$ . We recall that

$$e(\mathcal{V}', \mathbf{c}, \boldsymbol{\mu}) := \sum_{j=1}^r (c_j - c_{j+1}) \mathbb{H}_{\mu_j}(V'_j) + \sum_{j=1}^r c_j \dim(V'_j/V'_{j-1}).$$

*Remarks.* (a) It is easy to see that  $\gamma_k(\mathcal{V})$  always exists by considering the following example with  $c = 0$ . Take  $c_1 = 1$  and  $c_2 = \dots = c_{r+1} = 0$  and recall that  $\dim(V_1/V_0) = 1$ . Suppose that  $V_1 = \text{Span}(\mathbf{1}, \omega)$  with  $\omega \in \{0, 1\}^k$ . Thus,  $e(\mathcal{V}, \mathbf{c}, \boldsymbol{\mu}) = 1$  for any choice of  $\boldsymbol{\mu}$ . If  $V'_1 = V_1$  then likewise we have  $e(\mathcal{V}', \mathbf{c}, \boldsymbol{\mu}) = 1$ , and if  $V'_1 = V_0$  then  $e(\mathcal{V}', \mathbf{c}, \boldsymbol{\mu}) = \mathbb{H}_{\mu_1}(V_0)$ . Now  $V_0 + \mathbf{1}$ ,  $V_0 + \omega$  and  $V_0 + (\mathbf{1} - \omega)$  are three different cosets. Taking

$$\mu_1(\mathbf{1}) = \mu_1(\omega) = \mu_1(\mathbf{1} - \omega) = 1/3$$

we have  $e(\mathcal{V}', \mathbf{c}, \boldsymbol{\mu}) = \log 3$ . Thus, (3.4) holds. As we shall see in this section, this choice of  $\mu_1$  is the optimal choice for a very general class of flags, including those of interest to us.

(b) A simple compactness argument shows that the supremum is realised, that is, there is a choice of  $\mathbf{c}$  and  $\boldsymbol{\mu}$  satisfying the entropy condition 3.4 and with  $c_{r+1} = \gamma_k(\mathcal{V})$ .

(c) As long as we can show that  $\gamma_k > 0$  (which will be taken care of in Part IV), we can always find an optimal system  $(\mathcal{V}, \mathbf{c}, \boldsymbol{\mu})$  that also has  $c_j > c_{j+1}$  for each  $j$  (cf. Lemma 5.4(a)).

### 7.1 A restricted optimisation problem

It turns out to be very useful to consider a restricted variant of the problem in which the entropy condition (7.1) is only required to be satisfied for certain “basic” subflags  $\mathcal{V}'$ , rather than all of them.

**Definition 7.1** (*Basic subflag*) Given a flag  $\mathcal{V} : \langle \mathbf{1} \rangle = V_0 \leq V_1 \leq \dots \leq V_r$ , the basic subflags  $\mathcal{V}'_{\text{basic}(m)}$  are the ones in which  $V'_i = V_{\min(m,i)}$ , for  $m = 0, 1, \dots, r - 1$  (note that when  $m = r$  we recover  $\mathcal{V}$  itself).

Here is the restricted version of Problem 3.7. Recall that a flag is nondegenerate if the top space  $V_r$  is not contained in any of the subspaces  $\{x \in \mathbb{R}^k : x_i = x_j\}$ . The restriction to nondegenerate flags ensures that the subsets  $A_1, \dots, A_k$  in our main problem are distinct.

**Problem 7.2** Let  $\mathcal{V}$  be a nondegenerate flag of distinct spaces in  $\mathbb{Q}^k$ . Define  $\gamma_k^{\text{res}}(\mathcal{V})$  to be the supremum of all constants  $c \geq 0$  for which there are measures  $\mu_1, \dots, \mu_r$  such that  $\text{Supp}(\mu_i) \subset V_i$ , and parameters

$$1 = c_1 \geq \dots \geq c_{r+1} = c$$



such that the restricted entropy condition

$$e(\mathcal{V}'_{\text{basic}(m)}; \mathbf{c}, \boldsymbol{\mu}) \geq e(\mathcal{V}; \mathbf{c}, \boldsymbol{\mu}) \tag{7.2}$$

holds for all  $m = 0, 1, \dots, r - 1$ .

It is clear that

$$\gamma_k^{\text{res}}(\mathcal{V}) \geq \gamma_k(\mathcal{V}). \tag{7.3}$$

In general there is absolutely no reason to suppose that the two quantities are equal, since after all the restricted entropy condition (7.2) apparently only captures a small portion of the full condition (7.1).

Our reason for studying the restricted problem is that we do strongly believe that

$$\sup_{\mathcal{V} \text{ nondegenerate}} \gamma_k^{\text{res}}(\mathcal{V}) = \sup_{\mathcal{V} \text{ nondegenerate}} \gamma_k(\mathcal{V}) = \gamma_k.$$

One might think of this unproven assertion, on an intuitive level, in two (roughly equivalent) ways:

- for those flags optimal for Problem 3.7, the critical cases of (7.1) are those for which  $\mathcal{V}'$  is basic;
- for those flags optimal for Problem 3.7, and for the critical choice of the  $c_i, \mu_i$ , the restricted condition (7.2) in fact implies the more general condition (7.1).

### 7.2 The $\rho$ -equations, optimal measures and optimal parameters

The definitions and constructions of this section will appear unmotivated at first sight. They are forced upon us by the analysis of Sect. 7.5 below.

Let the flag  $\mathcal{V}$  be fixed.

It is convenient to call the intersection of a coset  $x + V_i$  with the cube  $\{0, 1\}^k$  a *cell at level  $i$* , and to denote the cells at various levels by the letter  $C$ . (The terminology comes from the fact it can be useful to think of  $V_i$  defining a  $\sigma$ -algebra (partition) on  $\{0, 1\}^k$ , the equivalence relation being given by  $\omega \sim \omega'$  iff  $\omega - \omega' \in V_i$ : however, we will not generally use the language of  $\sigma$ -algebras in what follows.)

If  $C$  is a cell at level  $i$ , then it will be a union of cells  $C'$  at level  $i - 1$ . These cells we call the children of  $C$ , and we write  $C \rightarrow C'$ .

Let  $\boldsymbol{\rho} = (\rho_1, \dots, \rho_{r-1})$  be real parameters in  $(0, 1)$ , and for each cell  $C$  define functions  $f^C(\boldsymbol{\rho})$  by the following recursive recipe:

- If  $C$  has level 0, then  $f^C(\rho) = 1$ ;
- If  $C$  has level  $i$ , then

$$f^C(\rho) = \sum_{C \rightarrow C'} f^{C'}(\rho)^{\rho_{i-1}}, \tag{7.4}$$

with the convention that  $\rho_0 = 0$ .

Write

$$\Gamma_i = V_i \cap \{0, 1\}^k$$

for the cell at level  $i$  which contains  $\mathbf{0}$ . Note that

$$\{\mathbf{0}, \mathbf{1}\} = \Gamma_0 \subset \Gamma_1 \subset \dots \subset \Gamma_r.$$

**Definition 7.3** ( $\rho$ -equations) The  $\rho$ -equations are the system of equations

$$f^{\Gamma_{j+1}}(\rho) = (f^{\Gamma_j}(\rho))^{\rho_j} e^{\dim(V_{j+1}/V_j)}, \quad j = 1, 2, \dots, r - 1. \tag{7.5}$$

We say that they have a solution if they are satisfied with  $\rho_1, \dots, \rho_{r-1} \in (0, 1)$ .

*Example* Figure 1 illustrates these definitions for the so-called *binary flag* in  $\mathbb{Q}^4$ , which will be a key object of study from Sect. 9 onwards. Here

$$V_1 = \{(x_1, x_2, x_3, x_4) \in \mathbb{Q}^4 : x_1 = x_2, x_3 = x_4\}$$

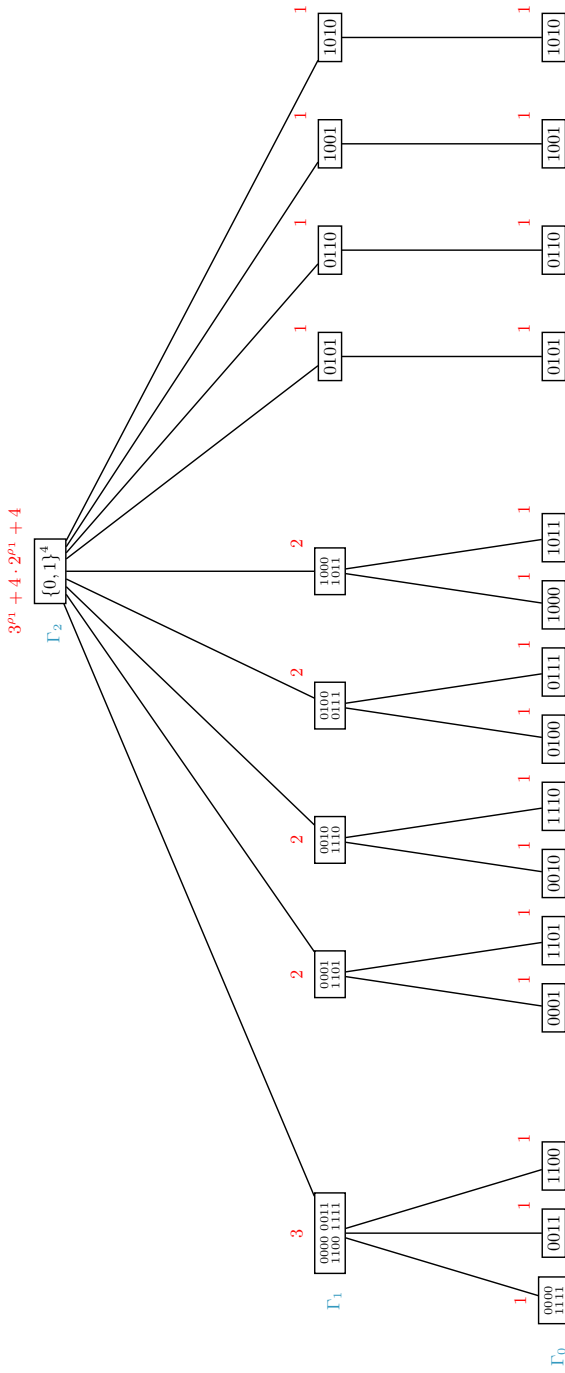
and  $V_2 = \mathbb{Q}^4$ . The  $\rho$ -equations consist of the single equation  $f^{\Gamma_2}(\rho) = (f^{\Gamma_1}(\rho))^{\rho_1} e^2$ , that is to say  $3^{\rho_1} + 4 \cdot 2^{\rho_1} + 4 = 3^{\rho_1} e^2$ . This has the unique solution  $\rho_1 \approx 0.306481$ .

In general the  $\rho$ -equations may or may not have a solution, but for flags  $\mathcal{V}$  of interest to us, it turns out that they have a unique such solution. In this case, we make the following definition.

**Definition 7.4** (*Optimal measures*) Suppose that  $\mathcal{V}$  is a flag for which the  $\rho$ -equations have a solution. Then the corresponding *optimal measure on  $\mu^*$  on  $\{0, 1\}^k$  with respect to  $\mathcal{V}$*  is defined as follows: we set  $\mu^*(\Gamma_r) = 1$ , and

$$\frac{\mu^*(C')}{\mu^*(C)} = \frac{f^{C'}(\rho)^{\rho_{i-1}}}{f^C(\rho)} \tag{7.6}$$

for any cell  $C$  at level  $i \geq 1$  and any child  $C \rightarrow C'$ . We also set  $\mu^*(\mathbf{0}) = \mu^*(\mathbf{1}) = \mu^*(\Gamma_0)/2$ . Lastly, we define the restrictions  $\mu_j^*(\omega) := \mu^*(\Gamma_j)^{-1} \mu^*(\omega) 1_{\omega \in \Gamma_j}$  for  $j = 1, 2, \dots, r$  (thus  $\mu_r^* = \mu^*$ ). We



**Fig. 1** the tree structure corresponding to the binary flag  $\langle \mathbf{1} \rangle = V_0 \leq V_1 \leq V_2 \leq \mathbb{Q}^4$ . Values of  $f^C(\rho)$  are given in red

call these<sup>7</sup> *optimal measures* (on  $\{0, 1\}^k$ , with respect to  $\mathcal{V}$ ). Finally, we write  $\mu^* = (\mu_1^*, \mu_2^*, \dots, \mu_r^*)$ .

*Remark 7.1* (a) By taking telescoping products of (7.6) for  $i = r, r - 1, \dots, 0$ , we see that  $\mu^*$  is uniquely defined on all cells at level 0, and these are the cell  $\{\mathbf{0}, \mathbf{1}\}$  and singletons  $\{\omega\}$  for all  $\omega \in \{0, 1\}^k \setminus \{\mathbf{0}, \mathbf{1}\}$ . Since we also specified  $\mu^*(\mathbf{0}) = \mu^*(\mathbf{1}) = \mu^*(\Gamma_0)/2$ , we see that  $\mu^*(\omega)$  is completely and uniquely determined by these rules, for all  $\omega$ . In particular, the  $\rho$ -equations (7.5) are equivalent to

$$\frac{\mu^*(\Gamma_j)}{\mu^*(\Gamma_{j+1})} = e^{-\dim(V_{j+1}/V_j)} \quad \text{for } j = 1, \dots, r - 1,$$

and thus

$$\mu_j^*(\Gamma_m) = e^{-\dim(V_j/V_m)} \quad (j \geq m \geq 1). \tag{7.7}$$

In addition, we have

$$\mu^*(\Gamma_0) = \mu^*(\Gamma_1) \cdot \frac{1}{f^{\Gamma_1}(\rho)} = \frac{e^{-\dim(V_1/V_r)}}{|\Gamma_1| - 1}. \tag{7.8}$$

(b) By construction, the measures  $\mu_j^*$  satisfy statements (d) and (e) of Lemma 5.3 for all  $j$ :

$$\text{Supp}(\mu_j) = \Gamma_j \quad \text{and} \quad \mu_j(\omega) = \mu_j(\mathbf{1} - \omega) \quad \forall \omega. \tag{7.9}$$

(c) At the moment, the term “optimal measure” is just a name. We will establish the sense in which (in situations of interest) the measures  $\mu_j^*$  are optimal in Proposition 7.7 below.

(d) Note that  $\mu^*$  and  $\mu^*$  are two different (but closely related) objects. The former is an  $r$ -tuple of measures  $\mu_j^*$ , all of which are induced from the single measure  $\mu^*$ .

**Definition 7.5** (*Optimal parameters*) Suppose that  $\mathcal{V}$  is a flag for which the  $\rho$ -equations have a solution. Let  $\mu^*$  be the corresponding optimal measure on  $\{0, 1\}^k$  with respect to  $\mathcal{V}$ . Suppose additionally that

$$\mathbb{H}_{\mu_{m+1}^*}(V_m) \neq \dim(V_{m+1}/V_m) \tag{7.10}$$

---

<sup>7</sup> Note that we have not said that the  $\rho_i$  are unique. However, in cases of interest to us this will turn out to be the case.

for  $m = 0, 1, \dots, r - 1$ . Then the corresponding *optimal parameters* with respect to  $\mathcal{V}$  and the solution  $\rho$  are the unique choice of

$$\mathbf{c}^* : 1 = c_1^* > c_2^* > \dots > c_{r+1}^* > 0,$$

if it exists, such that

$$e(\mathcal{V}'_{\text{basic}(m)}, \mu^*, \mathbf{c}^*) = e(\mathcal{V}, \mu^*, \mathbf{c}^*) \quad \text{for } m = 0, 1, \dots, r - 1. \tag{7.11}$$

The Eq. (7.11), written out in full, are

$$\sum_{j=m+1}^r (c_j^* - c_{j+1}^*) \mathbb{H}_{\mu_j^*}(V_m) = \sum_{j=m+1}^r c_j^* \dim(V_j/V_{j-1}) \tag{7.12}$$

$m = 0, 1, \dots, r - 1$ .

By (7.10), this uniquely determines  $c_{m+1}^* \in \mathbb{R}$  in terms of  $c_{m+2}^*, \dots, c_{r+1}^*$ . Hence, we recursively determine  $c_1, \dots, c_r$  in terms of  $c_{r+1}$ . Since we must further have  $c_1 = 1$ , this implicitly determines  $c_{r+1}$  as well, and thus the entire vector  $\mathbf{c}^*$ .

**Remark.** By Lemma 5.3 (ii), a stronger form of the condition (7.10) is required in order for the entropy gap condition to hold, and so in practice this assumption is not at all restrictive.

We conclude this subsection with a characterization of the optimal measure  $\mu^*$  and parameters  $\mathbf{c}^*$ . Given an  $r$ -step flag  $\mathcal{V}$ , there is an associated rooted tree  $\mathcal{T}(\mathcal{V})$ , which captures the structure of the cells at different levels  $0, \dots, r - 1$ . In particular, this tree always has exactly  $2^k - 1$  leaves at level  $0$ , corresponding to the cell  $\Gamma_0 = \{\mathbf{0}, \mathbf{1}\}$  and the singletons  $\{\omega\}$  for each  $\omega \in \{0, 1\}^k \setminus \{\mathbf{0}, \mathbf{1}\}$ .

**Lemma 7.6** *The optimal constant  $\gamma_k^{res}(\mathcal{V})$ , associated measures  $\mu_i^*(C)$  and optimal parameters  $c_i^*$  depend only on the tree  $\mathcal{T}(\mathcal{V})$  and the sequence of dimensions  $\dim(V_j)$ ,  $0 \leq j \leq r$ .*

*Proof* Let  $\mathcal{V}$  and  $\tilde{\mathcal{V}}$  be different flags with the same tree structure, that is,  $\mathcal{T}(\mathcal{V})$  is isomorphic to  $\mathcal{T}(\tilde{\mathcal{V}})$ , and with the same sequence of dimensions  $\dim(V_j)$  and  $\dim(V'_j)$ . By an easy induction on the level and the definition of  $f^C(\rho)$ , if  $C \in \mathcal{T}(\mathcal{V})$  and  $\tilde{C} \in \mathcal{T}(\tilde{\mathcal{V}})$  correspond, we find that  $f^C(\rho) = f^{\tilde{C}}(\rho)$ . The statements now follow from Definitions 7.4 and 7.5. □

### 7.3 Solution of the optimisation problem: statement

Here is the main result of this section, which explains the introduction of the various concepts above, as well as their names.

**Proposition 7.7** *Suppose that  $\mathcal{V} : \mathbf{1} = V_0 \leq V_1 \leq \dots \leq V_r \leq \mathbb{Q}^k$  is a nondegenerate flag such that  $\dim(V_1/V_0) = 1$  and the  $\rho$ -equations have a solution. Let  $\mu^*$  be the corresponding optimal measures, and suppose that the corresponding optimal parameters  $\mathbf{c}^*$  exist. Then*

$$\gamma_k^{\text{res}}(\mathcal{V}) = (\log 3 - 1) / \left( \log 3 + \sum_{i=1}^{r-1} \frac{\dim(V_{i+1}/V_i)}{\rho_1 \cdots \rho_i} \right). \tag{7.13}$$

Moreover, the optimal measures  $\mu^*$  and optimal parameters  $\mathbf{c}^*$  provide the solution to Problem 7.2; in particular,  $c_{r+1}^*$  is precisely the right-hand side of (7.13).

For this result to be of any use, we need methods for establishing, for flags  $\mathcal{V}$  of interest, that the  $\rho$ -equations have a solution, and also that the optimal parameters exist. The former is a very delicate matter, highly dependent on the specific structure of the flags of interest. Once this is sorted out, the latter problem is less serious, at least in situations relevant to us.

### 7.4 Linear forms in entropies

In the next section we will prove Proposition 7.7. In this section we isolate some lemmas from the proof.

Let  $\mathcal{V} : \langle \mathbf{1} \rangle = V_0 \leq \dots \leq V_r \leq \mathbb{Q}^k$  be a flag. We use the terminology of cells  $C$  at level  $i$ , introduced at the beginning of Sect. 7.2.

**Lemma 7.8** *Let  $\mathbf{y} = (y_0, \dots, y_{r-1})$  be real numbers with the property that all the partial sums  $y_{<i} := y_0 + \dots + y_{i-1}$  are positive. If  $C$  is a cell (at some level  $i$ ), then we write*

$$h^C(\mathbf{y}) := \sup_{\text{Supp}(\mu_C) \subset C} \left( \sum_{0 \leq m < r} y_m \mathbb{H}_{\mu_C}(V_m) \right), \tag{7.14}$$

where the supremum is over all probability measures  $\mu_C$  supported on  $C$ .

(a) The quantities  $h^C(\mathbf{y})$  are completely determined by the following rules:

- If  $C$  has level 0, then  $h^C(\mathbf{y}) = 0$ ;
- If  $C$  has level  $i$ , then

$$h^C(\mathbf{y}) = y_{<i} \log \left( \sum_{C': C \rightarrow C'} e^{h^{C'}(\mathbf{y})/y_{<i}} \right). \tag{7.15}$$

(b) For any  $C$ , the maximum in (7.14) occurs for a unique measure  $\mu_{C,\mathbf{y}}^*$ . Furthermore, all of the  $\mu_{C,\mathbf{y}}^*$  are restrictions of the “top” measure

$\mu_{\mathbf{y}}^* := \mu_{\Gamma_{r,\mathbf{y}}}^*$ , that is to say  $\mu_{C,\mathbf{y}}^*(x) = \mu_{\mathbf{y}}^*(x)/\mu_{\mathbf{y}}^*(C)$  for all  $x \in C$ , and

$$\frac{\mu_{\mathbf{y}}^*(C')}{\mu_{\mathbf{y}}^*(C)} = \frac{e^{h^{C'}(\mathbf{y}/y_{<i})}}{e^{h^C(\mathbf{y}/y_{<i})}}. \tag{7.16}$$

*Remark* As will be apparent from the proof, we do not use the linear structure of the cells  $C$  (that is, the fact that they come from cosets). We leave it to the reader to formulate a completely general version of this lemma in which the cells at level  $i$  are the atoms in a  $\sigma$ -algebra  $\mathcal{F}_i$ , with  $\mathcal{F}_i$  being a refinement of  $\mathcal{F}_{i+1}$  for all  $i$ .

*Proof* We prove both parts simultaneously. Let us temporarily write  $\tilde{h}^C(\mathbf{y})$  for the function defined by (7.15), thus the aim is to prove that  $h^C(\mathbf{y}) = \tilde{h}^C(\mathbf{y})$ , where  $h^C(\mathbf{y})$  is defined in (7.14). We do this by induction on  $i$ , the  $i = 0$  case being trivial since, in this case, all the entropies  $\mathbb{H}_{\mu_C}(V_m)$  are zero because each cell of level 0 lies in some coset mod  $V_0$ , and thus in the same coset mod  $V_m$  for  $m = 0, 1, \dots, r - 1$ .

Suppose now that we know the result for cells of level  $i - 1$ . Note that both  $h^C$  and  $\tilde{h}^C$  satisfy a homogeneity property

$$\tilde{h}^C(t\mathbf{y}) = t\tilde{h}^C(\mathbf{y}), \quad h^C(t\mathbf{y}) = th^C(\mathbf{y}).$$

This is obvious for  $h^C$ , and can be proven very easily for  $\tilde{h}^C$  by induction. Therefore we may assume that  $y_{<i} = 1$ . This does not affect the measure  $\mu_{\mathbf{y}}^*$ , which does not depend on the scaling of the parameters  $y_m$ .

Suppose that  $C$  is a cell at level  $i$ . A probability measure  $\mu_C$  on  $C$  is completely determined by probability measures  $\mu_{C'}$  on the children  $C'$  of  $C$  (at level  $i - 1$ ) together with the probabilities  $\mu_C(C')$ , which must sum to 1, with the relation being that  $\mu_{C'}(x) = \mu_C(x)/\mu_C(C')$  for  $x \in C'$ .

Suppose that  $0 \leq m < i$ . Let the random variables  $X, Y$  be random cosets of  $V_m, V_{i-1}$  respectively, sampled according to the measure  $\mu_C$ . Then  $X$  determines  $Y$  and so, by Lemma B.5,  $\mathbb{H}(X, Y) = \mathbb{H}(X)$ . The chain rule for entropy, Lemma B.4, then yields

$$\mathbb{H}(X) = \mathbb{H}(Y) + \sum_y \mathbb{P}(Y = y)\mathbb{H}(X|Y = y).$$

Translated back to the language we are using, this implies that

$$\mathbb{H}_{\mu_C}(V_m) = \mathbb{H}_{\mu_C}(V_{i-1}) + \sum_{C'} \mu_C(C')\mathbb{H}_{\mu_{C'}}(V_m).$$

Therefore

$$\sum_{0 \leq m < i} y_m \mathbb{H}_{\mu_C}(V_m) = \mathbb{H}_{\mu_C}(V_{i-1}) + \sum_{C'} \mu_C(C') \sum_{0 \leq m < i} y_m \mathbb{H}_{\mu_{C'}}(V_m).$$

(Here we used our assumption that  $y_{<i} = 1$ .) Since  $\mathbb{H}_{\mu_C}(V_m) = 0$  for  $m \geq i$ , and  $\mathbb{H}_{\mu_{C'}}(V_m) = 0$  for  $m \geq i - 1$ , we may extend the sums over all  $m \in \{0, 1, \dots, r - 1\}$  thereby obtaining

$$\sum_{0 \leq m < r} y_m \mathbb{H}_{\mu_C}(V_m) = \mathbb{H}_{\mu_C}(V_{i-1}) + \sum_{C'} \mu_C(C') \sum_{0 \leq m < r} y_m \mathbb{H}_{\mu_{C'}}(V_m).$$

Since the  $\mu_{C'}$  can be arbitrary probability measures, and  $\mathbb{H}_{\mu_C}(V_{i-1})$  depends only on the value of  $\mu_C(C')$ , it follows from the inductive hypothesis that

$$h^C(\mathbf{y}) = \sup_{\mu_C} \left( \sum_{0 \leq m < r} y_m \mathbb{H}_{\mu_C}(V_m) \right) \tag{7.17}$$

$$= \sup_{\mu_C(C'), \mu_{C'}} \left( \mathbb{H}_{\mu_C}(V_{i-1}) + \sum_{C'} \mu_C(C') \sum_{0 \leq m < r} y_m \mathbb{H}_{\mu_{C'}}(V_m) \right) \tag{7.18}$$

$$= \sup_{\mu_C(C')} \left( \mathbb{H}_{\mu_C}(V_{i-1}) + \sum_{C'} \mu_C(C') \tilde{h}^{C'}(\mathbf{y}) \right), \tag{7.19}$$

with equality when going from (7.18) to (7.19) when  $\mu_{C'} = \mu_{C', \mathbf{y}}^*$  for all  $C'$ . Applying Lemma B.3 with the  $p_j$  being the  $\mu_C(C')$  and the  $a_j$  being the  $\tilde{h}^{C'}(\mathbf{y})$ , and noting that  $\mathbb{H}_{\mu_C}(V_{i-1}) = \mathbb{H}(\mathbf{p})$  (where  $\mathbf{p} = (p_1, p_2, \dots)$ ), it follows that

$$\sup_{\mu_C(C')} \left( \mathbb{H}_{\mu_C}(V_{i-1}) + \sum_{C'} \mu_C(C') \tilde{h}^{C'}(\mathbf{y}) \right) = \log \left( \sum_{C': C \rightarrow C'} e^{\tilde{h}^{C'}(\mathbf{y})} \right) = \tilde{h}^C(\mathbf{y}). \tag{7.20}$$

In addition, Lemma B.3 implies that equality occurs in (7.20) precisely when  $p_j = e^{a_j} / \sum_i e^{a_i}$ , that is to say when

$$\mu_C(C') = \frac{e^{\tilde{h}^C(\mathbf{y})}}{\sum_{C': C \rightarrow C'} e^{\tilde{h}^C(\mathbf{y})}} = \frac{\mu_{\mathbf{y}}^*(C')}{\mu_{\mathbf{y}}^*(C)}.$$

(Here we used again that  $y_{<i} = 1$ .) Recalling that  $\mu_{C'} = \mu_{C', \mathbf{y}}^*$  for all  $C'$ , we see that the measure  $\mu_C$  for which equality occurs in (7.17) is the restriction of  $\mu_{\mathbf{y}}^* = \mu_{\Gamma_r, \mathbf{y}}^*$  to  $C$ . This completes the inductive step.  $\square$



### 7.5 Solution of the optimisation problem: proof

This section is devoted to the proof of Proposition 7.7. Strictly speaking, for our main theorems we only need a lower bound on  $\gamma_k^{\text{res}}(\mathcal{V})$ , and for this it suffices to show that  $c_{r+1}^*$  is given by the right-hand side of (7.13). This could, in principle, be phrased as a calculation, but it would look complicated and unmotivated. Instead, we present it in the way we discovered it, by showing that the RHS of (7.13) is an *upper* bound on  $\gamma_k^{\text{res}}(\mathcal{V})$ , and then observing that equality does occur when  $\mu = \mu^*$  is the optimal measure (Definition 7.4) and  $\mathbf{c} = \mathbf{c}^*$  the optimal parameters (Definition 7.5). We establish this upper bound using the duality argument from linear programming and Lemma 7.8.

To ease the notation, we use the shorthand  $d_i := \dim(V_i)$  throughout this subsection. Let us, then, consider the restricted optimisation problem, namely Problem 7.2. The condition (7.2) may be rewritten as

$$\sum_{j=m+1}^r (c_j - c_{j+1})(\mathbb{H}_{\mu_j}(V_m) + d_m - d_j) \geq c_{r+1}(d_r - d_m) \quad (7.21)$$

for  $m = 0, 1, \dots, r - 1$ . Therefore for any choice of “dual variables”  $\mathbf{y} = (y_0, y_1, \dots, y_{r-1}), y_0, \dots, y_{r-1} \geq 0$ , we have

$$\sum_{m=0}^{r-1} y_m \sum_{j=m+1}^r (c_j - c_{j+1})(\mathbb{H}_{\mu_j}(V_m) + d_m - d_j) \geq c_{r+1} \sum_{m=0}^{r-1} y_m (d_r - d_m), \quad (7.22)$$

which, upon rearranging, gives

$$\sum_{j=1}^r (c_j - c_{j+1})E_j(\mathbf{y}) + c_{r+1}E_{r+1}(\mathbf{y}) \geq c_{r+1}. \quad (7.23)$$

where

$$E_j(\mathbf{y}) := \sum_{m=0}^{j-1} y_m (\mathbb{H}_{\mu_j}(V_m) + d_m - d_j)$$

for  $j = 1, \dots, r$ , and

$$E_{r+1}(\mathbf{y}) := 1 - \sum_{m=0}^{r-1} y_m (d_r - d_m).$$

Since the  $c_j - c_{j+1}$ ,  $j = 1, \dots, r$ , and  $c_{r+1}$  are nonnegative and sum to 1, this implies that

$$c_{r+1} \leq \min_{y_i \geq 0 \forall i} \max\{E_1(\mathbf{y}), \dots, E_r(\mathbf{y}), E_{r+1}(\mathbf{y})\}. \tag{7.24}$$

By Lemma 7.8, this implies that

$$c_{r+1} \leq \min_{y_i \geq 0 \forall i} \max\{E'_1(\mathbf{y}), \dots, E'_r(\mathbf{y}), E_{r+1}(\mathbf{y})\}, \tag{7.25}$$

where

$$\begin{aligned} E'_j(\mathbf{y}) &:= h^{\Gamma_j}(\mathbf{y}) + \sum_{m=0}^{j-1} y_m(d_m - d_j) \\ &= \sum_{m=0}^{j-1} y_m(\mathbb{H}_{\mu_{\Gamma_j, \mathbf{y}}}^*(V_m) + d_m - d_j), \end{aligned} \tag{7.26}$$

for  $j = 1, \dots, r$ , and  $\mu_{\Gamma_j, \mathbf{y}}^*$  is the measure  $\nu$  supported on  $\Gamma_j = V_j \cap \{0, 1\}^k$  for which the sum  $\sum_m y_m \mathbb{H}_\nu(V_m)$  is maximal, as defined in Lemma 7.8.

Now we specify a choice of  $\mathbf{y}$ . To do this, we make a change of variables, defining  $\rho_i = y_{<i}/y_{<i+1}$ . Note that for fixed  $y_0 > 0$ , choices of  $y_1, \dots, y_{r-1} > 0$  are in one-to-one correspondence with choices of  $\rho_1, \dots, \rho_{r-1}$  with  $0 < \rho_i < 1$ . We must then have that

$$\log f^C(\boldsymbol{\rho}) = h^C(\mathbf{y}/y_{<i}) = \frac{1}{y_{<i}} h^C(\mathbf{y}) = \frac{\rho_1 \cdots \rho_{i-1}}{y_0} h^C(\mathbf{y}) \tag{7.27}$$

for the cells  $C$  at level  $i$ , which may easily be proven by induction on the level  $i$ , using the defining equations for the  $h^C$  and  $f^C$  (see (7.15), (7.4) respectively).

Now choose the  $\rho_i$  to satisfy the  $\rho$ -equations (7.5). In virtue of (7.27), the  $j$ -th  $\rho$ -equation

$$f^{\Gamma_{j+1}}(\boldsymbol{\rho}) = (f^{\Gamma_j}(\boldsymbol{\rho}))^{\rho_j} e^{d_{j+1}-d_j}$$

with  $j \in \{1, 2, \dots, r - 1\}$  is equivalent to

$$E'_j(\mathbf{y}) = E'_{j+1}(\mathbf{y}), \tag{7.28}$$

with  $E'_j(\mathbf{y})$  defined as in (7.26) above.

Recall that  $d_1 - d_0 = \dim(V_1/V_0) = 1$ . Thus, if we choose

$$y_0 := 1 / \left( \log 3 + \sum_{i=1}^{r-1} \frac{d_{i+1} - d_i}{\rho_1 \dots \rho_i} \right),$$

a short calculation confirms that

$$E_{r+1}(\mathbf{y}) = E'_1(\mathbf{y}) = y_0(\log 3 - 1). \tag{7.29}$$

With this choice of  $\mathbf{y}$  we therefore have, from (7.28) with  $j = 1, \dots, r - 1$ , (7.29) and (7.25),

$$c_{r+1} \leq E'_1(\mathbf{y}) = (\log 3 - 1) / \left( \log 3 + \sum_{i=1}^{r-1} \frac{d_{i+1} - d_i}{\rho_1 \dots \rho_i} \right). \tag{7.30}$$

In the above analysis, the  $\mu_i$  and the  $c_i$  were arbitrary subject to the conditions of Problem 7.2, thus  $\text{Supp}(\mu_i) \subset V_i$  and  $1 = c_1 > c_2 > \dots > c_{r+1}$ . Therefore, recalling the definition of  $\gamma_k^{\text{res}}(\mathcal{V})$  (see Problem 7.2), we have proven that

$$\gamma_k(\mathcal{V}) \leq \gamma_k^{\text{res}}(\mathcal{V}) \leq (\log 3 - 1) / \left( \log 3 + \sum_{i=1}^{r-1} \frac{d_{i+1} - d_i}{\rho_1 \dots \rho_i} \right).$$

Proposition 7.7 asserts that equality occurs in this bound when  $c_j = c_j^*$  and  $\mu_j = \mu_j^*$ , where  $\mathbf{c}^* = (c_1^*, \dots, c_{r+1}^*)$  are the optimal parameters defined in Definition 7.5, and  $\mu^*$  and its restrictions  $\mu_j^*$  are the optimal measures defined in Definition 7.4. To establish this, we must go back through the argument showing that equality occurs at every stage with these choices.

First note that (7.21) is equivalent (as we stated at the time) to  $e(\mathcal{V}'_{\text{basic}(m)}, \mathbf{c}, \boldsymbol{\mu}) \geq e(\mathcal{V}, \mathbf{c}, \boldsymbol{\mu})$ . The fact that equality occurs here when  $\mathbf{c} = \mathbf{c}^*$  and  $\boldsymbol{\mu} = \boldsymbol{\mu}^*$  is essentially the definition of the optimal parameters  $\mathbf{c}^*$  (Definition 7.5). That equality occurs in (7.22) and (7.23) is then automatic.

Working from the other end of the proof, the choice of  $\mathbf{y}$  was made so that  $E'_1(\mathbf{y}) = \dots = E'_r(\mathbf{y}) = E_{r+1}(\mathbf{y})$ . We claim that, with this choice of  $\mathbf{y}$ ,

$$\mu^* = \mu_{\mathbf{y}}^*. \tag{7.31}$$

By (7.16), it suffices to check that

$$\frac{\mu^*(C')}{\mu^*(C)} = \frac{e^{h^{C'}(\mathbf{y}/y_{<i})}}{e^{h^C(\mathbf{y}/y_{<i})}}.$$

This follows immediately from (7.6) and (7.27).

Since  $\mu_j^*$  is defined to be the restriction of  $\mu^*$  to  $\Gamma_j$ , it follows from (7.31) that  $\mu_j^* = \mu_{\Gamma_j, \mathbf{y}}^*$ , and hence that  $E_j(\mathbf{y}) = E'_j(\mathbf{y})$  for  $j = 1, \dots, r$ .

Thus all  $2r + 1$  of the quantities  $E'_j(\mathbf{y})$  ( $j = 1, \dots, r$ ) and  $E_j(\mathbf{y})$  ( $j = 1, \dots, r + 1$ ) are equal. It follows from this and the fact that equality occurs in (7.23) that equality occurs in (7.24), (7.25) and (7.30) as well. This concludes the proof of Proposition 7.7. □

## 8 The strict entropy condition

### 8.1 Introduction

Fix an  $r$ -step, nondegenerate flag  $\mathcal{V}$ . In the previous section, we studied a restricted optimization problem (Problem 7.2) asking for the supremum of  $c_{r+1}$  when ranging over all systems  $(\mathcal{V}, \mathbf{c}, \boldsymbol{\mu})$  satisfying the “restricted entropy condition”

$$e(\mathcal{V}'_{\text{basic}(m)}, \mathbf{c}, \boldsymbol{\mu}) \geq e(\mathcal{V}, \mathbf{c}, \boldsymbol{\mu}) \quad (m = 0, 1, \dots, r - 1). \tag{8.1}$$

The aim of the present section is two-fold: we wish to establish, under general conditions, that an “optimal system” with respect to (8.1) satisfies the more general entropy condition

$$e(\mathcal{V}', \mathbf{c}, \boldsymbol{\mu}) \geq e(\mathcal{V}, \mathbf{c}, \boldsymbol{\mu}) \quad (\text{all } \mathcal{V}' \leq \mathcal{V}). \tag{8.2}$$

In addition, we want to show that if we slightly perturb such a system, we may guarantee the strict entropy condition (3.5), which is a version of (8.2) with strict inequalities for all proper subflags  $\mathcal{V}'$  of  $\mathcal{V}$ .

Before stating our result, we need to define the notion of the automorphism group of a flag.

**Definition 8.1** (*Automorphism group*) For a permutation  $\sigma \in S_k$  and  $\omega = (\omega_1, \dots, \omega_k) \in \mathbb{Q}^k$ , denote by  $\sigma\omega$  the usual coordinate permutation action  $\sigma\omega = (\omega_{\sigma(1)}, \dots, \omega_{\sigma(k)})$ . The *automorphism group*  $\text{Aut}(\mathcal{V})$  is the group of all  $\sigma$  that satisfy  $\sigma V_i = V_i$  for all  $i$ .

**Proposition 8.2** *Let  $\mathcal{V}$  be an  $r$ -step, nondegenerate flag of distinct spaces. Assume that the  $\rho$ -equations (7.5) have a solution, and define the optimal measures  $\mu^*$  on  $\{0, 1\}^k$  as in Definition 7.4. Furthermore, assume that:*

- (a) *no intermediate subspace is fixed by  $\text{Aut}(\mathcal{V})$ , that is to say there is no space  $W$  that is invariant under the action of  $\text{Aut}(\mathcal{V})$  and such that  $V_{i-1} < W < V_i$  (the inclusions being strict);*

(b) the optimal parameters  $\mathbf{c}^*$  exist and they are distinct and positive, that is to say the system of Eq. (7.12) has a unique solution  $\mathbf{c}^*$  satisfying  $1 = c_1^* > c_2^* > \dots > c_{r+1}^* > 0$ ;

(c) the following “positivity inequalities” hold:

(i)  $\mathbb{H}_{\mu_{m+1}^*}(V_m) > \dim(V_{m+1}/V_m)$  for  $0 \leq m \leq r - 1$ ;

(ii)  $\mathbb{H}_{\mu_i^*}(V_{m-1}) - \mathbb{H}_{\mu_i^*}(V_m) < \dim(V_m/V_{m-1})$  for  $1 \leq m < i \leq r$ .

Then, for every  $\varepsilon > 0$ , there exists a perturbation  $\tilde{\mathbf{c}}$  of  $\mathbf{c}^*$  such that  $1 = \tilde{c}_1 > \tilde{c}_2 > \dots > \tilde{c}_{r+1} \geq c_{r+1} - \varepsilon$  and such that we have the strict entropy condition

$$e(\mathcal{V}', \tilde{\mathbf{c}}, \boldsymbol{\mu}^*) > e(\mathcal{V}, \tilde{\mathbf{c}}, \boldsymbol{\mu}^*) \quad \text{for all proper subflags } \mathcal{V}' \leq \mathcal{V}. \tag{8.3}$$

We assume throughout the rest of the section that (a), (b) and (c) of Proposition 8.2 are satisfied, and we now fix the system  $(\mathcal{V}, \mathbf{c}^*, \boldsymbol{\mu}^*)$ . For notational brevity in what follows, we write

$$e(\mathcal{V}') := e(\mathcal{V}', \mathbf{c}^*, \boldsymbol{\mu}^*).$$

Our strategy is as follows. First, we show the weaker “unperturbed” statement that

$$e(\mathcal{V}') \geq e(\mathcal{V}) \quad \text{for all subflags } \mathcal{V}' \leq \mathcal{V}, \tag{8.4}$$

noting that we have strict inequality for certain subflags  $\mathcal{V}'$  along the way. Then, in Sect. 8.8, we show how to perturb  $\mathbf{c}^*$  to  $\tilde{\mathbf{c}}$  so that the strict inequality (8.3) is satisfied. We also sketch a second way of effecting the perturbation which is in a sense more robust, but which in essence requires a perturbation of the whole proof of (8.4).

### 8.2 Analysis of non-basic flags

We turn now to the task of proving (8.4). We will prove it for progressively wider sets of subflags  $\mathcal{V}'$ , each time using the previous statement. In order, we will prove it for subflags  $\mathcal{V}'$  which we call:

(a) *semi-basic*: flags

$$\mathcal{V}' : V_0 \leq V_1 \leq V_2 \leq \dots \leq V_{m-1} \leq \dots \leq V_{m-1} \leq V_m \leq \dots \leq V_m$$

with  $m \geq 1$  (that is,  $\mathcal{V}'$  is like a basic flag, but there can be more than one copy of  $V_{m-1}$ );

(b) *standard*: each  $V'_i$  is one of the spaces  $V_j$ ;

- (c) *invariant*: this means that  $\sigma V'_i = V'_i$  for all automorphisms  $\sigma \in \text{Aut}(\mathcal{V})$  and all  $i$ ;
- (d) *general subflags*, i.e. we assume no restriction on the  $V'_i$  other than that  $V'_i \leq V_i$ .

Note that a semi-basic flag is standard, a standard flag is invariant, and of course an invariant flag is general.

We introduce some notation for standard flags. Let  $J \subset \mathbb{N}_0^r$  be the set of all  $r$ -tuples  $\mathbf{j} = (j_1, \dots, j_r)$  such that  $j_1 \leq \dots \leq j_r$  and  $j_i \leq i$  for all  $i$ . Then we define the flag  $\mathcal{V}_{\mathbf{j}}' = \mathcal{V}'_{(j_1, \dots, j_r)}$  to be the one with  $V'_i = V_{j_i}$ . This is a standard flag, and conversely every standard flag is of this form. If we define

$$\text{basic}(m) := (1, 2, \dots, m - 1, m, \dots, m)$$

then  $\text{basic}(m) \in J$ , and  $\mathcal{V}'_{\text{basic}(m)}$  agrees with our previous notation.

### 8.3 Semi-basic subflags

In this subsection we prove the following result, establishing that (8.4) holds for semi-basic subflags, and with strict inequality for those which are not basic.

**Lemma 8.3** (Assuming that (a), (b) and (c) of Proposition 8.2 hold) we have  $e(\mathcal{V}') > e(\mathcal{V})$  for all non-basic, semi-basic flags  $\mathcal{V}'$ .

We begin by setting a small amount of notation for semi-basic flags. We note that the idea of a semi-basic flag, which looks rather *ad hoc*, will only be used here and in Sect. 8.5.

**Definition 8.4** (Semi-basic flags that are not basic) Suppose that  $1 \leq m \leq r - 1$  and that  $m \leq s \leq r - 1$ . Then we define the element  $\text{semi}(m, s) \in J$  to be  $\mathbf{j} = (1, 2, \dots, m - 1, m - 1, \dots, m, \dots, m)$  such that  $j_i = i$  for  $i \leq m - 1$ ,  $j_i = m - 1$  for  $m \leq i \leq s$  and  $j_i = m$  for  $i > s$ .

It is convenient and natural to extend the notation to  $s = m - 1$  and  $s = r$ , by defining

$$\text{semi}(m, r) = \text{basic}(m - 1), \quad \text{semi}(m, m - 1) = \text{basic}(m). \tag{8.5}$$

One can think of the semi-basic flags as interpolating between the basic flags.

*Example* When  $r = 3$  there are three semi-basic flags  $\mathcal{V}_{\mathbf{j}}$  that are not basic, corresponding to

$$\begin{aligned} \mathbf{j} &= \text{semi}(1, 1) = (0, 1, 1), \\ \mathbf{j} &= \text{semi}(1, 2) = (0, 0, 1), \\ \mathbf{j} &= \text{semi}(2, 2) = (1, 1, 2). \end{aligned}$$

*Proof of Lemma 8.3* Assume that  $\mathcal{V}'$  is semi-basic but not basic. We will show that

$$e(\mathcal{V}'_{\text{semi}(m,s)}) > e(\mathcal{V}'_{\text{semi}(m,s+1)}) \tag{8.6}$$

for  $m \leq s \leq r - 1$ . Since  $\mathcal{V}'_{\text{semi}(m,r)} = \mathcal{V}'_{\text{basic}(m-1)}$  is basic, this establishes Lemma 8.3.

To prove (8.6), we simply compute that

$$\begin{aligned} & e(\mathcal{V}'_{\text{semi}(m,s)}) - e(\mathcal{V}'_{\text{semi}(m,s+1)}) \\ &= (c_{s+1}^* - c_{s+2}^*) [\mathbb{H}_{\mu_{s+1}}(V_m) - \mathbb{H}_{\mu_{s+1}}(V_{m-1}) + \dim(V_m/V_{m-1})] \end{aligned}$$

when  $m \leq s \leq r - 2$ , and

$$\begin{aligned} & e(\mathcal{V}'_{\text{semi}(m,r-1)}) - e(\mathcal{V}'_{\text{semi}(m,r)}) \\ &= (c_r^* - c_{r+1}^*) [\mathbb{H}_{\mu_r}(V_m) - \mathbb{H}_{\mu_r}(V_{m-1}) + \dim(V_m/V_{m-1})] \\ & \quad + \dim(V_m/V_{m-1})c_{r+1}^*. \end{aligned}$$

In both cases, the result follows from part (ii) of condition(c) of Proposition 8.2; in the second case, we also need to use our assumption that  $c_{r+1}^* \geq 0$ .  $\square$

### 8.4 Submodularity inequalities

To proceed further, we make heavy use of a submodularity property of the expressions  $e(\cdot)$ .

Suppose that  $\mathcal{V}'$ ,  $\tilde{\mathcal{V}}'$  are two subflags of  $\mathcal{V}$ . We can define the *sum*  $\mathcal{V}' + \tilde{\mathcal{V}}'$  and *intersection*  $\mathcal{V}' \cap \tilde{\mathcal{V}}'$  by

$$(\mathcal{V}' + \tilde{\mathcal{V}}')_i := V'_i + \tilde{V}'_i$$

and

$$(\mathcal{V}' \cap \tilde{\mathcal{V}}')_i := V'_i \cap \tilde{V}'_i.$$

Both of these are indeed subflags of  $\mathcal{V}$ .

**Lemma 8.5** *We have*

$$e(\mathcal{V}') + e(\tilde{\mathcal{V}}') \geq e(\mathcal{V}' + \tilde{\mathcal{V}}') + e(\mathcal{V}' \cap \tilde{\mathcal{V}}').$$

*Proof* We first note that the entropies  $\mathbb{H}_\mu(W)$  satisfy a submodularity inequality. Namely, if  $W_1, W_2$  are subspaces of  $\mathbb{Q}^k$  and  $\mu$  is a probability measure then

$$\mathbb{H}_\mu(W_1) + \mathbb{H}_\mu(W_2) \geq \mathbb{H}_\mu(W_1 \cap W_2) + \mathbb{H}_\mu(W_1 + W_2). \quad (8.7)$$

To prove this, consider the following three random variables:

- $X$  is a random coset of  $W_1 + W_2$ , sampled according to the measure  $\mu$ ;
- $Y$  is a random coset of  $W_1$ , sampled according to the measure  $\mu$ ;
- $Z$  is a random coset of  $W_2$ , sampled according to the measure  $\mu$ .

Then, more-or-less by definition,

$$\mathbb{H}(X) = \mathbb{H}_\mu(W_1 + W_2), \quad \mathbb{H}(Y) = \mathbb{H}_\mu(W_1), \quad \mathbb{H}(Z) = \mathbb{H}_\mu(W_2).$$

Note also that  $Y$  determines  $X$  and so  $\mathbb{H}(Y) = \mathbb{H}(X, Y)$ , and similarly  $\mathbb{H}(Z) = \mathbb{H}(X, Z)$ . Finally,  $(Y, Z)$  uniquely defines a random coset of  $W_1 \cap W_2$ , and so

$$\mathbb{H}_\mu(W_1 \cap W_2) = \mathbb{H}(Y, Z) = \mathbb{H}(X, Y, Z).$$

The inequality to be proven, (8.7) is therefore equivalent to

$$\mathbb{H}(X, Y) + \mathbb{H}(X, Z) \geq \mathbb{H}(X, Y, Z) + \mathbb{H}(X),$$

which is a standard entropy inequality (Lemma B.6; usually known as “submodularity of entropy” or “Shannon’s inequality” in the literature).

Lemma 8.5 is essentially an immediate consequence of (8.7) and the formula

$$\dim(W_1) + \dim(W_2) = \dim(W_1 \cap W_2) + \dim(W_1 + W_2).$$

(It is very important that this formula holds with *equality*, as compared to (8.7), which holds only with an inequality.)  $\square$

This has the following immediate corollary when applied to standard subflags. Here, the max and min are taken coordinatewise.

**Corollary 8.6** *Suppose that  $\mathbf{j}_1, \mathbf{j}_2 \in J$ . Then*

$$e(\mathcal{V}'_{\mathbf{j}_1}) + e(\mathcal{V}'_{\mathbf{j}_2}) \geq e(\mathcal{V}'_{\max(\mathbf{j}_1, \mathbf{j}_2)}) + e(\mathcal{V}'_{\min(\mathbf{j}_1, \mathbf{j}_2)})$$



### 8.5 Standard subflags

Now we extend the result of the Sect. 8.3 to all standard subflags.

**Lemma 8.7** (Assuming that (a), (b) and (c) of Proposition 8.2 hold) we have  $e(\mathcal{V}') > e(\mathcal{V})$  for all standard, non-basic subflags  $\mathcal{V}' \leq \mathcal{V}$ .

*Proof* Let  $\mathbf{j} \in J$  with  $\mathbf{j}$  non-basic, and let  $\mathcal{V}' = \mathcal{V}'_{\mathbf{j}}$ . Then  $r \geq 3$ , since when  $r \leq 2$  all standard flags are basic. We proceed by induction on  $\|\mathbf{j}\|_{\infty}$ , the case  $\|\mathbf{j}\|_{\infty} = 1$  being trivial, since then  $\mathcal{V}$  is semibasic and we may invoke Lemma 8.3. Now suppose we have proved  $e(\mathcal{V}') > e(\mathcal{V})$  for all non-basic standard flags  $\mathcal{V}' = \mathcal{V}'_{\mathbf{j}}$  with  $\|\mathbf{j}\|_{\infty} < m$ , and let  $\mathbf{j} \in J$  with  $\|\mathbf{j}\|_{\infty} = m$ . We apply Corollary 8.6 with  $\mathbf{j}_1 = \mathbf{j}$  and  $\mathbf{j}_2 = \text{basic}(j_r - 1)$ . Noting that  $\max(\mathbf{j}, \text{basic}(j_r - 1)) = \text{semi}(j_r, s)$ , where  $s$  is the largest index in  $\mathbf{j}$  such that  $j_s < j_r$ , we see that

$$e(\mathcal{V}'_{\mathbf{j}}) + e(\mathcal{V}'_{\text{basic}(j_r-1)}) \geq e(\mathcal{V}'_{\mathbf{j}_*}) + e(\mathcal{V}'_{\text{semi}(j_r,s)}), \tag{8.8}$$

where

$$\mathbf{j}_* := \min(\mathbf{j}, \text{basic}(j_r - 1)).$$

Suppose that both of the flags on the right of (8.8) are basic. If  $\text{semi}(j_r, s)$  is basic then it must be  $\text{basic}(j_r)$ , which means that  $s = j_r - 1$ . But then  $\mathbf{j}_* = (j_1, \dots, j_s, j_r - 1, \dots, j_r - 1)$  which, if it is basic, must be  $\text{basic}(j_r - 1)$ ; this then implies that  $j_i = i$  for  $1 \leq i \leq s$ , and hence that  $\mathbf{j} = \text{basic}(j_r)$ , a contradiction. Thus, at least one of the two flags  $\mathbf{j}_*$ ,  $\text{semi}(j_r, s)$  on the right of (8.8) is not basic. Since  $\|\mathbf{j}_*\|_{\infty} < \|\mathbf{j}\|_{\infty} = m$ , the induction hypothesis together with Lemma 8.3 implies that  $e(\mathcal{V}') > e(\mathcal{V})$ , as desired.  $\square$

### 8.6 Invariant subflags

Now we extend our results to all invariant flags, but now without the strict inequality.

**Lemma 8.8** (Assuming that (a), (b) and (c) of Proposition 8.2 hold) we have  $e(\mathcal{V}') \geq e(\mathcal{V})$  for all invariant subflags  $\mathcal{V}' \leq \mathcal{V}$ .

*Proof* We associate a pair  $(i, \ell)$ ,  $i \geq \ell$ , of positive integers to  $\mathcal{V}'$ , which we call the *signature*, in the following manner. If  $\mathcal{V}'$  is standard, then set  $(i, \ell) = (-1, -1)$ . Otherwise, let  $i$  be maximal so that  $V'_i$  is not a standard space  $V_t$ , and then let  $\ell$  be minimal such that  $V'_i \leq V_{\ell}$ . The fact that  $\ell \leq i$  is immediate from the definition of a subflag. We put a partial ordering on signatures as follows:  $(i', \ell') \leq (i, \ell)$  iff  $i' < i$ , or if  $i' = i$  and  $\ell' \leq \ell$ . We

proceed by induction on the pair  $(i, \ell)$  with respect to this ordering, the case  $(i, \ell) = (-1, -1)$  handled by Lemma 8.7.

For the inductive step, suppose  $\mathcal{V}'$  is nonstandard with signature  $(i, \ell)$ . By submodularity,

$$e(\mathcal{V}') + e(\mathcal{V}'_{\text{basic}(\ell-1)}) \geq e(\mathcal{V}_1) + e(\mathcal{V}_2), \tag{8.9}$$

where

$$\mathcal{V}_1 = \mathcal{V}' \cap \mathcal{V}'_{\text{basic}(\ell-1)}, \quad \mathcal{V}_2 = \mathcal{V}' + \mathcal{V}'_{\text{basic}(\ell-1)}.$$

Suppose that  $\mathcal{V}_1, \mathcal{V}_2$  have signatures  $(i_1, \ell_1), (i_2, \ell_2)$ , respectively. We show that

$$(i_1, \ell_1) \not\geq (i, \ell) \quad \text{and} \quad (i_2, \ell_2) \not\geq (i, \ell). \tag{8.10}$$

Both  $\mathcal{V}_1$  and  $\mathcal{V}_2$  are invariant flags. Thus, if (8.10) holds, then both flags on the right-hand side of (8.9) have strictly smaller signature than  $\mathcal{V}'$ , and the lemma follows by induction.

Finally, we prove (8.10). Note that if  $j > i$ , then  $V'_j$  is a standard space  $V_m$  and thus so are  $(\mathcal{V}_1)_j$  and  $(\mathcal{V}_2)_j$ . In particular,  $i_1 \leq i$  and  $i_2 \leq i$ . We have that  $(\mathcal{V}_2)_i$  contains  $V_{\ell-1}$ , is not equal to  $V_{\ell-1}$ , and is contained in  $V_\ell$ . But  $(\mathcal{V}_2)_i$  is invariant, and hence by our assumption that (a) of Proposition 8.2 holds,  $(\mathcal{V}_2)_i = V_\ell$ . Consequently,  $i_2 < i$  if  $\mathcal{V}_2$  is nonstandard. In the case that  $\mathcal{V}_1$  is nonstandard, we also have that  $\ell_1 < \ell$  because every space in the flag  $\mathcal{V}_1$  is contained in  $V_{\ell-1}$ . This proves (8.10).  $\square$

### 8.7 General subflags

In this section we establish (8.4), that is to say the inequality  $e(\mathcal{V}') \geq e(\mathcal{V})$  for all subflags  $\mathcal{V}'$ , of course subject to our standing assumption that (a), (b) and (c) of Proposition 8.2 hold. We need a simple lemma about the action of the automorphism group  $\text{Aut}(\mathcal{V})$  on subflags.

**Lemma 8.9** *Let  $\sigma \in \text{Aut}(\mathcal{V})$  and let  $\mathcal{V}'$  be a subflag of  $\mathcal{V}$ . Then one may define a new subflag  $\sigma(\mathcal{V}')$ , setting  $\sigma(\mathcal{V}')_i := \sigma(V'_i)$ . Moreover,  $e(\sigma(\mathcal{V}')) = e(\mathcal{V}')$ .*

*Proof* Since  $\mathcal{V}'$  is a subflag,  $V'_i \leq V_i$ . Applying  $\sigma$ , and recalling that  $V_i$  is invariant under  $\sigma$ , we see that  $\sigma(V'_i) \leq V_i$ . Therefore  $\sigma(\mathcal{V}')$  is also a subflag. To see that  $e(\sigma(\mathcal{V}')) = e(\mathcal{V}')$ , recall Lemma 7.6, which implies that  $\mu_i$  is invariant under  $\sigma$ , since the trees  $\mathcal{T}(\mathcal{V}')$  and  $\mathcal{T}(\sigma(\mathcal{V}'))$  are isomorphic and we have  $\dim(V'_j) = \dim(\sigma(V'_j))$  for all  $j$ . It follows that, for any subspace  $W \leq \mathbb{Q}^k$ ,

$$\begin{aligned}
 \mathbb{H}_{\mu_i}(\sigma(W)) &= - \sum_x \mu_i(x) \log \mu_i(\sigma(W) + x) \\
 &= - \sum_y \mu_i(\sigma(y)) \log \mu_i(\sigma(W) + y) \\
 &= - \sum_y \mu_i(y) \log \mu_i(W + y) \\
 &= \mathbb{H}_{\mu_i}(W).
 \end{aligned}$$

This completes the proof of the lemma. □

*Proof of (8.4)* Let  $m$  be the minimum of  $e(\mathcal{V}')$  over all subflags  $\mathcal{V}' \leq \mathcal{V}$ , and among the flags with  $e(\mathcal{V}') = m$ , take the one with  $\sum_i \dim V'_i$  minimal. Let  $\sigma \in \text{Aut}(\mathcal{V})$  be an arbitrary automorphism. By Lemma 8.9,  $e(\mathcal{V}') = e(\sigma(\mathcal{V}'))$ , and hence submodularity implies that

$$2e(\mathcal{V}') \geq e(\mathcal{V}' + \sigma(\mathcal{V}')) + e(\mathcal{V}' \cap \sigma(\mathcal{V}')). \tag{8.11}$$

In particular, we have  $e(\mathcal{V}' \cap \sigma(\mathcal{V}')) = m$  (and also  $e(\mathcal{V}' + \sigma(\mathcal{V}')) = e(\mathcal{V})$ , but we will not need this). Moreover, by the minimality of  $\sum_i \dim V'_i$ ,

$$\sum_i \dim(V'_i \cap \sigma(V'_i)) = \sum_i \dim V'_i,$$

which means that  $\mathcal{V}'$  is invariant. Invoking Lemma 8.8, we conclude that  $m = e(\mathcal{V}') \geq e(\mathcal{V})$ . □

### 8.8 The strict entropy condition

In this section we complete the proof of Proposition 8.2 by showing how to perturb (8.4) to the desired strict inequality (8.3).

*First argument.* Consider first the collection  $\mathcal{U}$  of all subflags  $\mathcal{V}'$  which satisfy, for some  $1 \leq j \leq r - 1$ , the relations

$$V'_i = V_i \quad (i \neq j), \quad V_{j-1} \leq V_{j'} < V_j.$$

These are flags which differ from  $\mathcal{V}$  in exactly one space. Our first task will be to establish the *strict* inequality

$$e(\mathcal{V}') > e(\mathcal{V}) \tag{8.12}$$

for all  $\mathcal{V}' \in \mathcal{U}$ , by elaborating upon the argument of the previous subsection. We already know that  $e(\mathcal{V}') \geq e(\mathcal{V})$ , so suppose as a hypothesis for contra-

diction that  $e(\mathcal{V}') = e(\mathcal{V})$  for some  $\mathcal{V}' \in \mathcal{U}$ . Amongst all such flags, take one with minimal  $\sum \dim(V'_i)$ . By submodularity, we have (8.11) and hence  $e(\mathcal{V}' \cap \sigma(\mathcal{V}')) = e(\mathcal{V})$  for any automorphism  $\sigma \in \text{Aut}(\mathcal{V})$ . But

$$\mathcal{V}' \cap \sigma(\mathcal{V}') = (V_1, \dots, V_{j-1}, V'_j \cap \sigma(V'_j), V_{j+1}, \dots, V_r)$$

is evidently in  $\mathcal{U}$  as well, and by our minimality assumption it follows that  $\dim(V'_j \cap \sigma(V'_j)) = \dim(V'_j)$ . Thus,  $\mathcal{V}'$  is invariant, and by assumption (a) of Proposition 8.2, it follows that  $V'_j = V_{j-1}$ . Thus,  $\mathcal{V}'$  is a standard flag, which is not basic since  $j \leq r - 1$ . Hence,  $e(\mathcal{V}') > e(\mathcal{V})$  by Lemma 8.7. This contradiction establishes (8.12).

Let  $1 \leq j \leq r - 1$  and let  $V$  be a space satisfying  $V_{j-1} \leq V < V_j$ . Let  $\mathcal{V}'$  be the subflag  $\langle \mathbf{1} \rangle = V_0 \leq \dots V_{j-1} \leq V \leq V_{j+1} \leq \dots \leq V_r$ . Then one easily computes that

$$e(\mathcal{V}') - e(\mathcal{V}) = (c_j - c_{j+1})(\mathbb{H}_{\mu_j}(V) - \dim(V_j/V)),$$

and so (8.12) implies that

$$\mathbb{H}_{\mu_j}(V) > \dim(V_j/V). \tag{8.13}$$

Now let  $\varepsilon > 0$  be sufficiently small and consider the perturbation  $\tilde{\mathbf{c}}$  given by

$$\tilde{c}_1 = 1, \quad \tilde{c}_j = c_j^* - \frac{1}{2} \sum_{\ell=1}^{j-1} \varepsilon^\ell \quad (2 \leq j \leq r + 1).$$

Evidently,  $1 = \tilde{c}_1 > \tilde{c}_2 > \dots > \tilde{c}_{r+1} \geq c_{r+1}^* - \varepsilon$ , as needed. For any proper subflag  $\mathcal{V}' \leq \mathcal{V}$ ,

$$\begin{aligned} & e(\mathcal{V}', \tilde{\mathbf{c}}, \mu^*) - e(\mathcal{V}, \tilde{\mathbf{c}}, \mu^*) \\ &= e(\mathcal{V}') - e(\mathcal{V}) + \frac{1}{2} \sum_{j=1}^r \varepsilon^j (\mathbb{H}_{\mu_j}(V'_j) - \dim(V_j/V'_j)) \\ & \quad + \frac{1}{2} (\varepsilon + \varepsilon^2 + \dots + \varepsilon^r) \dim(V_r/V'_r). \end{aligned}$$

Let  $J = \min\{j : V'_j \neq V_j\}$ . If  $J = r$ , then  $\dim(V_r/V'_r) \geq 1$  and the right side above is at least  $\varepsilon/2 + O(\varepsilon^r)$ , which is positive for small enough  $\varepsilon$ . If  $J \leq r - 1$ , then  $V_{J-1} \leq V'_J < V_J$  and we see that the right side above is at least

$$e(\mathcal{V}') - e(\mathcal{V}) + \varepsilon^J (\mathbb{H}_{\mu_J}(V'_J) - \dim(V_J/V'_J)) + O(\varepsilon^{J+1}),$$

which is also positive for sufficiently small  $\varepsilon$  by (8.4) and (8.12).

*Second argument.* We now sketch a second approach to the proof of Proposition 8.2. The idea is to introduce a small perturbation of our fundamental quantity  $e()$ , namely

$$e_\lambda(\mathcal{V}', \mathbf{c}, \boldsymbol{\mu}) := \lambda \sum_{j=1}^r (c_{j+1} - c_j) \mathbb{H}_{\mu_j}(V'_j) + \sum_{j=1}^r c_j \dim(V'_j/V'_{j-1}),$$

where  $\lambda \approx 1$ . Note that  $e_1(\mathcal{V}', \mathbf{c}, \boldsymbol{\mu}) = e(\mathcal{V}', \mathbf{c}, \boldsymbol{\mu})$ , and also that  $e_\lambda(\mathcal{V}, \mathbf{c}, \boldsymbol{\mu})$  does not depend on  $\lambda$ , since all the entropies  $\mathbb{H}_{\mu_j}(V_j)$  vanish. Define the  $\lambda$ -perturbed optimal parameters  $\mathbf{c}^*(\lambda)$  to be the unique solution to the  $\lambda$ -perturbed version of (7.11), that is to say the equations

$$e_\lambda(\mathcal{V}'_{\text{basic}(m)}, \mathbf{c}^*(\lambda), \boldsymbol{\mu}) = e_\lambda(\mathcal{V}, \mathbf{c}^*(\lambda), \boldsymbol{\mu}), m = 0, 1, \dots, r - 1.$$

By a continuity argument, these exist for  $\lambda$  sufficiently close to 1 and they satisfy  $\lim_{\lambda \rightarrow 1} \mathbf{c}^*(\lambda) = \mathbf{c}^*(1) = \mathbf{c}^*$ .

Now, assume that  $\lambda$  is close enough to 1 so that

$$1 = c_1^*(\lambda) > c_2^*(\lambda) > \dots > c_{r+1}^*(\lambda) > 0$$

and we have the following “positivity inequalities”:

- (i)  $\lambda \mathbb{H}_{\mu_{m+1}^*}(V_m) > \dim(V_{m+1}/V_m)$  for  $0 \leq m \leq r - 1$ ;
- (ii)  $\lambda \cdot (\mathbb{H}_{\mu_i^*}(V_{m-1}) - \mathbb{H}_{\mu_i^*}(V_m)) < \dim(V_m/V_{m-1})$  for  $1 \leq m < i \leq r$ .

These conditions can be clearly guaranteed by a continuity argument and our assumption that they hold when  $\lambda = 1$ . For a parameter  $\lambda$  satisfying (i) and (ii) above, the proof of (8.4) holds verbatim for the  $\lambda$ -perturbed quantities  $e_\lambda$ , allowing one to conclude that

$$e_\lambda(\mathcal{V}', \mathbf{c}^*(\lambda), \boldsymbol{\mu}) \geq e_\lambda(\mathcal{V}, \mathbf{c}^*(\lambda), \boldsymbol{\mu})$$

for all subflags  $\mathcal{V}'$  of  $\mathcal{V}$ .

Now suppose that  $\lambda < 1$ . Then we have

$$e(\mathcal{V}', \mathbf{c}, \boldsymbol{\mu}^*) \geq e_\lambda(\mathcal{V}', \mathbf{c}, \boldsymbol{\mu}^*),$$

with equality if and only if  $\mathcal{V}' = \mathcal{V}$  because  $\text{Supp}(\mu_j^*) = V_j \cap \{0, 1\}^k$  for all  $j$ . Therefore if  $\mathcal{V}'$  is a proper subflag of  $\mathcal{V}$  we have

$$\begin{aligned} e(\mathcal{V}', \mathbf{c}^*(\lambda), \boldsymbol{\mu}^*) &> e_\lambda(\mathcal{V}', \mathbf{c}^*(\lambda), \boldsymbol{\mu}^*) \\ &\geq e_\lambda(\mathcal{V}, \mathbf{c}^*(\lambda), \boldsymbol{\mu}^*) = e(\mathcal{V}, \mathbf{c}^*(\lambda), \boldsymbol{\mu}^*). \end{aligned}$$

Taking  $\tilde{\mathbf{c}} = \mathbf{c}^*(\lambda)$  for  $\lambda$  sufficiently close to 1, Proposition 8.2 follows.

## Part IV. Binary systems

### 9 Binary systems and a lower bound for $\beta_k$

In this section we define certain special flags  $\mathcal{V}$  on  $\mathbb{Q}^k$ ,  $k = 2^r$ , which we call the *binary systems of order  $r$* . It is these systems which lead to the lower bound on  $\beta_k$  given in Theorem 2, which is one of the main results of the paper.

In this section we will define these flags (which is easy) and state their basic properties. The proofs of these properties, some of which are quite lengthy, are deferred to Sect. 10.

We are then in a position to prove part of one of our main theorems, Theorem 2 (a), which we do in Sect. 9.2.

For the convenience of the reader, recall us here the three parts of Theorem 2, as stated at the end of Sect. 1.3:

- (a) Showing that for every  $r \geq 1$ ,  $\beta_{2^r} \geq \theta_r$  for a certain explicitly defined constant  $\theta_r$ ;
- (b) Showing that  $\lim_{r \rightarrow \infty} \theta_r^{1/r}$  exists;
- (c) Showing that (1.1) has a unique solution  $\rho \in [0, 1/3]$  and that  $\rho = 2 \lim_{r \rightarrow \infty} \theta_r^{1/r}$ .

#### 9.1 Binary flags and systems: definitions and properties

**Definition 9.1** (*Binary flag of order  $r$* ) Let  $k = 2^r$  be a power of two. Identify  $\mathbb{Q}^k$  with  $\mathbb{Q}^{\mathcal{P}[r]}$  (where  $\mathcal{P}[r]$  means the power set of  $[r] = \{1, \dots, r\}$ ) and define a flag  $\mathcal{V}$ ,  $\langle \mathbf{1} \rangle = V_0 \leq V_1 \leq \dots \leq V_r = \mathbb{Q}^{\mathcal{P}[r]}$ , as follows:  $V_i$  is the subspace of all  $(x_S)_{S \subset [r]}$  for which  $x_S = x_{S \cap [i]}$  for all  $S \subset [r]$ .

*Remark* We have  $\dim(V_i) = 2^i$ , and  $V_r = \mathbb{Q}^{\mathcal{P}[r]}$ , so the system is trivially nondegenerate. Note that we have been using the letter  $r$  to denote the number of  $V_i$  in the flag  $\mathcal{V}$ , throughout the paper. It just so happens that, in this example, this is the same  $r$  as in the definition of  $k = 2^r$ .

One major task is to show that optimal measures and optimal parameters, as described in Sect. 7, may be defined on the binary flags. Since we will be seeing them so often, let us write down the  $\rho$ -equations (7.5) for the binary flags explicitly:

$$f^{\Gamma_{j+1}}(\rho) = f^{\Gamma_j}(\rho)^{\rho_j} e^{2^j}, \quad j = 1, 2, \dots \quad (9.1)$$

**Proposition 9.2** *Let  $\mathcal{V}$  be the binary flag of order  $r$ . Then*

- (a) *the  $\rho$ -equations (9.1) have a solution with  $0 < \rho_i < 1$  for  $i \geq 1$ , and consequently we may define the optimal measures  $\mu^*$  on  $\{0, 1\}^k$  as in Definition 7.4;*
- (b) *the optimal parameters  $\mathbf{c}^*$  (in the sense of Definition 7.5) exist.*

We call the binary flag  $\mathcal{V}$  (of order  $r$ ) together with the additional data of the optimal measures  $\mu = \mu^*$  and optimal parameters  $\mathbf{c} = \mathbf{c}^*$ , the binary system (of order  $r$ ). We caution that for fixed  $i$  (such as  $i = 2$ ) the parameters  $c_i$  do depend on  $r$ , although not very much.

The second major task is to show that the binary systems satisfy the entropy condition (3.4), or more accurately that arbitrarily small perturbations of them satisfy the strict entropy condition (3.5). In the last section we provided a tool for doing this in somewhat general conditions, namely Proposition 8.2. That proposition has four conditions, (a), (b), (c)(i) and (c)(ii) which must be satisfied. Of these, (b) (the existence of the optimal parameters  $\mathbf{c}^*$ ) has already been established, assuming the validity of Proposition 9.2. We state the other three conditions separately as lemmas.

**Lemma 9.3** *Suppose that  $V_{i-1} \leq W \leq V_i$  and that  $W$  is invariant under  $\text{Aut}(\mathcal{V})$ . Then  $W$  is either  $V_{i-1}$  or  $V_i$ . Thus, the binary flags satisfy Proposition 8.2 (a).*

**Lemma 9.4** *We have  $\mathbb{H}_{\mu_{m+1}^*}(V_m) > 2^m$  for  $0 \leq m \leq r - 1$ . Thus, the binary flags satisfy Proposition 8.2 (c)(i).*

**Lemma 9.5** *We have  $\mathbb{H}_{\mu_i^*}(V_{m-1}) - \mathbb{H}_{\mu_i^*}(V_m) < 2^{m-1}$  for  $1 \leq m < i \leq r$ . Thus, the binary flags satisfy Proposition 8.2 (c)(ii).*

The proofs of these various facts are given in Sect. 10.

## 9.2 Proof of Theorems 2 (a) and 7

We are now in a position to complete the proof of Theorem 2 (a), modulo the results stated above. First, we define the constants  $\theta_r$ .

**Definition 9.6** Let  $\rho_1, \rho_2, \dots$  be the solution to the  $\rho$ -equations (9.1) for the binary flag. Then we define

$$\theta_r := (\log 3 - 1) / \left( \log 3 + \sum_{i=1}^{r-1} \frac{2^i}{\rho_1 \dots \rho_i} \right).$$

*Proof of Theorem 2 (a)* By Proposition 7.7,  $\theta_r$  is equal to  $c_{r+1}^*$ , where  $\mathbf{c}^*$  are the optimal parameters on the binary flag  $\mathcal{V}$  of order  $r$ , the existence of which is Proposition 9.2 (b) above.

Fix  $\delta \in (0, \theta_r/2]$ . By Proposition 8.2 (the hypotheses of which are satisfied by Lemma 9.3, Proposition 9.2 (b) and Lemmas 9.4 and 9.5), there exists a perturbation  $\tilde{\mathbf{c}}$  of  $\mathbf{c}^*$  such that

$$1 = \tilde{c}_1 > \tilde{c}_2 > \dots > \tilde{c}_{r+1} \geq c_{r+1}^* - \delta = \theta_r - \delta > 0$$

and  $(\mathcal{V}, \tilde{\mathbf{c}}, \boldsymbol{\mu}^*)$  satisfies the strict entropy condition (3.5). By Lemma 5.2, there exists some  $\varepsilon > 0$  such that the “entropy gap” condition (5.1) holds. Finally, by Remark 7.1 (b), we have that  $\text{Supp}(\mu_j^*) = \Gamma_j$  for all  $j$ . Hence, Proposition 5.5 implies that  $\beta_{2^r} \geq \tilde{c}_{r+1} = \theta_r - \delta$ . Since  $\delta$  is arbitrary, this proves Theorem 2 (a). □

*Proof of Theorem 7* The upper bound  $\beta_k \leq \gamma_k$  is established in Sect. 4. The lower bound  $\beta_k \geq \tilde{\gamma}_k$  follows by Lemma 5.3, Proposition 5.5 and the fact that there exists at least one system satisfying the strict entropy condition (3.5), as per the proof of Theorem 2 (a) above. □

### 9.3 Remarks on Theorem 2 (b)

Theorem 2 (b) is a problem of a combinatorial and analytic nature which can be considered more-or-less completely independently of the first three parts of the paper.

To get a feel for it, and a sense of why it is difficult, let us write down the first two  $\rho$ -equations (9.1) for the binary flags. The equation with  $j = 1$  is

$$f^{\Gamma_2}(\rho) = f^{\Gamma_1}(\rho)^{\rho_1} e^2. \tag{9.2}$$

This has the numerical solution  $\rho_1 \approx 0.306481$ .

To write down the  $\rho$ -equation for  $j = 2$ , one must compute  $f^{\Gamma_3}(\rho)$ , and without any additional theory the only means we have to do this is to draw the full tree structure for the binary flag  $\mathcal{V}$  of order 3 (on  $\mathbb{Q}^8$ ). This is a tractable exercise and one may confirm that

$$f^{\Gamma_3}(\rho) = (3^{\rho_1} + 4 \cdot 2^{\rho_1} + 4)^{\rho_2} + 8(2 \cdot 2^{\rho_1} + 4)^{\rho_2} + 16 \cdot 4^{\rho_2} + 8 \cdot (2^{\rho_1} + 2)^{\rho_2} + 32 \cdot 2^{\rho_2} + 16.$$

The  $\rho$ -equation with  $j = 2$  is then

$$f^{\Gamma_3}(\rho) = f^{\Gamma_2}(\rho)^{\rho_2} e^4,$$



where (recall from Fig. 1)  $f^{\Gamma_2}(\rho) = 3^{\rho_1} + 4 \cdot 2^{\rho_1} + 4$ . This may be solved numerically, with the value  $\rho_2 \approx 0.2796104\dots$ , using Mathematica.

Such a numerical procedure, however, is already quite an unappetising prospect if one wishes to compute  $\rho_3$ .

Consequently, we must develop more theory to understand the  $\rho_i$  and to prove Theorem 2 (b). This is the task of the last two sections of the paper.

### 10 Binary systems: proofs of the basic properties

In this section, we prove the various statements in Sect. 9.1.

We begin, in Sect. 10.2, by proving Lemma 9.3. This is a relatively simple and self-contained piece of combinatorics.

In Sect. 10.3 we introduce the concept of *genotype*, which allows us to describe the tree structure induced on  $\{0, 1\}^k$  by the binary flag  $\mathcal{V}$ . In Sect. 10.4 we show how to compute the quantities  $f^C(\rho)$  in terms of the genotype.

We are then, in Sect. 10.5, in a position to prove Proposition 9.2 (a), guaranteeing that the  $\rho_i$  exist and allowing us to define the optimal measures  $\mu^*$ .

In Sect. 10.6 we establish the two entropy inequalities, Lemmas 9.4 and 9.5.

Finally, in Sect. 10.7 we prove Proposition 9.2 (b), which confirms the existence of the optimal parameters  $\mathbf{c}^*$ .

#### 10.1 Basic terminology

Throughout the section,  $\mathcal{V}$  will denote the binary flag or order  $r$ , as defined in Definition 9.1. That is, we take  $k = 2^r$ , identify  $\mathbb{Q}^k$  with  $\mathbb{Q}^{\mathcal{P}[r]}$ , and take  $V_i$  to be the subspace of all  $(x_S)_{S \subset \mathcal{P}[r]}$  for which  $x_S = x_{S \cap [i]}$  for all  $S \subset [r]$ .

In addition, we will write  $\mathbf{0}_j, \mathbf{1}_j$  for the vectors in  $\{0, 1\}^{\mathcal{P}[j]}$  consisting of all 0s (respectively all 1s). We call these (or any multiples of them) *constant vectors*.

Finally, we introduce the notion of a *block* of a vector

$$x = (x_S)_{S \subset [r]} \in \mathbb{Q}^{\mathcal{P}[r]}.$$

For each  $A \subset [i]$  we consider the  $2^{r-i}$ -tuple

$$x(A, i) := (x_{A \cup A'})_{A' \subset [i+1, \dots, r]}.$$

We call these the *i-blocks* of  $x$ .

*Remark 10.1* (a) One should note carefully that the *i*-blocks are strings of length  $2^{r-i}$ . In this language,  $V_i$  is the space of vectors  $x$ , all of whose *i*-blocks are constant.

(b) If we put together the coordinates of the *i*-blocks  $x(A, i)$  and  $x(A \Delta \{i\}, i)$ , then we obtain the  $(i - 1)$ -block  $x(A \cap [i - 1], i - 1)$ .

In order to visualize the structure of the flag  $\mathcal{V}$  and of the partition of  $\{0, 1\}^{\mathcal{P}[r]}$  by the cosets of  $V_j$ , it will be often useful to write elements of  $\{0, 1\}^{\mathcal{P}[r]}$  as strings of 0s and 1s of length  $2^r$ . When we do this we use the *reverse binary order*, which is the one induced from  $\mathbb{N}$  via the map  $f(S) = \sum_{s \in S} 2^{r-s}$ .

*Example 10.2* For concreteness, let us consider the case  $r = 3$ . In this case, the ordering of the coordinates of  $x$  is

$$(x_{\emptyset}, x_{\{3\}}, x_{\{2\}}, x_{\{2,3\}}, x_{\{1\}}, x_{\{1,3\}}, x_{\{1,2\}}, x_{\{3\}}). \tag{10.1}$$

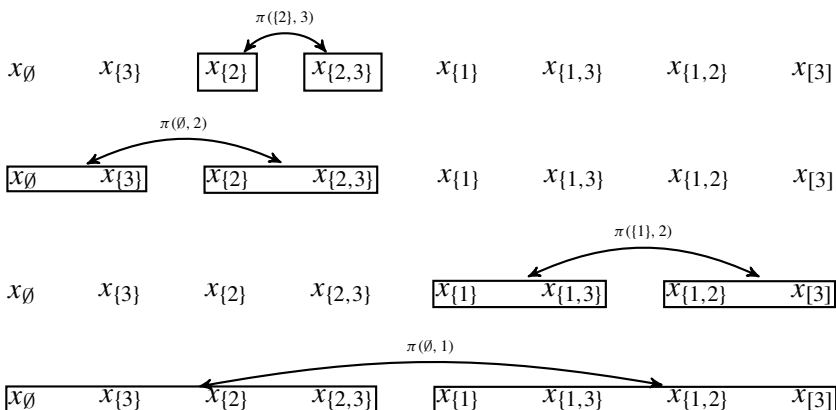
If  $x = 01001110$  then its 2-blocks are 01, 00, 11, 10, and its 1-blocks are 0100, 1110.

### 10.2 Automorphisms of the binary system

*Proof of Lemma 9.3* We begin by defining some permutations of  $\mathcal{P}[r]$  for which, we claim, the corresponding coordinate permutations give elements of  $\text{Aut}(\mathcal{V})$ . Suppose that  $1 \leq j \leq r$  and that  $A \subset [j - 1]$ . Then we may consider the permutation  $\pi(A, j)$  defined by

$$\pi(A, j)(S) = \begin{cases} S \Delta \{j\} & \text{if } S \cap [j - 1] = A, \\ S & \text{otherwise.} \end{cases}$$

To visualize the action of this permutation on the coordinates of a vector  $x$ , it is useful to order its coordinates as we explained above. The action of  $\pi(A, j)$  is then to permute the two adjacent  $j$ -blocks  $x(A, j)$  and  $x(A \sqcup \{j\}, j)$ , which together form the  $(j - 1)$ -block  $x(A, j - 1)$ , as per Remark 10.1(b). More concretely, below are some examples of the action of the permutations  $\pi(A, j)$  in the setting of Example 10.2:



If the readers wish, they may translate the arguments below in the above more visual language.

*Claim.*  $\pi(A, j)$  preserves  $V_i$  for all  $i$ , and therefore  $\pi(A, j) \in \text{Aut}(\mathcal{V})$ .

*Proof* Suppose that  $x = (x_S)_{S \subset [r]} \in V_i$  and let us write for simplicity  $\pi$  instead of  $\pi(A, j)$ .

Suppose first that  $j > i$ . Then  $\pi(S) \cap [i] = S \cap [i]$  for all  $S$ , and so

$$x_{\pi(S)} = x_{\pi(S) \cap [i]} = x_{S \cap [i]} = x_S.$$

where the first and last steps used the fact that  $x \in V_i$ . Thus the claim follows in this case.

Suppose that  $j \leq i$ . Let  $t > i$ . Then the conditions  $(S \Delta \{t\}) \cap [j - 1] = A$  and  $S \cap [j - 1] = A$  are equivalent. Hence, if  $S \cap [j - 1] = A$ , then we find that

$$x_{\pi(S \Delta \{t\})} = x_{S \Delta \{t\} \Delta \{j\}} = x_{S \Delta \{j\}} = x_{\pi(S)},$$

where we used that  $x \in V_i$  and that  $t > i$  at the second step. Similarly, if  $S \cap [j - 1] \neq A$ , then

$$x_{\pi(S \Delta \{t\})} = x_{S \Delta \{t\}} = x_S = x_{\pi(S)}.$$

In all cases, we have found that  $x_{\pi(S \Delta \{t\})} = x_{\pi(S)}$ . Since this is true for all  $t > i$ ,  $\pi(x)$  indeed lies in  $V_i$ . This completes the proof of the claim.  $\square$

Suppose now that  $W$  is an invariant subspace of  $\mathcal{V}$  satisfying the inclusions  $V_{i-1} < W \leq V_i$ . We want to conclude that  $W = V_i$ . To accomplish this, we introduce some auxiliary notation.

For each  $A \subset [i - 1]$ , we consider the vector  $y^A = (y^A_S)_{S \subset [r]} \in V_i$  that is uniquely determined by the relations  $y^A_A = 1$ ,  $y^A_{A \cup \{i\}} = -1$  and  $y^A_S = 0$  for all other  $S \subset [i]$ . There are  $2^{i-1}$  such vectors  $y^A$ . They are mutually orthogonal, hence linearly independent. In addition, together with  $V_{i-1}$ , they generate all of  $V_i$ . Since  $V_{i-1} < W \leq V_i$ , there must exist  $A \subset [i - 1]$  such that  $y^A \in W$ .

Now, it is easy to check that for any  $j < i$  and any  $A \subset [i - 1]$ , we have

$$\pi(A \cap [j - 1], j)y^A = y^{A \Delta \{j\}}.$$

From the above relation and the invariance of  $W$  under  $\text{Aut}(\mathcal{V})$ , it is clear that if  $W$  contains at least one vector  $y^A$  with  $A \subset [i - 1]$ , then it contains all such vectors. Since we also know that  $V_{i-1} \leq W \leq V_i$ , we must have that  $W = V_i$ , which completes the proof of Lemma 9.3.  $\square$

*Remark* A minor elaboration of the above argument in fact allows one to show that the subspaces of  $\mathbb{Q}^{\mathcal{P}[r]}$  invariant under  $\text{Aut}(\mathcal{V})$  are the  $V_i$ , the orthogonal complements of  $V_{i-1}$  in  $V_i$ , and all direct sums of these spaces. However, we will not need the classification in this explicit form.

### 10.3 Cell structure and genotype

The cosets of  $V_i$  partition  $\{0, 1\}^{\mathcal{P}[r]}$  into sets which we call the *cells at level  $i$* . Our first task is to describe these explicitly.

Consider  $\omega, \omega' \in \{0, 1\}^{\mathcal{P}[r]}$ . It is easy to see that  $\omega - \omega' \in V_i$  (and so  $\omega, \omega'$  lie in the same cell at level  $i$ ) if and only if for every  $A \subset [i]$  one of the following is true:

- (a) Both  $\omega(A, i)$  and  $\omega'(A, i)$  are constant blocks (that is, they both lie in  $\{\mathbf{0}_{r-i}, \mathbf{1}_{r-i}\}$ ).
- (b)  $\omega(A, i) = \omega'(A, i)$ , and neither of these blocks is constant (that is, neither is  $\mathbf{0}_{r-i}$  nor  $\mathbf{1}_{r-i}$ ).

Thus a cell at level  $i$  is completely specified by the *positions*  $A$  of its constant  $i$ -blocks, and the *values*  $\omega(A, i)$  (for an arbitrary  $\omega \in C$ ) of its non-constant  $i$ -blocks.

*Example* With  $r = 3$  and  $\omega = 01001110$ , the level 2 cell that contains  $\omega$  is the set

$$\{\omega, 01111110, 01000010, 01000010\}.$$

Its constant 2-blocks are at  $A = \{2\}$  and  $A = \{1\}$ . Its non-constant 2-blocks are at  $A = \emptyset$  (taking the value  $\omega(A, 2) = 01$ ) and at  $A = \{1, 2\}$  (taking the value  $\omega(A, 2) = 10$ ). The level 1 cell containing  $\omega$  is just  $\{\omega\}$ .

The positions of the constant  $i$ -blocks play an important role, and we introduce the name *genotype* to describe these.<sup>8</sup>

**Definition 10.1** (*Genotype*) If  $C$  is a cell at level  $i$ , its *genotype*  $g(C) \subset \mathcal{P}[i]$  is defined to be the collection of  $A \subset [i]$  for which  $\omega(A, i) \in \{\mathbf{0}_{r-i}, \mathbf{1}_{r-i}\}$  for all  $\omega \in C$ . We refer to any subset of  $\mathcal{P}[i]$  as an  *$i$ -genotype*. If  $g, g'$  are two  $i$ -genotypes, then we write  $g \leq g'$  to mean the same as  $g \subseteq g'$ . We write  $|g|$  for the cardinality of  $g$ .

*Example* If  $C$  is the cell at level 2 containing  $\omega = 01001110$ , the genotype  $g(C)$  is equal to  $\{\{2\}, \{1\}\}$ . (We have listed these sets in the reverse binary ordering once again.)

<sup>8</sup> The term genotype is appropriate, as each component in  $g$  acts like recessive gene with respect to child cells.

**Definition 10.2** (*Consolidations*) If  $g$  is an  $i$ -genotype, then its *consolidation* is the  $(i - 1)$ -genotype  $g^*$  defined by  $g^* := \{A' \subset [i - 1] : A' \in g, A' \cup \{i\} \in g\}$  (cf. Remark 10.1 (b)).

Let us pause to note the easy inequality

$$\frac{1}{2}|g| \geq |g^*| \geq |g| - 2^{i-1}, \tag{10.2}$$

valid for all  $i$ -genotypes.

The genotype is intimately connected to the cell structure on  $\{0, 1\}^k$  induced by  $\mathcal{V}$ , as the following lemma shows.

**Lemma 10.3** *We have the following statements.*

- (a) *If  $C$  is a cell, we have  $|C| = 2^{|g(C)|}$ .*
- (b) *Suppose that  $g$  is an  $i$ -genotype. There are  $(2^{2^{r-i}} - 2)^{2^i - |g|}$  cells (at level  $i$ ) with  $g(C) = g$ .*
- (c) *If  $g(C) = g$ , and if  $C'$  is a child of  $C$ , then  $g(C') \leq g^*$ . In particular,  $|g(C')| \leq \frac{1}{2}|g(C)|$ .*
- (d) *Suppose that  $g(C) = g$ . Suppose that  $g'$  is an  $(i - 1)$ -genotype and that  $g' \leq g^*$ . Then number of children  $C'$  of  $C$  with  $g(C') = g'$  is  $2^{|g| - |g^*| - |g'|}$ .*
- (e) *Suppose that  $C$  is a cell at level  $i$  with  $g(C) = g$ . Then the number of children of  $C$  (at level  $i - 1$ ) is  $2^{|g| - 2|g^*|} 3^{|g^*|}$ .*

*Proof* (a) This is almost immediate: for each of the  $A \subset g(C)$  of constant blocks, there are two choices ( $\mathbf{0}_{r-i}$  or  $\mathbf{1}_{r-i}$ ) for  $\omega(A, i)$ .

(b) To determine  $C$  completely (given  $g$ ), one must specify the value of each of  $2^i - |g|$  non-constant  $i$ -blocks. For each such block, there are  $2^{2^{r-i}} - 2$  possible non-constant values.

(c) A set  $A' \subset [i - 1]$  can only possibly be the position of a constant block in some child cell of  $C$  if both  $A'$  and  $A' \cup \{i\}$  are the positions of constant blocks in  $C$ , or in other words  $A', A' \cup \{i\} \in g$ , which is precisely what it means for  $A'$  to lie in  $g^*$ .

Note that the child cell  $C'$  containing  $\omega$  only *does* have a constant  $(i - 1)$ -block at position  $A'$  if  $\omega(A', i) = \omega(A' \cup \{i\}, i)$ , which may or may not happen.

The second statement is an immediate consequence of the first and (10.2).

(d) Let  $A \in g$ . We say that  $A$  is *productive* if  $A' := A \cap [i - 1] \in g^*$ , or equivalently if  $A'$  and  $A' \cup \{i\}$  both lie in  $g$  (or, more succinctly,  $A \Delta \{i\} \in g$ ). These are the positions which can give rise to constant  $(i - 1)$ -blocks in children of  $C$ . There are  $2|g^*|$  such positions, coming in  $|g^*|$  pairs. To create a child  $C'$  with genotype  $g'$ , we have a binary choice at  $|g^*| - |g'|$  of these pairs: at each of them either  $\omega(A', i) = \mathbf{0}_{r-i}$  and  $\omega(A' \cup \{i\}, i) = \mathbf{1}_{r-i}$ , or the other way

around. There are  $|g| - 2|g^*|$  non-productive positions  $A \in g$ , and for each of these there is also a binary choice, either  $\omega(A, i) = \mathbf{0}_{r-i}$  or  $\omega(A, i) = \mathbf{1}_{r-i}$ . The total number of choices is therefore  $2^{|g^*| - |g'|} \times 2^{|g| - 2|g^*|}$ , which is exactly as claimed.

(e) This is immediate from part (d), upon summing over  $g' \subseteq g^*$ .  $\square$

#### 10.4 The $f^C(\rho)$ and genotype

We begin by recalling from (7.4) the definition of the functions  $f^C(\rho)$ . Here  $\rho = (\rho_1, \dots, \rho_{r-1})$  is a sequence of parameters, and we define  $\rho_0 = 0$ . If  $C$  has level 0, we set  $f^C(\rho) = 1$ , whilst for  $C$  at level  $i \geq 1$  we apply the recursion

$$f^C(\rho) = \sum_{C \rightarrow C'} f^{C'}(\rho)^{\rho_{i-1}}.$$

**Proposition 10.4** *The quantities  $f^C$  depend only on the genotype of  $C$ , and thus for any  $i$ -genotype  $g$  we may define  $F(g) := f^C(\rho)$ , where  $C$  is any cell with  $g(C) = g$ . We have the recursion*

$$F(g) = \sum_{g' \leq g^*} 2^{|g| - |g^*| - |g'|} F(g')^{\rho_{i-1}}. \quad (10.3)$$

*Remark* The  $F(g)$  depend on  $\rho$ , as well as on  $i$  (where  $g$  is an  $i$ -genotype) but we suppress explicit mention of this. For example, it should be clear from context that  $g$  on the left is an  $i$ -genotype, but the sum on the right is over  $(i - 1)$ -genotypes, since  $g^*$  is an  $(i - 1)$ -genotype by definition.

*Proof* This is a simple induction on the level  $i$  using the definition of the  $f^C(\rho)$ , and parts (c) and (d) of Lemma 10.3.  $\square$

Let us pause to record two corollaries which we will need later.

**Corollary 10.5** *Suppose that  $g_1, g_2$  are two  $i$ -genotypes with  $g_1 \leq g_2$ . Then  $F(g_1) \leq F(g_2)$ .*

*Proof.* Note that  $g_1^* \leq g_2^*$ , and also that  $|g_1| - |g_1^*| \leq |g_2| - |g_2^*|$ , since

$$|g| - |g^*| = |g^*| + \#\{A \in \mathcal{P}[i - 1] : \#\{A, A \cup \{i\}\} \cap g = 1\}.$$

Hence, by two applications of Proposition 10.4,

$$\begin{aligned}
 F(g_1) &= 2^{|g_1|-|g_1^*|} \sum_{g' \leq g_1^*} 2^{-|g'|} F(g')^{\rho_{i-1}} \\
 &\leq 2^{|g_2|-|g_2^*|} \sum_{g' \leq g_2^*} 2^{-|g'|} F(g')^{\rho_{i-1}} = F(g_2).
 \end{aligned}$$

□

Recall that  $\Gamma_i$  is the cell at level  $i$  containing  $\mathbf{0}$ . Note that  $g(\Gamma_i) = \mathcal{P}[i]$ .

**Corollary 10.6** *If  $C \neq \Gamma_i$  is a cell of level  $i$ , then  $f^C(\rho) < f^{\Gamma_i}(\rho)$ .*

*Proof* This is simply the special case  $g_2 = \mathcal{P}[i]$  of the preceding corollary. The inequality is strict because if  $g < \mathcal{P}[i]$ , then  $g^* < \mathcal{P}[i - 1]$ . □

### 10.5 Existence of the $\rho_i$

In this section we prove Proposition 9.2 (a), which asserts that for the binary flags there is a unique solution  $\rho = (\rho_1, \rho_2, \dots)$  to the  $\rho$ -equations (9.1). In fact, we will prove the following more general fact which treats the  $j$ th  $\rho$ -equation in isolation, irrespective of whether the earlier ones have already been solved.

**Proposition 10.7** *Let  $j \in \mathbb{N}$  and let  $\rho_1, \dots, \rho_{j-1} \in (0, 1)$ . Then there is a unique  $\rho_j \in (0, 1)$  such that the  $j$ th  $\rho$ -equation for the binary flag,  $f^{\Gamma_{j+1}}(\rho) = e^{2^j} f^{\Gamma_j}(\rho)^{\rho_j}$ , is satisfied.*

*Remark* We will prove in the next section (Lemma 11.2) that for the solution  $\rho_1, \rho_2, \dots$  to the full set of  $\rho$ -equations we have  $\rho_j \leq \rho_1 = 0.30648 \dots$  for all  $j$ . For a table of numerical values of the  $\rho_j$ , see Table 1 in Sect. 12.

Before beginning the proof of Proposition 10.7, we isolate a lemma.

**Lemma 10.8** *Fix a  $(j - 1)$ -genotype  $g'$ . Then*

$$\sum_{g: g^* \geq g'} 2^{-|g^*|} = 2^{-2^{j-1}} 7^{2^{j-1}-|g'|},$$

where the sum is over all  $j$ -genotypes  $g$ .

*Proof* In order to determine  $g$ , we must determine for each  $A \subset [j - 1]$  whether  $A$  and/or  $A \cup \{j\}$  lie in  $g$ . Since we are only summing over  $g$  whose consolidation  $g^*$  contains  $g'$ , we must have that  $A$  and  $A \cup \{j\}$  belong to  $g$  for all  $A \in g'$ , so the membership of  $A$  and  $A \cup \{j\}$  to  $g$  is fully determined for all  $A \in g'$ . For any  $A \subset [j - 1]$  with  $A \notin g'$ , we have four choices, according to

whether  $A \in g$  and whether  $A \cup \{j\}$ . If both of these conditions hold, then we further have  $A \in g^*$ ; in the other three cases, we have  $A \notin g^*$ . We conclude that

$$\sum_{g: g^* \geq g'} 2^{-|g^*|} = \prod_{A \in g'} 2^{-1} \prod_{A \notin g'} (1 \cdot 2^{-1} + 3 \cdot 2^{-0}) = 2^{-2^{j-1}} 7^{2^{j-1} - |g'|}.$$

This completes the proof. □

*Proof of Proposition 10.7* For  $j = 1$ , the equation to be satisfied is  $3^{\rho_1} + 4 \cdot 2^{\rho_1} + 4 = e^2 3^{\rho_1}$ . It may easily be checked numerically that this has a unique solution  $\rho_1 \approx 0.306481 \dots$  in  $(0, 1)$ . One may also proceed analytically as follows. Define

$$\begin{aligned} G(x) = G_1(x) &:= e^2 3^x - (3^x + 4 \cdot 2^x + 4) \\ &= 3^x (e^2 - (1 + 4 \cdot (2/3)^x + 4/3^x)), \end{aligned}$$

In particular, the roots of  $G$  are in correspondence with the roots of  $H(x) = e^2 - (1 + 4 \cdot (2/3)^x + 4/3^x)$ . This is clearly a continuous and strictly increasing function. In addition,  $H(0) = e^2 - 9 < 0$  and  $H(1) = e^2 - 5 > 0$ . Thus,  $H$  has a unique root  $\rho_1 \in (0, 1)$ , and so does  $G$ .

Now assume  $j \geq 2$ . It turns out that much the same argument works, although the details are more elaborate. Assume that  $0 < \rho_i < 1$  for  $1 \leq i < j$ . Define

$$G(x) := G_j(x) = e^{2^j} (f^{\Gamma_j}(\rho))^x - f^{\Gamma_{j+1}}(\rho_1, \dots, \rho_{j-1}, x).$$

Proposition 10.4 implies that

$$\begin{aligned} G(x) &= e^{2^j} (F(\mathcal{P}[j]))^x - \sum_g 2^{2^j - |g|} F(g)^x \\ &= F(\mathcal{P}[j])^x \cdot H(x), \end{aligned} \tag{10.4}$$

where

$$H(x) = e^{2^j} - 2^{2^j} \sum_g 2^{-|g|} (F(g)/F(\mathcal{P}[j]))^x$$

and the sums over  $g$  run over all genotypes  $g \subset \mathcal{P}[j]$  at level  $j$ . Since (by an easy induction)  $F(\mathcal{P}[j]) > 0$ , it follows that  $G$  and  $H$  have the same roots. The latter is a continuous and strictly increasing function because Corollary 10.6 implies that  $F(g)/F(\mathcal{P}[j]) \leq 1$ , with equality only when  $g = \mathcal{P}[j]$ .



Moreover,  $H(0) = e^{2^j} - 3^{2^j} < 0$ . Therefore to complete the proof it suffices to show that  $H(1) > 0$ .

To show this, we use (10.4). First note that

$$F(\mathcal{P}[j]) = (\sqrt{2})^{2^j} \sum_{g'} 2^{-|g'|} F(g')^{\rho_{j-1}}, \tag{10.5}$$

where the sum is over all genotypes  $g'$  of level  $(j - 1)$ .

Next, by Proposition 10.4 and Lemma 10.8 we have

$$\begin{aligned} \sum_{g \subset \mathcal{P}[j]} 2^{-|g|} F(g) &= \sum_g 2^{-|g^*|} \sum_{g' \leq g^*} 2^{-|g'|} F(g')^{\rho_{j-1}} \\ &= \sum_{g' \subset \mathcal{P}[j-1]} 2^{-|g'|} F(g')^{\rho_{j-1}} \sum_{g: g^* \geq g'} 2^{-|g^*|} \\ &= (7/2)^{2^{j-1}} \sum_{g'} 14^{-|g'|} F(g')^{\rho_{j-1}}. \end{aligned} \tag{10.6}$$

Putting (10.4), (10.5) and (10.6) together we obtain

$$\begin{aligned} H(1) \cdot F(\mathcal{P}[j]) &= (e\sqrt{2})^{2^j} \sum_{g'} 2^{-|g'|} F(g')^{\rho_{j-1}} \\ &\quad - (\sqrt{14})^{2^j} \sum_{g'} 14^{-|g'|} F(g')^{\rho_{j-1}}. \end{aligned}$$

Since  $e^2 > 7$ , we have  $\sqrt{14} < e\sqrt{2}$ , and thus  $H(1) > 0$ . This completes the proof. □

### 10.6 Entropy inequalities for the binary systems

We begin with a lemma which will be used a few times in what follows.

**Lemma 10.9** *Let  $C'$  be one of the children of  $\Gamma_i$ , thus  $C'$  is a cell at level  $(i - 1)$ . Then*

$$\mu_i(C') \leq \mu_i(\Gamma_{i-1}) = e^{-2^{i-1}},$$

and equality occurs only when  $C' = \Gamma_{i-1}$ .

*Proof* We showed in Corollary 10.6 that  $f^{C'}(\rho) < f^{\Gamma_{i-1}}(\rho)$ , for any choice of  $\rho = (\rho_1, \dots, \rho_{r-1})$ , and for any child  $C'$  of  $\Gamma_i$  with  $C' \neq \Gamma_{i-1}$ . Now that we know that the  $\rho$ -equations have a solution, it follows immediately from

the definition of the optimal measures  $\mu^*$  in (7.6), applied with  $C = \Gamma_i$ , that  $\mu_i(C') < \mu_i(\Gamma_{i-1})$ , again for any child  $C'$  of  $\Gamma_i$  with  $C' \neq \Gamma_{i-1}$ . Finally, observe that  $\mu_i(\Gamma_{i-1}) = e^{-2^{i-1}}$  by (7.7).  $\square$

*Proof of Lemma 9.4* This follows almost immediately from Lemma 10.9 with  $i = m + 1$ . Indeed since  $\mu_{m+1}(C) \leq e^{-2^m}$  for all cells  $C$  at level  $m$ , with equality only for  $C = \Gamma_m$ , we have

$$\mathbb{H}_{\mu_{m+1}}(V_m) = \sum_C \mu_{m+1}(C) \log \frac{1}{\mu_{m+1}(C)} > 2^m \sum_C \mu_{m+1}(C) = 2^m.$$

This concludes the proof.  $\square$

*Proof of Lemma 9.5* Let  $\mu = \mu_i$  with  $m < i \leq r$ . We must show that

$$\mathbb{H}_\mu(V_{m-1}) - \mathbb{H}_\mu(V_m) < 2^{m-1}. \tag{10.7}$$

Let  $C$  denote a cell at level  $m$  and  $C'$  a child of  $C$  at level  $(m - 1)$ . In addition, let the notations  $g(C)$  and  $g(C)^*$  refer to the genotype of  $C$  and its consolidation, as defined in Definitions 10.1 and 10.2. By the definition of entropy, Lemma 10.3 (e), and the concavity of  $L(x) = -x \log x$  we find that

$$\begin{aligned} \mathbb{H}_\mu(V_{m-1}) - \mathbb{H}_\mu(V_m) &= \sum_C \mu(C) \sum_{C'} L\left(\frac{\mu(C')}{\mu(C)}\right) \\ &\leq \sum_C \mu(C) \log(\#C') \\ &= \sum_C \mu(C) \log \left[ 2^{|g(C)|} (3/4)^{|g(C)^*|} \right]. \end{aligned} \tag{10.8}$$

Now by (10.2) we have  $|g(C)^*| \geq |g(C)| - 2^{m-1}$ , whence

$$2^{|g(C)|} (3/4)^{|g(C)^*|} \leq 2^{|g(C)|} (3/4)^{|g(C)| - 2^{m-1}} = (3/2)^{|g(C)|} (4/3)^{2^{m-1}}. \tag{10.9}$$

Since we also have that  $|g(C)| \leq 2^m$ , we infer that

$$2^{|g(C)|} (3/4)^{|g(C)^*|} \leq 3^{2^{m-1}}. \tag{10.10}$$

This and (10.8) already imply the bound

$$\mathbb{H}_\mu(V_{m-1}) - \mathbb{H}_\mu(V_m) \leq 2^{m-1} \log 3,$$

which is only very slightly weaker than Lemma 9.5.

To make the crucial extra saving, write  $S$  for the union of all cells  $C$  at level  $m$  with  $|g(C)| > \frac{3}{4}2^m$ . We claim that

$$\mu(S) < \frac{1}{2}. \tag{10.11}$$

We postpone the proof of this inequality momentarily and show how to use it to complete the proof of Lemma 9.5.

Observe that if  $C$  is not one of the cells making up  $S$ , that is to say if  $|g(C)| \leq \frac{3}{4}2^m$ , then

$$\begin{aligned} \log \left[ 2^{|g(C)|} (3/4)^{|g(C)^*|} \right] &\leq \log \left[ (3/2)^{|g(C)|} (4/3)^{2^{m-1}} \right] \\ &\leq \left( \frac{3}{2} \log(3/2) + \log(4/3) \right) 2^{m-1} \\ &\leq 0.9 \cdot 2^{m-1}, \end{aligned}$$

where we used (10.9) to obtain the first inequality. Assuming the claim (10.11), it follows from this, (10.8) and (10.10) that

$$\begin{aligned} \mathbb{H}_\mu(V_{m-1}) - \mathbb{H}_\mu(V_m) &\leq 2^{m-1}(\log 3)\mu(S) + 0.9 \cdot 2^{m-1}(1 - \mu(S)) \\ &< 2^{m-1}, \end{aligned}$$

which is the statement of Lemma 9.5.

It remains to prove (10.11). Recall that  $1 \leq m < i \leq r$ .

When  $1 \leq m \leq 2$ , the only integer in  $(\frac{3}{4}2^m, 2^m]$  is  $2^m$ . Hence, if a cell  $C$  at level  $m$  satisfies the inequality  $|g(C)| > \frac{3}{4}2^m$ , we must have  $|g(C)| = 2^m$ . The only cell with this property is  $\Gamma_m$ . Since we have  $\mu(\Gamma_m) = e^{2^m-2^i} \leq e^{-1}$  by (7.7), our claim (10.11) follows in this case.

Assume now that  $m \geq 3$ . Let  $\tilde{S}$  be the union of all children  $\tilde{C}$  of  $\Gamma_i$  (thus these are cells at level  $i - 1 \geq m$ ) which contain a cell  $C$  in  $S$ . By repeated applications of Lemma 10.3 (c) we have  $|g(\tilde{C})| > 2^{i-1-m}(\frac{3}{4}2^m) = \frac{3}{4}2^{i-1}$  for any such  $\tilde{C}$ . Lemma 10.3 (d), applied with  $C = \Gamma_i$ , implies that the number of such cells  $\tilde{C}$  is at most

$$\sum_{h > (3/4)2^{i-1}} \binom{2^{i-1}}{h} 2^{2^{i-1}-h} \leq 2^{\frac{1}{4}2^{i-1}} 2^{2^{i-1}} = 2^{(5/4)2^{i-1}}.$$

By Lemma 10.9 and our assumption that  $i - 1 \geq m \geq 3$ , it follows that

$$\mu(S) \leq \mu(\tilde{S}) \leq (2^{5/4}/e)^{2^{i-1}} < 0.35.$$

This completes the proof of the claim (10.11) and hence of Lemma 9.5. □

## 10.7 Existence of the optimal parameters $c^*$

*Proof of Proposition 9.2 (b)* We have  $\text{Supp}(\mu_j^*) = \Gamma_j$  by Remark 7.1 (b), and hence  $|\text{Supp}(\mu_j^*)| = 2^{2^j}$  by Lemma 5.1. By Lemma B.2, when  $j \geq m + 2$  we deduce the inequality

$$\mathbb{H}_{\mu_j^*}(V_m) \leq \log |\text{Supp}(\mu_j^*)| \leq 2^j \log 2 < 2^j - 2^m. \quad (10.12)$$

Now recall (Definition 7.5) that the optimal parameters should satisfy the conditions (7.12) (which are the fully written out version of (7.11)). We wish to show that there is a solution with  $1 = c_1^* > c_2^* > \dots > c_{r+1}^* > 0$ . Rearranging (7.12) and recalling  $\dim(V_j) = 2^j$ , we find that

$$\begin{aligned} & (c_{m+1}^* - c_{m+2}^*)(\mathbb{H}_{\mu_{m+1}^*}(V_m) - 2^m) \\ &= \sum_{j=m+2}^r (2^j - 2^m - \mathbb{H}_{\mu_j^*}(V_m))(c_j^* - c_{j+1}^*) + (2^r - 2^m)c_{r+1}^* \end{aligned}$$

for  $0 \leq m \leq r - 1$ . By Lemma 9.4 and (10.12), we may apply a downwards induction on  $m = r - 1, r - 2, \dots$  to solve these equations with  $0 < c_{r+1}^* < c_r^* < \dots < c_1^*$ . Rescaling, we may additionally ensure that  $c_1^* = 1$ .  $\square$

## 11 The limit of the $\rho_i$

In the last section we showed that there is a unique solution  $\boldsymbol{\rho} = (\rho_1, \rho_2, \dots)$  to the  $\boldsymbol{\rho}$ -equations (9.1) for the binary system with  $0 < \rho_j < 1$  for all  $j$ . In this section, we show that the limit  $\lim_{j \rightarrow \infty} \rho_j$  exists.

**Proposition 11.1**  $\boldsymbol{\rho} = \lim_{j \rightarrow \infty} \rho_j$  exists.

### 11.1 $\rho_1$ is the largest $\rho_j$

The estimates required in the proof of Proposition 11.1 are rather delicate, and to make them usable for our purposes we need the following a priori bound on the  $\rho_j$ .

**Lemma 11.2** For all  $j \geq 1$ , we have  $\rho_j \leq \rho_1 = 0.30648 \dots$

The reader should recall the notion of genotype  $g$  (Definition 10.1) and of the function  $F(g)$  (Proposition 10.4).

The next lemma is a stronger version of Corollary 10.5, whose proof uses that result as an ingredient.

**Lemma 11.3** For any  $j \geq 1$  and  $g_1 \leq g_2$  at level  $j$ , we have

$$\frac{F(g_1)}{F(g_2)} \leq \left(\frac{1}{2}\right)^{|g_2|-|g_1|} \left(\frac{4}{3}\right)^{|g_2^*|-|g_1^*|}.$$

*Proof* We have

$$\begin{aligned} F(g_2) &= 2^{|g_2|-|g_2^*|} \sum_{g \leq g_1^*} 2^{-|g|} \sum_{g' \leq g_2^* \setminus g_1^*} 2^{-|g'|} F(g \cup g')^{\rho_{j-1}} \\ &\hspace{15em} \text{(by Proposition 10.4)} \\ &\geq 2^{|g_2|-|g_2^*|} \sum_{g \leq g_1^*} 2^{-|g|} \sum_{g' \leq g_2^* \setminus g_1^*} 2^{-|g'|} F(g)^{\rho_{j-1}} \\ &\hspace{15em} \text{(by Corollary 10.5)} \\ &= 2^{|g_2|-|g_2^*|} \sum_{g \leq g_1^*} 2^{-|g|} F(g)^{\rho_{j-1}} (3/2)^{|g_2^*|-|g_1^*|} \\ &\hspace{15em} \text{(by the binomial theorem)} \\ &= F(g_1) 2^{|g_2|-|g_1|} (3/4)^{|g_2^*|-|g_1^*|} \\ &\hspace{15em} \text{(by Proposition 10.4).} \end{aligned}$$

This concludes the proof. □

*Proof of Lemma 11.2* We begin by observing that

$$\sum_{g \leq \mathcal{P}[j]} c_1^{|g|} c_2^{|g^*|} = \prod_{A \subset [j-1]} \left( \sum_{a,b \in \{0,1\}} c_1^{a+b} c_2^{ab} \right) = (1 + 2c_1 + c_1^2 c_2)^{2^{j-1}}. \tag{11.1}$$

The  $\rho$ -equations (9.1), translated into the language of genotypes, are  $F(\mathcal{P}[j + 1]) = e^{2^j} F(\mathcal{P}[j])^{\rho_j}$ . Therefore, by Proposition 10.4 (with  $g = \mathcal{P}[j + 1]$ ) followed by Lemma 11.3 (with  $g_2 = \mathcal{P}[j]$ ), we have

$$\begin{aligned} e^{2^j} F(\mathcal{P}[j])^{\rho_j} &= F(\mathcal{P}[j + 1]) = 2^{2^j} \sum_{g \leq \mathcal{P}[j]} 2^{-|g|} F(g)^{\rho_j} \\ &\leq 2^{2^j} \sum_{g \leq \mathcal{P}[j]} 2^{-|g|} F(\mathcal{P}[j])^{\rho_j} \left[ (1/2)^{2^j-|g|} (4/3)^{2^{j-1}-|g^*|} \right]^{\rho_j} \\ &= 2^{2^j} (1/3)^{2^{j-1} \rho_j} F(\mathcal{P}[j])^{\rho_j} \sum_{g \leq \mathcal{P}[j]} 2^{(\rho_j-1)|g|} (3/4)^{\rho_j |g^*|}. \end{aligned}$$

Dividing through by  $F(\mathcal{P}[j])^{\rho_j}$ , and applying (11.1) with  $c_1 = 2^{\rho_j-1}$  and  $c_2 = (3/4)^{\rho_j}$ , we find that

$$\begin{aligned} e^{2^j} &\leq (4/3^{\rho_j})^{2^{j-1}} (1 + 2^{\rho_j} + 2^{2\rho_j-2}(3/4)^{\rho_j})^{2^{j-1}} \\ &= (4/3^{\rho_j} + 4(2/3)^{\rho_j} + 1)^{2^{j-1}}. \end{aligned}$$

Therefore

$$3^{\rho_j} e^2 \leq 4 + 4 \cdot 2^{\rho_j} + 3^{\rho_j}.$$

However, the first  $\rho$ -equation (9.2) is precisely that

$$3^{\rho_1} e^2 = 4 + 4 \cdot 2^{\rho_1} + 3^{\rho_1}.$$

The result follows immediately (using the monotonicity of the function  $1 + 4(2/3)^t + 4(1/3)^t$  - see the proof of Proposition 10.7).  $\square$

## 11.2 Preamble to the proof

In this section, we set up some notation and structure necessary for the proof of Proposition 11.1. Since we wish to let  $r \rightarrow \infty$ , it is convenient to embed all binary  $r$ -step systems into a universal infinite binary system. To this end, and with a slight abuse of notation, we let

$$V_j = \{(x_A)_{A \subset \mathcal{P}(\mathbb{N})} : x_A \in \mathbb{Q} \text{ and } x_A = x_{A \cap [j]} \text{ for all } A \subset \mathcal{P}(\mathbb{N})\}$$

for  $j = 0, 1, \dots$ . Clearly,  $V_j \simeq \mathbb{Q}^{2^j}$  for all  $j$ , and the flag

$$\mathcal{V}^r : V_0 \leq V_1 \leq \dots \leq V_r$$

is isomorphic to the flag of the  $r$ -step binary system.

In this notation, we have

$$\Gamma_j = \{\omega \in \Omega : \omega \equiv \mathbf{0} \pmod{V_j}\} \text{ for } j = 0, 1, \dots,$$

where

$$\Omega = \{\omega = (\omega_A)_{A \subset \mathcal{P}(\mathbb{N})} : \omega_A \in \{0, 1\} \text{ for all } A \subset \mathcal{P}(\mathbb{N})\}$$

is the discrete unit cube. We further set

$$\Gamma_\infty = \bigcup_{j=0}^\infty \Gamma_j.$$

Lastly, for each  $j \geq 0$ , we say that  $C$  is a cell at level  $j$  if  $C \subset \Gamma_\infty$  and there exists some  $x = (x_A)_{A \subset \mathcal{P}(\mathbb{N})}$  such that  $x_A \in \mathbb{Q}$  for all  $A$  and  $C = \Omega \cap (x + V_j)$ . We may easily check that the collection of cells lying in  $\Gamma_r$  forms the tree corresponding to the  $r$ -step binary system.

We may now define the functions  $f^C$  for our infinite binary flag. It is convenient to reverse the indices in  $f^C$ . Specifically, let  $\mathbf{x} = (x_1, x_2, \dots) \in [0, 1]^{\mathbb{N}}$ . If  $C$  is a cell at level  $j \geq 0$ , then we define

$$\psi^C(\mathbf{x}) := \log f^C(x_{j-1}, \dots, x_1).$$

In particular,  $\psi^C(\mathbf{x}) = 0$  when  $j = 0$ , and  $\psi^C(\mathbf{x}) = \log |C \setminus \{\mathbf{0}\}|$  when  $j = 1$ .

In the special case  $C = \Gamma_j$  we define also

$$\phi_j(\mathbf{x}) = 2^{-j} \psi^{\Gamma_j}(\mathbf{x}) = 2^{-j} \log f^{\Gamma_j}(x_{j-1}, \dots, x_1).$$

Thus  $\phi_1(\mathbf{x}) = \frac{1}{2} \log 3$  and  $\phi_2(\mathbf{x}) = \frac{1}{4} \log(3^{x_1} + 4 \cdot 2^{x_1} + 4)$ .

Note that  $\psi^C, \phi_j$  are increasing in each variable. Moreover we have the following simple bounds.

**Lemma 11.4** (Simple bounds) *We have  $\frac{1}{2} \log 3 \leq \phi_j(\mathbf{x}) < \log 2$ .*

*Proof* For the upper bound, note that  $f^{\Gamma_j}(\mathbf{x}) \leq f^{\Gamma_j}(\mathbf{1})$ . By the definition of  $f^C$  (see (7.4)), we have that  $f^{\Gamma_j}(\mathbf{1})$  is equal to the number of children of  $\Gamma_j$  at level 0, which, in turn, is equal to  $2^{2^j} - 1$ . This proves the claimed upper bound on  $\phi_j(\mathbf{x})$ .

For the lower bound, observe that  $f^{\Gamma_j}(\mathbf{x}) \geq f^{\Gamma_j}(\mathbf{0})$ . Using again the definition of  $f^C$ , we find that  $f^{\Gamma_j}(\mathbf{0})$  equals the number of children of  $\Gamma_j$  at level  $j - 1$ . Thus  $f^{\Gamma_j}(\mathbf{0}) = 3^{2^{j-1}}$  by Lemma 10.3. This proves the claimed lower bound of  $\phi_j(\mathbf{x})$ , thus completing the proof of the lemma.  $\square$

The  $\rho$ -equations (9.1) may be expressed in terms of the  $\phi_j$  in the following simple form:

$$\phi_{j+1}(\rho_j, \rho_{j-1}, \dots) = \frac{1}{2} (\rho_j \phi_j(\rho_{j-1}, \rho_{j-2}, \dots) + 1). \tag{11.2}$$

### 11.3 Product structure of cells and self-similarity of the functions $\phi_j$

There is a natural bijection  $\pi : \mathbb{Q}^{\mathcal{P}(\mathbb{N})} \times \mathbb{Q}^{\mathcal{P}(\mathbb{N})} \rightarrow \mathbb{Q}^{\mathcal{P}(\mathbb{N})}$  defined by  $\pi((x, x')) = y$ , where  $y_A = x_{A-1}$  and  $y_{\{1\} \cup A} = x'_{A-1}$ , for all  $A \subset \{2, 3, \dots\}$ . Here, we write  $A - 1$  for the set  $\{a - 1 : a \in A\}$ . There is a finite version of this map that can be visualized as a concatenation map. For each  $r$ , let  $\pi_r : \mathbb{Q}^{\mathcal{P}[r-1]} \times \mathbb{Q}^{\mathcal{P}[r-1]} \rightarrow \mathbb{Q}^{\mathcal{P}[r]}$  defined by  $\pi((x, x')) = y$ , where  $y_A = x_{A-1}$  and  $y_{\{1\} \cup A} = x'_{A-1}$ , for all  $A \subset \{2, 3, \dots, r\}$ . If we place the coordinates of  $x$  and  $x'$  in reverse binary order, as per the map  $\{2, \dots, r\} \supset A \rightarrow \sum_{a \in A} 2^{r-a} \in \{0, 1, \dots, 2^{r-1} - 1\}$ , then  $\pi_r$  is the concatenation map that generates  $y$  by placing first all coordinates of  $x$ , followed by all coordinates of  $x'$ .

Now one may easily check that  $\pi(V_{j-1} \times V_{j-1}) = V_j$  for all  $j = 1, 2, \dots$ . Therefore if  $C_1, C_2$  are two cells at level  $(j - 1)$  in the infinite binary system, then  $\pi(C_1 \times C_2)$  is a cell at level  $j$ , and conversely every cell of level  $j$  is of this form. The children  $C'$  of  $C$  are precisely  $\pi(C'_1 \times C'_2)$  where  $C_1 \rightarrow C'_1, C_2 \rightarrow C'_2$ .

The product structure established above manifests itself in a self-similarity property  $\phi_j \approx \phi_{j-1}$ . In this section, we will establish the following precise version of this.

**Proposition 11.5** *Let  $\alpha \in (0, 1]$  and consider a vector  $\mathbf{x} = (x_1, x_2, \dots) \in [0, \alpha]^{\mathbb{N}}$ . In addition, let  $C = \pi(C_1 \times C_2)$  be a cell of level  $j \geq 2$ . Then we have*

$$\psi^{C_1}(\mathbf{x}) + \psi^{C_2}(\mathbf{x}) \leq \psi^C(\mathbf{x}) \leq \psi^{C_1}(\mathbf{x}) + \psi^{C_2}(\mathbf{x}) + \alpha^{j-1} \log 2. \quad (11.3)$$

*In particular, taking  $C = \Gamma_j = \pi(\Gamma_{j-1} \times \Gamma_{j-1})$ , we have*

$$\phi_{j-1}(\mathbf{x}) \leq \phi_j(\mathbf{x}) \leq \phi_{j-1}(\mathbf{x}) + (\alpha/2)^j \frac{\log 2}{\alpha}. \quad (11.4)$$

*Proof* We proceed by induction on  $j$ . When  $j = 2$ , we proceed by hand. Notice that at level 1, there are three different types of cells, having 4, 2 and 1 elements, respectively. There is only one cell with 4 elements, the cell  $\Gamma_1$ ; it splits into three cells at level 0: one with two elements, and two unicells (singletons). All other cells at level 1 split into unicells at level 0. Hence, at level 2, there are six different types of cells  $C = \pi(C_1 \times C_2)$  corresponding to the six possibilities for the unordered pair  $\{|C_1|, |C_2|\}$ . Their subcells are in 1-1 correspondence with the cells  $\pi(C'_1 \times C'_2)$ , where  $C'_1$  is a subcell of  $C_1$  (at level 0) and  $C'_2$  is a subcell of  $C_2$  (also at level 0).

The three cases with  $\max(|C_1|, |C_2|) \leq 2$  are trivial, because we then have that all the cells at level 1 are unicells, and thus we readily find that  $f^C = f^{C_1} f^{C_2} = |C_1| \cdot |C_2|$ .



The two other cases with  $|C_1| \leq 2$  and  $|C_2| = 4$  (so that  $C_2 = \Gamma_1$ ) are only slightly harder: if  $|C_1| = 2$ , then  $f^C(\mathbf{x}) = 2 \cdot 2^{x_1} + 4$ ,  $f^{C_1} = 2$ ,  $f^{C_2} = 3$  and so the desired inequalities are  $\log 6 \leq \log(2 \cdot 2^{x_1} + 4) \leq \log 6 + x_1 \log 2$ , which are immediately seen to be true for all  $x_1 \geq 0$ . Similarly, if  $|C_1| = 1$ , then  $f^C(\mathbf{x}) = 2^{x_1} + 2$ ,  $f^{C_1} = 1$ ,  $f^{C_2} = 3$ , and so the desired inequalities are  $\log 3 \leq \log(2^{x_1} + 2) \leq \log 3 + x_1 \log 2$ , which are again true for all  $x_1 \geq 0$ .

A little trickier is the case  $|C_1| = |C_2| = 3$ , corresponding to  $C = \Gamma_2 = \pi(\Gamma_1 \times \Gamma_1)$ . In this case  $f^C(\mathbf{x}) = 3^{x_1} + 4 \cdot 2^{x_1} + 4$ ,  $f^{C_1} = f^{C_2} = 3$ , so the desired inequalities are  $2 \log 3 \leq \log(3^x + 4 \cdot 2^x + 4) \leq 2 \log 3 + x \log 2$ . The lower bound is evident. For the upper bound, we must equivalently show that  $g(x) := 5 \cdot 2^x - 3^x - 4 \geq 0$  for  $x \in [0, 1]$ . Since  $g(0) = 0$  and  $g'(x) = 5 \log 2 \cdot 2^x - \log 3 \cdot 3^x > 0$  for  $x \leq 1$ , the desired inequality follows.

Now suppose that  $j \geq 3$ , and assume the result is true for cells at level  $(j - 1)$ . By the recursive definition of  $f^C$ , if  $C$  is a cell at level  $j$ , we have the recurrence

$$e^{\psi^C(\mathbf{x})} = \sum_{C \rightarrow C'} e^{x_1 \psi^{C'}(T\mathbf{x})}, \tag{11.5}$$

where  $T\mathbf{x}$  denotes the *shift operator*

$$T\mathbf{x} = (x_2, x_3, \dots).$$

For the upper bound, note that

$$e^{\psi^C(\mathbf{x})} = \sum_{C \rightarrow C'} e^{x_1 \psi^{C'}(T\mathbf{x})} \leq \sum_{\substack{C_1 \rightarrow C'_1 \\ C_2 \rightarrow C'_2}} e^{x_1(\psi^{C'_1}(T\mathbf{x}) + \psi^{C'_2}(T\mathbf{x}) + \alpha^{j-2} \log 2)}.$$

Recalling that  $x_1 \leq \alpha$ , we conclude that

$$\begin{aligned} e^{\psi^C(\mathbf{x})} &\leq 2^{\alpha^{j-1}} \left( \sum_{C_1 \rightarrow C'_1} e^{x_1 \psi^{C'_1}(T\mathbf{x})} \right) \left( \sum_{C_2 \rightarrow C'_2} e^{x_1 \psi^{C'_2}(T\mathbf{x})} \right) \\ &= 2^{\alpha^{j-1}} e^{\psi^{C_1}(\mathbf{x})} e^{\psi^{C_2}(\mathbf{x})}. \end{aligned}$$

The lower bound is proven similarly. The result thus follows. □

### 11.4 Derivatives and the limit of the $\rho_i$

Because of the implicit definition of the parameters  $\rho_i$ , the self-similarity property (11.4) is not enough for us by itself. We will also require the following (rather ad hoc) derivative bounds.

Here, and in what follows,  $\partial_m F(y_1, \dots) := \frac{\partial F}{\partial y_m}(y_1, \dots)$ , that is to say the derivative of the function  $F$  with respect to its  $m$ th variable. Thus, for instance,

$$\partial_m \psi^C(T\mathbf{x}) = \frac{\partial}{\partial x_{m+1}} [\psi^C(T\mathbf{x})]. \tag{11.6}$$

**Proposition 11.6** *Set  $\Delta_m := \sup_{j \geq 2} \sup_{\mathbf{x} \in [0, 0.31]^{\mathbb{N}}} |\partial_m \phi_j(\mathbf{x})|$ . Then  $\Delta_1 < 0.17$ ,  $\Delta_2 < 0.05$ ,  $\sum_{m \geq 3} \Delta_m < 0.01$  and  $\Delta_m \ll 0.155^m$ .*

The proof of this proposition is given in Sect. 11.5. Let us now show how this proposition, together with (11.4), implies Proposition 11.1.

*Proof of Proposition 11.1* Write  $\varepsilon_i := \rho_{i+1} - \rho_i$ ,  $i = 1, 2, 3, \dots$ . The  $\rho$ -equation at level  $(j + 1)$  is

$$\phi_{j+2}(\rho_{j+1}, \rho_j, \dots) = \frac{1}{2}(\rho_{j+1}\phi_{j+1}(\rho_j, \rho_{j-1}, \dots) + 1)$$

by (11.2). Recall that that  $\rho_j \leq \rho_1 \leq 0.31$  for all  $j$ , by Lemma 11.2. Hence, two applications of (11.4) (with  $\alpha = 0.31$ ) yield the asymptotic formula

$$\phi_{j+1}(\rho_{j+1}, \rho_j, \dots) = \frac{1}{2}(\rho_{j+1}\phi_j(\rho_j, \rho_{j-1}, \dots) + 1) + O(0.155^j).$$

Subtracting (11.2), the  $\rho$ -equation at level  $j$ , from this gives

$$\begin{aligned} & \phi_{j+1}(\rho_{j+1}, \rho_j, \dots) - \phi_{j+1}(\rho_j, \rho_{j-1}, \dots) \\ &= \frac{\rho_{j+1}}{2}(\phi_j(\rho_j, \rho_{j-1}, \dots) - \phi_j(\rho_{j-1}, \rho_{j-2}, \dots)) \\ & \quad + \frac{\varepsilon_j}{2}\phi_j(\rho_j, \rho_{j-1}, \dots) + O(0.155^j). \end{aligned} \tag{11.7}$$

Now by the mean value theorem,

$$|\phi_{j+1}(\rho_{j+1}, \rho_j, \dots) - \phi_{j+1}(\rho_j, \rho_{j-1}, \dots)| \leq \Delta_1|\varepsilon_j| + \dots + \Delta_j|\varepsilon_1| \tag{11.8}$$

and

$$|\phi_j(\rho_j, \rho_{j-1}, \dots) - \phi_j(\rho_{j-1}, \rho_{j-2}, \dots)| \leq \Delta_1|\varepsilon_{j-1}| + \dots + \Delta_{j-1}|\varepsilon_1|. \tag{11.9}$$

Therefore, from (11.7), the triangle inequality and the fact that

$$\frac{\rho_{j+1}}{2} \leq \frac{\rho_1}{2} \leq 0.155,$$

we have

$$\begin{aligned}
 |\varepsilon_j| & \left( \frac{1}{2} \phi_j(\rho_j, \rho_{j-1}, \dots) - \Delta_1 \right) \\
 & \leq (\Delta_2 + 0.155\Delta_1)|\varepsilon_{j-1}| + (\Delta_3 + 0.155\Delta_2)|\varepsilon_{j-2}| \\
 & \quad + \dots + O(0.155^j).
 \end{aligned}
 \tag{11.10}$$

Now by Lemma 11.4 and Proposition 11.6,

$$\frac{1}{2} \phi_j(\rho_j, \rho_{j-1}, \dots) - \Delta_1 > \frac{1}{4} \log 3 - 0.17 > 0.104.$$

Also, by Proposition 11.6 we have

$$(\Delta_2 + 0.155\Delta_1) + (\Delta_3 + 0.155\Delta_2) + \dots < 0.096.$$

Assuming that  $j \geq j_0$  with  $j_0$  large enough, (11.10) implies a bound

$$|\varepsilon_j| \leq c_1|\varepsilon_{j-1}| + c_2|\varepsilon_{j-2}| + \dots + c_{j-1}|\varepsilon_1| + 2^{-j}, \tag{11.11}$$

where  $c_1, c_2, \dots$  are fixed nonnegative constants with  $\sum_i c_i < \frac{0.096}{0.104} < 0.93$  and, by Proposition 11.6,  $c_i \leq 2^{-i}$  for all  $i \geq i_0$  for some  $i_0$ . It is convenient to assume that  $i_0, j_0 \geq 10$ , which we clearly may.

We claim that (11.11) implies exponential decay of the  $\varepsilon_j$ , which of course immediately implies Theorem 11.1. To see this, take  $\delta \in (0, \frac{1}{4})$  so small that  $0.94(1 - \delta)^{-i_0} < 0.99$ , and then take  $A \geq 100$  large enough that  $|\varepsilon_j| \leq A(1 - \delta)^j$  for all  $j \leq j_0$ . We claim that the same bound holds for all  $j$ , which follows immediately by induction using (11.11) provided one can show that

$$\sum_{i \geq 1} c_i(1 - \delta)^{-i} + \frac{1}{A} \left( \frac{1}{2(1 - \delta)} \right)^j < 1 \tag{11.12}$$

for  $j \geq j_0$ . Since  $\delta < \frac{1}{2}$  and  $A \geq 100$ , it is enough to show that

$$\sum_{i \geq 1} c_i(1 - \delta)^{-i} < 0.99.$$

The contribution to this sum from  $i \leq i_0$  is at most  $0.93(1 - \delta)^{-i_0}$ , whereas the contribution from  $i > i_0$  is (by summing the geometric series) at most

$$\sum_{i > i_0} 2^{-i}(1 - \delta)^{-i} < 2 \cdot 2^{-i_0}(1 - \delta)^{-i_0} < 0.01(1 - \delta)^{-i_0}.$$

Therefore the desired bound follows from our choice of  $\delta$ . □

### 11.5 Self-similarity for derivatives

Our remaining task is to prove Proposition 11.6. Once again we use self-similarity of the  $\phi_j$ , but now for their derivatives, the key point being that  $\partial_m \phi_j \approx \partial_m \phi_{j-1}$ . Here is a precise statement.

**Proposition 11.7** *Suppose that  $C = \pi(C_1 \times C_2)$  is cell at level  $j \geq 1$ . Let  $\alpha \in [0, 1)$  and  $m \geq 1$ , and suppose that  $\mathbf{x} \in [0, \alpha]^{\mathbb{N}}$ . Then we have*

$$0 \leq \partial_m \psi^C(\mathbf{x}) \leq 2^{\sum_{i=1}^m \alpha^{j-i}} (\partial_m \psi^{C_1}(\mathbf{x}) + \partial_m \psi^{C_2}(\mathbf{x}) + \alpha^{j-2} \log 2).$$

In particular, taking  $C = \Gamma_j = \pi(\Gamma_{j-1} \times \Gamma_{j-1})$ , we have

$$0 \leq \partial_m \phi_j(\mathbf{x}) \leq 2^{\sum_{i=1}^m \alpha^{j-i}} \left( \partial_m \phi_{j-1}(\mathbf{x}) + \left(\frac{\alpha}{2}\right)^j \frac{\log 2}{\alpha^2} \right). \tag{11.13}$$

*Proof* The lower bound follows by noticing that  $\psi^C$  is increasing in each variable. For the upper bound, we may assume that  $m \leq j - 1$ , for when  $m \geq j$ ,  $\partial_m \phi_j(\mathbf{x})$  is identically zero. We proceed by induction on  $m$ , first establishing the case  $m = 1$ . Differentiating (11.5) gives

$$e^{\psi^C(\mathbf{x})} \partial_1 \psi^C(\mathbf{x}) = \sum_{C \rightarrow C'} \psi^{C'}(T\mathbf{x}) e^{x_1 \psi^{C'}(T\mathbf{x})}.$$

By two applications of the upper bound in Proposition 11.5 (applied to  $C' = \pi(C'_1 \times C'_2)$ ), we obtain

$$\begin{aligned} e^{\psi^C(\mathbf{x})} \partial_1 \psi^C(\mathbf{x}) &\leq 2^{\alpha^{j-1}} \sum_{\substack{C_1 \rightarrow C'_1 \\ C_2 \rightarrow C'_2}} (\psi^{C'_1}(T\mathbf{x}) + \psi^{C'_2}(T\mathbf{x}) + \alpha^{j-2} \log 2) \\ &\quad \times e^{x_1(\psi^{C'_1}(T\mathbf{x}) + \psi^{C'_2}(T\mathbf{x}))}. \end{aligned} \tag{11.14}$$

On the other hand, for  $i = 1, 2$  we get by differentiating the recurrence

$$e^{\psi^{C_i}(\mathbf{x})} = \sum_{C_i \rightarrow C'_i} e^{x_1 \psi^{C'_i}(T\mathbf{x})} \tag{11.15}$$

with respect to  $x_1$  that

$$e^{\psi^{C_i}(\mathbf{x})} \partial_1 \psi^{C_i}(\mathbf{x}) = \sum_{C_i \rightarrow C'_i} \psi^{C'_i}(T\mathbf{x}) e^{x_1 \psi^{C'_i}(T\mathbf{x})}. \tag{11.16}$$

Substituting (11.15) and (11.16) into (11.14) gives

$$e^{\psi^C(\mathbf{x})} \partial_1 \psi^C(\mathbf{x}) \leq 2^{\alpha^{j-1}} (\partial_1 \psi^{C_1}(\mathbf{x}) + \partial_1 \psi^{C_2}(\mathbf{x}) + \alpha^{j-2} \log 2) e^{\psi^{C_1}(\mathbf{x}) + \psi^{C_2}(\mathbf{x})}.$$

Finally, Proposition 11.5 implies that  $e^{\psi^{C_1}(\mathbf{x}) + \psi^{C_2}(\mathbf{x})} \leq e^{\psi^C(\mathbf{x})}$ . Dividing both sides by  $e^{\psi^C(\mathbf{x})}$  gives the result when  $m = 1$ .

Now suppose that  $m \geq 2$ . Differentiating (11.5) with respect to  $x_m$  and applying (11.6) gives

$$e^{\psi^C(\mathbf{x})} \partial_m \psi^C(\mathbf{x}) = \sum_{C \rightarrow C'} x_1 e^{x_1 \psi^{C'}(T\mathbf{x})} \partial_{m-1} \psi^{C'}(T\mathbf{x}). \tag{11.17}$$

By the inductive hypothesis, if  $C' = \pi(C'_1 \times C'_2)$  we have

$$\partial_{m-1} \psi^{C'}(T\mathbf{x}) \leq 2^{\sum_{i=2}^m \alpha^{j-i}} \left( \partial_{m-1} \psi^{C'_1}(T\mathbf{x}) + \partial_{m-1} \psi^{C'_2}(T\mathbf{x}) + \alpha^{j-3} \log 2 \right). \tag{11.18}$$

Also, by the upper bound in Proposition 11.5, we have

$$\psi^{C'}(T\mathbf{x}) \leq \psi^{C'_1}(T\mathbf{x}) + \psi^{C'_2}(T\mathbf{x}) + \alpha^{j-2} \log 2. \tag{11.19}$$

Substituting (11.18) and (11.19) into (11.17) and using the assumption that  $0 \leq x_1 \leq \alpha$  gives

$$\begin{aligned} e^{\psi^C(\mathbf{x})} \partial_m \psi^C(\mathbf{x}) &\leq 2^{\sum_{i=1}^m \alpha^{j-i}} \\ &\times \sum_{\substack{C_1 \rightarrow C'_1 \\ C_2 \rightarrow C'_2}} x_1 \left[ \partial_{m-1} \psi^{C'_1}(T\mathbf{x}) + \partial_{m-1} \psi^{C'_2}(T\mathbf{x}) + \alpha^{j-3} \log 2 \right] \\ &\times e^{x_1(\psi^{C'_1}(T\mathbf{x}) + \psi^{C'_2}(T\mathbf{x}))}. \end{aligned} \tag{11.20}$$

Now, differentiating the recurrence (11.15) with respect to  $x_m$  (using (11.6)) gives, for  $i = 1, 2$ ,

$$e^{\psi^{C_i}(\mathbf{x})} \partial_m \psi^{C_i}(\mathbf{x}) = \sum_{C_i \rightarrow C'_i} x_1 e^{x_1 \psi^{C'_i}(T\mathbf{x})} \partial_{m-1} \psi^{C'_i}(T\mathbf{x}). \tag{11.21}$$

Substituting (11.15) and (11.21) into (11.20), and using once again that  $x_1 \leq \alpha$ , gives

$$\begin{aligned} e^{\psi^C(\mathbf{x})} \partial_m \psi^C(\mathbf{x}) e^{\psi^C(\mathbf{x})} &\leq 2^{\sum_{i=1}^m \alpha^{j-i}} \left( \partial_m \psi^{C_1}(\mathbf{x}) + \partial_m \psi^{C_2}(\mathbf{x}) + \alpha^{j-2} \log 2 \right) \\ &\times e^{\psi^{C_1}(\mathbf{x}) + \psi^{C_2}(\mathbf{x})}. \end{aligned}$$

Again, Proposition 11.5 implies that  $e^{\psi^{C_1}(\mathbf{x})+\psi^{C_2}(\mathbf{x})} \leq e^{\psi^C(\mathbf{x})}$ , and so by dividing both sides by  $e^{\psi^C(\mathbf{x})}$ , we obtain the stated result.  $\square$

Before proving Proposition 11.6, we isolate a lemma.

**Lemma 11.8** *For  $0 \leq x_1 \leq 0.31$  we have  $0 \leq 4\partial_1\phi_2(\mathbf{x}) \leq 0.481$ .*

*Proof* We have  $e^{4\phi(\mathbf{x})} = 3^{x_1} + 4 \cdot 2^{x_1} + 4$ , and thus

$$4\partial_1\phi_2(\mathbf{x}) = \frac{\log 3 \cdot 3^{x_1} + \log 2 \cdot 4 \cdot 2^{x_1}}{3^{x_1} + 4 \cdot 2^{x_1} + 4}.$$

The lemma is therefore equivalent to

$$\frac{1}{4}(\log 3 - 0.481)3^{x_1} + (\log 2 - 0.481)2^{x_1} \leq 0.481.$$

The left-hand side here is increasing in  $x_1$  and, when  $x_1 = 0.31$ , it is equal to  $0.480052\dots$   $\square$

*Proof of Proposition 11.6* Henceforth, set  $\alpha := 0.31$  and fix two integers  $m \geq 1$  and  $j \geq 2$ . Our goal is to bound  $\partial_m\phi(\mathbf{x})$  uniformly for  $\mathbf{x} \in [0, \alpha]^{\mathbb{N}}$ . We may assume that  $j \geq m + 1$ , as  $\partial_m\phi_j(\mathbf{x}) = 0$  when  $j \leq m$ .

Now, let us define

$$A_m := 2^{1+\alpha+\dots+\alpha^{m-1}} \quad \text{and} \quad B_m := 2^{\frac{1+\alpha+\dots+\alpha^{m-1}}{1-\alpha}}.$$

Then, if we apply (11.13)  $\ell$  times, we obtain

$$\begin{aligned} 0 \leq \partial_m\phi_j(\mathbf{x}) &\leq A_m^{\alpha^{j-m}+\dots+\alpha^{j-m-(\ell-1)}} \partial_m\phi_{j-\ell}(\mathbf{x}) \\ &\quad + \frac{\log 2}{\alpha^2} \sum_{k=0}^{\ell-1} A_m^{\alpha^{j-m}+\dots+\alpha^{j-m-k}} \left(\frac{\alpha}{2}\right)^{j-k} \\ &\leq B_m^{\alpha^{j-m-(\ell-1)}} \partial_m\phi_{j-\ell}(\mathbf{x}) + \frac{\log 2}{\alpha^2} \sum_{k=0}^{\ell-1} B_m^{\alpha^{j-m-k}} \left(\frac{\alpha}{2}\right)^{j-k} \\ &\leq B_m^{\alpha^{j-m-\ell+1}} \left( \partial_m\phi_{j-\ell}(\mathbf{x}) + \frac{\log 2}{\alpha^2} \left(\frac{\alpha}{2}\right)^{j-\ell+1} \frac{1}{1-\alpha/2} \right). \end{aligned} \tag{11.22}$$

Here, we observed that all the  $B_m^{\alpha^t}$  terms in (11.22) have  $t \geq s + 1 - m$ ; bounding them all above by  $B_m^{\alpha^{s+1-m}}$  then allowed us to sum a geometric series.

Let us fix some  $s \in \{1, 2, \dots, m + 1\}$  independent of  $j$ . Then the number  $j - s$  lies in  $\{0, 1, \dots, j - 1\}$ . Hence, applying (11.22) with  $\ell = j - s$ , and then taking the supremum over all  $j \geq m + 1$  and all  $\mathbf{x} \in [0, \alpha]^{\mathbb{N}}$ , we find that

$$\Delta_m \leq B_m^{\alpha^{s+1-m}} \left( \sup_{\mathbf{x} \in [0, \alpha]^{\mathbb{N}}} |\partial_m \phi_s(\mathbf{x})| + \frac{\log 2}{\alpha^2} \left(\frac{\alpha}{2}\right)^{s+1} \frac{1}{1 - \alpha/2} \right). \tag{11.23}$$

When  $m = 1$ , we take  $s = 2$ . Then Lemma 11.8 and relation (11.23) give

$$\Delta_1 \leq 2^{\alpha^2/(1-\alpha)} \left( \frac{0.481}{4} + \frac{\alpha \log 2}{8(1 - \alpha/2)} \right) < 0.17,$$

as required. When  $m \geq 2$ , we take  $s = m$ . Then  $\partial_m \phi_s \equiv 0$  and so (11.23) degenerates to

$$\Delta_m \leq B_m^{\alpha} \frac{\log 2}{\alpha^2} \left(\frac{\alpha}{2}\right)^{m+1} \frac{1}{1 - \alpha/2}. \tag{11.24}$$

This gives  $\Delta_2 < 0.05$ , and also confirms that  $\Delta_m \ll 0.155^m$ . To bound  $\sum_{m \geq 3} \Delta_m$  we use (11.24) and the uniform bound  $B_m \leq 2^{1/(1-\alpha)^2}$ , obtaining

$$\sum_{m \geq 3} \Delta_m \leq \frac{\alpha^2 \log 2}{16(1 - \alpha/2)^2} 2^{\alpha/(1-\alpha)^2} < 0.01.$$

This completes the proof of Proposition 11.6. □

## 12 Calculating the $\rho_i$ and $\rho$

In this section we conclude our analysis of the parameters  $\rho_1, \rho_2, \dots$  for the binary flags. The situation so far is that we have shown that these parameters exist, are unique and lie in  $(0, 0.31)$ . Moreover, their limit  $\rho = \lim_{i \rightarrow \infty} \rho_i$  exists (Proposition 11.1).

None of this helps with actually computing the limit numerically or giving any kind of closed form for it, and the objective of this section is to provide tools for doing that. We prove two main results, Propositions 12.1 and 12.2 below. Recall the convention that  $\rho_0 = 0$ .

**Table 1** Table of  $\rho_j$ 

$j$	$\rho_j$	$j$	$\rho_j$
1	0.3064810093305	7	0.2812113502101
2	0.2796104150767	8	0.2812113496729
3	0.2813005404710	9	0.2812113496974
4	0.2812067224539	10	0.2812113496963
5	0.2812115789381	11	0.2812113496964
6	0.2812113387071	12	0.2812113496964

**Proposition 12.1** Recall the convention that  $\rho_0 = 0$ . Define a sequence  $(a_{i,j})_{i \geq 1, 1 \leq j \leq i+1}$  by the relations  $a_{i,1} = 2$ ,  $a_{i,2} = 2 + 2^{\rho_{i-1}}$  and

$$a_{i,j} = a_{i,j-1}^2 + a_{i-1,j-1}^{\rho_{i-1}} - a_{i-1,j-2}^{2\rho_{i-1}} \quad (3 \leq j \leq i+1). \quad (12.1)$$

Then

$$a_{i,i+1} = a_{i-1,i}^{\rho_{i-1}} e^{2^{i-1}} \quad \text{for } i = 2, 3, \dots \quad (12.2)$$

In practice, these relations are enough to calculate the  $\rho_j$  to high precision. Indeed, a short computer program produced the data in Table 1. (We suppress any discussion of the numerical precision of our routines.)

Using Proposition 12.1 we may obtain the following reasonably satisfactory description of  $\rho$ , which is equivalent to the statement of Theorem 2 (c).

**Proposition 12.2** For each  $t \in (0, 1)$ , define a sequence  $a_j(t)$  by

$$\begin{aligned} a_1(t) &= 2, a_2(t) = 2 + 2^t, \\ a_j(t) &= a_{j-1}(t)^2 + a_{j-1}(t)^t - a_{j-2}(t)^{2t} \quad (j \geq 3). \end{aligned} \quad (12.3)$$

Then the limit  $\rho = \lim_{i \rightarrow \infty} \rho_i$  is a solution (in the variable  $t$ ) to the equation

$$\frac{1}{1-t/2} = \lim_{j \rightarrow \infty} \frac{\log a_j(t)}{2^{j-2}}. \quad (12.4)$$

Furthermore,  $\rho$  is the unique solution to (12.4) in the interval  $0 \leq t \leq 1/3$ .

*Remark.* This is easily seen to be equivalent to Theorem 2 (c), but we have introduced  $t$  as a dummy variable since  $\rho$  now has the specific meaning  $\rho = \lim_{i \rightarrow \infty} \rho_i$ , and this will avoid confusion in the proof.

Before starting the proofs of Propositions 12.1 and 12.2, let us pause to observe a simple link between the sequences  $a_{i,j}$  and  $a_j(t)$  defined in (12.1) and (12.3) respectively.



**Lemma 12.3** For each fixed  $j \geq 1$ , the limit  $\lim_{i \rightarrow \infty} a_{i,j}$  exists and equals  $a_j(\rho)$ .

*Proof* The existence of the limit follows by induction on  $j$ , using Proposition 11.1, noting that the result is trivial for  $j = 1$  and immediate from Proposition 11.1 when  $j = 2$ . The fact that the limit equals  $a_j(\rho)$  then follows immediately by letting  $i \rightarrow \infty$  in (12.1) and comparing with (12.3).  $\square$

**12.1 Product formula for  $f^C(\rho)$  and a double recursion for the  $\rho_i$**

Proposition 12.1 is a short deduction from a product formula for  $F(g)$ , or equivalently for  $f^C(\rho)$ , given in Proposition 12.5 below. Whilst it would be a stretch to say that this formula is of independent interest, it is certainly a natural result to prove in the context of our work.

Before we state the formula, the reader should recall the notion of genotype  $g$  (Definition 10.1) and of the function  $F(g)$  (Proposition 10.4). We require the following further small definition.

**Definition 12.4 (Defects)** Let  $i, m \in \mathbb{Z}_{\geq 0}$  and let  $g$  be an  $i$ -genotype.

(a) If  $m \leq i$ , then we define the  $m$ th consolidation

$$g^{(m)} := \{A' \subset [i - m] : A' \cup X \in g \text{ for all } X \subset \{i - m + 1, \dots, i\}\}.$$

Otherwise, if  $m \geq i + 1$ , then by convention we define  $g^{(m)}$  to be empty.

(b) For  $m \geq 1$ , we set

$$\Delta^m(g) := |g^{(m-1)}| - 2|g^{(m)}|.$$

*Remark* Note that  $g^{(0)} = g$ ,  $g^{(1)} = g^*$  and  $g^{(m)} = (g^{(m-1)})^*$ . It is easy to see that  $\Delta^m(g)$  is always a nonnegative integer. Observe that  $\Delta^{i+1}(g) = 0$  unless  $g = \mathcal{P}[i]$ , in which case  $\Delta^{i+1}(g) = 1$ , and that  $\Delta^m(g) = 0$  whenever  $m > i + 1$ .

**Proposition 12.5** Let  $i \in \mathbb{N}$  and suppose that  $g$  is an  $i$ -genotype. Then

$$F(g) = \prod_{m=1}^{i+1} a_{i,m}^{\Delta^m(g)},$$

with the  $a_{i,m}$  defined as in Proposition 12.1 above.

*Proof of Proposition 12.1, given Proposition 12.5* Note that we have  $\Delta^m(\mathcal{P}[i]) = 1_{m=i+1}$  for  $1 \leq m \leq i + 1$ . Together with Proposition 12.5, this implies that  $F(\mathcal{P}[i]) = a_{i,i+1}$ . Thus  $f^{\Gamma_i}(\rho) = F(\mathcal{P}[i]) = a_{i,i+1}$ . The Eq. (12.2) is then an immediate consequence of the  $\rho$ -equations (9.1).  $\square$

Before turning to the proof of Proposition 12.5, we isolate a couple of lemmas from the proof.

**Lemma 12.6** *Let  $\alpha \in \mathbb{R}$  and  $i \in \mathbb{N}$ . Let  $g$  be an  $i$ -genotype, and suppose that  $k$  is an  $(i - 1)$ -genotype with  $k \leq g^*$ . Then*

$$\sum_{\substack{g' \leq g \\ (g')^* = k}} \alpha^{|g'|} = (1 + \alpha)^{\Delta^1(g)} (1 + 2\alpha)^{|g^*| - |k|} \alpha^{2|k|}.$$

*Proof* We have  $g = \{A \subset [i - 1] : A \in g\} \cup \{A \subset [i - 1] : A \cup \{i\} \in g\}$ . Hence, if we let

$$X = \{A \subset [i - 1] : A \in g, A \cup \{i\} \notin g\}$$

and

$$Y = \{A \subset [i - 1] : A \notin g, A \cup \{i\} \in g\},$$

then we have  $|g| = 2|g^*| + |X| + |Y|$ , and thus  $\Delta^1(g) = |X| + |Y|$ .

Now, in order to choose  $g' \leq g$  with  $(g')^* = k$ , we must decide independently for each  $A \subset [i - 1]$  whether  $A \in g'$  and/or  $A \cup \{i\} \in g'$ . The condition that  $g' \leq g$  means that if  $A \notin g$  (resp. if  $A \cup \{i\} \notin g$ ), then we are forced to have  $A \notin g'$  (resp.  $A \cup \{i\} \notin g'$ ). Let us now examine all admissible options for the conditions “ $A \in g'$ ” and “ $A \cup \{i\} \in g'$ ”:

- $A \in k$ : since  $(g')^* = k$ , we are forced to have  $A, A \cup \{i\} \in g'$ .
- $A \in g^* \setminus k$ : we know in this case that  $A, A \cup \{i\} \in g$ , so the condition  $g' \leq g$  imposes no further restrictions on the membership of  $A$  and of  $A \cup \{i\}$  in  $g'$ . On the other hand, we know that  $A \notin k = (g')^*$ , and thus at most one out of  $A$  and of  $A \cup \{i\}$  may belong to  $g'$ .
- $A \in X$ : the condition  $g' \leq g$  implies the restriction that  $A \cup \{i\} \notin g'$ , and we may then choose freely among the two options of having  $A \in g'$  or  $A \notin g'$ .
- $A \in Y$ : the condition  $g' \leq g$  implies the restriction that  $A \notin g'$ , and we may then choose freely among the two options of having  $A \cup \{i\} \in g'$  or  $A \cup \{i\} \notin g'$ .

By the above discussion, we have

$$\sum_{\substack{g' \leq g \\ (g')^* = k}} \alpha^{|g'|} = \alpha^{2|k|} \prod_{A \in g^* \setminus k} (1 + \alpha + \alpha) \prod_{A \in X} (1 + \alpha) \prod_{A \in Y} (1 + \alpha).$$

Since  $|X| + |Y| = \Delta^1(g)$ , the proof is complete. □

For  $\mathbf{a} = (a_1, a_2, \dots)$ , and for some ( $i$ -)genotype  $g$ , write

$$P_{\mathbf{a}}(g) := \prod_{m=1}^{i+1} a_m^{\Delta^m(g)}. \tag{12.5}$$

(Note that the  $a_m$  here are just parameters, not related to the recursion (12.3), which does not feature in this subsection.) If  $\theta \in \mathbb{R}_{>0}$ , define

$$\Phi_{\theta, \mathbf{a}}(g) := \sum_{g' \leq g} \theta^{|g|-|g'|} P_{\mathbf{a}}(g'). \tag{12.6}$$

**Lemma 12.7** *We have the functional equation*

$$\Phi_{\theta, \mathbf{a}}(g) = (\theta + a_1)^{\Delta^1(g)} \Phi_{\theta^2+2a_1\theta, T\mathbf{a}}(g^*).$$

As before,  $T\mathbf{a}$  denotes the shift operator  $T\mathbf{a} = (a_2, a_3, \dots)$ .

*Proof* Using the relation  $P_{\mathbf{a}}(g') = a_1^{\Delta^1(g')} P_{T\mathbf{a}}((g')^*)$ , we have

$$\begin{aligned} \Phi_{\theta, \mathbf{a}}(g) &= \theta^{|g|} \sum_{g' \leq g} \left(\frac{a_1}{\theta}\right)^{|g'|} \left(\frac{1}{a_1^2}\right)^{|(g')^*|} P_{T\mathbf{a}}((g')^*) \\ &= \theta^{|g|} \sum_{k \leq g^*} \left(\frac{1}{a_1^2}\right)^{|k|} P_{T\mathbf{a}}(k) \sum_{\substack{g' \leq g \\ (g')^* = k}} \left(\frac{a_1}{\theta}\right)^{|g'|}. \end{aligned}$$

The result now follows from Lemma 12.6 and a routine short calculation.  $\square$

We are now in a position to prove Proposition 12.5.

*Proof of Proposition 12.5* Let  $a_{i,m}$  be as in the statement of Proposition 12.5, and write  $\mathbf{a}_i = (a_{i,1}, a_{i,2}, \dots)$ . In the notation introduced above (cf. (12.5)) the claim of Proposition 12.5 is then that

$$F(g) = P_{\mathbf{a}_i}(g). \tag{12.7}$$

We proceed by induction on  $i$ . Let us first consider the base case when  $i = 1$ .

- If  $g = \mathcal{P}[1]$ , we have  $F(g) = f^{\Gamma_1}(\rho) = 3$ . On the other hand,  $P_{\mathbf{a}_1}(\mathcal{P}[1]) = a_{1,2} = 3$  in this case by the convention that  $\rho_0 = 0$ .
- If  $g \subsetneq \mathcal{P}[1]$ , then  $g^* = \emptyset$  and thus  $\Delta^1(g) = |g|$  and  $\Delta^2(g) = 0$ . So we conclude that  $P_{\mathbf{a}_1}(g) = 2^{|g|}$ . On the other hand, for all such genotypes, the corresponding cell contains  $2^{|g|}$  elements that all split into unicells at level 0. Consequently,  $F(g) = 2^{|g|} = P_{\mathbf{a}_1}(g)$  in this case too.

Next, suppose that we have the result for  $(i - 1)$ -genotypes for some  $i \geq 2$ , and let  $g$  be an  $i$ -genotype. We know from (10.3) that

$$F(g) = \sum_{g' \leq g^*} 2^{|g| - |g^*| - |g'|} F(g')^{\rho_{i-1}}.$$

By the induction hypothesis, we have  $F(g')^{\rho_{i-1}} = P_{\mathbf{a}_{i-1}^{\rho_{i-1}}}(g')$  for all  $g' \leq g^*$ , where  $\mathbf{a}_{i-1}^{\rho_{i-1}}$  is shorthand for  $(a_{i-1,1}^{\rho_{i-1}}, a_{i-1,2}^{\rho_{i-1}}, \dots)$ . Hence, it follows immediately that

$$F(g) = 2^{\Delta^1(g)} \Phi_{2, \mathbf{a}_{i-1}^{\rho_{i-1}}}(g^*). \tag{12.8}$$

with  $\Phi$  defined in (12.6). The fact that the right-hand side of (12.8) is a product  $P_*(g)$  is now clear by an iterated application of Lemma 12.7. To get a handle on exactly which product, suppose that the result of applying Lemma 12.7  $j - 1$  times is that

$$F(g) = \left( \prod_{m=1}^j b_{i,m}^{\Delta^m(g)} \right) \Phi_{\theta_{i,j}, T^{j-1}(\mathbf{a}_{i-1}^{\rho_{i-1}})}(g^{(j)}). \tag{12.9}$$

Thus  $b_{i,1} = \theta_{i,1} = 2$ , and we have the relations

$$b_{i,j+1} = \theta_{i,j} + a_{i-1,j}^{\rho_{i-1}} \tag{12.10}$$

and

$$\theta_{i,j+1} = \theta_{i,j}^2 + 2a_{i-1,j}^{\rho_{i-1}}\theta_{i,j} \tag{12.11}$$

for  $j \in \{1, \dots, i\}$ . We claim that  $b_{i,j} = a_{i,j}$  for all  $j \leq i + 1$ . This will complete the proof of Proposition 12.5, because we may then apply (12.9) with  $j = i + 1$  to show that

$$F(g) = \left( \prod_{m=1}^{i+1} a_{i,m}^{\Delta^m(g)} \right) \Phi_{\theta_{i,i+1}, T^{i+1}(\mathbf{a}_{i-1}^{\rho_{i-1}})}(g^{(i+1)}) = \prod_{m=1}^{i+1} a_{i,m}^{\Delta^m(g)}$$

because  $g^{(i+1)} = \emptyset$  for all  $i$ -genotypes  $g$ .

Let us now prove our claim that  $b_{i,j} = a_{i,j}$  for all  $j \leq i + 1$ . We shall use induction on  $j$ . We have  $b_{i,1} = 2 = a_{i,1}$ . In addition,  $b_{i,2} = 2 + 2^{\rho_{i-1}} = a_{i,2}$  by (12.10) with  $j = 1$  and by the fact that  $\theta_{i,1} = 2$ . Now, assume that we have

proven that  $b_{i,j} = a_{i,j}$  for some  $j \in \{2, \dots, i\}$ . Relation (12.11) applied with  $j - 1$  in place of  $j$  implies that

$$\theta_{i,j} + a_{i-1,j-1}^{2\rho_{i-1}} = (\theta_{i,j-1} + a_{i-1,j-1}^{\rho_{i-1}})^2.$$

The right-hand side equals  $b_{i,j}^2 = a_{i,j}^2$  by applying (12.10) followed by the induction hypothesis. Thus,  $\theta_{i,j} = a_{i,j}^2 - a_{i-1,j-1}^{2\rho_{i-1}}$ . Inserting this relation into (12.10) and using the recursive formula (12.1) shows that  $b_{i,j+1} = a_{i,j+1}$ . This completes the inductive step and thus the proof of Proposition 12.5.  $\square$

### 12.2 A single recurrence for $\rho$

In this section we deduce Proposition 12.2 from Proposition 12.1 by a limiting argument.

To carry this out, we will need the following fairly crude estimates for the  $a_{i,j}$  and the  $a_j(t)$ , defined in (12.1) and (12.3) respectively.

**Lemma 12.8** *We have*

$$a_{i,j+1} \leq a_{i,j}^2 \text{ for } 1 \leq j \leq i \tag{12.12}$$

and

$$3^{2^{j-2}} \leq a_{i,j} \leq a_{i,2}^{2^{j-2}} \leq 4^{2^{j-2}} \text{ for } 2 \leq j \leq i + 1. \tag{12.13}$$

*Proof* Since  $\rho_{i-1} < 1$  for all  $i \geq 1$  (cf. Lemma 11.2), we have  $a_{i,2} < 4 = a_{i,1}^2$ . Hence, the inequality (12.12) follows from a simple induction using (12.1).

Using another simple induction, we readily confirm the inequality  $a_{i,j} \leq a_{i,2}^{2^{j-2}}$  in (12.13).

For the lower bound in (12.13), we know from (12.10) and (12.11) and from the fact that  $b_{i,j} = a_{i,j}$  for all  $j \leq i + 1$  that

$$a_{i,j+1} = \theta_{i,j} + a_{i-1,j}^{\rho_{i-1}} \tag{12.14}$$

and that

$$\theta_{i,j+1} = \theta_{i,j}^2 + 2a_{i-1,j}^{\rho_{i-1}}\theta_{i,j} \tag{12.15}$$

for  $j \in \{1, \dots, i\}$ . By a simple induction, these formulas imply that  $a_{i,j} > 1$  and  $\theta_{i,j} > 0$  for all  $j \leq i + 1$ , and thus  $\theta_{i,j+1} + 1 \geq (\theta_{i,j} + 1)^2$  for  $j = 1, 2, \dots, i$ . By yet another induction, we find  $\theta_{i,j} \geq 3^{2^{j-1}} - 1$ . Finally, the lower bound on the  $a_{i,j}$  in (12.13) follows from this and (12.14).  $\square$

**Lemma 12.9** *Let  $t \in (0, 1)$ . We have*

$$a_{j+1}(t) \leq a_j(t)^2 \quad \text{for } j \geq 1 \quad (12.16)$$

and

$$3^{2^{j-2}} \leq a_j(t) \leq a_2(t)^{2^{j-2}} \leq 4^{2^{j-2}} \quad \text{for } j \geq 2. \quad (12.17)$$

*Proof* The inequality (12.16) follows from a simple induction using (12.3), and the upper bound in (12.17) follows with a further induction.

For the lower bound, we first set up relations analogous to (12.14) and (12.15), defining  $\theta_j(t)$  for  $j \geq 1$  via the relation

$$a_{j+1}(t) = \theta_j(t) + a_j(t)^t. \quad (12.18)$$

We then note that we also have

$$\theta_{j+1}(t) = \theta_j(t)^2 + 2a_j(t)^t \theta_j(t). \quad (12.19)$$

Indeed, on the one hand, we have

$$\theta_{j+1}(t) = a_{j+2}(t) - a_{j+1}(t)^t = a_{j+1}(t)^2 - a_j(t)^{2t}$$

by (12.3). On the other hand,

$$\theta_j(t)^2 + 2a_j(t)^t \theta_j(t) = (\theta_j(t) + a_j(t))^2 - a_j(t)^{2t} = a_{j+1}(t)^2 - a_j(t)^{2t}$$

by (12.18).

Having proven (12.19), we now proceed analogously to the proof of Lemma 12.8. We have  $a_j(t) > 1$  and  $\theta_j(t) > 0$  for all  $j \geq 1$ , by a simple induction using (12.18) and (12.19). Therefore, from (12.19), we have that

$$\theta_{j+1}(t) + 1 \geq (\theta_j(t) + 1)^2.$$

By induction, this implies that  $\theta_j(t) \geq 3^{2^{j-1}} - 1$ . Finally, the lower bound on the  $a_j(t)$  in (12.17) follows from this and (12.18).  $\square$

We are now in a position to prove that the relation

$$\frac{1}{1 - t/2} = \lim_{j \rightarrow \infty} \frac{\log a_j(t)}{2^{j-2}} \tag{12.20}$$

holds with  $t = \rho$ , which is one of the main statements of Proposition 12.2. Iterating (12.2) gives

$$\begin{aligned} a_{i,i+1} &= \exp(2^{i-1}) a_{i-1,i}^{\rho_{i-1}} = \exp(2^{i-1} + \rho_{i-1} 2^{i-2}) a_{i-2,i-1}^{\rho_{i-2} \rho_{i-1}} = \dots \\ &= \exp\left(2^{i-1} + \sum_{j=1}^{i-2} (\rho_{i-j} \dots \rho_{i-1}) 2^{i-j-1}\right) a_{1,2}^{\rho_1 \dots \rho_{i-1}}. \end{aligned}$$

By Proposition 11.1, we have  $\rho_i \rightarrow \rho$ . In addition, by Lemma 11.2, we have  $0 \leq \rho_i \leq \rho_1 < 0.31$  for all  $i$ . Thus, taking limits as  $i \rightarrow \infty$  gives

$$\lim_{i \rightarrow \infty} \frac{\log a_{i,i+1}}{2^{i-1}} = 1 + \frac{\rho}{2} + \left(\frac{\rho}{2}\right)^2 + \dots = \frac{1}{1 - \rho/2}. \tag{12.21}$$

We now derive another expression for the left-hand side of (12.21). A telescoping argument gives

$$\frac{\log a_{i,i+1}}{2^{i-1}} = \log 4 + \sum_{j=1}^i \frac{1}{2^{j-1}} \log \left( \frac{a_{i,j+1}}{a_{i,j}^2} \right). \tag{12.22}$$

The terms on the right-hand side of (12.22) are rapidly decreasing. Indeed, by (12.12) we have  $1 \geq a_{i,j+1}/a_{i,j}^2$  for all  $j \geq 1$ . On the other hand, by (12.1) (with  $j$  replaced by  $j + 1$  there) and by (12.13), we have

$$\frac{a_{i,j+1}}{a_{i,j}^2} \geq 1 - \frac{a_{i-1,j-1}^{2\rho_1}}{a_{i,j}^2} = 1 + O\left(\left(\frac{2^{\rho_1-1}}{3}\right)^{2^{j-1}}\right).$$

for all  $j \in \{2, \dots, i\}$ . Since  $\rho_{i-1} \leq \rho_1 \leq 0.31$ , we have  $2^{\rho_1-1}/3 < 1/2$ . In conclusion,

$$\log \left( \frac{a_{i,j+1}}{a_{i,j}^2} \right) = O(2^{-2^{j-1}}) \tag{12.23}$$

for all  $j \in \{1, \dots, i\}$ . By a simple limiting argument using relation (12.22) and Lemma 12.3, we thus find that

$$\lim_{i \rightarrow \infty} \frac{\log a_{i,i+1}}{2^{i-1}} = \log 4 + \sum_{j=1}^{\infty} \frac{1}{2^{j-1}} \log \left( \frac{a_{j+1}(\rho)}{a_j(\rho)^2} \right) = \lim_{j \rightarrow \infty} \frac{\log a_j(\rho)}{2^{j-2}}.$$

Here, we used (12.23) to bound the terms with  $j$  large. Comparing this with (12.21) confirms that indeed (12.20) is satisfied with  $t = \rho$ .

We turn now to the final statement in Proposition 12.2, the statement that (12.20) has a unique solution in  $t \in [0, \frac{1}{3}]$  (which must, by the above discussion, be  $\rho$ ). This is a purely analytic problem. Write

$$W_j(t) := \frac{1}{1-t/2} - \frac{\log a_j(t)}{2^{j-2}}, \quad W(t) := \lim_{j \rightarrow \infty} W_j(t).$$

We must show that there is only one solution to  $W(t) = 0$ . We already know  $W(\rho) = 0$ , so it would suffice to show that  $W$  is strictly increasing in  $[0, 1/3]$ . This would certainly follow if we could show that

$$W_j(t') - W_j(t) \geq \frac{1}{6}(t' - t)$$

for all  $j \geq 2$  and all  $0 \leq t \leq t' \leq 1/3$ . Since the derivative of  $\frac{1}{1-t/2}$  is bounded below by  $\frac{1}{2}$  on  $[0, \frac{1}{3}]$ , it is enough to establish the derivative bound

$$\frac{d}{dt} \left( \frac{\log a_j(t)}{2^{j-2}} \right) \leq \frac{1}{3}$$

for all  $j \geq 2$  and all  $t \in (0, \frac{1}{3})$ . The remainder of the section is devoted to proving this bound, which it is convenient to write in the form

$$\ell_j(t) \leq \frac{1}{3} \cdot 2^{j-2}, \tag{12.24}$$

where  $\ell_j(t) := a'_j(t)/a_j(t)$ .

We begin by observing that, since  $t \in (0, \frac{1}{3})$ , we have  $a_2(t) \leq 2 + 2^{1/3}$  and so we may upgrade the upper bound in (12.17) to

$$a_j(t) \leq (2 + 2^{1/3})^{2^{j-2}} \tag{12.25}$$

for  $j \geq 2$ . Note also that, by induction using (12.18) and (12.19), both  $a_j(t)$  and  $\theta_j(t)$  are increasing functions of  $t$ . In particular,  $a_j(t)$  is an increasing function of  $t$  so the derivative  $a'_j(t)$  is positive.

Differentiating (12.3) gives

$$a'_{j+1} = 2a_j a'_j + (a'_j \log a_j - 2a_j^{2t} \log a_{j-1}) + t a_j^t \frac{a'_j}{a_j} - 2t a_j^{2t} \frac{a'_{j-1}}{a_{j-1}}, \tag{12.26}$$

where here and in the next few lines we have omitted the argument ( $t$ ) from the functions for brevity. The term in parentheses is non-positive by (12.16), and



the final term  $-2ta_{j-1}^{2t} \frac{a'_{j-1}}{a_{j-1}}$  is negative since the derivative  $a'_{j-1}$  is positive. It follows from (12.26) that

$$a'_{j+1} < 2a_j a'_j + ta_j^t \frac{a'_j}{a_j}.$$

A little computation using (12.3) shows that this may equivalently be written as

$$\ell_{j+1} < 2\ell_j \left( \frac{1}{1 + a_j^{t-2} - a_{j-1}^{2t} a_j^{-2}} + \frac{ta_j^t}{2a_{j+1}} \right), \tag{12.27}$$

where we used our notation  $\ell_j = a'_j/a_j$ .

Denote

$$\xi_j := \sup_{t \in [0, \frac{1}{3}]} \left( \frac{1}{1 + a_j(t)^{t-2} - a_{j-1}(t)^{2t} a_j(t)^{-2}} + \frac{ta_j(t)^t}{2a_{j+1}(t)} \right). \tag{12.28}$$

Then (12.27) implies that  $\ell_{j+1}(t) < 2\ell_j(t)\xi_j$  for all  $t \in [0, 1/3]$  and all  $j \geq 2$ . Telescoping this inequality gives

$$\ell_j(t) \leq (\ell_2(t)\xi_2\xi_3 \cdots \xi_{j-1}) \cdot 2^{j-2}.$$

We have

$$\ell_2(t) = \frac{2^t \log 2}{2 + 2^t} \leq \frac{\log 2}{1 + 2^{2/3}} < 0.268$$

for all  $t \in [0, 1/3]$ . Hence, in order to obtain the desired bound (12.24), it is enough to show

$$\xi_2\xi_3 \cdots \xi_{j-1} < 1.2. \tag{12.29}$$

The  $\xi_i$  tend to 1 exceptionally rapidly, and crude bounds (together with a little computation) turn out to suffice, as follows.

First, by (12.17) and the fact that  $a_2(t)^{2-t} = (2 + 2^t)^{2-t} \leq 9$  for  $t \in [0, 1]$  (a calculus exercise), we have

$$a_j(t)^{t-2} \geq (a_2(t)^{t-2})^{2^{j-2}} \geq 9^{-2^{j-2}} \quad \text{for } j \geq 2. \tag{12.30}$$

Second, by the lower bound in (12.17) and by (12.25) we have

$$a_{j-1}(t)^{2t} a_j(t)^{-2} \leq ((2 + 2^{1/3})^{2^{j-3}})^{2/3} (3^{2^{j-2}})^{-2} < 6^{-2^{j-2}} \quad \text{for } j \geq 3.$$

We may also check by hand that  $a_1(t)^{2t}/a_2(t)^2 = (2^{1-t} + 1)^{-2} < 1/6$  for all  $t \in [0, 1/3]$ . Hence,

$$a_{j-1}(t)^{2t} a_j(t)^{-2} < 6^{-2^{j-2}} \quad \text{for } j \geq 2. \tag{12.31}$$

Third, again by the lower bound in (12.17) and by (12.25), we have

$$\frac{a_j(t)^t}{a_{j+1}(t)} \leq \frac{((2 + 2^{1/3})2^{j-2})^{1/3}}{3^{2^{j-1}}} \leq \left(\frac{1}{6}\right)^{2^{j-2}} \quad \text{for } j \geq 2. \tag{12.32}$$

Substituting (12.30), (12.31) and (12.32) into the definition (12.28) gives

$$\xi_j \leq \frac{1}{1 + (\frac{1}{9})^{2^{j-2}} - (\frac{1}{6})^{2^{j-2}}} + \left(\frac{1}{6}\right)^{1+2^{j-2}} \quad \text{for } j \geq 2.$$

Using this bound, one may check the bound  $\prod_{j=2}^\infty \xi_j \leq 10/9$ , which is stronger than the desired bound (12.29), on a pocket calculator or even by hand. For example, we have  $\xi_2 \xi_3 \leq \frac{46751495}{42169248}$  and can use a very crude bounds for the higher terms. Since  $\frac{1}{1-x} + \frac{x}{6} \leq e^{2x}$  for  $0 \leq x \leq 0.1$ , taking  $x = 6^{-2^{j-2}}$  gives

$$\xi_j \leq \exp(2 \cdot 6^{-2^{j-2}})$$

for  $j \geq 4$ . Therefore

$$\prod_{j=4}^\infty \xi_j < \exp\left(2 \sum_{i=4}^\infty \frac{1}{6^i}\right) = e^{2/(5 \cdot 6^3)} < 1.002.$$

This concludes the proof of the final statement in Proposition 12.2.

### 12.3 Proof of parts (b) and (c) of Theorem 2

To conclude this paper, we complete the proof of parts (b) and (c) of Theorem 2, as defined in the end of Sect. 1.3. In fact, all of the ingredients have already been assembled and we must simply remark on how they fit together.

First, recall from Definition 9.6 that

$$\theta_r = (\log 3 - 1) / \left( \log 3 + \sum_{i=1}^{r-1} \frac{2^i}{\rho_1 \cdots \rho_i} \right).$$

Now, it is an easy exercise to see that if  $x_1, x_2, \dots$  is a sequence of positive real numbers for which  $x = \lim_{i \rightarrow \infty} x_i$  exists and is positive, then

$$\lim_{r \rightarrow \infty} \left( \sum_{i=1}^r x_1 \cdots x_i \right)^{1/r} = \max(x, 1).$$

Applying this with  $x_i = 2/\rho_i$  gives, by Proposition 11.1, that

$$\lim_{r \rightarrow \infty} \theta_r^{1/r} = \frac{\rho}{2}.$$

This, together with Proposition 12.2, completes the proof of Theorem 2.

**Acknowledgements** This collaboration began at the MSRI program on Analytic Number Theory, which took place in the first half of 2017 and which was supported by the National Science Foundation under Grant No. DMS-1440140. All three authors are grateful to MSRI for allowing us the opportunity to work together. The project was completed during a visit of KF and DK to Oxford in the first half of 2019. Both authors are grateful to the University of Oxford for its hospitality. KF is supported by the National Science Foundation Grants DMS-1501982 and DMS-1802139. In addition, his stay at Oxford in early 2019 was supported by a Visiting Fellowship at Magdalen College Oxford. BG is supported by a Simons Investigator Grant, which also funded DK’s visit to Oxford. DK is also supported by the Courtois Chair II in fundamental research, by the Natural Sciences and Engineering Research Council of Canada (RGPIN-2018-05699) and by the Fonds de recherche du Québec - Nature et technologies (2019-PR-256442 and 2022-PR-300951).

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## Appendices

### Appendix A. Some probabilistic lemmas

Throughout this section,  $\mathbf{A} \subset \mathbb{N}$  will be a random set, with  $\mathbb{P}(i \in \mathbf{A}) = 1/i$  and these choices being independent for different values of  $i$ .

**Lemma A.1** *For any finite subset  $B \subset \mathbb{Z}_{\geq 4}$  and any  $k \in \mathbb{Z}_{\geq 0}$ , we have*

$$\left( 1 - \frac{2k^2(\sum_{m \in B} 1/(m-1))^{-2}}{\min B} \right) M \leq \mathbb{P}(\#\mathbf{A} \cap B = k) \leq M,$$

where

$$M = \frac{1}{k!} \left( \sum_{m \in B} \frac{1}{m-1} \right)^k \prod_{m \in B} \left( 1 - \frac{1}{m} \right).$$

*Proof* The result follows by a standard inclusion-exclusion argument. We have

$$\begin{aligned} \mathbb{P}(\#(\mathbf{A} \cap B) = k) &= \sum_{\substack{a_1, \dots, a_k \in B \\ a_1 < \dots < a_k}} \frac{1}{a_1 \cdots a_k} \prod_{\substack{m \in B \\ m \notin \{a_1, \dots, a_k\}}} \left( 1 - \frac{1}{m} \right) \\ &= \prod_{m \in B} \left( 1 - \frac{1}{m} \right) \sum_{\substack{a_1, \dots, a_k \in B \\ a_1 < \dots < a_k}} \frac{1}{(a_1 - 1) \cdots (a_k - 1)} \leq M. \end{aligned}$$

For the lower bound, we note that

$$\begin{aligned} &\frac{1}{k!} \left( \sum_{m \in B} \frac{1}{m-1} \right)^k - \sum_{\substack{a_1, \dots, a_k \in B \\ a_1 < \dots < a_k}} \frac{1}{(a_1 - 1) \cdots (a_k - 1)} \\ &= \frac{1}{k!} \sum_{\substack{a_1, \dots, a_k \in B \\ \exists i < j \text{ with } a_i = a_j}} \frac{1}{(a_1 - 1) \cdots (a_k - 1)} \\ &\leq \frac{1}{k!} \binom{k}{2} \left( \sum_{a \in B} \frac{1}{(a-1)^2} \right) \left( \sum_{a \in B} \frac{1}{(a-1)} \right)^{k-2}. \end{aligned}$$

Since  $\sum_{a \in B} 1/(a-1)^2 < 1/(\min B - 2)^2 \leq 4/(\min B)^2$ , the proof is complete. □

**Lemma A.2** *Uniformly for  $B \subset \mathbb{N}$  with  $\lambda := \sum_{m \in B} 1/m \geq 1$  and  $0 \leq \varepsilon \leq 1$ , we have*

$$\mathbb{P}\left( \left| \#(\mathbf{A} \cap B) - \lambda \right| > \varepsilon \lambda \right) \ll \exp(-\varepsilon^2 \lambda / 3).$$

*Proof* This follows by the upper bound in Lemma A.1 with standard bounds on the tails of the Poisson distribution, e.g. Norton’s bounds [15, Theorem 09]. □

**Lemma A.3** *For any  $x > 0$  and finite set  $B \subset \mathbb{N}$ ,*

$$\mathbb{E}x^{\#(\mathbf{A} \cap B)} \leq \exp\left( (x-1) \sum_{j \in B} \frac{1}{j} \right).$$

*Proof* The random variable  $\#(\mathbf{A} \cap B)$  is the sum of independent Bernoulli random variables and thus

$$\mathbb{E}x^{\#(\mathbf{A} \cap B)} = \prod_{j \in B} \left( 1 + \frac{x - 1}{j} \right).$$

Note that all factors are positive because  $x > 0$ . The lemma now follows from the inequality  $1 + y \leq e^y$ , valid for all real  $y$ . □

**Lemma A.4** *Let  $k \in \mathbb{N}$ , and let  $B$  and  $G$  be finite sets such that  $B \subset G \subset \mathbb{Z}_{\geq 4}$  and*

$$|B| = k \leq \frac{\sqrt{\min(G)}}{2} \sum_{m \in G} \frac{1}{m}.$$

*Then*

$$\begin{aligned} & \mathbb{P}(\mathbf{A} \cap G = B \mid \#(\mathbf{A} \cap G) = k) \\ &= \frac{k!(1 + O(\frac{k^2(\sum_{m \in G} 1/m)^{-2}}{\min(G)}))}{(\sum_{m \in G} 1/(m - 1))^k} \prod_{b \in B} \frac{1}{b} \prod_{m \in G} \left( 1 - \frac{1}{m} \right). \end{aligned}$$

*Proof* Since  $|B| = k$ , we have

$$\mathbb{P}(\mathbf{A} \cap G = B \mid \#(\mathbf{A} \cap G) = k) = \frac{\mathbb{P}(\mathbf{A} \cap G = B)}{\mathbb{P}(\#(\mathbf{A} \cap G) = k)}.$$

The denominator is estimated using Lemma A.1, whereas for the numerator we simply note that

$$\mathbb{P}(\mathbf{A} \cap G = B) = \prod_{b \in B} \frac{1}{b} \prod_{m \in G \setminus B} \left( 1 - \frac{1}{m} \right) = \prod_{b \in B} \frac{1}{b - 1} \prod_{m \in G} \left( 1 - \frac{1}{m} \right).$$

This completes the proof of the lemma. □

**Lemma A.5** *Given  $0 < c < 1$  and  $D \geq e^{100/c}$ , the probability that  $\mathbf{A} \subset (D^c, D]$  satisfies*

$$\left| \#(\mathbf{A} \cap (D^\alpha, D^\beta]) - (\beta - \alpha) \log D \right| \leq (\log D)^{3/4} \quad (c \leq \alpha \leq \beta \leq 1) \tag{A.1}$$

*is*  $\geq 1 - O(e^{-(1/4)(\log D)^{1/2}})$ .

*Proof* It suffices to bound the probability that

$$\left| \#\mathbf{A} \cap (D^\alpha, D^\beta] - (\beta - \alpha) \log D \right| \geq (\log D)^{3/4} - 2 \tag{A.2}$$

whenever  $\alpha \log D, \beta \log D \in \mathbb{N}$ . The random variable  $N = N(\alpha, \beta) := \#\mathbf{A} \cap (D^\alpha, D^\beta]$  is the sum of Bernoulli random variables and has expectation  $\mathbb{E}N = M + O(1)$ , where

$$M = (\beta - \alpha) \log D.$$

By Lemma A.3,  $\mathbb{E}\lambda^N \leq e^{(\lambda-1)\mathbb{E}N}$ . Thus, for  $y = (\log D)^{3/4}$  and  $\lambda_j = 1 + (-1)^j \frac{y}{\log D}$  we have

$$\begin{aligned} \mathbb{P}(N \geq M + y) &\leq \mathbb{E}\lambda_2^{N-M-y} \ll \lambda_2^{-M-y} e^{(\lambda_2-1)M} \ll e^{-(1/3)(\log D)^{1/2}}, \\ \mathbb{P}(N \leq M - y) &\leq \mathbb{E}\lambda_1^{N-M+y} \ll \lambda_1^{-M+y} e^{(\lambda_1-1)M} \ll e^{-(1/3)(\log D)^{1/2}}. \end{aligned}$$

Summing over all possible  $\alpha, \beta$  completes the proof. □

**Lemma A.6** *Uniformly for  $X \geq 2$  and  $K \geq 2$  we have*

$$\sum_{a \in \mathbf{A} \cap [2, X]} a \leq KX$$

with probability  $\geq 1 - e^{2-K}$ .

*Proof* We use Chernoff’s inequality, often called Rankin’s trick in this context:

$$\begin{aligned} \mathbb{P}\left(\sum_{a \in \mathbf{A} \cap [2, X]} a > KX\right) &\leq e^{-K} \sum_{A' \subset [2, X]} \mathbb{P}(\mathbf{A} \cap [2, X] = A') e^{\frac{1}{X} \sum_{a \in A'} a} \\ &= e^{-K} \sum_{A' \subset [2, X]} \prod_{\substack{2 \leq a \leq X \\ a \notin A'}} \left(1 - \frac{1}{a}\right) \prod_{a \in A'} \frac{e^{a/X}}{a} \\ &= e^{-K} \prod_{2 \leq a \leq X} \left(1 - \frac{1}{a}\right) \left(1 + \frac{e^{a/X}}{a-1}\right) \\ &= e^{-K} \prod_{2 \leq a \leq X} \left(1 + \frac{e^{a/X} - 1}{a}\right) \\ &\leq e^{-K} (1 + 2/X)^X \leq e^{2-K} \end{aligned}$$

because  $e^t \leq 1 + 2t$  for all  $t \in [0, 1]$ . This concludes the proof. □

**Lemma A.7** *Let  $\eta \in [0, 1]$  and let  $J_1, \dots, J_d \subset \mathbb{N}$  be mutually disjoint intervals. Suppose that  $X \subset J_1 \times \dots \times J_d$  is a set of size  $\eta \prod_i \max J_i$ . If  $\min_i |J_i|$  is sufficiently large in terms of  $\eta$  and  $d$ , then with probability  $\geq (\eta/4)^d$ , there are distinct elements  $a_i \in \mathbf{A}$  with  $(a_1, \dots, a_d) \in X$ .*

*Proof* Let  $M_i = \max J_i$  for each  $i$ . We will prove the lemma by induction on  $d$ .

The case  $d = 1$  follows by direct calculation: Suppose that  $X \subset J_1$  has size  $\geq \eta M_1$ . Then

$$\mathbb{P}(\mathbf{A} \cap X = \emptyset) = \prod_{n \in X} (1 - 1/n) \leq (1 - 1/M_1)^{\eta M_1} \leq e^{-\eta} \leq 1 - \eta/2.$$

Let us now assume we have proven the lemma for  $d - 1$  intervals, and let us prove it for  $d$  intervals  $J_1, \dots, J_d$ . For each  $j_1 \in J_1$ , we set

$$X_{j_1} := \{(j_2, \dots, j_d) \in J_2 \times \dots \times J_d : (j_1, j_2, \dots, j_d) \in X\}.$$

Let  $Y = \{j_1 \in J_1 : |X_{j_1}| \geq (\eta/2)M_1\}$ . Then  $|Y| \geq (\eta/2)M_1$ , because otherwise we would have  $|X| < \eta \prod_i M_i$ , a contradiction to our hypotheses. By the case  $d = 1$  (just described),  $\mathbf{A} \cap Y$  is nonempty with probability  $\geq \eta/4$ . Fix some  $a_1 \in \mathbf{A} \cap Y$ . Then, by the inductive hypothesis and the fact that the  $J_i$  are disjoint, with probability  $\geq (\eta/4)^{d-1}$ , independent of the choice of  $a_1$ , there are elements  $a_i \in \mathbf{A} \cap J_i, i = 2, \dots, d$  with  $(a_2, \dots, a_d) \in X_{a_1}$ , and therefore  $(a_1, \dots, a_d) \in X$ . The disjointness of the  $J_i$  of course guarantees that the  $a_i$  are all distinct. This completes the proof.  $\square$

**Lemma A.8** *If  $X_j, Y_j$  live on the same discrete probability space for  $1 \leq j \leq k$ , and furthermore  $X_1, \dots, X_k$  are independent, and  $Y_1, \dots, Y_k$  are also independent, then*

$$d_{\text{TV}}((X_1, \dots, X_k), (Y_1, \dots, Y_k)) \leq \sum_{j=1}^k d_{\text{TV}}(X_j, Y_j),$$

*Proof* We begin with the following identity

$$a_1 \dots a_m - b_1 \dots b_m = \sum_{j=1}^m (a_j - b_j) \prod_{i < j} a_i \prod_{i > j} b_i.$$

Denoting  $\Omega$  the domain of  $(X_1, \dots, X_m)$ , and writing  $a_i = \mathbb{P}(X_i = \omega_i)$ ,  $b_i = \mathbb{P}(Y_i = \omega_i)$ , we then have

$$\begin{aligned} & d_{TV}((X_1, \dots, X_m), (Y_1, \dots, Y_m)) \\ &= \frac{1}{2} \sum_{(\omega_1, \dots, \omega_m) \in \Omega} |\mathbb{P}(X_j = \omega_j, 1 \leq j \leq m) - \mathbb{P}(Y_j = \omega_j, 1 \leq j \leq m)| \\ &= \frac{1}{2} \sum_{(\omega_1, \dots, \omega_m) \in \Omega} |a_1 \dots a_m - b_1 \dots b_m| \\ &\leq \frac{1}{2} \sum_{j=1}^m \sum_{\omega_j} |a_j - b_j| \sum_{\omega_i (i \neq j)} \prod_{i < j} a_i \prod_{i > j} b_i \\ &= \frac{1}{2} \sum_{j=1}^m \sum_{\omega_j} |a_j - b_j| \\ &= \sum_{j=1}^m d_{TV}(X_j, Y_j). \end{aligned}$$

□

### Appendix B. Basic properties of entropy

The notion of entropy plays a key role in our paper. In this appendix we record the key facts about it that we need. Proofs may be found in many places. One convenient resource is [1].

If  $X$  is a random variable taking values in a finite set then we define

$$\mathbb{H}(X) := - \sum_x \mathbb{P}(X = x) \log(\mathbb{P}(X = x)),$$

where the log is to base  $e$  and the summation runs over the range of  $X$ .

If  $\mathbf{p} = (p_1, \dots, p_n)$  is a vector of probabilities (that is, if  $p_1, \dots, p_n \geq 0$  and  $p_1 + \dots + p_n = 1$ ), then we write

$$\mathbb{H}(\mathbf{p}) := - \sum_{i=1}^n p_i \log p_i.$$

There should be no danger of confusing the two slightly different usages.

Our first lemma gives a simple upper bound for multinomial coefficients in terms of entropies.

**Lemma B.1** *Let  $n, n_1, \dots, n_k$  be non-negative integers with  $\sum n_i = n$ . Then*

$$\frac{n!}{n_1! \dots n_k!} \leq e^{\mathbb{H}(\mathbf{p})n},$$

where  $\mathbf{p} = (p_1, \dots, p_k)$  with  $p_i := n_i/n$ .



Proof. The right-hand side is  $(n/n_1)^{n_1} \dots (n/n_k)^{n_k}$ . Now simply observe that

$$\begin{aligned} & \frac{n!}{(n_1)! \dots (n_k)!} (n_1/n)^{n_1} \dots (n_k/n)^{n_k} \\ & \leq \sum_{k_1 + \dots + k_m = n} \frac{n!}{k_1! \dots k_m!} (n_1/n)^{k_1} \dots (n_k/n)^{k_m} = 1. \end{aligned} \quad \square$$

Our next lemma is a simple and well-known upper bound for the entropy.

**Lemma B.2** *Let  $X$  be a random variable taking values in a set of size  $N$ . Then  $\mathbb{H}(X) \leq \log N$ .*

*Proof* This follows immediately from the convexity of the function  $L(x) = -x \log x$  and Jensen’s inequality. See [1, Lemma 14.6.1 (i)].  $\square$

The next lemma is simple and has no doubt appeared elsewhere, but we do not know an explicit reference. In its statement, we use the notation  $\langle \mathbf{a}, \mathbf{p} \rangle = \sum_{i=1}^n a_i p_i$ .

**Lemma B.3** *Let  $\mathbf{p} = (p_1, \dots, p_n)$  be a vector of probabilities, and let  $\mathbf{a} = (a_1, \dots, a_n)$  be a vector of real numbers. Then*

$$\mathbb{H}(\mathbf{p}) + \langle \mathbf{a}, \mathbf{p} \rangle \leq \log \left( \sum_{j=1}^n e^{a_j} \right),$$

and equality occurs if and only if  $p_j = e^{a_j} / \sum_{i=1}^n e^{a_i}$  for all  $j$ .

*Proof* Let us begin by recalling that if  $t_1, \dots, t_n > 0$  are such that

$$t_1 + \dots + t_n = 1,$$

then the concavity of the logarithm implies that

$$t_1 \log x_1 + \dots + t_n \log x_n \leq \log(t_1 x_1 + \dots + t_n x_n) \tag{B.1}$$

for all  $x_1, \dots, x_n > 0$ . In addition, equality occurs in (B.1) if and only if  $x_1 = \dots = x_n$ . One may also prove this fact by induction on  $n$ , and by noticing that the case  $n = 2$  is equivalent to having  $u^t \leq tu + 1 - t$  for all  $u > 0$  and all  $t \in (0, 1)$ , with equality occurring if and only if  $u = 1$ .

Let us now prove the lemma. If  $p_j = 1$  for some  $j$ , then  $\mathbb{H}(\mathbf{p}) + \langle \mathbf{a}, \mathbf{p} \rangle = a_j$ . If  $n = 1$ , then this is equal to  $\log(\sum_{i=1}^n e^{a_i})$ , whereas if  $n \geq 2$ , then we have

$a_j < \log(\sum_{i=1}^n e^{a_i})$ , so that the lemma holds in both cases. Assume now that  $p_j \in (0, 1)$  for all  $j$ . We then have

$$\mathbb{H}(\mathbf{p}) + \langle \mathbf{a}, \mathbf{p} \rangle = \sum_{j=1}^n p_j \log(e^{a_j}/p_j).$$

We may then use (B.1) with  $t_j = p_j$  and  $x_j = e^{a_j}/p_j$  to complete the proof of the lemma.  $\square$

The next lemma, known as the *chain rule for entropy*, is nothing more than a short computation.

**Lemma B.4** *Let  $X, Y$  be random variables taking values in finite sets. Then*

$$\mathbb{H}(X, Y) = \mathbb{H}(Y) + \sum_y \mathbb{P}(Y = y) \mathbb{H}(X|Y = y).$$

*Remark* The sum over  $y$  is usually written  $\mathbb{H}(X|Y)$  and called the *conditional entropy*.

We will apply the preceding result together with the following observation.

**Lemma B.5** *Suppose that  $X, Y$  are random variables with finite ranges and that  $Y$  is a deterministic function of  $X$ . Then  $\mathbb{H}(X, Y) = \mathbb{H}(X)$ .*

*Proof* This follows from Lemma B.4 with the role of  $X$  and  $Y$  reversed, since all the entropies  $\mathbb{H}(Y|X = x)$  are zero.  $\square$

The next result, known as the submodularity property of entropy, is a crucial ingredient in our paper.

**Lemma B.6** . *Let  $X, Y, Z$  be any random variables taking values in finite sets. Then*

$$\mathbb{H}(X, Y) + \mathbb{H}(X, Z) \geq \mathbb{H}(X, Y, Z) + \mathbb{H}(X).$$

*Proof* This is [1, Lemma 14.6.1 (iv)].  $\square$

## Appendix C. Maier–Tenenbaum flags

The purpose of this appendix is to say a little more about the bound (3.12), which corresponds in the language of this paper to [22, Theorem 1.4]. Numerically, this bound is  $\tilde{\gamma}_{2^r} \gg (0.12885796477 \dots)^r$ , which is a little weaker than the bound leading to Theorem 2, which is  $\tilde{\gamma}_{2^r} \gg (0.140605674848 \dots)^r$ . What

is interesting, however, is that the flags  $\mathcal{V}$  which lead to (3.12) are completely different to the binary flags which have been the main focus of our paper. The fact that these very different flags – the “Maier–Tenenbaum flags” – lead to a result which appears to be within 10 % of optimal suggests that they will have a key role to play in any future upper bound arguments for these questions.

**Definition C.1** (*Maier–Tenenbaum flag of order  $r$* ) Let  $k = 2^r$  be a power of two. Identify  $\mathbb{Q}^k$  with  $\mathbb{Q}^{\mathcal{P}[r]}$  and define a flag  $\mathcal{V}$ ,  $\langle \mathbf{1} \rangle = V_0 \leq V_1 \leq \dots \leq V_r \leq \mathbb{Q}^{\mathcal{P}[r]}$ , as follows:  $V_i = \text{Span}(\mathbf{1}, \omega^1, \dots, \omega^i)$ , where  $\omega^i_S = 1_{i \in S}$  for  $S \subset [r]$ .

*Remark* We have  $\dim(V_i) = i + 1$  and in particular  $V_r$  is much smaller than  $\mathbb{Q}^k$ , in contrast to the situation for binary systems. We leave it to the reader to check that  $\mathcal{V}$  is nondegenerate.

Recall that  $\mathcal{V}$  gives rise to a tree structure, with the cells at level  $i$  being the intersections of cosets  $x + V_i$  with the cube  $\{0, 1\}^k$  (cf. Sect. 7.2). It is easy to check that this tree structure has a very simple form, with the cell  $\Gamma_i = V_i \cap \{0, 1\}^k$  being  $\{\mathbf{0}, \mathbf{1}, \omega^1, \mathbf{1} - \omega^1, \dots, \omega^i, \mathbf{1} - \omega^i\}$ , this dividing into three children at level  $i - 1$ ; the cell  $\Gamma_{i-1}$  together with two singletons,  $\{\omega^i\}$  and  $\{\mathbf{1} - \omega^i\}$ .

The recursive definition of the quantities  $f^C(\rho)$  (see (7.4)) therefore becomes  $f^{\Gamma_1}(\rho) = 3$ ,

$$f^{\Gamma_{j+1}}(\rho) = f^{\Gamma_j}(\rho)^{\rho_j} + 2. \tag{C.1}$$

In addition, the  $\rho$ -equations (7.5) become

$$f^{\Gamma_{j+1}}(\rho) = e(f^{\Gamma_j}(\rho))^{\rho_j}. \tag{C.2}$$

On the one hand, iterating (C.2) yields that

$$\log f^{\Gamma_j}(\rho) = \rho_1 \dots \rho_{j-1} \log 3 + \sum_{i=0}^{j-2} \rho_{j-1} \dots \rho_{j-i}$$

for all  $j \geq 1$ . On the other hand, combining (C.1) and (C.2), we find that

$$\rho_j \log f^{\Gamma_j}(\rho) = \log 2 - \log(e - 1),$$

and thus

$$\rho_1 \dots \rho_j \log 3 + \sum_{i=0}^{j-2} \rho_j \rho_{j-1} \dots \rho_{j-i} = \log 2 - \log(e - 1)$$

for all  $j \geq 1$ . Hence, we obtain the formulas

$$\rho_1 = \frac{\log 2 - \log(e - 1)}{\log 3}, \quad \rho_2 = \rho_3 = \dots = \frac{\log 2 - \log(e - 1)}{\log 2 + 1 - \log(e - 1)} =: \kappa.$$

Let us also note that the above discussion implies that

$$\log f^{\Gamma_j}(\rho) = \frac{\log 2 - \log(e - 1)}{\rho_j} = \begin{cases} \log 3 & \text{if } j = 1, \\ \log 2 - \log(e - 1) + 1 & \text{if } j \geq 2. \end{cases} \tag{C.3}$$

Now, assuming that the conditions of Proposition 7.7 hold, we therefore have

$$\begin{aligned} \gamma_k^{\text{res}}(\mathcal{V}) &= (\log 3 - 1) / \left( \log 3 + \frac{1}{\rho_1} \left( 1 + \frac{1}{\kappa} + \dots + \frac{1}{\kappa^{r-2}} \right) \right) \\ &= \left( 1 - \frac{1}{\log 3} \right) \kappa^{r-1}. \end{aligned}$$

Now it can be shown by explicit calculation that the conditions of Proposition 7.7 do hold. The optimal measures  $\mu_i^*$  are all induced from the measure  $\mu^*$  in which

$$\mu^*(\omega^j) = \mu^*(\mathbf{1} - \omega^j) = \mu^*(\Gamma_j) \cdot \frac{1}{f^{\Gamma_j}(\rho)} = \begin{cases} \frac{1}{3} e^{1-r} & \text{if } j = 1, \\ \frac{e-1}{2e} e^{j-r} & \text{if } j \geq 2. \end{cases}$$

In addition, we have

$$\mu^*(\mathbf{0}) = \mu^*(\mathbf{1}) = \frac{\mu^*(\Gamma_0)}{2} = \frac{1}{6} e^{1-r}.$$

We may then prove by a slightly lengthy computation whose details we leave to the reader that the optimal parameters  $\mathbf{c}^*$  are given by

$$c_1^* = 1, \quad c_j^* = \frac{1}{\kappa^2} \left( \frac{e - \kappa}{e - 1} \right) \left( 1 - \frac{1}{\log 3} \right) \kappa^j, \quad c_{r+1}^* = \left( 1 - \frac{1}{\log 3} \right) \kappa^{r-1}.$$

It can also be shown that  $\gamma_k^{\text{res}}(\mathcal{V}) = \gamma_k(\mathcal{V})$ , by showing that the full entropy condition (3.6) follows from the restricted conditions (7.11). This is a little involved, but a fairly direct inductive argument can be made to work and this

is certainly less subtle than the arguments of Sect. 8. In this way one may establish the bound

$$\gamma_{2^r} \geq \left(1 - \frac{1}{\log 3}\right) \left(\frac{\log 2 - \log(e-1)}{\log 2 + 1 - \log(e-1)}\right)^{r-1} \gg (0.131810543\dots)^r. \tag{C.4}$$

Finally, a relatively routine perturbative argument yields the same bound for  $\tilde{\gamma}_{2^r}$ .

It will be noted that (C.4) is strictly stronger than (3.12), the bound obtained in [22]. This is because, in essence, Maier and Tenenbaum chose slightly suboptimal measures and parameters on the system  $\mathcal{V}$ , roughly corresponding to  $\mu(\omega^j) \sim 3^{j-r-1}$ , which then leads to  $c_j \sim \left(\frac{1-1/\log 3}{1-1/\log 27}\right)^j$ .

### References

1. Alon, N., Spencer, J.H.: The Probabilistic Method. Wiley Series in Discrete Mathematics and Optimization, fourth edn. John Wiley & Sons Inc, Hoboken, NJ (2016)
2. Arratia, R., Barbour, A.D., Tavaré, S.: On random polynomials over finite fields. In: Mathematical Proceedings of the Cambridge Philosophical Society, vol. 114, pp. 347–368 (1993)
3. Arratia, R., Tavaré, S.: The cycle structure of random permutations. *Ann. Probab.* **20**, 1567–1591 (1992)
4. Elliott, P.D.T.A.: Probabilistic Number Theory. I. Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Science], vol. 239. Springer-Verlag, New York (1979). (Mean-value theorems)
5. Elliott, P.D.T.A.: Probabilistic Number Theory. II. Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], vol. 240. Springer-Verlag, New York (1980). (Central limit theorems)
6. Erdős, P., Hall, R.R.: The propinquity of divisors. *Bull. Lond. Math. Soc.* **11**, 304–307 (1979)
7. Erdős, P., Nicolas, J.-L.: Répartition des nombres superabondants. *Bull. Soc. Math. Fr.* **103**, 65–90 (1975)
8. Erdős, P., Nicolas, J.-L.: Méthodes probabilistes et combinatoires en théorie des nombres. *Bull. Sci. Math. (2)* **100**, 301–320 (1976)
9. Erdős, P.: On the density of some sequences of integers. *Bull. Am. Math. Soc.* **54**, 685–692 (1948)
10. Erdős, P.: On some applications of probability to analysis and number theory. *J. Lond. Math. Soc.* **39**, 692–696 (1964)
11. Ford, K.: Joint Poisson distribution of prime factors in sets. *Math. Proc. Camb. Philos. Soc.* **173**, 189–200 (2022)
12. Hall, R.R., Tenenbaum, G.: On the average and normal orders of Hooley’s  $\Delta$ -function. *J. Lond. Math. Soc. (2)* **25**, 392–406 (1982)
13. Hall, R.R., Tenenbaum, G.: The average orders of Hooley’s  $\Delta_r$ -functions. *Mathematika* **31**, 98–109 (1984)
14. Hall, R.R., Tenenbaum, G.: The average orders of Hooley’s  $\Delta_r$ -functions. II. *Compos. Math.* **60**, 163–186 (1986)
15. Hall, R.R., Tenenbaum, G.: Divisors. Cambridge Tracts in Mathematics, vol. 90. Cambridge University Press, Cambridge (1988)

16. Hooley, C.: On a new technique and its applications to the theory of numbers. *Proc. Lond. Math. Soc.* (3) **38**, 115–151 (1979)
17. Koukoulopoulos, D.: Localized factorizations of integers. *Proc. Lond. Math. Soc.* **101**, 392–426 (2010)
18. Koukoulopoulos, D.: On the number of integers in a generalized multiplication table. *J. Reine Angew. Math.* **689**, 33–99 (2014)
19. Levin, D.A., Peres, Y., Wilmer, E.L.: *Markov Chains and Mixing Times*. American Mathematical Society, Providence, RI (2009). With a chapter by James G. Propp and David B. Wilson
20. Maier, H., Tenenbaum, G.: On the set of divisors of an integer. *Invent. Math.* **76**, 121–128 (1984)
21. Maier, H., Tenenbaum, G.: On the normal concentration of divisors. *J. Lond. Math. Soc.* (2) **31**, 393–400 (1985)
22. Maier, H., Tenenbaum, G.: On the normal concentration of divisors. II. *Math. Proc. Camb. Philos. Soc.* **147**, 513–540 (2009)
23. Tenenbaum, G.: Sur la concentration moyenne des diviseurs. *Comment. Math. Helv.* **60**, 411–428 (1985)
24. Tenenbaum, G.: Fonctions  $\Delta$  de Hooley et applications. In: *Séminaire de théorie des nombres, Paris 1984–85*. Progress in Mathematics, vol. 63, pp. 225–239. Birkhäuser, Boston, MA (1986)
25. Tenenbaum, G.: Crible d’ératosthène et modèle de Kubilius. In: *Number Theory in Progress (Zakopane-Kościelisko, 1997)*, vol. 2, pp. 1099–1129. de Gruyter, Berlin (1999)
26. Tenenbaum, G.: Some of Erdős’ unconventional problems in number theory, thirty-four years later. In: *Erdős centennial. Bolyai Society Mathematical Studies*, vol. 25, pp. 651–681. János Bolyai Mathematical Society, Budapest (2013)

**Publisher’s Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.