

AN EXTENSION OF THE ERDŐS–TETALI THEOREM

CHRISTIAN TÁFULA

ABSTRACT. Given a sequence $\mathcal{A} = \{a_0 < a_1 < a_2 \dots\} \subseteq \mathbb{N}$, let $r_{\mathcal{A},h}(n)$ denote the number of ways n can be written as the sum of h elements of \mathcal{A} . Fixing $h \geq 2$, we show that if f is a suitable real function (namely: locally integrable, O -regularly varying and of positive increase) satisfying

$$x^{1/h} \log(x)^{1/h} \ll f(x) \ll \frac{x^{1/(h-1)}}{\log(x)^\varepsilon} \text{ for some } \varepsilon > 0,$$

then there must exist $\mathcal{A} \subseteq \mathbb{N}$ with $|\mathcal{A} \cap [0, x]| = \Theta(f(x))$ for which $r_{\mathcal{A},h+\ell}(n) = \Theta(f(n)^{h+\ell}/n)$ for all $\ell \geq 0$. Furthermore, for $h = 2$ this condition can be weakened to $x^{1/2} \log(x)^{1/2} \ll f(x) \ll x$.

The proof is somewhat technical and the methods rely on ideas from regular variation theory, which are presented in an appendix with a view towards the general theory of additive bases. We also mention an application of these ideas to Schnirelmann's method.

1. INTRODUCTION

Denote by \mathbb{N} the set of natural numbers with 0. Given a sequence $\mathcal{A} = \{a_0 < a_1 < a_2 \dots\} \subseteq \mathbb{N}$ and an integer $h \geq 2$, the h -fold sumset $h\mathcal{A}$ is the set $\{\sum_{i=1}^h k_i : k_1, \dots, k_h \in \mathcal{A}\}$. We say that \mathcal{A} is an *additive basis* when there is $h \geq 2$ such that $\mathbb{N} \setminus h\mathcal{A}$ is empty, and the least such h is the *order* of \mathcal{A} . An h -basis is an additive basis of order h . The *representation functions* $r_{\mathcal{A},h}(n)$ and $s_{\mathcal{A},h}(x)$ count the number of solutions of $k_1 + k_2 + \dots + k_h = n$ and $k_1 + k_2 + \dots + k_h \leq x$ resp., where $h \geq 1$ is fixed and $k_i \in \mathcal{A}$, considering permutations in the sense of the formal series:

$$(1.1) \quad \left(\sum_{a \in \mathcal{A}} z^a \right)^h = \sum_{n \geq 0} r_{\mathcal{A},h}(n) z^n, \quad \frac{(\sum_{a \in \mathcal{A}} z^a)^h}{1 - z} = \sum_{n \geq 0} s_{\mathcal{A},h}(n) z^n.$$

When $h = 1$, we denote $r_{\mathcal{A},1}(n)$, $s_{\mathcal{A},1}(n)$ by $\mathbb{1}_{\mathcal{A}}(n)$ and $A(n)$ resp..

An h -basis is said to be *economical*¹ when $r_{\mathcal{A},h}(n) \ll_\varepsilon n^\varepsilon$, that is, $r_{\mathcal{A},h}(n) = O(n^\varepsilon)$ for every $\varepsilon > 0$. The study of such objects dates back to S. Sidon, who in the 1930s inquired about the existence of economical 2-basis (cf. Erdős [4]). A positive answer was given by Erdős in the 1950s by means of probabilistic (therefore non-constructive) methods, showing the existence of $\mathcal{A} \subseteq \mathbb{N}$ with $r_{\mathcal{A},2}(n) = \Theta(\log(n))$.

2010 *Mathematics Subject Classification*. Primary 11B13, 11B34; Secondary 05D40.

Key words and phrases. Erdős–Tetali theorem, economical bases, probabilistic method, representation functions, O -regular variation.

¹There appears to be conflicting descriptions of *economical* and *thin* h -bases in the literature. We believe that the most satisfactory definitions are the following: an h -basis $\mathcal{A} \subseteq \mathbb{N}$ is *thin* when $A(x) = \Theta(x^{1/h})$ (cf. Nathanson [15]), and *economical* when $r_{\mathcal{A},h}(n) \ll_\varepsilon n^\varepsilon$ (cf. Chapter III, p. 111 of Halberstam & Roth [10]). Note that not all economical bases are thin, and, a priori, thin bases are not necessarily economical. In spite of that, economical h -basis always satisfy $A(x) = x^{1/h+o(1)}$.

A constructive proof was later found in the 1990s (cf. Kolountzakis [12]). Also in the 1990s, Erdős and Tetali [6] settled the $h \geq 3$ case, showing for every such h the existence of an h -basis $\mathcal{A} \subseteq \mathbb{N}$ with $r_{\mathcal{A},h}(n) = \Theta(\log(n))$. The strategy behind their proof can be outlined in three main steps:

I. Define a random sequence $\omega \subseteq \mathbb{N}$ by

$$\Pr(n \in \omega) = \frac{f(n)}{n+1},$$

where $f(x) := Mx^{1/h} \log(x)^{1/h}$ for some large constant $M > 1$;

II. Show that the expected value of $r_{\omega,h}(n)$ is $\Theta(f(n)^h/n)$;

III. Show that $r_{\omega,h}(n)$ concentrates around its mean. That is,

$$\Pr(\exists c_1, c_2 > 0 : \forall n \in \mathbb{N}, c_1 \mathbb{E}(r_{\omega,h}(n)) \leq r_{\omega,h}(n) \leq c_2 \mathbb{E}(r_{\omega,h}(n))) = 1.$$

The third step is by far the most involved. The case $h = 2$ is rather special, for $r_{\omega,2}(n)$ can be written as a sum of independent Bernoulli trials, hence the third step may be tackled by using Chernoff-type bounds.² This is not the case for $h \geq 3$. The way Erdős and Tetali worked around this limitation was by showing that the events counted by $r_{\omega,h}(n)$ have low correlation. Vu's alternative proof in Section 5 of [20] uses the same principle, albeit with a much more powerful two-sided concentration inequality. In this paper we take each of these aforementioned steps and extend their reach, being relatively faithful to the methods from Erdős & Tetali [6]. Our studies culminate in the following:

Main Theorem. *Let $h \geq 2$ be fixed and $x_0 \in \mathbb{R}_+$ arbitrarily large. For any locally integrable, positive real function $f : [x_0, +\infty) \rightarrow \mathbb{R}_+$ such that*

$$(i) \int_{x_0}^x \frac{f(t)}{t} dt = \Theta(f(x)) ,$$

$$(ii) x^{1/h} \log(x)^{1/h} \ll f(x) \ll \frac{x^{1/(h-1)}}{\log(x)^\varepsilon} \text{ for some } \varepsilon > 0;$$

there is an h -basis \mathcal{A} with $A(x) = \Theta(f(x))$ such that

$$r_{\mathcal{A},h+\ell}(n) = \Theta\left(\frac{f(n)^{h+\ell}}{n}\right), \text{ for every } \ell \geq 0.$$

Furthermore, when $h = 2$ the " $\log(n)^\varepsilon$ " term can be dropped.

One way of thinking about the condition (ii) is that while $x^{1/h} \log(x)^{1/h}$ works as a natural lower bound for the methods used by Erdős and Tetali,³ $x^{1/(h-1)}$ is a natural upper bound. When $h \geq 3$, however, the term $\log(n)^\varepsilon$ pops up because $r_{\mathcal{B},h-1}(n)$ is not necessarily bounded. These thoughts shall be made clear in our study of Step III, which constitutes Section 5. Condition (i) relates to Steps I and II. In Appendix A we show that when $f(x) = A(x)$ for some $\mathcal{A} \subseteq \mathbb{N}$, this condition is satisfied if and only if \mathcal{A} is an *O-regular plus* (OR_+) *sequence*, which are sequences such that $A(2x) = O(A(x))$ and $a_{2n} = O(a_n)$. This appendix is used to explain the idea behind the regularity conditions we assume on sequences in order to achieve

²As in Section 8.6, p. 139 of Alon & Spencer [1].

³So much so that Erdős [4] inquired on the existence of 2-bases \mathcal{B} with $r_{\mathcal{B},2}(n) \neq o(\log(n))$, a question that naturally extends to h -bases in general and is still open.

our result, with a view towards the general theory of sequences and additive bases, and some applications are given. In Section 4 we present a natural way of using an OR_+ sequence \mathcal{A} to induce a probability measure in the space of sequences so that the expected value of $r_{\omega,h}(n)$ is $\Theta(A(n)^h/n)$ for all $h \geq 1$. We will argue that these sequences are the only kind for which one can draw such conclusion.

Notation. Our use of the asymptotic notations $\Theta, \asymp, O, \ll, o, \sim$ is standard, as well as the “floor” and “ceiling” functions $\lfloor \cdot \rfloor, \lceil \cdot \rceil$. Unless otherwise specified, asymptotic symbols are used assuming $n \rightarrow +\infty$ through \mathbb{N} and $x \rightarrow +\infty$ through \mathbb{R} . When we say “ $P(n)$ holds for all large n ” we mean that there is $n_0 \in \mathbb{N}$ such that $P(n)$ holds for all $n \geq n_0$. Whenever we define a positive real function $f : [x_0, +\infty) \rightarrow \mathbb{R}_+$, the $x_0 \in \mathbb{R}_+$ is to be assumed arbitrarily large; this is just to accommodate functions such as $\log \log \log(x)/\log \log(x)$ without worrying about small x .

We use \Pr for the probability measure, \mathbb{E} for expectation and Var for variance. Given a probability space $(\Omega, \mathcal{F}, \Pr)$, we denote the indicator function of an event $E \in \mathcal{F}$ by \mathcal{I}_E . A random variable (abbreviated r.v.) $X : \Omega \rightarrow \mathbb{R}$ is said to be a *Bernoulli trial* when $X(\Omega) \subseteq \{0, 1\}$. The conditional expectation of a r.v. X given an event $E \in \mathcal{F}$ is denoted by $\mathbb{E}(X|E) := \Pr(E)^{-1} \cdot \int_E X \, d\Pr$.

Finally, whenever we write an asymptotic sign with a superscripted “a.s.” we mean that the limit in the definition of that sign holds *almost surely*, i.e. for a subset of Ω with complement having measure 0. Please refer to the beginning of Appendix A for the notation regarding *sequences*.

2. PRELIMINARIES

This section gathers basic preliminary concepts and results that will be used in our arguments; for this reason, everything is unnumbered. The reader may skim through it first and come back as necessary.

2.1. Probabilistic tools. Consider a probability space $(\Omega, \mathcal{F}, \Pr)$. Given a collection \mathcal{X} of r.v.s, denote by $\sigma(\mathcal{X})$ the σ -algebra generated by \mathcal{X} , which is the smallest σ -algebra contained in \mathcal{F} such that all r.v.s $X \in \mathcal{X}$ are measurable. Given a sequence $(X_n)_{n \geq 1}$ of r.v.s, consider an event

$$E \in \bigcap_{N \geq 1} \sigma((X_n)_{n \geq N}).$$

When the sequence consists of mutually independent r.v.s, this event is independent of any finite subsequence and is called a *tail event* for the sequence. Such events have either probability 0 or 1.⁴ These observations will be relevant for our discussion in Section 3. In what follows, we present the three essential lemmas from probability theory that shall be used in our considerations.

Disjointness lemma (Lemma 8.4.1, p. 135 of Alon & Spencer [1]). *Let $\mathcal{E} = \{E_i : i \in I\}$ be a (not necessarily finite) family of events and define $S := \sum_{E \in \mathcal{E}} \mathcal{I}_E$. If*

⁴Kolmogorov’s zero–one law (Theorem 2 in Section IV.6, p. 124 of Feller [8]).

$\mathbb{E}(S) < +\infty$, then for all $k \geq 1$ the following holds:

$$\Pr(\mathcal{D} \subseteq \mathcal{E} \text{ disfam}^5 : |\mathcal{D}| = k) \leq \sum_{\substack{\mathcal{J} : \mathcal{J} \subseteq \mathcal{E} \\ \text{disfam}, |\mathcal{J}|=k}} \Pr\left(\bigwedge_{E \in \mathcal{J}} E\right) \leq \frac{\mathbb{E}(S)^k}{k!}.$$

Chernoff bounds (Theorem 1.8, p. 11 of Tao & Vu [18]). *Let x_1, \dots, x_n be mutually independent r.v.s with $|x_i| \leq \lambda$, $\forall i \leq n$ for some fixed constant $\lambda > 0$. Taking $S := x_1 + \dots + x_n$, if $\mathbb{E}(S) > 0$ then, for every $\varepsilon > 0$,*

$$\Pr\left(\left|\frac{S}{\mathbb{E}(S)} - 1\right| \geq \varepsilon\right) \leq 2e^{-\min\{\varepsilon/2, (\varepsilon/2)^2\} \cdot \mathbb{E}(S)/\lambda}.$$

Correlation inequality (Boppana & Spencer [3]). *Suppose $\Omega = \{\omega_1, \omega_2, \dots\}$ is finite and let $\mathcal{R} \subseteq \Omega$ be a random subset given by $\Pr(\omega_k \in \mathcal{R}) = p_k$, these events being mutually independent. Let $\mathfrak{s}_1, \dots, \mathfrak{s}_n \subseteq \Omega$ be different subsets of Ω and, respectively, let E_1, \dots, E_n denote the events “ $\mathfrak{s}_i \subseteq \mathcal{R}$ ”. Furthermore, assume $\Pr(E_i) \leq 1/2$ for $1 \leq i \leq n$. Then:*

$$\prod_{i=1}^n \Pr(\bar{E}_i) \leq \Pr\left(\bigwedge_{i=1}^n \bar{E}_i\right) \leq \left(\prod_{i=1}^n \Pr(\bar{E}_i)\right) e^{2\Delta},$$

where

$$\Delta = \sum_{\substack{1 \leq i < j \leq n \\ E_i \cap E_j \neq \emptyset}} \Pr(E_i \wedge E_j)$$

and \bar{E}_i is the complement of E_i .

It is worth noting that, in Section 5, almost all the of techniques used may be broken down into several instances of the disjointness lemma followed by an application of the Borel–Cantelli lemma.⁶

2.2. Exact representation functions. Fix an integer $h \geq 2$. Let $n \in \mathbb{N}$ and consider an arbitrary sequence $\mathcal{A} \subseteq \mathbb{N}$. Denote by

$$R_h(n; \mathcal{A}) := \{(k_1, \dots, k_h) \in \mathcal{A}^h : k_1 + \dots + k_h = n\}$$

the set of h -representations of n in \mathcal{A} , writing just $R_h(n)$ when $\mathcal{A} = \mathbb{N}$. We say an h -representation $\mathfrak{R} = (k_1, \dots, k_h) \in R_h(n; \mathcal{A})$ is *exact* when $k_i \neq k_j$ whenever $i \neq j$. Denote the set of exact h -representations by

$$\text{ER}_h(n; \mathcal{A}) := \{\mathfrak{R} \in R_h(n; \mathcal{A}) : \mathfrak{R} \text{ is exact}\} / \mathfrak{S}_h,$$

where \mathfrak{S}_h is the symmetric group in h symbols; that is, we consider two exact representations to be equal when one can be obtained by a permutation of the other, hence $\mathfrak{R} \in \text{ER}_h(n; \mathcal{A})$ can be thought of as a set $\{k_1, \dots, k_h\}$ rather than as an h -tuple. Similarly, write simply $\text{ER}_h(n)$ when $\mathcal{A} = \mathbb{N}$. Two h -representations $\mathfrak{R}, \mathfrak{R}' \in R_h(n; \mathcal{A})$ are *disjoint* when $\mathfrak{R} \cap \mathfrak{R}' = \emptyset$. Thus we define two auxiliary representation functions.

⁵A *disjoint family* (abbreviated *disfam*) is a mutually independent subfamily of \mathcal{E} .

⁶cf. Section VIII.3, p. 200 of Feller [7].

- The *exact representation function*:

$$\rho_{\mathcal{A},h}(n) := |\text{ER}_h(n; \mathcal{A})|;$$

- The *maxdisfam exact representation function*:

$$\widehat{\rho}_{\mathcal{A},h}(n) := \max \left\{ |\mathcal{C}| : \mathcal{C} \subseteq \text{ER}_h(n; \mathcal{A}) \text{ is a maximal disjoint family of exact } h\text{-representations.} \right\}.$$

While $r_{\mathcal{A},h}$ is more convenient for induction purposes, $\rho_{\mathcal{A},h}$ is easier to work with in probabilistic settings. This is because on the one hand, one has equations (A.1), while on the other, $\rho_{\mathcal{A},h}$ may be written as

$$\rho_{\mathcal{A},h}(n) = \sum_{\substack{k_1 + \dots + k_h = n \\ k_1 < \dots < k_h}} \mathbb{1}_{\mathcal{A}}(k_1) \cdots \mathbb{1}_{\mathcal{A}}(k_h).$$

The $\widehat{\rho}_{\mathcal{A},h}$ function is more abstract. It is useful in probabilistic settings when accompanied by the disjointness lemma. If $\mathcal{C}_n \subseteq \text{ER}_h(n; \mathcal{A})$ is a maximal disjoint family of exact h -representations, then

$$\rho_{\mathcal{A},h}(n) \leq \sum_{k \in \bigcup \mathcal{C}_n} \rho_{\mathcal{A},h-1}(n-k) \leq h \cdot \widehat{\rho}_{\mathcal{A},h}(n) \cdot \left(\max_{k \in \bigcup \mathcal{C}_n} \rho_{\mathcal{A},h-1}(n-k) \right),$$

for $|\bigcup \mathcal{C}_n| \leq h \cdot \widehat{\rho}_{\mathcal{A},h}(n)$. This allows us to derive some rough upper bounds for $\rho_{\mathcal{A},h}$ by induction, but is a fairly limited method since $\widehat{\rho}_{\mathcal{A},h}(n) \leq n/h$.

Remark. The reader familiar with hypergraphs may interpret these functions as counting hyperedges in $\mathcal{H}_{\mathcal{A},h}^{(n)} = (\mathcal{V}, \mathcal{E})$, where $\mathcal{V} := \mathcal{A} \cap \{0, 1, \dots, n\}$ and $\mathcal{E} := \text{ER}_h(n; \mathcal{A})$. Thus $\widehat{\rho}_{\mathcal{A},h}(n)$ is the cardinality of a maximum matching in $\mathcal{H}_{\mathcal{A},h}^{(n)}$, which when used with the maximum degree $\Delta_1(\mathcal{H}) := \max_{u \in \mathcal{V}} |\{e \in \mathcal{E} : u \in e\}|$ provides some rough upper bounds for $|\mathcal{E}|$. This is a pictorial analogy to keep in mind, but we will not need to make explicit use of it in this paper. For a closely related problem which requires extensive use of this graph-theoretical language, refer to Warnke [21].

Another very useful restriction of the set of representations concerns *lower bounded representations*. Given $x_0 \in \mathbb{R}_+$, consider the partitions

$$\begin{aligned} \text{R}_h(n; \mathcal{A})|_{\geq x_0} &:= \{\mathfrak{R} \in \text{R}_h(n; \mathcal{A}) : \forall k \in \mathfrak{R}, k \geq x_0\}, \\ \text{R}_h(n; \mathcal{A})\mathfrak{b}_{x_0} &:= \{\mathfrak{R} \in \text{R}_h(n; \mathcal{A}) : \exists k \in \mathfrak{R}, k < x_0\}. \end{aligned}$$

This is distinguishing those h -representations for which the smallest term is at least as large as x_0 from those for which some of its elements are smaller than it. In view of what has been discussed, $\text{ER}_h(n; \mathcal{A})|_{\geq x_0}$, $\text{ER}_h(n; \mathcal{A})\mathfrak{b}_{x_0}$, $\rho_{\mathcal{A},h}(n)|_{\geq x_0}$, $\rho_{\mathcal{A},h}(n)\mathfrak{b}_{x_0}$, $\widehat{\rho}_{\mathcal{A},h}(n)|_{\geq x_0}$ and $\widehat{\rho}_{\mathcal{A},h}(n)\mathfrak{b}_{x_0}$ are defined accordingly.

2.3. Regular variation. Fundamental to the calculations done in this paper are the notions of O -regular variation (Subsection A.2) and Matuszewska indices ((A.2) in Subsection A.3). These are introduced in Appendix A, with a view towards the theory of *sequences* (in the sense of Halberstam & Roth [10]) and the general study of additive bases. The concepts necessary for the proof of the Main Theorem are the definitions of OR, PI and OR₊ sequences (Propositions A.2, A.5 and A.8, resp.), as well as what we called *the OR–PI lemma* (Lemma A.7).

3. THE SPACE OF SEQUENCES (STEP I)

Denote by $\mathcal{S} := \{\mathcal{A} \subseteq \mathbb{N} : \mathcal{A} \text{ is infinite}\}$ the *space of integer sequences*. Say a real sequence $(\alpha_n)_{n \geq 0}$ is a *sequence of probabilities* when $0 \leq \alpha_n \leq 1$ for all $n \in \mathbb{N}$, and say it is *proper* when $\sum_{n \geq 0} \alpha_n$ diverges. A proper sequence of probabilities induces a probability measure on \mathcal{S} ; we denote by $\langle \mathcal{S}, (\alpha_n)_n \rangle$ the space $(\mathcal{S}, \mathcal{F}, \text{Pr})$ that satisfies the following:

- (i) The events $E_n := \{\omega : n \in \omega\}$ ⁷ are measurable and $\text{Pr}(E_n) = \alpha_n$;
- (ii) $\{E_0, E_1, E_2, \dots\}$ is a collection of mutually independent events;
- (iii) \mathcal{F} is the σ -algebra induced by the collection $\{E_0, E_1, E_2, \dots\}$.

The requirement for $(\alpha_n)_{n \geq 0}$ to be proper is a necessary and sufficient condition to ensure that this is a well-defined probability space according to the Borel–Cantelli lemmas. Indeed, once infinite Cartesian products of probability spaces are established,⁸ consider a sequence of independent Bernoulli trials with probability α_n of success, then remove those events in which only a finite number of successes occur and identify it with $\sigma((\mathbb{1}_\omega(n))_{n \geq 0})$.

Remark. As a side note, the reason we chose to work with \mathcal{S} instead of the entire power set of \mathbb{N} is that we will only deal with infinite sequences. It also has the nice one-to-one correspondence:

$$\begin{aligned} \mathcal{S} &\longrightarrow \mathbb{R}/\mathbb{Z} \\ \mathcal{A} &\longmapsto \left[\sum_{a \in \mathcal{A}} 2^{-(a+1)} \right] \end{aligned}$$

thus probability measures on \mathcal{S} correspond naturally to certain measures on the unit circle. The only $\langle \mathcal{S}, (\alpha_n)_n \rangle$ corresponding to a translation-invariant measure in the circle is the one with $\alpha_n = 1/2$ for all $n \in \mathbb{N}$, which coincides with the Borel measure on \mathbb{R}/\mathbb{Z} .

Instances of this construction appear in numerous works, most notably in Erdős & Rényi [5]. For a classical treatment on related combinatorial and number-theoretical applications, see Chapter III of Halberstam & Roth [10]. Essentially, the idea behind these probability measures is to study sequences with certain prescribed rates of growth. Indeed, it is a consequence of Kolmogorov’s three-series theorem (cf. Section IX.9 of Feller [8]) that if $(X_n)_{n \geq 0}$ is a sequence of independent r.v.s with $\mathbb{E}(X_n) = 0$ and $\sum_{n \geq 0} \text{Var}(X_n) < +\infty$, then $\sum_{n \geq 0} X_n$ converges almost surely. The idea of prescribing rates of growth has then the precise meaning implied by the following variant of the strong law of large numbers.⁹

Theorem 3.1 (Strong law). *If $(\alpha_n)_{n \geq 0}$ is a proper sequence of probabilities, then in $\langle \mathcal{S}, (\alpha_n)_n \rangle$ the following holds:*

$$W(x) \stackrel{\text{a.s.}}{\sim} \mathbb{E}(W(x)),$$

⁷Here and throughout this paper we denote by ω ($\omega \in \mathcal{S}$) a random sequence following the probability distribution induced by $(\alpha_n)_{n \geq 0}$.

⁸This may be found in several sources, most notably in Halmos [11] (cf. Section 38).

⁹This is a well-known variant of the strong law, as it is mentioned in Erdős & Rényi [5]. A similar statement is present in Chapter III, §11 of Halberstam & Roth [10], but with some unnecessary additional hypotheses. Given the fundamental nature of this result in relation to $\langle \mathcal{S}, (\alpha_n)_n \rangle$, we offer it a short proof just for the sake of completeness.

where $W(x) := |\omega \cap [0, x]|$.

Proof. First, notice that $\mathbb{E}(W(x)) = \sum_{n \leq x} \alpha_n$. Letting n_0 be the smallest number for which $\alpha_{n_0} > 0$, define the r.v.s

$$X_n := \frac{\mathbb{1}_\omega(n) - \alpha_n}{\sum_{k \leq n} \alpha_k},$$

for all $n \geq n_0$, with $X_n \equiv 0$ for $n < n_0$. Since these are just linear transformations being applied to each $\mathbb{1}_\omega(n)$, the r.v.s X_n are still mutually independent. By routine calculations, it follows that $\mathbb{E}(X_n) = 0$ and $\text{Var}(X_n) = (\alpha_n - \alpha_n^2) / (\sum_{k \leq n} \alpha_k)^2$. Note that

$$\frac{\alpha_n - \alpha_n^2}{(\sum_{k \leq n} \alpha_k)^2} \leq \frac{1}{\sum_{k < n} \alpha_k} - \frac{1}{\sum_{k \leq n} \alpha_k},$$

hence we may telescope the sum $\sum_{n \geq n_0} \text{Var}(X_n)$ and thus it converges.

By the aforementioned theorem of Kolmogorov, it follows that $\sum_{n \geq 0} X_n$ converges almost surely to, say, $C \geq 0$. Then, for each $n \geq -1$, define $\varepsilon_n := C - \sum_{k \leq n} X_k$. Note that $X_k = \varepsilon_{k-1} - \varepsilon_k$, thus by partial summation we get

$$\begin{aligned} W(x) - \sum_{n \leq x} \alpha_n &= \sum_{n \leq x} X_n \mathbb{E}(W(n)) \\ &= \sum_{n \leq x} \varepsilon_n \alpha_{n+1} + C \cdot \mathbb{E}(W(0)) - \varepsilon_{\lfloor x \rfloor} \cdot \mathbb{E}(W(x)). \end{aligned}$$

Since $\varepsilon_n \stackrel{\text{a.s.}}{=} o(1)$, the last term in the RHS is a.s. $o(\mathbb{E}(W(x)))$. Recall that $\sum_{n \leq x} \alpha_n$ diverges, thus the first term is also a.s. $o(\mathbb{E}(W(x)))$, hence the theorem is proved. \square

From the many types of tail events in $\langle \mathcal{S}, (\alpha_n)_n \rangle$ one then might study, we focus on the behavior of the r.v.s $r_{\omega, h}(n)$ and $s_{\omega, h}(n)$.

4. THE FUNDAMENTAL LEMMA (STEP II)

Recall that our goal in this step is to describe $\mathbb{E}(r_{\omega, h}(n))$ in a simple way. If we want to have $|\omega \cap [0, x]| \stackrel{\text{a.s.}}{=} \Theta(f(x))$ in $\langle \mathcal{S}, (\alpha_n)_n \rangle$ for some f , by the strong law it is necessary and sufficient that $\sum_{n \leq x} \alpha_n = \Theta(f(x))$. If we assume f to be differentiable, one could then choose α_n to be something like $f'(n)$, so that everything is expressed in terms of f . Alternatively, in view of the OR–PI lemma (Lemma A.7), we could just assume f to be O -regularly varying and having positive increase, so that α_n can be chosen to be something like $f(n)/n$. This is actually preferable, for it can be formulated in a very clean way:

Definition (The space $\mathcal{S}_{\mathcal{A}}$). When \mathcal{A} is an OR_+ sequence, let $\mathcal{S}_{\mathcal{A}}$ denote the space $\langle \mathcal{S}, (\alpha_n)_n \rangle$ with

$$\alpha_n := \frac{A(n)}{n+1}.$$

This definition only makes sense when \mathcal{A} is OR_+ (Proposition A.8),¹⁰ for it is the only case in which this construction yields $|\omega \cap [0, x]| \stackrel{\text{a.s.}}{\asymp} A(x)$. Not only that, but

¹⁰Hence whenever we write $\mathcal{S}_{\mathcal{A}}$ in this paper, \mathcal{A} is being assumed OR_+ .

since \mathcal{A} is, in particular, OR (Proposition A.2), it follows that $s_{\omega,h}(x) \stackrel{\text{a.s.}}{\asymp} A(x)^h$ for all $h \geq 1$. If one has a suitable function f and still wants $\Theta(f(x))$ as the prescribed rate of growth, it is then sufficient to obtain a sequence $\mathcal{A} \subseteq \mathbb{N}$ with $A(x) = \Theta(f(x))$. Moreover, in this setting, the expected result holds. We first need the following lemma.

Lemma 4.1. *Let $\mathcal{A} \subseteq \mathbb{N}$, $h \geq 1$, $n \in \mathbb{N}$ and $k \geq 0$. Then*

$$r_{\mathcal{A},h}(n) \leq r_{\mathcal{A} \cup \{k\},h}(n) \leq r_{\mathcal{A},h}(n) + \binom{h}{\lfloor h/2 \rfloor} \sum_{\ell=1}^{h-1} r_{\mathcal{A},h-\ell}(n - \ell k) + \delta_{n/h}(k),$$

where $\delta_{n/h}(k) = 1$ if $k = n/h$ and 0 otherwise.

Proof. The LHS inequality is obvious. For the other side we have:

$$\begin{aligned} r_{\mathcal{A} \cup \{k\},h}(n) &= \sum_{k_1+k_2+\dots+k_h=n} \mathbb{1}_{\mathcal{A} \cup \{k\}}(k_1) \mathbb{1}_{\mathcal{A} \cup \{k\}}(k_2) \dots \mathbb{1}_{\mathcal{A} \cup \{k\}}(k_h) \\ &= \sum_{k_1+k_2+\dots+k_h=n} \mathbb{1}_{\mathcal{A}}(k_1) \mathbb{1}_{\mathcal{A}}(k_2) \dots \mathbb{1}_{\mathcal{A}}(k_h) + \sum_{\substack{k_1+k_2+\dots+k_h=n \\ k \in \{k_1, \dots, k_h\} \\ k_j \neq k \iff k_j \in \mathcal{A}}} 1. \end{aligned}$$

Thus, assuming $k \notin \mathcal{A}$,

$$r_{\mathcal{A} \cup \{k\},h}(n) = r_{\mathcal{A},h}(n) + \sum_{\ell=1}^{h-1} \binom{h}{\ell} r_{\mathcal{A},h-\ell}(n - \ell k) + \delta_{n/h}(k),$$

and the RHS inequality follows. \square

Lemma 4.2 (Fundamental lemma, r form). *In $\mathcal{S}_{\mathcal{A}}$ the following holds:*

$$\mathbb{E}(r_{\omega,h}(n)) = \Theta\left(\frac{A(n)^h}{n}\right), \quad \forall h \geq 1.$$

Proof. Recall the recursive formulas in (A.1). We have

$$\mathbb{E}(r_{\omega,h}(n)) = \sum_{k \leq n} \mathbb{E}(r_{\omega,h-1}(k) \mathbb{1}_{\omega}(n - k)).$$

We claim that it is sufficient to show that

$$(4.1) \quad \sum_{k \leq n} \mathbb{E}(r_{\omega,h-1}(k) \mathbb{1}_{\omega}(n - k)) \asymp \sum_{k \leq n} \mathbb{E}(r_{\omega,h-1}(k)) \mathbb{E}(\mathbb{1}_{\omega}(n - k)).$$

Indeed, since the claim of our lemma holds for $h = 1$ by definition, one may apply induction to see that

$$\begin{aligned} (4.2) \quad & \sum_{k \leq n} \mathbb{E}(r_{\omega,h-1}(k)) \mathbb{E}(\mathbb{1}_{\omega}(n - k)) \\ & \asymp \sum_{1 \leq k \leq n-1} \frac{A(k)^{h-1}}{k} \frac{A(n - k)}{n - k} \\ & = \sum_{1 \leq k \leq n/2} \frac{A(k)^{h-1}}{k} \frac{A(n - k)}{n - k} + \sum_{n/2 < k \leq n-1} \frac{A(k)^{h-1}}{k} \frac{A(n - k)}{n - k}. \end{aligned}$$

Then, using that \mathcal{A} is OR_+ , for the “ \gg ” side we have

$$\begin{aligned} \sum_{1 \leq k \leq n-1} \frac{A(k)^{h-1}}{k} \frac{A(n-k)}{n-k} &\gg \sum_{n/2 < k \leq n-1} \frac{A(k)^{h-1}}{k} \frac{A(n-k)}{n-k} \\ &\gg \frac{A(n)^{h-1}}{n} \sum_{n/2 < k \leq n-1} \frac{A(n-k)}{n-k} \\ &= \frac{A(n)^{h-1}}{n} \sum_{1 \leq k < n/2} \frac{A(k)}{k}, \end{aligned}$$

which by Lemma A.7 is $\gg A(n)^h/n$. For the “ \ll ” side

$$\begin{aligned} \sum_{1 \leq k \leq n/2} \frac{A(k)^{h-1}}{k} \frac{A(n-k)}{n-k} + \sum_{n/2 < k \leq n-1} \frac{A(k)^{h-1}}{k} \frac{A(n-k)}{n-k} \\ \ll \frac{A(n)}{n} \sum_{1 \leq k \leq n/2} \frac{A(k)^{h-1}}{k} + \frac{A(n)^{h-1}}{n} \sum_{n/2 < k \leq n-1} \frac{A(n-k)}{n-k}, \end{aligned}$$

which, again by Lemma A.7, is $\ll A(n)^h/n$.

Hence, we just need to prove (4.1). Note that the case $h = 2$ holds, since the $\mathbb{1}_\omega(n)$ are mutually independent over n ; hence we apply induction for $h \geq 3$. The r.v.s $r_{\omega, h-1}(k)$ and $\mathbb{1}_\omega(n-k)$ are not mutually independent in general, but since $\mathbb{1}_\omega(n-k)$ is a Bernoulli trial, we at least have:

$$\mathbb{E}(r_{\omega, h-1}(k) \mathbb{1}_\omega(n-k)) = \mathbb{E}(\mathbb{1}_\omega(n-k)) \mathbb{E}(r_{\omega, h-1}(k) \mid \mathbb{1}_\omega(n-k) = 1).$$

Applying Lemma 4.1 to ω , $h-1$, k and taking $n-k$,

$$\begin{aligned} &\mathbb{E}(\mathbb{1}_\omega(n-k)) \mathbb{E}(r_{\omega, h-1}(k)) \\ (4.3) \quad &\leq \mathbb{E}(\mathbb{1}_\omega(n-k)) \mathbb{E}(r_{\omega, h-1}(k) \mid \mathbb{1}_\omega(n-k) = 1) \\ &\leq \mathbb{E}(\mathbb{1}_\omega(n-k)) \left(\mathbb{E}(r_{\omega, h-1}(k)) + \delta_{k/(h-1)}(n-k) + \right. \\ &\quad \left. + \binom{h-1}{\lfloor \frac{h-1}{2} \rfloor} \sum_{\ell=1}^{h-2} \mathbb{E}(r_{\omega, h-\ell-1}(\ell k - (\ell-1)n)) \right). \end{aligned}$$

The apparent problem in this estimate is the summation on the upper bound. We show that it ends up being negligible. For $h \geq 3$ we have

$$\begin{aligned} &\sum_{k \leq n} \mathbb{E}(\mathbb{1}_\omega(n-k)) \left(\binom{h-1}{\lfloor \frac{h-1}{2} \rfloor} \sum_{\ell=1}^{h-2} \mathbb{E}(r_{\omega, h-\ell-1}(\ell k - (\ell-1)n)) \delta_{k/(h-1)}(n-k) \right) \\ &= \binom{h-1}{\lfloor \frac{h-1}{2} \rfloor} \sum_{\ell=1}^{h-2} \sum_{k \leq n} \mathbb{E}(r_{\omega, h-\ell-1}(\ell k - (\ell-1)n)) \mathbb{E}(\mathbb{1}_\omega(n-k)) + \\ &\quad + \sum_{k \leq n} \delta_{(h-1)n/h}(k) \mathbb{E}(\mathbb{1}_\omega(n-k)) \end{aligned}$$

$$\begin{aligned}
&\ll \sum_{\ell=1}^{h-2} \sum_{k \leq n} \mathbb{E}(r_{\omega, h-2}(\ell k - (\ell-1)n)) \mathbb{E}(\mathbb{1}_{\omega}(n-k)) + \frac{A(n)}{n} \\
&= \sum_{\ell=1}^{h-2} \sum_{k \leq n} \mathbb{E}(r_{\omega, h-2}(n - \ell k)) \mathbb{E}(\mathbb{1}_{\omega}(k)) + \frac{A(n)}{n}.
\end{aligned}$$

Since “ $n - \ell k \geq 0$ ” \iff “ $k \leq n/\ell$ ”, this summation can be subdivided as follows:

$$\begin{aligned}
&\sum_{k \leq n} \mathbb{E}(r_{\omega, h-2}(n - \ell k)) \mathbb{E}(\mathbb{1}_{\omega}(k)) \\
&\asymp \sum_{1 \leq k \leq \frac{n}{2\ell}} \frac{A(n - \ell k)^{h-2}}{n - \ell k} \frac{A(k)}{k} + \sum_{\frac{n}{2\ell} < k \leq \frac{n}{\ell} - 1} \frac{A(n - \ell k)^{h-2}}{n - \ell k} \frac{A(k)}{k} \\
&\ll \frac{A(n)^{h-2}}{n} \sum_{1 \leq k \leq \frac{n}{2\ell}} \frac{A(k)}{k} + \frac{A(n)}{n} \sum_{\frac{n}{2\ell} < k \leq \frac{n}{\ell} - 1} \frac{A(n - \ell k)^{h-2}}{n - \ell k},
\end{aligned}$$

therefore, by Lemma A.7, this is $O(A(n)^{h-1}/n)$ for all $\ell \geq 2$. With this, we conclude from (4.3) that

$$\begin{aligned}
&\sum_{k \leq n} \mathbb{E}(r_{\omega, h-1}(k)) \mathbb{E}(\mathbb{1}_{\omega}(n-k)) \\
(4.4) \quad &\leq \sum_{k \leq n} \mathbb{E}(r_{\omega, h-1}(k) \mathbb{1}_{\omega}(n-k)) \\
&\leq \sum_{k \leq n} \mathbb{E}(r_{\omega, h-1}(k)) \mathbb{E}(\mathbb{1}_{\omega}(n-k)) + O\left(\frac{A(n)^{h-1}}{n}\right),
\end{aligned}$$

hence (4.1) follows, and our proof is complete. \square

Two variations of this lemma will be important to us, one for exact (ρ) and other for lower bounded exact ($\rho|_{\geq}$) representation functions. When referenced along the text, consider them in conjunction with Lemma 4.2.

Lemma 4.3 (Fundamental lemma, ρ form). *In $\mathcal{S}_{\mathcal{A}}$ the following holds:*

$$\mathbb{E}(\rho_{\omega, h}(n)) = \frac{1}{h!} \mathbb{E}(r_{\omega, h}(n)) + O\left(\frac{A(n)^{h-1}}{n}\right), \quad \forall h \geq 2.$$

Proof. All we need to do is show that

$$\mathbb{E}(|\{\mathfrak{R} \in R_h(n; \omega) : \mathfrak{R} \text{ is non-exact}\}|) = O\left(\frac{A(n)^{h-1}}{n}\right).$$

This is clearly true for $h = 2$, thus we may focus on $h \geq 3$. For this case, consider the non-probabilistic estimate:

$$|\{\mathfrak{R} \in R_h(n; \omega) : \mathfrak{R} \text{ is non-exact}\}| \leq \binom{h}{2} \sum_{k \leq n/2} r_{\omega, h-2}(n-2k) \mathbb{1}_{\omega}(k).$$

To analyze this summation, let $\mathcal{N}_{\omega,h}(n) := \sum_{k \leq n/2} r_{\omega,h-1}(n-2k) \mathbb{1}_{\omega}(k)$. We are going to prove that

$$(4.5) \quad \mathbb{E}(\mathcal{N}_{\omega,h}(n)) = \Theta \left(\frac{A(n)^h}{n} \right), \quad \forall h \geq 2.$$

Just as in (4.3) on the proof of Lemma 4.2, we apply Lemma 4.1 for ω , $h-1$, $n-2k$ and k to obtain:

$$\begin{aligned} & \mathbb{E}(\mathbb{1}_{\omega}(k)) \mathbb{E}(r_{\omega,h-1}(n-2k)) \\ & \leq \mathbb{E}(\mathbb{1}_{\omega}(k)) \mathbb{E}(r_{\omega,h-1}(n-2k) \mid \mathbb{1}_{\omega}(k) = 1) \\ & \leq \mathbb{E}(\mathbb{1}_{\omega}(k)) \left(\mathbb{E}(r_{\omega,h-1}(n-2k)) + \delta_{(n-2k)/(h-1)}(k) + \right. \\ & \quad \left. + \left(\frac{h-1}{\lfloor \frac{h-1}{2} \rfloor} \right) \sum_{\ell=1}^{h-2} \mathbb{E}(r_{\omega,h-\ell-1}(n-(\ell+2)k)) \right). \end{aligned}$$

By calculations similar to the ones in Lemma 4.2 it follows

$$\begin{aligned} \mathbb{E}(\mathcal{N}_{\omega,h}(n)) &= \sum_{k \leq n/2} \mathbb{E}(r_{\omega,h-1}(n-2k) \mathbb{1}_{\omega}(k)) \\ &= \sum_{k \leq n/2} \mathbb{E}(r_{\omega,h-1}(n-2k)) \mathbb{E}(\mathbb{1}_{\omega}(k)) + O \left(\frac{A(n)^{h-1}}{n} \right). \end{aligned}$$

Finally, by using Lemma 4.2,

$$\begin{aligned} & \sum_{k \leq n/2} \mathbb{E}(r_{\omega,h-1}(n-2k)) \mathbb{E}(\mathbb{1}_{\omega}(k)) \\ & \asymp \sum_{1 \leq k \leq \frac{n}{4}} \frac{A(n-2k)^{h-1}}{n-2k} \frac{A(k)}{k} + \sum_{\frac{n}{4} < k \leq \frac{n}{2}-1} \frac{A(n-2k)^{h-1}}{n-2k} \frac{A(k)}{k}, \end{aligned}$$

thus (4.5) follows from Lemma A.7. \square

Lemma 4.4 (Fundamental Lemma, $\rho|_{\geq}$ form). *Given $h \geq 2$, for every $\varepsilon > 0$ there is $v_h(\varepsilon) > 0$ such that in $\mathcal{S}_{\mathcal{A}}$ the following holds:*

$$\mathbb{E}(\rho_{\omega,h}(n)|_{\geq \varepsilon n}) \geq (1 - v_h(\varepsilon)) \mathbb{E}(\rho_{\omega,h}(n)) \text{ for all large } n.$$

Furthermore, if $\varepsilon = \varepsilon(n) \rightarrow 0^+$ as $n \rightarrow +\infty$, then for every positive $\gamma < \mathcal{M}_*(A)$ one can take $v_h(\varepsilon)$ in a manner that satisfies

$$(4.6) \quad v_h(\varepsilon) \ll \varepsilon^{\gamma/2^{h-2}} \quad (\text{as } n \rightarrow +\infty),$$

where $\mathcal{M}_*(A)$ is the lower Matuszewska index (cf. (A.2)) of A .

Proof. Recall that $\rho_{\omega,h}(n)|_{\geq \varepsilon n} = \rho_{\omega,h}(n) - \rho_{\omega,h}(n)|_{< \varepsilon n}$ is counting those exact h -representations in which at least one term is smaller than εn . All we need to do is show that $\mathbb{E}(\rho_{\omega,h}(n)|_{\geq \varepsilon n}) \leq v_{h-1}(\varepsilon) \mathbb{E}(\rho_{\omega,h}(n))$.

An exact h -representation $\mathfrak{R} = \{k_1, \dots, k_h\}$ of n which has $k_1 < \varepsilon n$ must have $k_h > \frac{(1-\varepsilon)}{h-1}n$, therefore

$$\begin{aligned}
 \mathbb{E}(\rho_{\omega,h}(n) \mathfrak{b}_{\varepsilon n}) &= \mathbb{E} \left(\sum_{\substack{k_1+\dots+k_h=n \\ k_1 < k_2 < \dots < k_h \\ k_1 < \varepsilon n}} \mathbb{1}_{\omega}(k_1) \cdots \mathbb{1}_{\omega}(k_h) \right) \\
 &= \sum_{\substack{k_1+\dots+k_h=n \\ k_1 < k_2 < \dots < k_h \\ k_1 < \varepsilon n}} \frac{A(k_1)}{k_1+1} \cdots \frac{A(k_h)}{k_h+1} \\
 (4.7) \quad &< \frac{h-1}{(1-\varepsilon)} \frac{A(n)}{n + \frac{1}{1-\varepsilon}} \sum_{\substack{k_1+\dots+k_{h-1} \leq \frac{h-2-\varepsilon}{h-1}n \\ k_1 < k_2 < \dots < k_{h-1} \\ k_1 < \varepsilon n}} \frac{A(k_1)}{k_1+1} \cdots \frac{A(k_{h-1})}{k_{h-1}+1},
 \end{aligned}$$

where the last sum can be estimated by

$$\begin{aligned}
 &\sum_{\ell \leq \frac{h-2-\varepsilon}{h-1}n} \sum_{\substack{k_1+\dots+k_{h-1}=\ell \\ k_1 < k_2 < \dots < k_{h-1} \\ k_1 < \varepsilon n}} \frac{A(k_1)}{k_1+1} \cdots \frac{A(k_{h-1})}{k_{h-1}+1} \\
 &\leq \sum_{\ell \leq \varepsilon^{1/2}n} \mathbb{E}(\rho_{\omega,h-1}(\ell)) + \sum_{\varepsilon^{1/2}n < \ell \leq n} \sum_{\substack{k_1+\dots+k_{h-1}=\ell \\ k_1 < k_2 < \dots < k_{h-1} \\ k_1 < \varepsilon^{1/2}\ell}} \frac{A(k_1)}{k_1+1} \cdots \frac{A(k_{h-1})}{k_{h-1}+1} \\
 (4.8) \quad &= \sum_{\ell \leq \varepsilon^{1/2}n} \mathbb{E}(\rho_{\omega,h-1}(\ell)) + \sum_{\varepsilon^{1/2}n < \ell \leq n} \mathbb{E}(\rho_{\omega,h-1}(\ell) \mathfrak{b}_{\varepsilon^{1/2}\ell}).
 \end{aligned}$$

By Lemma 4.3, there are $t_h, T_h > 0$ such that, for large n ,

$$t_h \frac{A(n)^h}{n} \leq \mathbb{E}(\rho_{\omega,h}(n)) \leq T_h \frac{A(n)^h}{n}.$$

In addition, recall that \mathcal{A} is OR_+ , thus its lower Matuszewska index $\mathcal{M}_*(A)$ is positive. Therefore for every $0 < \gamma < \mathcal{M}_*(A)$ there must be some $M = M_\gamma \in \mathbb{R}_+$ for which $A(\varepsilon n) \leq M\varepsilon^\gamma A(n)$ for all large n . Using that, we can prove our statement by induction.

When $h = 2$, by Lemma A.7, there exists $C \in \mathbb{R}_+$ such that, for large n ,

$$\begin{aligned}
 \mathbb{E}(\rho_{\omega,2}(n) \mathfrak{b}_{\varepsilon n}) &= \sum_{k < \varepsilon n} \frac{A(k)}{k+1} \frac{A(n-k)}{n-k+1} \\
 &\leq \frac{A(n)}{(1-\varepsilon)n+1} \sum_{k < \varepsilon n} \frac{A(k)}{k+1} \leq C \cdot \frac{A(n)}{n} A(\varepsilon n),
 \end{aligned}$$

thus, since $A(\varepsilon n) \leq M\varepsilon^\gamma A(n)$, our statement applies for $h = 2$, including the upper bound for $v_2(\varepsilon)$ in (4.6) for $\varepsilon = \varepsilon(n) \rightarrow 0^+$. Supposing it holds for $h-1$, from (4.8) we get

$$\sum_{\ell \leq \varepsilon^{1/2}n} \mathbb{E}(\rho_{\omega,h-1}(\ell)) + \sum_{\varepsilon^{1/2}n < \ell \leq n} \mathbb{E}(\rho_{\omega,h-1}(\ell) \mathfrak{b}_{\varepsilon^{1/2}\ell})$$

$$\begin{aligned}
&\leq CT_{h-1}M^{h-1}\varepsilon^{(h-1)\gamma/2}A(n)^{h-1} + T_{h-1}v_{h-1}(\varepsilon^{1/2})\sum_{\ell\leq n}\frac{A(\ell)^{h-1}}{\ell} \\
&\leq CT_{h-1}\left(M^{h-1}\varepsilon^{(h-1)\gamma/2} + v_{h-1}(\varepsilon^{1/2})\right)A(n)^{h-1},
\end{aligned}$$

for large n and some $C \in \mathbb{R}_+$ that comes from Lemma A.7. Substituting this in (4.7) yields

$$\mathbb{E}(\rho_{\omega,h}(n)b_{\varepsilon n}) < \frac{h-1}{(1-\varepsilon)}\frac{n}{n+\frac{1}{1-\varepsilon}}CT_{h-1}\left(M^{h-1}\varepsilon^{(h-1)\gamma/2} + v_{h-1}(\varepsilon^{1/2})\right)\frac{A(n)^h}{n}.$$

We have that $n/(n+\frac{1}{1-\varepsilon}) \sim 1$ as $n \rightarrow +\infty$ and every other term in this coefficient is either bounded or independent of ε , except for $\varepsilon^{(h-1)\gamma/2}$ and $v_{h-1}(\varepsilon^{1/2})$. Since the induction step includes (4.6), we deduce that these two terms must remain bounded, and if $\varepsilon = \varepsilon(n) \rightarrow 0^+$ then they must vanish like $O(\varepsilon^{\gamma/2^{h-2}})$. The proof is then complete. \square

We finish this section with an important remark about Lemma 4.2.

Remark 4.5. Let $g : [x_0, +\infty) \rightarrow \mathbb{R}_+$ be a positive real function satisfying the conditions of Lemma A.7. Based on this lemma, for each $2 \leq \ell \leq h$ let c_ℓ be some positive constant for which

$$\sum_{x_0 < k \leq n-x_0} \frac{g(k)^{\ell-1}}{k} \frac{g(n-k)}{n-k} > c_\ell \frac{g(n)}{n}$$

holds for all large n . In Lemma 4.2, in view of (4.4), it is possible to deduce from the equations in (4.2) that if \mathcal{A} is such that $A(n) > Mg(n)$ for some $M > 0$ and all large n , then

$$\mathbb{E}(r_{\omega,h}(n)) > M^h \left(\prod_{\ell=2}^h c_h \right) \frac{g(n)^h}{n}, \quad \forall \text{ suff. large } n.$$

Note that the same holds when inequalities are reversed. These observations are key for the proof of the Main Theorem.

5. CONCENTRATION (STEP III)

Now that the general setup is established, we will start working with the conditions from the Main Theorem. We start with the two lemmas that constitute the most technical parts of the proof.

5.1. Lemmas. The following results have a similar theme concerning the condition $A(x) \ll x^{1/(h-1)}$. As mentioned in Section 2, our methods are naturally limited by the fact that $\widehat{\rho}_{\omega,h}(n) \leq n/h$; this is basically why $x^{1/(h-1)}$ pops up, and the purpose of the following lemmas is to make the most out of this restraint. The first lemma says that non-exact h -representations are, almost surely, conveniently scarce in $\mathcal{S}_{\mathcal{A}}$ for such \mathcal{A} .

Lemma 5.1. *Let $h \geq 2$ be fixed. If $A(x) \ll x^{1/(h-1)}$, then in $\mathcal{S}_{\mathcal{A}}$ the following holds:*

$$\rho_{\omega,h-1}(n) \stackrel{\text{a.s.}}{=} O\left(\frac{\log(n)}{\log \log(n)}\right),$$

$$(5.1) \quad |\{\mathfrak{R} \in \text{R}_h(n; \omega) : \mathfrak{R} \text{ is non-exact}\}| \stackrel{\text{a.s.}}{=} O\left(\frac{\log(n)}{\log \log(n)}\right).$$

Furthermore,

$$\rho_{\omega, h}(n) \stackrel{\text{a.s.}}{\ll} \sum_{k \in \bigcup \mathcal{C}_n} \widehat{\rho}_{\omega, h-1}(n-k),$$

where $\mathcal{C}_n \subseteq \text{ER}_h(n; \omega)$ is a maximal disjoint family.

Proof. We start by proving the first two estimates. Just as in Lemma 4.3, we have the non-probabilistic estimate

$$|\{\mathfrak{R} \in \text{R}_\ell(n; \omega) : \mathfrak{R} \text{ is non-exact}\}| \leq \binom{\ell}{2} \mathcal{N}_{\omega, \ell-1}(n)$$

for $\ell \geq 3$, where $\mathcal{N}_{\omega, \ell}(n) = \sum_{k \leq n/2} r_{\omega, \ell-1}(n-2k) \mathbb{1}_\omega(k)$. Note that this is counting the number of solutions to

$$(5.2) \quad 2x_1 + x_2 + \dots + x_\ell = n, \quad x_i \in \omega.$$

Consider now $\widehat{\mathcal{N}}_{\omega, \ell}(n)$ to be the cardinality of the largest maximal collection of disjoint *multisets*¹¹ $\{k_1, k_2, \dots, k_\ell\} \subseteq \omega$ satisfying (5.2). We may then deduce the following non-probabilistic inequalities:

$$(5.3) \quad r_{\omega, \ell}(n) \leq \ell! \ell \left(\max_{k \leq n} r_{\omega, \ell-1}(k) \right) \widehat{\rho}_{\omega, \ell}(n) + \binom{\ell}{2} \mathcal{N}_{\omega, \ell-1}(n)$$

$$(5.4) \quad \mathcal{N}_{\omega, \ell}(n) \leq \ell! \ell^2 \left(\max_{k \leq n} \{\mathcal{N}_{\omega, \ell-1}(k)\} + \max_{k \leq n} \{r_{\omega, \ell-1}(k)\} \right) \widehat{\mathcal{N}}_{\omega, \ell}(n).$$

Indeed, (5.3) follows by noticing that when $\mathcal{C} \subseteq \text{ER}_\ell(n; \omega)$ is a maximal collection of pairwise disjoint representations with $|\mathcal{C}| = \widehat{\rho}_{\omega, \ell}(n)$, we have $|\bigcup \mathcal{C}| = \ell \cdot \widehat{\rho}_{\omega, \ell}(n)$ and $\mathfrak{R} \cap (\bigcup \mathcal{C}) \neq \emptyset$ for every $\mathfrak{R} \in \text{ER}_\ell(n; \omega)$; thus the first term bounds the number of exact ℓ -representations, whereas the latter bounds the number of non-exact ones. The reasoning for (5.4) is analogous.

To prove our lemma, we just need to show that $\mathcal{N}_{\omega, h-1}(n) \stackrel{\text{a.s.}}{=} O\left(\frac{\log(n)}{\log \log(n)}\right)$. The estimate in (5.4) allows us to break this into the following three items:

- (i) $r_{\omega, h-2}(n) \stackrel{\text{a.s.}}{=} O(1)$;
- (ii) $\mathcal{N}_{\omega, h-2}(n) \stackrel{\text{a.s.}}{=} O(1)$;
- (iii) $\widehat{\mathcal{N}}_{\omega, h-1}(n) \stackrel{\text{a.s.}}{=} O(\log(n)/\log \log(n))$.

The first two are dealt with by induction. Since $r_{\omega, 2}(n) \ll \widehat{\rho}_{\omega, 2}(n) + 1$ and $\mathcal{N}_{\omega, 2}(n) \ll \widehat{\mathcal{N}}_{\omega, 2}(n)$, to show (i) and (ii) it is sufficient to, in view of (5.3) and (5.4), show that:

- (i)* $\widehat{\rho}_{\omega, \ell}(n) \stackrel{\text{a.s.}}{=} O(1)$, for all $2 \leq \ell \leq h-2$;
- (ii)* $\widehat{\mathcal{N}}_{\omega, \ell}(n) \stackrel{\text{a.s.}}{=} O(1)$, for all $2 \leq \ell \leq h-2$.

At this point, we have managed to reduce all our problems to estimates involving maxdisfam-type counting functions, thus the next natural step is to use the disjointness lemma! We now prove (i)*, (ii)* and (iii).

¹¹A *multiset* is a set that allows multiple instances of an element.

• **Items (i)* and (ii)*:** Starting with (ii)*, from (4.5) in Lemma 4.3 we know that $\mathbb{E}(\mathcal{N}_{\omega,\ell}(n)) = \Theta(A(n)^\ell/n)$. Thus, as the function $\widehat{\mathcal{N}}_{\omega,\ell}(n)$ is counting mutually independent events, applying the disjointness lemma yields, in view of having $A(x) \ll x^{1/(h-1)}$,

$$\Pr\left(\widehat{\mathcal{N}}_{\omega,\ell}(n) \geq M\right) \leq \frac{\mathbb{E}(\mathcal{N}_{\omega,\ell}(n))^M}{M!} \ll \left(\frac{A(n)^{h-2}}{n}\right)^M \ll n^{-M/(h-1)}.$$

Therefore, if we take $M > 2(h-1)$, it follows from the Borel–Cantelli lemma that $\widehat{\mathcal{N}}_{\omega,\ell}(n) \stackrel{\text{a.s.}}{=} O(1)$ for all ℓ considered. The case (i)* for $\widehat{\rho}_{\omega,\ell}(n)$ goes analogously.

• **Item (iii):** Once again we apply the disjointness lemma, but this time we shall take $M = M_n$ to be a slowly growing function in n . First, we have

$$(5.5) \quad \Pr\left(\widehat{\mathcal{N}}_{\omega,h-1}(n) \geq M_n\right) \leq \frac{\mathbb{E}(\mathcal{N}_{\omega,h-1}(n))^{M_n}}{M_n!}.$$

Let $M_n := \lceil 2\log(n)/\log\log(n) \rceil$ for $n > e^e$. Since M_n is unbounded and non-decreasing, we have that for every $\varepsilon > 0$ there is $n_0 \in \mathbb{N}$ such that

$$(1 - \varepsilon)M_n \log(M_n) \leq \log(M_n!) \leq (1 + \varepsilon)M_n \log(M_n), \quad \forall n \geq n_0.$$

Thus, keeping in mind that $\mathbb{E}(\mathcal{N}_{\omega,h-1}(n)) < C$ for some $C > 0$ and all large n , we may apply “ $-\log$ ” in both sides of (5.5) to obtain, for large n ,

$$\begin{aligned} -\log\left(\Pr\left(\widehat{\mathcal{N}}_{\omega,\ell}(n) \geq M_n\right)\right) &\geq \log(M_n!) - M_n \log(C) \\ &\geq (1 - \varepsilon)M_n \log(M_n) \\ &\geq (1 - \varepsilon) \frac{2\log(n)}{\log\log(n)} \log\left(\frac{2\log(n)}{\log\log(n)}\right) \\ &\geq 2(1 - \varepsilon) \log(n) - o(\log(n)). \end{aligned}$$

Therefore $\Pr(\widehat{\mathcal{N}}_{\omega,h-1}(n) \geq M_n) \leq n^{-2(1-\varepsilon)+o(1)}$. Choosing $\varepsilon < 1/2$, we conclude from the Borel–Cantelli lemma that $\widehat{\mathcal{N}}_{\omega,h-1}(n) \stackrel{\text{a.s.}}{=} O(M_n)$, thus completing the proof of (5.1). As before, the reasoning for $\rho_{\omega,h-1}(n)$ is analogous.

Finally, given $\ell \geq 2$, consider $\mathcal{C}_n^{(\ell)} \subseteq \text{ER}_\ell(n; \omega)$ a maximal disjoint family of exact ℓ -representations of n . By definition, we will have

$$\rho_{\omega,\ell}(n) \leq \sum_{k \in \bigcup \mathcal{C}_n^{(\ell)}} \rho_{\omega,\ell-1}(n - k).$$

From item (i) we know that $\rho_{\omega,\ell-1}(n) \stackrel{\text{a.s.}}{=} O(1)$ for $2 \leq \ell \leq h-1$, therefore, since $|\bigcup \mathcal{C}_n^{(\ell)}| = \ell \cdot \widehat{\rho}_{\omega,\ell}(n)$, the above implies $\rho_{\omega,\ell}(n) \stackrel{\text{a.s.}}{\ll} \widehat{\rho}_{\omega,\ell}(n)$ for such ℓ , thus the last estimate stated in the lemma follows. \square

The second lemma is very similar to Lemma 11 from Erdős & Tetali [6]; the goal is to show that we will have “ Δ ” arbitrarily small in our applications of the correlation inequality. By an abuse of notation, given $\mathfrak{R} \in \text{ER}_h(n)$ we shall also use “ \mathfrak{R} ” to denote the event “ $\mathfrak{R} \in \text{ER}_h(n; \omega)$ ”.

Lemma 5.2. *Let $h \geq 2$ be fixed. For every $\varepsilon > 0$ there is $M_0 \in \mathbb{R}_+$ such that, if $A(x) < M^{-1}x^{1/(h-1)}$ for some $M > M_0$ and all sufficiently large x , then in $\mathcal{S}_{\mathcal{A}}$ the following holds:*

$$\sum_{\substack{\mathfrak{R}, \mathfrak{R}' \in \text{ER}_h(n) \\ 1 \leq |\mathfrak{R} \cap \mathfrak{R}'| \leq h-2}} \Pr(\mathfrak{R} \wedge \mathfrak{R}') < \varepsilon$$

for all sufficiently large n .

Proof. Rewriting this sum as

$$\sum_{\substack{\mathfrak{R}, \mathfrak{R}' \in \text{ER}_h(n) \\ 1 \leq |\mathfrak{R} \cap \mathfrak{R}'| \leq h-2}} \Pr(\mathfrak{R} \wedge \mathfrak{R}') = \sum_{\ell=1}^{h-2} \sum_{\substack{\mathfrak{R}, \mathfrak{R}' \in \text{ER}_h(n) \\ |\mathfrak{R} \cap \mathfrak{R}'| = \ell}} \Pr(\mathfrak{R} \wedge \mathfrak{R}'),$$

we just need to estimate each of the terms for ℓ fixed. Taking $\mathfrak{R} \cap \mathfrak{R}'$ with $|\mathfrak{R} \cap \mathfrak{R}'| = \ell$, say we have

$$\mathfrak{R} = \{x_1, \dots, x_\ell, y_1, \dots, y_{h-\ell}\} \text{ and } \mathfrak{R}' = \{x_1, \dots, x_\ell, z_1, \dots, z_{h-\ell}\},$$

with $\sum_{i=1}^\ell x_i = k$ and $y_i \neq z_j$ for all i, j . Letting \sum^* denote the sum over exact representations and abusing the notation again by writing “ $\Pr(x_i)$ ” to denote “ $\Pr(\mathbb{1}_\omega(x_i) = 1)$ ”, we have:

$$\begin{aligned} & \sum_{\substack{\mathfrak{R}, \mathfrak{R}' \in \text{ER}_h(n) \\ |\mathfrak{R} \cap \mathfrak{R}'| = \ell}} \Pr(\mathfrak{R} \wedge \mathfrak{R}') \\ &= \sum_{k \leq n} \sum_{\substack{x_1 + \dots + x_\ell = k \\ y_1 + \dots + y_{h-\ell} = n-k \\ z_1 + \dots + z_{h-\ell} = n-k}}^* \Pr(x_1) \cdots \Pr(x_\ell) \Pr(y_1) \cdots \Pr(y_{h-\ell}) \Pr(z_1) \cdots \Pr(z_{h-\ell}) \\ &\leq \sum_{k \leq n} \left(\sum_{x_1 + \dots + x_\ell = k}^* \Pr(x_1) \cdots \Pr(x_\ell) \right) \left(\sum_{y_1 + \dots + y_{h-\ell} = n-k}^* \Pr(y_1) \cdots \Pr(y_{h-\ell}) \right)^2 \\ &= \sum_{k \leq n} \mathbb{E}(\rho_{\omega, \ell}(k)) \mathbb{E}(\rho_{\omega, h-\ell}(n-k))^2. \end{aligned}$$

Since $A(n) < M^{-1}n^{1/(h-1)}$ and $\mathbb{E}(\rho_{\omega, \ell}(n)) \leq \ell!^{-1} \mathbb{E}(r_{\omega, \ell}(n))$ for all large n , the observation from Remark 4.5 allows us to derive

$$\begin{aligned} & \sum_{k \leq n} \mathbb{E}(\rho_{\omega, \ell}(k)) \mathbb{E}(\rho_{\omega, h-\ell}(n-k))^2 \\ (5.6) \quad & < M^{-\ell(h-\ell)^2} \frac{C_\ell(C_{h-\ell})^2}{\ell!(h-\ell)!^2} \sum_{1 \leq k \leq n-1} \frac{k^{\ell/(h-1)}}{k} \left(\frac{(n-k)^{(h-\ell)/(h-1)}}{n-k} \right)^2, \end{aligned}$$

for some $C_\ell, C_{h-\ell} > 0$ not depending on \mathcal{A} , for all large n . Note that for all ℓ ranging from 1 to $h-2$, which is the range being considered,

$$\max_{1 \leq k \leq n-1} \left(\frac{(n-k)^{(h-\ell)/(h-1)}}{n-k} \right) = 1,$$

hence

$$\begin{aligned} M^{-\ell(h-\ell)^2} \frac{C_\ell(C_{h-\ell})^2}{\ell!(h-\ell)!^2} \sum_{1 \leq k \leq n-1} \frac{k^{\ell/(h-1)}}{k} \left(\frac{(n-k)^{(h-\ell)/(h-1)}}{n-k} \right)^2 \\ \leq M^{-\ell(h-\ell)^2} \frac{C_\ell(C_{h-\ell})^2}{\ell!(h-\ell)!^2} \sum_{1 \leq k \leq n-1} \frac{k^{\ell/(h-1)}}{k} \frac{(n-k)^{(h-\ell)/(h-1)}}{n-k}. \end{aligned}$$

The nature of the above convolution is analogous to that of (4.2) and the ones described in Remark 4.5, thus it is not difficult to see that this sum is bounded as $n \rightarrow +\infty$ by, say, some $T > 0$. Going back to (5.6), we then arrive at

$$\sum_{\substack{\mathfrak{R}, \mathfrak{R}' \in \text{ER}_h(n) \\ 1 \leq |\mathfrak{R} \cap \mathfrak{R}'| \leq h-2}} \Pr(\mathfrak{R} \wedge \mathfrak{R}') < M^{-\ell(h-\ell)^2} \frac{C_\ell(C_{h-\ell})^2}{\ell!(h-\ell)!^2} T$$

for all large n , where M is the only term depending on \mathcal{A} . Since M can be chosen as big as desirable, our lemma follows. \square

We now prove the Main Theorem.

5.2. Proof of the Main Theorem. Fix $h \geq 2$ an integer. Take a positive and locally integrable real function $f : [x_0, +\infty) \rightarrow \mathbb{R}_+$ such that

$$x^{1/h} \log(x)^{1/h} \ll f(x) \ll x^{1/(h-1)} \quad \text{and} \quad \int_{x_0}^x \frac{f(t)}{t} dt = \Theta(f(x)).$$

Since $f(x) \ll x$, let $K := \limsup_x f(x)/x$. Choose $\epsilon > 0$ and let x^* be such that $f(x) \leq (K + \epsilon)x$ for all $x \geq x^*$. Define then, for every $n \geq 0$,

$$\begin{cases} \alpha_n := 1, & \text{if } n < x^*; \\ \alpha_n := (K + \epsilon)^{-1} f(n)/n, & \text{if } n \geq x^*. \end{cases}$$

By Lemma A.7, f is in particular almost increasing, thus

$$\int_{x_0}^x \frac{f(t)}{t} dt \asymp \sum_{x_0 < n \leq x} \frac{f(n)}{n},$$

which, by the same lemma, is $\Theta(f(x))$. From strong law of large numbers it then follows that in $\langle \mathcal{S}, (\alpha_n)_n \rangle$ it holds $|\omega \cap [0, x]| \stackrel{\text{a.s.}}{\asymp} f(x)$. Since this is a non-empty probability space, there must be at least some $\mathcal{C} \subseteq \mathbb{N}$ for which $C(x) = \Theta(f(x))$.

Based on this, let us construct a more convenient sequence \mathcal{A} . We know from $x^{1/h} \log(x)^{1/h} \ll C(x) \ll x^{1/(h-1)}$ that there must exist $q_1, q_2 > 0$ for which $\exists n_0 \in \mathbb{N} : \forall n \geq n_0$ the following holds:

$$(5.7) \quad q_1 \cdot x^{1/h} \log(x)^{1/h} \leq C(x) \leq q_2 \cdot x^{1/(h-1)}$$

With this in mind, take:

$$(5.8) \quad M_1, M_2 > 1 \quad \text{large real constants to be specified later.}$$

Let $n_1 \in \mathbb{N}$ be large enough so that, for all $n \geq n_1$,

$$M_1 n^{1/h} \log(n)^{1/h} + 1 > M_1 (n+1)^{1/h} \log(n+1)^{1/h},$$

$$\frac{1}{M_2} n^{1/(h-1)} + 1 > \frac{1}{M_2} (n+1)^{1/(h-1)}.$$

The following construction is then well-defined:

- Define $\mathcal{A}_{n_1} := \{1, 2, \dots, n_1\}$.
- For each integer $k > n_1$, define \mathcal{A}_k in the following manner:
 - In case $k \notin \mathcal{C}$, check whether

$$\frac{|\mathcal{A}_{k-1} \cap [0, k]|}{k^{1/h} \log(k)^{1/h}} \leq M_1.$$

If so, take $\mathcal{A}_k := \mathcal{A}_{k-1} \cup \{k\}$; otherwise, take $\mathcal{A}_k := \mathcal{A}_{k-1}$.

- In case $k \in \mathcal{C}$, check whether

$$\frac{|(\mathcal{A}_{k-1} \cup \{k\}) \cap [0, k]|}{k^{1/(h-1)}} \geq \frac{1}{M_2}.$$

If so, take $\mathcal{A}_k := \mathcal{A}_{k-1}$; otherwise, take $\mathcal{A}_k := \mathcal{A}_{k-1} \cup \{k\}$.

Finally, let $\mathcal{A} := \bigcup_{k \geq n_1} \mathcal{A}_k$. From the symmetric nature of this construction and from (5.7), it is possible to deduce that if M_1, M_2 are large enough so that $M_1 M_2 > q_1/q_2$, then \mathcal{A} must satisfy

$$(5.9) \quad \limsup_{n \rightarrow +\infty} \frac{A(n)}{C(n)} \leq \frac{M_1}{q_1} \quad \text{and} \quad \liminf_{n \rightarrow +\infty} \frac{A(n)}{C(n)} \geq \frac{1}{q_2 M_2},$$

which implies $A(x) = \Theta(C(x))$. More importantly, \mathcal{A} satisfies:

$$(5.10) \quad M_1 \cdot n^{1/h} \log(n)^{1/h} \leq A(n) \leq \frac{1}{M_2} \cdot n^{1/(h-1)}$$

from some point onward.

Before proceeding, we highlight the important stuff that shall be used throughout the proof of our next results in the following remark.

Remark. If the constants M_1, M_2 from (5.8) are large enough – i.e. satisfying $M_1 M_2 > q_1/q_2$ where q_1, q_2 are the constants from (5.7), then we can construct an OR_+ sequence \mathcal{A} with $A(n) = \Theta(f(n))$ satisfying (5.10) for all sufficiently large n .

In short, to prove the Main Theorem stated at the introduction we first show that $\widehat{\rho}_{\omega, h}(n)$ has a.s. the right rate of growth, then $\rho_{\omega, h}(n)$, and then $r_{\omega, h+\ell}(n)$ for all $\ell \geq 0$.

Lemma 5.3. *In the above construction of \mathcal{A} , if $M_1 > 1$ is sufficiently large, then in $\mathcal{S}_{\mathcal{A}}$ the following holds:*

$$\widehat{\rho}_{\omega, h}(n) \stackrel{\text{a.s.}}{=} \Theta\left(\frac{A(n)^h}{n}\right).$$

Proof. We divide the proof into two parts.

- **Part 1:** Showing $\widehat{\rho}_{\omega, h}(n) \stackrel{\text{a.s.}}{\ll} A(n)^h/n$.

Let $T > 0$ be a large integer. By the disjointness lemma,

$$\Pr(\widehat{\rho}_{\omega,h}(n) \geq T \lceil \mathbb{E}(\rho_{\omega,h}(n)) \rceil) \leq \frac{\mathbb{E}(\rho_{\omega,h}(n))^{T \lceil \mathbb{E}(\rho_{\omega,h}(n)) \rceil}}{(T \lceil \mathbb{E}(\rho_{\omega,h}(n)) \rceil)!},$$

thus, as for all $k \geq 1$ it holds $k! > k^k e^{-k}$,

$$\Pr(\widehat{\rho}_{\omega,h}(n) \geq T \lceil \mathbb{E}(\rho_{\omega,h}(n)) \rceil) \leq \left(\frac{e}{T} \frac{\mathbb{E}(\rho_{\omega,h}(n))}{\lceil \mathbb{E}(\rho_{\omega,h}(n)) \rceil} \right)^{T \lceil \mathbb{E}(\rho_{\omega,h}(n)) \rceil}.$$

Since $A(n) \geq M_1 n^{1/h} \log(n)^{1/h}$, in view of Lemma 4.3 and Remark 4.5 we have $\mathbb{E}(\rho_{\omega,h}(n)) > \alpha M_1^h \log(n)$ for all large n , where $\alpha > 0$ is some constant (not depending on \mathcal{A} !). Thus, taking $T > \max\{e^2, 2\alpha^{-1} M_1^{-h}\}$ yields

$$\left(\frac{e}{T} \frac{\mathbb{E}(\rho_{\omega,h}(n))}{\lceil \mathbb{E}(\rho_{\omega,h}(n)) \rceil} \right)^{T \lceil \mathbb{E}(\rho_{\omega,h}(n)) \rceil} \ll \left(\frac{1}{e} \right)^{2 \log(n) + o(\log(n))} \ll n^{-2+o(1)},$$

hence it follows from the Borel–Cantelli lemma that $\widehat{\rho}_{\omega,h}(n) \stackrel{\text{a.s.}}{\ll} A(n)^h/n$.

• **Part 2:** Showing $\widehat{\rho}_{\omega,h}(n) \stackrel{\text{a.s.}}{\gg} A(n)^h/n$.

This part is basically Theorem 3 from Erdős & Tetali [6] with the added nuances from our hypotheses. Recalling our definition of lower bounded representations, it suffices to show that $\widehat{\rho}_{\omega,h}(n)|_{\geq n^{1/2}} \stackrel{\text{a.s.}}{\gg} A(n)^h/n$. From Lemma 4.4,

$$(5.11) \quad \mathbb{E}(\rho_{\omega,h}(n)|_{\geq n^{1/2}}) = (1 - o(1)) \mathbb{E}(\rho_{\omega,h}(n)).$$

Similar to what was done in Lemma 5.2, given $\mathfrak{R} \in \text{ER}_h(n)|_{\geq n^{1/2}}$ we abuse the notation by letting “ \mathfrak{R} ” denote the event “ $\mathfrak{R} \in \text{ER}_h(n; \omega)|_{\geq n^{1/2}}$ ”. In addition, let $\sum^{(k)}$ denote the sum over all k -sets of pairwise disjoint lower bounded (by $n^{1/2}$) exact representations of n . Letting $t > 0$ be a small real number, we then have:

$$\begin{aligned} & \Pr(\exists \mathcal{C} \subseteq \text{ER}_h(n; \omega)|_{\geq n^{1/2}} \text{ maxdisfam} : |\mathcal{C}| \leq t \mathbb{E}(\rho_{\omega,h}(n))) \\ &= \sum_{1 \leq k \leq t \mathbb{E}(\rho_{\omega,h}(n))} \Pr(\exists \mathcal{C} \subseteq \text{ER}_h(n; \omega)|_{\geq n^{1/2}} \text{ maxdisfam} : |\mathcal{C}| = k) \\ &= \sum_{1 \leq k \leq t \mathbb{E}(\rho_{\omega,h}(n))} \sum_{\substack{\{\mathfrak{R}_1, \dots, \mathfrak{R}_k\} \\ \text{pairwise} \\ \text{disjoint}}}^{(k)} \left(\Pr(\mathfrak{R}_1 \wedge \dots \wedge \mathfrak{R}_k) \cdot \Pr(\{\mathfrak{R}_1 \dots \mathfrak{R}_k\} \text{ is maximal}) \right) \end{aligned}$$

We may estimate both factors in the above summation separately. The first term can be estimated directly from the disjointness lemma:

$$(5.12) \quad \sum_{\substack{\{\mathfrak{R}_1, \dots, \mathfrak{R}_k\} \\ \text{pairwise} \\ \text{disjoint}}}^{(k)} \Pr(\mathfrak{R}_1 \wedge \dots \wedge \mathfrak{R}_k) \leq \frac{\mathbb{E}(\rho_{\omega,h}(n))^k}{k!}$$

For the second term we use the correlation inequality! Given $\mathfrak{R} \in \text{ER}_h(n)|_{\geq n^{1/2}}$, let “ $\bar{\mathfrak{R}}$ ” denote the event “ $\mathfrak{R} \notin \text{ER}_h(n; \omega)|_{\geq n^{1/2}}$ ”. Then:

$$\Pr(\{\mathfrak{R}_1, \dots, \mathfrak{R}_k\} \text{ is maximal}) = \Pr\left(\bigwedge_{\substack{\mathfrak{R} \in \text{ER}_h(n)|_{\geq n^{1/2}} \\ \mathfrak{R} \cap \mathfrak{R}_j = \emptyset (j=1, \dots, k)}} \bar{\mathfrak{R}}\right)$$

When $h = 2$ the events $\bar{\mathfrak{R}}$ are mutually independent, and when $h \geq 3$ we can choose n large enough so that $\Pr(\mathfrak{R}) \leq 1/2$ for every $\mathfrak{R} \in \text{ER}_h(n)|_{\geq n^{1/2}}$.¹² Therefore we can apply the correlation inequality as follows:

$$\begin{aligned} \Pr\left(\bigwedge_{\substack{\mathfrak{R} \in \text{ER}_h(n)|_{\geq n^{1/2}} \\ \mathfrak{R} \cap \mathfrak{R}_j = \emptyset (j=1, \dots, k)}} \bar{\mathfrak{R}}\right) &\leq \left(\prod_{\substack{\mathfrak{R} \in \text{ER}_h(n)|_{\geq n^{1/2}} \\ \mathfrak{R} \cap \mathfrak{R}_j = \emptyset (j=1, \dots, k)}} \Pr(\bar{\mathfrak{R}})\right) e^{2\Delta(n)} \\ &= \left(\prod_{\substack{\mathfrak{R} \in \text{ER}_h(n)|_{\geq n^{1/2}} \\ \mathfrak{R} \cap \mathfrak{R}_j = \emptyset (j=1, \dots, k)}} (1 - \Pr(\mathfrak{R}))\right) e^{2\Delta(n)} \\ &\leq \exp\left(-\sum_{\substack{\mathfrak{R} \in \text{ER}_h(n)|_{\geq n^{1/2}} \\ \mathfrak{R} \cap \mathfrak{R}_j = \emptyset (j=1, \dots, k)}} \Pr(\mathfrak{R})\right) e^{2\Delta(n)}, \end{aligned}$$

where $\Delta(n)$ can be taken to be

$$\Delta(n) = \sum_{\substack{\mathfrak{R}, \mathfrak{R}' \in \text{ER}_h(n) \\ 1 \leq |\mathfrak{R} \cap \mathfrak{R}'| \leq h-2}} \Pr(\mathfrak{R} \wedge \mathfrak{R}').$$

However, since $A(x) \leq M_2^{-1} x^{1/(h-1)}$, we know from Lemma 5.2 that $\Delta(n)$ is bounded, thus $e^{2\Delta(n)} < K$ for some constant $K > 0$. For the term inside exp, from (5.11),

$$\begin{aligned} \sum_{\substack{\mathfrak{R} \in \text{ER}_h(n)|_{\geq n^{1/2}} \\ \mathfrak{R} \cap \mathfrak{R}_j = \emptyset (j=1, \dots, k)}} \Pr(\mathfrak{R}) &\geq \mathbb{E}(\rho_{\omega, h}(n)|_{\geq n^{1/2}}) - \mathbb{E}\left(\sum_{\ell \in \bigcup\{\mathfrak{R}_1, \dots, \mathfrak{R}_k\}} \rho_{\omega, h-1}(n - \ell)\right) \\ &= (1 - o(1))\mathbb{E}(\rho_{\omega, h}(n)) - \sum_{\ell \in \bigcup\{\mathfrak{R}_1, \dots, \mathfrak{R}_k\}} \mathbb{E}(\rho_{\omega, h-1}(n - \ell)). \end{aligned}$$

Since $A(n) \ll n^{1/(h-1)}$, we have by Lemma 4.3 that $\mathbb{E}(\rho_{\omega, h-1}(n)) = O(1)$, therefore the summation in the last equality above is bounded above by $k \cdot hJ$ for some fixed $J > 0$. Since h, J are fixed and $k \leq t\mathbb{E}(\rho_{\omega, h}(n))$, we finally arrive at

$$(5.13) \quad \Pr(\{\mathfrak{R}_1, \dots, \mathfrak{R}_k\} \text{ is maximal}) \leq K e^{-(1-\varepsilon_t - o(1))\mathbb{E}(\rho_{\omega, h}(n))}$$

for all k -sets $\{\mathfrak{R}_1, \dots, \mathfrak{R}_k\}$ considered, where $\varepsilon_t := t \cdot hJ$, which goes to 0 as $t \rightarrow 0^+$.

¹²This passage illustrates the necessity of considering $\text{ER}_h(n)|_{\geq n^{1/2}}$ instead of $\text{ER}_h(n)$.

Keeping in mind that, again, $k! > k^k e^{-k}$ for all $k \geq 1$, plugging (5.12) and (5.13) into what we originally wanted to estimate yields

$$\begin{aligned} \Pr(\exists \mathcal{C} \subseteq \text{ER}_h(n; \omega)|_{\geq n^{1/2}} \text{ maxdisfam} : |\mathcal{C}| \leq t\mathbb{E}(\rho_{\omega,h}(n))) \\ \leq K e^{-(1-\varepsilon_t-o(1))\mathbb{E}(\rho_{\omega,h}(n))} \sum_{1 \leq k \leq t\mathbb{E}(\rho_{\omega,h}(n))} \frac{\mathbb{E}(\rho_{\omega,h}(n))^k}{k!} \\ \leq K e^{-(1-\varepsilon'_t-o(1))\mathbb{E}(\rho_{\omega,h}(n))} \sum_{1 \leq k \leq t\mathbb{E}(\rho_{\omega,h}(n))} \left(\frac{\mathbb{E}(\rho_{\omega,h}(n))}{k} \right)^k, \end{aligned}$$

where $\varepsilon'_t := \varepsilon_t + t$, still going to 0 with t . As a function of k , one can show that the maximum of $(\mathbb{E}(\rho_{\omega,h}(n))/k)^k$ is attained when $k = \mathbb{E}(\rho_{\omega,h}(n))/e$. Thus, choosing $t < 1/e$, we have

$$\begin{aligned} \sum_{1 \leq k \leq t\mathbb{E}(\rho_{\omega,h}(n))} \left(\frac{\mathbb{E}(\rho_{\omega,h}(n))}{k} \right)^k &\leq t\mathbb{E}(\rho_{\omega,h}(n)) \cdot t^{-t\mathbb{E}(\rho_{\omega,h}(n))} \\ &\leq t\mathbb{E}(\rho_{\omega,h}(n)) \cdot e^{-t \log(t)\mathbb{E}(\rho_{\omega,h}(n))}, \end{aligned}$$

leading us to, in view of $\mathbb{E}(\rho_{\omega,h}(n)) \ll n^{\frac{1}{h-1}}$ (Lemma 4.3),

$$\Pr(\exists \mathcal{C} \subseteq \text{ER}_h(n; \omega)|_{\geq n^{1/2}} \text{ m.d.f.}^{13} : |\mathcal{C}| \leq t\mathbb{E}(\rho_{\omega,h}(n))) \ll n^{\frac{1}{h-1}} e^{-(1-\varepsilon''_t-o(1))\mathbb{E}(\rho_{\omega,h}(n))},$$

where $\varepsilon''_t := \varepsilon'_t - t \log(t)$, which once more goes to 0 with t .

Repeating the argument at the end of Part 1, we know that for some $\beta > 0$ independent of \mathcal{A} we have $\mathbb{E}(\rho_{\omega,h}(n)) > \beta M_1^h \log(n)$ for all large n . Therefore, taking $t > 0$ small and *choosing* $M_1 > 1$ large enough so that $(1 - \varepsilon''_t)\beta M_1^h > 2 + \frac{1}{h-1}$, we get

$$\Pr(\exists \mathcal{C} \subseteq \text{ER}_h(n; \omega)|_{\geq n^{1/2}} \text{ maxdisfam} : |\mathcal{C}| \leq t\mathbb{E}(\rho_{\omega,h}(n))) \ll n^{-2+o(1)},$$

from which we apply the Borel–Cantelli lemma to conclude that $\widehat{\rho}_{\omega,h}(n)|_{\geq n^{1/2}} \stackrel{\text{a.s.}}{\gg} A(n)^h/n$, thus concluding our proof. \square

Notice that in order to yield the conclusion of our previous result we only had to deal with M_1 , which was done at the very end. We will only deal directly with M_2 in Theorem 5.6. Since $\widehat{\rho}_{\omega,h}(n) \ll \rho_{\omega,h}(n)$, to prove that $\rho_{\omega,h}(n)$ has a.s. the right rate of growth we just need to show that

$$\rho_{\omega,h}(n) \stackrel{\text{a.s.}}{\ll} \widehat{\rho}_{\omega,h}(n),$$

and this is what the next two lemmas will be aiming at. This is immediate for $h = 2$, thus the assumption $h \geq 3$ will be implicit. From Lemma 5.1 we know that if $\mathcal{C}_n \subseteq \text{ER}_\ell(n; \omega)$ is a maximal disjoint family of exact h -representations of n then

$$(5.14) \quad \rho_{\omega,h}(n) \stackrel{\text{a.s.}}{\ll} \sum_{k \in \bigcup \mathcal{C}_n} \widehat{\rho}_{\omega,h-1}(n-k).$$

In this lemma, we use the fact that $\mathbb{E}(\rho_{\omega,h-1}(n)) = O(1)$ to deduce that $\widehat{\rho}_{\omega,h-1}(n) \stackrel{\text{a.s.}}{=} O(\log(n)/\log \log(n))$. This fact, however, is not sufficient to conclude from (5.14) the

¹³m.d.f. = maxdisfam.

statement we want.¹⁴ We can, however, draw this conclusion for certain partitions of \mathbb{N} . Letting $0 < \varepsilon < h^{-2}$, consider

$$\mathbf{P}_1 := \{n \in \mathbb{N} : A(n) \leq n^{1/(h-1)-\varepsilon}\} \text{ and } \mathbf{P}_2 := \mathbb{N} \setminus \mathbf{P}_1.$$

Before studying these partitions, let us first split $\text{ER}_h(n; \omega)$ in view of Lemma 4.4. We know that \mathcal{A} is OR_+ , thus its lower Matuszewska index $\mathcal{M}_*(A)$ is positive. With that in mind, fix some $0 < \gamma < \mathcal{M}_*(A)$ and take $\delta > 0$ with

$$(5.15) \quad \delta < \frac{\varepsilon(h-1)}{1-\gamma(h-1)}.$$

Notice that the RHS is positive, for since $A(x) \ll x^{1/(h-1)}$ one must have $\mathcal{M}_*(A) \leq 1/(h-1)$. Now, split $\rho_{\omega,h}$ into $\rho_{\omega,h}(n) = \rho_{\omega,h}(n)|_{\geq n^{1-\delta}} + \rho_{\omega,h}(n)|_{< n^{1-\delta}}$, with $\mathcal{C}_n|_{\geq n^{1-\delta}}$ and $\mathcal{C}_n|_{< n^{1-\delta}}$ being maxdisfams in each corresponding set. The next lemma shows that $\hat{\rho}_{\omega,h}$ will almost surely dominate $\rho_{\omega,h}$ in \mathbf{P}_1 .

Lemma 5.4. *Let “ \ll_1 ” denote “ \ll as $n \rightarrow +\infty$ through \mathbf{P}_1 ”. In the above notation, following Lemma 5.3 (i.e. having $M_1 > 1$ sufficiently large), if \mathbf{P}_1 is infinite then in $\mathcal{S}_{\mathcal{A}}$ the following holds:*

$$\rho_{\omega,h}(n) \stackrel{\text{a.s.}}{\ll_1} \hat{\rho}_{\omega,h}(n).$$

Proof. We divide the proof into two parts. First, we show that $\rho_{\omega,h}(n)|_{< n^{1-\delta}}$ does not contribute to the growth rate of $\rho_{\omega,h}(n)$.

• **Part 1:** Showing $\rho_{\omega,h}(n)|_{< n^{1-\delta}} \stackrel{\text{a.s.}}{=} O(\hat{\rho}_{\omega,h}(n)/\log \log(n))$.

Recall that this quantity is counting those representations in which at least one member is less than $n^{1-\delta}$. From (4.6) in Lemma 4.4,

$$\mathbb{E}(\rho_{\omega,h}(n)|_{< n^{1-\delta}}) \ll n^{-\delta\gamma/2^{h-2}} \mathbb{E}(\rho_{\omega,h}(n)),$$

thus, proceeding as in Part 1 from Lemma 5.3, we apply the disjointness lemma to estimate $|\bigcup \mathcal{C}_n|_{< n^{1-\delta}}|$. Keeping in mind that for $k \geq 1$ it holds $k! > k^k e^{-k}$, for any integer $T > 0$ it follows that

$$\begin{aligned} \Pr \left(\hat{\rho}_{\omega,h}(n)|_{< n^{1-\delta}} \geq T \left\lceil \frac{\hat{\rho}_{\omega,h}(n)}{\log(n)} \right\rceil \right) &\leq \frac{\left(n^{-\delta\gamma/2^{h-2}} \mathbb{E}(\rho_{\omega,h}(n)) \right)^{T \left\lceil \frac{\hat{\rho}_{\omega,h}(n)}{\log(n)} \right\rceil}}{(T \lceil \hat{\rho}_{\omega,h}(n)/\log(n) \rceil)!} \\ &\leq \left(\frac{e \mathbb{E}(\rho_{\omega,h}(n))}{T \hat{\rho}_{\omega,h}(n)} \frac{n^{-\delta\gamma/2^{h-2}}}{1/\log(n)} \right)^{T \left\lceil \frac{\hat{\rho}_{\omega,h}(n)}{\log(n)} \right\rceil}. \end{aligned}$$

Since $\hat{\rho}_{\omega,h}(n) \stackrel{\text{a.s.}}{=} \Theta(\mathbb{E}(\rho_{\omega,h}(n)))$ by Lemma 5.3 and $\mathbb{E}(\rho_{\omega,h}(n)) \gg \log(n)$ by Lemma 4.3, let $\vartheta > 0$ such that $\log(n)n^{-\delta\gamma/2^{h-2}} \ll n^{-\vartheta}$ and consider T large enough so as to cancel out all constant terms inside the parentheses. This will yield

$$\left(\frac{e \mathbb{E}(\rho_{\omega,h}(n))}{T \hat{\rho}_{\omega,h}(n)} \frac{n^{-\delta\gamma/2^{h-2}}}{1/\log(n)} \right)^{T \left\lceil \frac{\hat{\rho}_{\omega,h}(n)}{\log(n)} \right\rceil} \ll \left(\log(n)n^{-\delta\gamma/2^{h-2}} \right)^{T \left\lceil \frac{\hat{\rho}_{\omega,h}(n)}{\log(n)} \right\rceil}$$

¹⁴In fact, this $O(\log(n)/\log \log(n))$ seems to be a strong barrier to this method. Not having $\mathbb{E}(\rho_{\omega,h-1}(n)) \ll n^{-\varepsilon}$ is also the reason we avoided the methods from Tao & Vu [18] (particularly Theorem 1.37, p. 35).

$$\ll n^{-\vartheta T \left\lceil \frac{\widehat{\rho}_{\omega,h}(n)}{\log(n)} \right\rceil}.$$

Therefore, picking T large enough to also have $\vartheta T \lceil \widehat{\rho}_{\omega,h}(n)/\log(n) \rceil > 2$ from some point onward, we conclude that $\widehat{\rho}_{\omega,h}(n) \mathfrak{b}_{n^{1-\delta}} \stackrel{\text{a.s.}}{\ll} \widehat{\rho}_{\omega,h}(n)/\log(n)$ from the Borel–Cantelli lemma. Finally, the same reasoning behind (5.14) may be used to derive

$$\rho_{\omega,h}(n) \mathfrak{b}_{n^{1-\delta}} \stackrel{\text{a.s.}}{\ll} \sum_{k \in \bigcup \mathcal{C}_n \mathfrak{b}_{n^{1-\delta}}} \widehat{\rho}_{\omega,h-1}(n-k) \mathfrak{b}_{n^{1-\delta}},$$

thus, since $\widehat{\rho}_{\omega,h-1}(n) \stackrel{\text{a.s.}}{=} O(\log(n)/\log \log(n))$ from Lemma 5.1, we get:

$$\begin{aligned} \rho_{\omega,h}(n) \mathfrak{b}_{n^{1-\delta}} &\stackrel{\text{a.s.}}{\ll} \left| \bigcup \mathcal{C}_n \mathfrak{b}_{n^{1-\delta}} \right| \cdot \max_{k \in \bigcup \mathcal{C}_n \mathfrak{b}_{n^{1-\delta}}} \left(\widehat{\rho}_{\omega,h-1}(n-k) \mathfrak{b}_{n^{1-\delta}} \right) \\ &\stackrel{\text{a.s.}}{\ll} \frac{\widehat{\rho}_{\omega,h}(n)}{\log(n)} \cdot \frac{\log(n)}{\log \log(n)} = \frac{\widehat{\rho}_{\omega,h}(n)}{\log \log(n)}, \end{aligned}$$

as required.

• **Part 2:** Showing $\rho_{\omega,h}(n)|_{\geq n^{1-\delta}} \stackrel{\text{a.s.}}{\ll}_1 \widehat{\rho}_{\omega,h}(n)$.

Recall that $\delta = \delta(\varepsilon, \gamma)$, i.e. it depends on both ε and γ , where $0 < \gamma < \mathcal{M}_*(A)$. By the definition of $\mathcal{M}_*(A)$, we know there is $m > 0$ such that for large x , for every $\lambda \geq 1$ the following holds:

$$A(\lambda x) > m\lambda^\gamma A(x).$$

Rearranging this, for large x we have

$$\frac{A(x^{1-\delta})^{h-1}}{x^{1-\delta}} < \frac{1}{m^{h-1}} x^{\delta(1-\gamma(h-1))} \frac{A(x)^{h-1}}{x},$$

thus, in view of Lemma 4.3,

$$\mathbb{E}(\rho_{\omega,h-1}(n^{1-\delta})) \ll n^{\delta(1-\gamma(h-1))} \mathbb{E}(\rho_{\omega,h-1}(n)).$$

The partition \mathbf{P}_1 was chosen because $\mathbb{E}(\rho_{\omega,h-1}(n)) \ll_1 n^{-\varepsilon(h-1)}$, and thus

$$\mathbb{E}(\rho_{\omega,h-1}(n^{1-\delta})) \ll_1 n^{\delta(1-\gamma(h-1))-\varepsilon(h-1)}.$$

Therefore, by our choice of δ in (5.15), we have that for some $\vartheta > 0$ it holds:

$$\mathbb{E}(\rho_{\omega,h-1}(k)) \ll k^{-\vartheta} \text{ as } k \rightarrow +\infty \text{ through } \mathbf{P}_1 \cup \{[n^{1-\delta}, n] : n \in \mathbf{P}_1\},$$

Using the same strategy used in items (i)* and (ii)* at the proof of Lemma 5.1, our argument to bound $\rho_{\omega,h}(n)|_{\geq n^{1-\delta}}$ will be an application of the disjointness lemma followed by Borel–Cantelli. From disjointness, for any integer $T > 2/\vartheta$ we will have

$$\Pr(\widehat{\rho}_{\omega,h-1}(k) \geq T) \leq \frac{\mathbb{E}(\rho_{\omega,h-1}(k))^T}{T!} \ll k^{-2}$$

as $k \rightarrow +\infty$ through $\mathbf{P}_1 \cup \{[n^{1-\delta}, n] : n \in \mathbf{P}_1\}$; by the Borel–Cantelli lemma, this implies $\widehat{\rho}_{\omega,h-1}(k) \stackrel{\text{a.s.}}{=} O(1)$ as k varies throughout this domain.

Now, just as in (5.14), we have

$$(5.16) \quad \rho_{\omega,h}(n)|_{\geq n^{1-\delta}} \stackrel{\text{a.s.}}{\ll} \sum_{k \in \bigcup \mathcal{C}_n|_{\geq n^{1-\delta}}} \widehat{\rho}_{\omega,h-1}(n-k)|_{\geq n^{1-\delta}}.$$

Since the restrictions on the RHS imply that the corresponding sum is occurring only over $n - k \geq n^{1-\delta}$, and also since $\widehat{\rho}_{\omega,h-1}(k)$ is conveniently limited on the domain needed, it follows that

$$\rho_{\omega,h}(n)|_{\geq n^{1-\delta}} \stackrel{\text{a.s.}}{\ll}_1 \left| \bigcup \mathcal{C}_n|_{\geq n^{1-\delta}} \right|,$$

which is obviously $O(\widehat{\rho}_{\omega,h}(n))$. \square

It now only remains \mathbf{P}_2 for us to deal with. We will, however, not work with it in its entirety; instead of breaking the aforementioned “ $\log(n)/\log \log(n)$ barrier”, we will work around it. Our method is inspired by the idea of *r-star matchings* described in Subsection 3.2 of Warnke [21].

First, to make the notation in the next lemma less charged, let

$$L(n) := \frac{\log(n)}{\log \log(n)}.$$

Next, fix some $\tilde{\varepsilon} > 0$ and define

$$\mathbf{P}_{21} := \{n \in \mathbf{P}_2 : A(n) \leq n^{1/(h-1)} L(n)^{-\tilde{\varepsilon}}\} \text{ and } \mathbf{P}_{22} := \mathbf{P}_2 \setminus \mathbf{P}_{21}.$$

We now use the same $\gamma < \mathcal{M}_*(A)$ from before to take

$$(5.17) \quad \kappa > 2^{h-1}/\gamma,$$

and thus split $\rho_{\omega,h}$ into $\rho_{\omega,h}(n) = \rho_{\omega,h}(n)|_{\geq nL(n)^{-\kappa}} + \rho_{\omega,h}(n)|_{< nL(n)^{-\kappa}}$, with $\mathcal{C}_n|_{\geq nL(n)^{-\kappa}}$ and $\mathcal{C}_n|_{< nL(n)^{-\kappa}}$ being maxdisfams in each corresponding set. Our next lemma will show that, on the occasion of \mathbf{P}_{22} being finite, $\widehat{\rho}_{\omega,h}$ will almost surely dominate $\rho_{\omega,h}$ in \mathbf{P}_{21} , and thus in the whole \mathbf{P}_2 . The finiteness of \mathbf{P}_{22} is a key hypothesis for the second part of the proof.

Lemma 5.5. *Let “ \ll_2 ” denote “ \ll as $n \rightarrow +\infty$ through \mathbf{P}_2 ”. In the above notation, following Lemma 5.3 (i.e. having $M_1 > 1$ sufficiently large), if \mathbf{P}_{22} is finite then in $\mathcal{S}_{\mathcal{A}}$ the following holds:*

$$\rho_{\omega,h}(n) \stackrel{\text{a.s.}}{\ll}_2 \widehat{\rho}_{\omega,h}(n).$$

Proof. Just like in Lemma 5.4, we divide the proof into two parts, the first of which shows that $\rho_{\omega,h}(n)|_{< nL(n)^{-\kappa}}$ does not contribute to the growth rate of $\rho_{\omega,h}(n)$.

• **Part 1:** Showing $\rho_{\omega,h}(n)|_{< nL(n)^{-\kappa}} \stackrel{\text{a.s.}}{\ll}_2 \widehat{\rho}_{\omega,h}(n)/L(n)^{1/2}$.

Recall that this quantity is counting those representations in which at least one member is less than $nL(n)^{-\kappa}$. From (4.6) in Lemma 4.4,

$$\mathbb{E}(\rho_{\omega,h}(n)|_{< nL(n)^{-\kappa}}) \ll L(n)^{-\kappa\gamma/2^{h-2}} \mathbb{E}(\rho_{\omega,h}(n)),$$

From our choice of κ in (5.17), the exponent of $L(n)$ may be taken to be -2 . Proceeding as in Part 1 from Lemma 5.4, we apply the disjointness lemma to estimate $|\bigcup \mathcal{C}_n|_{< nL(n)^{-\kappa}}|$. Keeping in mind that for $k \geq 1$ it holds $k! > k^k e^{-k}$, for any integer $T > 0$ it then follows that

$$\Pr \left(\widehat{\rho}_{\omega,h}(n)|_{< nL(n)^{-\kappa}} \geq T \left\lceil \frac{\widehat{\rho}_{\omega,h}(n)}{L(n)^{3/2}} \right\rceil \right) \leq \frac{(L(n)^{-2} \mathbb{E}(\rho_{\omega,h}(n)))^{T \left\lceil \frac{\widehat{\rho}_{\omega,h}(n)}{L(n)^{3/2}} \right\rceil}}{(T \lceil \widehat{\rho}_{\omega,h}(n)/L(n)^{3/2} \rceil)!}$$

$$\leq \left(\frac{e}{T} \frac{\mathbb{E}(\rho_{\omega,h}(n))}{\widehat{\rho}_{\omega,h}(n)} \frac{L(n)^{-2}}{L(n)^{-3/2}} \right)^{T \left\lceil \frac{\widehat{\rho}_{\omega,h}(n)}{L(n)^{3/2}} \right\rceil}.$$

By Lemma 5.3 we have $\widehat{\rho}_{\omega,h}(n) \stackrel{\text{a.s.}}{=} \Theta(\mathbb{E}(\rho_{\omega,h}(n)))$. Hence, considering T large enough to cancel out the constant terms inside the parentheses, we have

$$\left(\frac{e}{T} \frac{\mathbb{E}(\rho_{\omega,h}(n))}{\widehat{\rho}_{\omega,h}(n)} \frac{L(n)^{-2}}{L(n)^{-3/2}} \right)^{T \left\lceil \frac{\widehat{\rho}_{\omega,h}(n)}{L(n)^{3/2}} \right\rceil} \ll L(n)^{-\frac{1}{2}T \left\lceil \frac{\widehat{\rho}_{\omega,h}(n)}{L(n)^{3/2}} \right\rceil}$$

From Lemma 4.3 and the definition of \mathbf{P}_2 , we know that $\mathbb{E}(\rho_{\omega,h}(n)) \gg_2 n^\alpha$ for some $\alpha > 0$. One can then deduce that the sum of this probabilities for $n \in \mathbf{P}_2$ will converge, and thus from the Borel–Cantelli lemma we may conclude that $\widehat{\rho}_{\omega,h}(n) \mathfrak{b}_{nL(n)^{-\kappa}} \stackrel{\text{a.s.}}{\ll}_2 \widehat{\rho}_{\omega,h}(n)/L(n)^{3/2}$.

Finally, once again the reasoning behind (5.14) may be used to deduce

$$\rho_{\omega,h}(n) \mathfrak{b}_{nL(n)^{-\kappa}} \stackrel{\text{a.s.}}{\ll} \sum_{k \in \bigcup \mathcal{C}_n \mathfrak{b}_{nL(n)^{-\kappa}}} \widehat{\rho}_{\omega,h-1}(n-k) \mathfrak{b}_{nL(n)^{-\kappa}},$$

hence, since $\widehat{\rho}_{\omega,h-1}(n) \stackrel{\text{a.s.}}{=} O(L(n))$ from Lemma 5.1, we arrive at

$$\begin{aligned} \rho_{\omega,h}(n) \mathfrak{b}_{nL(n)^{-\kappa}} &\stackrel{\text{a.s.}}{\ll} \left| \bigcup \mathcal{C}_n \mathfrak{b}_{nL(n)^{-\kappa}} \right| \cdot \max_{k \in \bigcup \mathcal{C}_n \mathfrak{b}_{nL(n)^{-\kappa}}} \left(\widehat{\rho}_{\omega,h-1}(n-k) \mathfrak{b}_{nL(n)^{-\kappa}} \right) \\ &\stackrel{\text{a.s.}}{\ll}_2 \frac{\widehat{\rho}_{\omega,h}(n)}{L(n)^{3/2}} \cdot L(n) = \frac{\widehat{\rho}_{\omega,h}(n)}{L(n)^{1/2}}, \end{aligned}$$

as required.

- **Part 2:** If \mathbf{P}_{22} is finite, then $\rho_{\omega,h}(n)|_{\geq nL(n)^{-\kappa}} \stackrel{\text{a.s.}}{\ll}_2 \widehat{\rho}_{\omega,h}(n)$.

First, just as in (5.16), we have

$$(5.18) \quad \rho_{\omega,h}(n)|_{\geq nL(n)^{-\kappa}} \stackrel{\text{a.s.}}{\ll} \sum_{k \in \bigcup \mathcal{C}_n|_{\geq nL(n)^{-\kappa}}} \rho_{\omega,h-1}(n-k)|_{\geq nL(n)^{-\kappa}}$$

Given a big integer $T \geq 1$, split $\bigcup \mathcal{C}_n|_{\geq nL(n)^{-\kappa}}$ into the following two sets:

$$\begin{aligned} \mathcal{A}_n^{(T)} &:= \left\{ k \in \bigcup \mathcal{C}_n|_{\geq nL(n)^{-\kappa}} : \rho_{\omega,h-1}(n-k)|_{\geq nL(n)^{-\kappa}} < T \right\} \\ \mathcal{B}_n^{(T)} &:= \left\{ k \in \bigcup \mathcal{C}_n|_{\geq nL(n)^{-\kappa}} : \rho_{\omega,h-1}(n-k)|_{\geq nL(n)^{-\kappa}} \geq T \right\} \end{aligned}$$

Now, consider the set

$$\widetilde{\mathcal{B}}_n^{(T)} := \bigcup_{k \in \mathcal{B}_n^{(T)}} \text{ER}_{h-1}(n-k; \omega),$$

and the following maxdisfam-type counting function:

$$\widehat{\beta}^{(T)}(n) := \max \left\{ |\mathcal{D}| : \mathcal{D} \subseteq \widetilde{\mathcal{B}}_n^{(T)} \text{ is a maximal disjoint family} \right\}$$

From Lemma 5.1 we know $r_{\omega,h-1}(n) \stackrel{\text{a.s.}}{=} O(L(n))$. Thus, letting $\mathcal{D} \subseteq \widetilde{\mathcal{B}}_n^{(T)}$ be a maximal disjoint family, we have

$$|\mathcal{B}_n^{(T)}| \leq T^{-1} \cdot |\widetilde{\mathcal{B}}_n^{(T)}|$$

$$\begin{aligned}
&\leq T^{-1} \cdot \sum_{k \in \bigcup \mathcal{D}} \rho_{\omega, h-1}(n-k) \\
(5.19) \quad &\stackrel{\text{a.s.}}{\ll} \widehat{\beta}^{(T)}(n) \cdot L(n).
\end{aligned}$$

Knowing that $A(n) \leq n^{1/(h-1)} L(n)^{-\tilde{\varepsilon}}$ for all but finitely many n in \mathbf{P}_2 , and also that $L(nL(n)^{-\kappa}) \sim L(n)$, in view of Lemma 4.3 and Remark 4.5 one can see that $\mathbb{E}(\rho_{\omega, h-1}(n-k)|_{\geq nL(n)^{-\kappa}}) < \eta \cdot L(n)^{-\tilde{\varepsilon}(h-1)}$ for all k when n is sufficiently large, where $\eta > 0$ is some constant not depending on \mathcal{A} . Thus, for such large n , since $T! > T^T e^{-T}$, by the disjointness lemma:

$$\begin{aligned}
\Pr\left(\rho_{\omega, h-1}(n-k)|_{\geq nL(n)^{-\kappa}} \geq T\right) &\leq \frac{1}{T!} (\eta \cdot L(n)^{-\tilde{\varepsilon}(h-1)})^T \\
&\leq \left(\frac{\eta \cdot e}{T}\right)^T L(n)^{-T\tilde{\varepsilon}(h-1)}.
\end{aligned}$$

Keeping in mind again that $\rho_{\omega, h-1}(n) \stackrel{\text{a.s.}}{\ll} L(n)$, we can estimate $|\widetilde{\mathcal{B}}_n^{(T)}|$ by

$$|\widetilde{\mathcal{B}}_n^{(T)}| \leq J \cdot L(n) \sum_{k \in \bigcup \mathcal{C}_n |_{\geq nL(n)^{-\kappa}}} \mathcal{I}_{\{\rho_{\omega, h-1}(n-k)|_{\geq nL(n)^{-\kappa}} \geq T\}},$$

for some constant J and all large n . Hence

$$\mathbb{E}(|\widetilde{\mathcal{B}}_n^{(j)}|) \leq c_T \cdot L(n)^{1-T\tilde{\varepsilon}(h-1)} \cdot \frac{A(n)^h}{n},$$

where $c_T := J \cdot (\eta e/T)^T$. With this, we can then apply the disjointness lemma again, this time to estimate $\widehat{\beta}^{(T)}(n)$. We have

$$\begin{aligned}
\Pr\left(\widehat{\beta}^{(T)}(n) \geq \left\lceil \frac{A(n)^h}{n} \frac{1}{L(n)^3} \right\rceil\right) &\leq \frac{(c_T \cdot L(n)^{1-T\tilde{\varepsilon}(h-1)} A(n)^h/n)^{\left\lceil \frac{A(n)^h}{n} \frac{1}{L(n)^3} \right\rceil}}{\lceil A(n)^h/n \cdot L(n)^{-3} \rceil!} \\
&\leq (c'_T \cdot L(n))^{(4-T\tilde{\varepsilon}(h-1)) \left\lceil \frac{A(n)^h}{n} \frac{1}{L(n)^3} \right\rceil},
\end{aligned}$$

where $c'_T = c_T \cdot e$. Let T be large enough so that $4 - T\tilde{\varepsilon}(h-1) < 0$. Since the function in the exponent is at least n^α for some $\alpha > 0$ in \mathbf{P}_2 , it is not hard to see that the sum of the above probabilities for $n \in \mathbf{P}_2$ will converge. Therefore, from (5.19) and the Borel–Cantelli lemma we deduce

$$(5.20) \quad |\mathcal{B}_n^{(T)}| \stackrel{\text{a.s.}}{\ll}_2 \frac{A(n)^h}{n} \frac{1}{L(n)^2}$$

Going back to (5.18), notice that from $\rho_{\omega, h-1}(n) \stackrel{\text{a.s.}}{\ll} L(n)$ we get

$$\begin{aligned}
\sum_{k \in \mathcal{B}_n^{(T)}} \widehat{\rho}_{\omega, h-1}(n-k)|_{\geq nL(n)^{-\kappa}} &\ll |\mathcal{B}_n^{(T)}| \cdot L(n) \\
&\stackrel{\text{a.s.}}{\ll}_2 \frac{A(n)^h}{n} \frac{1}{L(n)},
\end{aligned}$$

thus, since $\bigcup \mathcal{C}_n|_{\geq nL(n)^{-\kappa}} = \mathcal{A}_n^{(T)} \cup \mathcal{B}_n^{(T)}$,

$$\sum_{k \in \bigcup \mathcal{C}_n|_{\geq nL(n)^{-\kappa}}} \widehat{\rho}_{\omega, h-1}(n-k)|_{\geq nL(n)^{-\kappa}} \stackrel{\text{a.s.}}{\ll}_2 T \cdot |\mathcal{A}_n^{(T)}| + \frac{A(n)^h}{n} \frac{1}{L(n)}.$$

Finally, we know $|\mathcal{A}_n^{(T)}| = O(\widehat{\rho}_{\omega, h}(n))$, thus, since from Lemma 5.3 we have $\frac{A(n)^h}{n} \frac{1}{L(n)} \stackrel{\text{a.s.}}{=} o(\widehat{\rho}_{\omega, h}(n))$, we conclude that $\rho_{\omega, h}(n)|_{\geq nL(n)^{-\kappa}} \stackrel{\text{a.s.}}{\ll}_2 \widehat{\rho}_{\omega, h}(n)$. \square

It then follows immediately from (5.1) in Lemma 5.1 that, when \mathbf{P}_{22} is finite, $r_{\omega, h}(n)$ has a.s. the right rate of growth. Since $\tilde{\varepsilon}$ can be chosen arbitrarily small in the definition of \mathbf{P}_{21} , the case $\ell = 0$ of the Main Theorem follows immediately.¹⁵ The next result shall deal with the further cases. We actually show something a little more general: if one could prove that $r_{\omega, h}(n)$ has a.s. the right growth order regardless of \mathbf{P}_{22} being finite or not, then, provided M_1, M_2 are large enough, $r_{\omega, h+\ell}(n)$ will also concentrate around its mean for all $\ell \geq 0$.

Theorem 5.6. *Suppose there is $M^\dagger \in \mathbb{R}_+$ such that, in the construction of \mathcal{A} , if $M_1, M_2 > M^\dagger$ then in $\mathcal{S}_{\mathcal{A}}$ the following holds:*

$$\rho_{\omega, h}(n) \stackrel{\text{a.s.}}{\ll} \widehat{\rho}_{\omega, h}(n).$$

Then there must be $M^\ddagger \in \mathbb{R}_+$ such that $M_1, M_2 > M^\ddagger$ implies

$$r_{\omega, h+\ell}(n) \stackrel{\text{a.s.}}{=} \Theta\left(\frac{A(n)^{h+\ell}}{n}\right), \quad \forall \ell \geq 0$$

Proof. Let us start with the observation that if $t_1, t_2 \geq h$ are such that

$$r_{\omega, t_1}(n) \stackrel{\text{a.s.}}{=} \Theta\left(\frac{A(n)^{t_1}}{n}\right) \text{ and } r_{\omega, t_2}(n) \stackrel{\text{a.s.}}{=} \Theta\left(\frac{A(n)^{t_2}}{n}\right),$$

then $r_{\omega, t_1+t_2}(n) \stackrel{\text{a.s.}}{=} \Theta(A(n)^{t_1+t_2}/n)$. In fact, from the recursive formulas that can be derived from (1.1), it follows that

$$\begin{aligned} r_{\omega, t_1+t_2}(n) &= \sum_{k \leq n} r_{\omega, t_1}(k) r_{\omega, t_2}(n-k) \\ &\stackrel{\text{a.s.}}{\asymp} \sum_{1 \leq k \leq n-1} \frac{A(k)^{t_1}}{k} \frac{A(n-k)^{t_2}}{n-k} \\ &= \sum_{1 \leq k \leq n/2} \frac{A(k)^{t_1}}{k} \frac{A(n-k)^{t_2}}{n-k} + \sum_{n/2 < k \leq n-1} \frac{A(k)^{t_1}}{k} \frac{A(n-k)^{t_2}}{n-k}, \end{aligned}$$

which may be dealt with in the same way we treated (4.2). Hence, for instance, if $M_1, M_2 > M^\dagger$ then $r_{\omega, th}(n) \stackrel{\text{a.s.}}{=} \Theta(A(n)^{th}/n)$ for every $t \geq 1$, for the case $t = 1$ follows from our hypothesis. With this in mind, to prove our theorem it is sufficient to show that $r_{\omega, h+\ell}(n) \stackrel{\text{a.s.}}{=} \Theta(A(n)^{h+\ell}/n)$ for $0 \leq \ell < h$, thus the remaining cases will follow from this observation.

¹⁵More precisely, it follows from the fact that $\mathcal{S}_{\mathcal{A}}$ is a non-empty probability space.

We know this is true for $\ell = 0$, so we shall proceed by induction on ℓ . Assuming it is valid for some $\ell \in \{0, 1, \dots, h-2\}$, we will deduce the case $\ell + 1$. Start by noticing that

$$\begin{aligned}
 r_{\omega, h+\ell+1}(n) &= \sum_{k \leq n} r_{\omega, h+\ell}(k) \mathbb{1}_{\omega}(n-k) \\
 &\stackrel{\text{a.s.}}{\asymp} \sum_{1 \leq k \leq n} \frac{A(k)^{h+\ell}}{k} \mathbb{1}_{\omega}(n-k) \\
 (5.21) \quad &= \sum_{T < k \leq n-T} \frac{A(k)^{h+\ell}}{k} \mathbb{1}_{\omega}(n-k) + O\left(\frac{A(n)^{h+\ell}}{n}\right),
 \end{aligned}$$

where $T > 0$ is some arbitrarily large constant to be determined. The main term from (5.21) is a sum of independent random variables bounded by $\max_{T < k \leq n-T} \{A(k)^{h+\ell}/k\}$, therefore, letting

$$X_n^{(\ell)} := \sum_{T < k \leq n-T} \frac{A(k)^{h+\ell}}{k} \mathbb{1}_{\omega}(n-k),$$

we can apply the Chernoff bounds stated in Section 2 to obtain:

$$(5.22) \quad \Pr\left(\left|\frac{X_n^{(\ell)}}{\mathbb{E}(X_n^{(\ell)})} - 1\right| \geq \frac{1}{2}\right) \leq 2e^{-\frac{1}{16}\mathbb{E}(X_n^{(\ell)})\left(\max_{T < k \leq n-T} \frac{A(k)^{h+\ell}}{k}\right)^{-1}}$$

Here enters our final step: to bound from below the absolute value of the exponent in (5.22). Since we know from the Lemma 4.2 that $\mathbb{E}(X_n^{(\ell)}) = \Theta(A(n)^{h+\ell+1}/n)$, it suffices to show this exponent is at least $2\log(n) + o(\log(n))$, then we may apply the Borel–Cantelli lemma and our theorem will follow.

We start by bounding $\mathbb{E}(X_n^{(\ell)})$ from below. We have

$$\begin{aligned}
 \mathbb{E}(X_n^{(\ell)}) &= \sum_{T < k \leq n-T} \frac{A(k)^{h+\ell}}{k} \mathbb{E}(\mathbb{1}_{\omega}(n-k)) \\
 &\geq \sum_{n/2 < k \leq n-T} \frac{A(k)^{h+\ell}}{k} \frac{A(n-k)}{n-k+1} \\
 &\geq \frac{A(n/2)^{h+\ell}}{n} \sum_{n/2 < k \leq n-T} \frac{A(n-k)}{n-k+1}
 \end{aligned}$$

Since \mathcal{A} is OR_+ , from Lemma A.7 we know there is some constant $d > 0$ for which

$$\mathbb{E}(X_n^{(\ell)}) \geq d \frac{A(n)A(n/2)^{h+\ell}}{n} \quad (\forall \text{ large } n).$$

Note that d does not depend on ℓ . More importantly, it also does not depend on T , as can be seen from (5.21). Now it remains us to bound the “max” term from above. Firstly, since $A(n)$ is non-decreasing, we have

$$\left(\max_{T < k \leq n-T} \frac{A(k)^{h+\ell}}{k}\right) \leq A(n)^{\ell+1} \cdot \left(\max_{T < k \leq n-T} \frac{A(k)^{h-1}}{k}\right).$$

Since $A(n) \leq M_2^{-1} n^{1/(h-1)}$ for large n , we may take T large enough so that

$$\left(\max_{T < k \leq n-T} \frac{A(k)^{h-1}}{k} \right) < \frac{2}{M_2^{h-1}}.$$

In view of these estimates, we then have, for all large n ,

$$(5.23) \quad \frac{1}{16} \mathbb{E}(X_n^{(\ell)}) \left(\max_{T < k \leq n-T} \frac{A(k)^{h+\ell}}{k} \right)^{-1} > d \frac{M_2^{h-1}}{32} \frac{A(n/2)^\ell}{A(n)^\ell} \frac{A(n/2)^h}{n}.$$

There are two problematic terms in this lower bound: $A(n/2)^\ell/A(n)^\ell$ & $A(n/2)^h/n$. For the first one we have to go back to our original OR_+ sequence \mathcal{C} . Since \mathcal{C} is in particular OR , we know there exists some small $\vartheta > 0$ for which $C(n/2)/C(n) > \vartheta$ for all large n . We also know from (5.9) that for every $\varepsilon > 0$ there is some $n_\varepsilon \in \mathbb{N}$ such that, for all $n > n_\varepsilon$,

$$\begin{aligned} C(n/2) &< (q_2 M_2 + \varepsilon) A(n/2) \\ C(n) &> \left(\frac{q_1}{M_1} - \varepsilon \right) A(n), \end{aligned}$$

being q_1, q_2 the constants from the construction of \mathcal{C} . Putting these two equations together we get that

$$\frac{A(n/2)}{A(n)} > \frac{q_1}{q_2} \frac{\vartheta}{M_1 M_2}$$

for all large n , thus, since ℓ ranges from 0 to $h-2$, we have the first problematic term bounded from below by $(M_1 M_2)^{-(h-2)} \left(\frac{q_1}{q_2} \vartheta \right)^{h-2}$. The lower bound to the second term being considered, $A(n/2)^h/n$, comes from the fact that $A(n) \geq M_1 n^{1/h} \log(n)^{1/h}$ for large n . This implies

$$\frac{A(n/2)^h}{n} \geq \frac{M_1^h}{2} \log(n) + O(1).$$

Finally, (5.23) then turns into

$$(5.24) \quad d \frac{M_2^{h-1}}{32} \frac{A(n/2)^\ell}{A(n)^\ell} \frac{A(n/2)^h}{n} > M_1^2 M_2 \frac{d}{64} \left(\frac{q_1}{q_2} \vartheta \right)^{h-2} \log(n) + O(1).$$

In conclusion, having M_1, M_2 large enough so that

$$M_1^2 M_2 > \frac{2}{\frac{d}{64} \left(\frac{q_1}{q_2} \vartheta \right)^{h-2}}$$

will imply that (5.23) is greater than $2 \log(n) + o(\log(n))$, so we may finally apply the Borel–Cantelli lemma in (5.22) to obtain, in view of (5.21),

$$r_{\omega, h+\ell+1}(n) \stackrel{\text{a.s.}}{\asymp} \frac{A(n)^{h+\ell+1}}{n},$$

and this concludes our proof. \square

Since the countable intersection of events with probability 1 also has probability 1, our Main Theorem follows from Lemma 5.4, Lemma 5.5, Theorem 5.6 and the fact that $\mathcal{S}_{\mathcal{A}}$ is non-empty.

6. FURTHER REMARKS

Remark 6.1. On the conditions of Theorem 5.6, one can extend the result from Lemma 5.1 by showing that

$$\binom{h+\ell}{2} \sum_{k \leq n/2} r_{\omega, h+\ell-2}(n-2k) \mathbb{1}_{\omega}(k) \stackrel{\text{a.s.}}{\ll} A(n)^{\ell} \frac{\log(n)}{\log \log(n)},$$

where the LHS bounds the number of non-exact $(h+\ell)$ -representations of n from above. Since we have $r_{\omega, h+\ell}(n) \stackrel{\text{a.s.}}{\gg} A(n)^{\ell} \log(n)^{1+\ell/h}$, it then follows that $\rho_{\omega, h+\ell}(n) \stackrel{\text{a.s.}}{\asymp} r_{\omega, h+\ell}(n)$. This fact implies that the Main Theorem must also apply to exact representation functions. Originally, Erdős and Tetali [6] worked exclusively with $\rho_{\omega, h}$.

Remark 6.2. Given two sequences $\mathcal{A}, \mathcal{L} \subseteq \mathbb{N}$, one may define their *composition* as $\mathcal{L}[\mathcal{A}] := \{\ell_{a_0}, \ell_{a_1}, \ell_{a_2}, \dots\}$. Since $|\mathcal{L}[\mathcal{A}] \cap [0, x]| = A(L(x) - 1)$, we can define, in view of Step II (Section 4), the *space of \mathcal{L} -subsequences* $\mathcal{L}[\mathcal{S}_{\mathcal{A}}]$ as being $\langle \mathcal{S}, (\alpha_n)_n \rangle$ with

$$\alpha_n := \frac{A(L(n) - 1)}{L(n)} \mathbb{1}_{\mathcal{L}}(n).$$

This is just a more direct way of considering $\mathcal{L}[\omega]$ with $\omega \in \mathcal{S}_{\mathcal{A}}$. Since \mathcal{A} is OR_+ , by the strong law of large numbers (Theorem 3.1) we have $|\omega \cap [0, x]| \stackrel{\text{a.s.}}{\asymp} A(L(x))$; in addition, $A(L(n))^{h \frac{r_{\mathcal{L}, h}(n)}{s_{\mathcal{L}, h}(n)}}$ is then the natural candidate for $\mathbb{E}(r_{\omega, h}(n))$.

It is a natural question to ask which kinds of regularity assumptions do we need to impose on \mathcal{L} to be able to draw nice conclusions from $\mathcal{L}[\mathcal{S}_{\mathcal{A}}]$. In the case of Waring bases, for example, one can find subbases with similar properties to those described by our Main Theorem (see Vu [19]).

APPENDIX A. O -REGULARITY IN SEQUENCES

In this appendix we motivate and characterize the concepts of OR , PI and OR_+ sequences. In spite of OR_+ sequences being an essential part of our proof, this section is only tangentially related to the main subject of our paper. In Propositions A.2, A.5 and A.8 one finds equivalent definitions of OR , PI and OR_+ sequences, resp.. In Subsection A.5 we show an application of these concepts in the non-probabilistic context of additive bases, culminating on the sketch of an essentially elementary proof that primes that split completely in a given number field constitute a basis.

Notation. Script letters $\mathcal{A}, \mathcal{B}, \mathcal{C} \dots$ denote *sequences*, which are infinite subset of \mathbb{N} . Roman capital letters $A, B, C \dots$ denote the *counting function* of the sequence represented by the corresponding script letter, as in $A(x) = |\mathcal{A} \cap [0, x]|$. Lowercase roman letters $a_0, a_1, a_2 \dots$ (resp. $b, c \dots$) denote the elements of \mathcal{A} (resp. $\mathcal{B}, \mathcal{C} \dots$) in order, with a_0 (resp. $b_0, c_0 \dots$) being its smallest element.

A.1. Generalities: from $A(x)$ to $s_{\mathcal{A}, h}(x)$. To motivate the study of certain regularity conditions on sequences, we need to state our goals clearly and get our priorities in order. First of all, one of the overarching goals in the study of additive bases is to *characterize all bases*. This is a problem with clear number-theoretical grounding, as in this direction we find the study of the *classical bases*: primes (Goldbach's conjecture), k th-powers (Waring's problem) and polygonal numbers (Cauchy–Fermat

polygonal number theorem). A thorough overview of these topics is covered by Nathanson [13].

One, however, might argue that this is more naturally understood as a *combinatorial* problem, in whichever sense of the word. In this direction we find all the research springing from *Sidon sequences*, such as the Erdős–Turán conjecture for additive bases and the very Erdős–Tetali theorem, as well as Erdős–Fuchs-type estimates for representation functions. The classical reference for this more abstract treatment of sequences is Halberstam & Roth [10], and the types of questions concerning us here are the ones outlined in the introduction of Chapter II, which we paraphrase:

- I. What is the relationship between the asymptotic behaviour of the counting function of a sequence and its representation functions?
- II. Which properties of classical bases used to derive that these are in fact bases are actually applicable to wider classes of sequences?
- III. Which techniques employed in the study of classical bases are generalizable to a broader context?

In this appendix we only deal with the representation functions coming from the formal series (1.1). Given a sequence $\mathcal{A} \subseteq \mathbb{N}$, let $n \in \mathbb{N}$ and $h \geq 2$ an integer. The following recursive formulas can be immediately deduced:

$$(A.1) \quad \begin{aligned} r_{\mathcal{A},h}(n) &= \sum_{k \leq n} r_{\mathcal{A},h-\ell}(k) r_{\mathcal{A},\ell}(n-k) \\ s_{\mathcal{A},h}(n) &= \sum_{k \leq n} r_{\mathcal{A},h-\ell}(k) s_{\mathcal{A},\ell}(n-k) = \sum_{k \leq n} s_{\mathcal{A},h-\ell}(k) r_{\mathcal{A},\ell}(n-k) \end{aligned}$$

For purposes of induction the case $\ell = 1$ is usually enough, and we denote $r_{\mathcal{A},1}(n)$ simply by $\mathbb{1}_{\mathcal{A}}(n)$. Subsection A.5 tries to provide an answer for questions II and III via a framework which mixes both Schnirelmann’s classical theory of sequences and Hardy–Littlewood’s circle method, while the remainder is dedicated to exploring question I and, from it, possible “classes of sequences” as mentioned in question II.

We start with the following lemma.

Lemma A.1. *For every sequence $\mathcal{A} \subseteq \mathbb{N}$,*

$$A(x/2)^h \leq s_{\mathcal{A},h}(x) \leq A(x)^h, \quad \forall h \geq 1.$$

Proof. The case $h = 1$ is clear, hence we proceed by induction. By the recursive formulas in (A.1),

$$\begin{aligned} s_{\mathcal{A},h}(n) &= \sum_{k \leq n} s_{\mathcal{A},h-1}(k) \mathbb{1}_{\mathcal{A}}(n-k) \\ &\leq s_{\mathcal{A},h-1}(n) \sum_{k \leq n} \mathbb{1}_{\mathcal{A}}(n-k) \\ &= s_{\mathcal{A},h-1}(n) \cdot A(n) \leq A(n)^h, \end{aligned}$$

and

$$s_{\mathcal{A},h}(n) \geq \sum_{n/2 \leq k \leq n} s_{\mathcal{A},h-1}(k) \mathbb{1}_{\mathcal{A}}(n-k)$$

$$\begin{aligned}
&\geq s_{\mathcal{A},h-1}(n/2) \sum_{n/2 \leq k \leq n} \mathbb{1}_{\mathcal{A}}(n-k) \\
&= s_{\mathcal{A},h-1}(n/2) \cdot A(n/2) \geq A(n/2)^h,
\end{aligned}$$

as required. \square

The reason why we initially focus on $s_{\mathcal{A},h}(n)$ instead of $r_{\mathcal{A},h}(n)$ is that, apart from being easier to deal with, it relates to $r_{\mathcal{A},h}(n)$ as being a sort of an “average” when one tries to employ the rough approximation

$$r_{\mathcal{A},h}(n) \approx \frac{s_{\mathcal{A},h}(n)}{n+1}.$$

Much of what we propose comes from this intuition. To put more precisely, we seek for the conditions necessary for this intuition to make sense.

A.2. OR sequences. Let us introduce some bits of regular variation theory. An extensive treatment on this topic can be found in Bingham, Goldie & Teugels [2]. We will only use the theory from Chapters 1 and 2.

Take $f : [x_0, +\infty) \rightarrow \mathbb{R}_+$ a positive real function. We say that f is

- *Slowly varying* if $f(\lambda x) \sim f(x)$ for all $\lambda > 0$;
- *Regularly varying* if $f(\lambda x) \sim \lambda^\rho f(x)$, for all $\lambda > 0$ and some $\rho \in \mathbb{R}$;
- *O-regularly varying* if $f(\lambda x) \asymp f(x)$ for all $\lambda > 0$.

The generality of these definitions lies on Karamata’s characterization theorem,¹⁶ which states that if $f(\lambda x) \sim g(\lambda)f(x)$ with $g(\lambda) \in (0, +\infty)$ for all $\lambda > 0$ and f is measurable, then f is regularly varying, i.e. there is $\rho \in \mathbb{R}$ such that $g(\lambda) = \lambda^\rho$. One can then promptly see how *O*-regular variation extends regular variation, and that is why we shall focus only on the former. We have mentioned in the introduction that a sequence $\mathcal{A} \subseteq \mathbb{N}$ is OR whenever $A(2x) = O(A(x))$. This is equivalent to A being an *O*-regularly varying function. To be more precise, we define an *OR sequence* to be a sequence satisfying any of the equivalent conditions of the following proposition.

Proposition A.2 (OR sequences). *Let $\mathcal{A} \subseteq \mathbb{N}$ be a sequence. The following are equivalent:*

- (i) A is *O-regularly varying*;
- (ii) $A(2x) = O(A(x))$;
- (iii) $s_{\mathcal{A},h}(x) = \Theta(A(x)^h)$ for all $h \geq 1$;
- (iv) $s_{\mathcal{A},h}(2x) = O(s_{\mathcal{A},h}(x))$ for all or at least some $h \geq 1$.¹⁷

Proof. (iii) \implies (iv) is clear, and (i) \implies (iii) follows from Lemma A.1.

We show now that (iv) \implies (ii). The case $h = 1$ is trivial, so suppose $h \geq 2$. From Lemma A.1 it follows $s_{\mathcal{A},h}(x) \leq A(x)^h \leq s_{\mathcal{A},h}(2x)$, therefore (iv) implies $s_{\mathcal{A},h}(x) \asymp A(x)^h$. By the recursive formulas (A.1),

$$A(2n)^h \ll s_{\mathcal{A},h}(2n)$$

¹⁶Theorem 1.4.1, p. 17 of Bingham et al. [2].

¹⁷By this we mean that if this holds for at least some $h \geq 1$, then it also holds for all $h \geq 1$. The same appears in items (iv) of Propositions A.5 and A.8.

$$\begin{aligned}
&= \sum_{k \leq 2n} r_{\mathcal{A},h-1}(k) s_{\mathcal{A},1}(2n-k) \\
&\leq \sum_{n < k \leq 2n} r_{\mathcal{A},h-1}(k) s_{\mathcal{A},1}(2n-k) \\
&\ll A(n) \sum_{n < k \leq 2n} r_{\mathcal{A},h-1}(k) \\
&\ll A(n) \cdot s_{\mathcal{A},h-1}(2n) \ll A(n) \cdot A(2n)^{h-1},
\end{aligned}$$

thus $A(2x) \ll A(x)$.

Finally, to show that (ii) \implies (i), note that $A(x)$ is non-decreasing, thus (ii) is equivalent to having $A(x) \asymp A(2x)$. Hence, for all $k \in \mathbb{Z}$,

$$A(x) \asymp A(2^k x),$$

which implies $A(x) \asymp A(\lambda x)$ for all $\lambda > 0$. \square

Keeping Lemma A.1 in mind, one can then say that OR sequences are the most well-behaved sequences in terms of growth order, for it is sufficient to have $A(x)$ in order to deduce the growth of $s_{\mathcal{A},h}$.

A.3. PI sequences. In spite of that, just assuming \mathcal{A} to be OR does not guarantee that a_n , when viewed as a function of n , is O -regularly varying. For this part we need another bit of regular variation theory. Still following Bingham et al. [2], a positive real function $f : [x_0, +\infty) \rightarrow \mathbb{R}_+$ is said to be *almost increasing* when there is $m > 0$ such that

$$f(y) \geq mf(x), \quad \forall y \geq x \geq x_0.$$

This is equivalent to saying that $f(x) \ll \inf_{y \geq x} f(y)$; *almost decreasing* functions can be defined in an analogous way, namely $f(x) \gg \sup_{y \geq x} f(y)$. Consider then the *upper* (\mathcal{M}^*) and *lower Matuszewska index* (\mathcal{M}_*) of f :

$$\begin{aligned}
(A.2) \quad \mathcal{M}^*(f) &:= \inf \{ \gamma \in \mathbb{R} : x^{-\gamma} f(x) \text{ is almost decreasing} \}, \\
\mathcal{M}_*(f) &:= \sup \{ \gamma \in \mathbb{R} : x^{-\gamma} f(x) \text{ is almost increasing} \}.
\end{aligned}$$

This definition is given in view of the almost-monotonicity theorem.¹⁸ These indices are preserved under the asymptotic sign “ \asymp ”, and both are finite if and only if f is O -regularly varying.¹⁹ Functions with $\mathcal{M}_*(f) > 0$ are said to have *positive increase*. Our motivation to study these indices comes from the next two lemmas. By “ $a_{[\cdot]}$ ” we shall be denoting the positive real function that takes $x \mapsto a_{[x]}$.

Lemma A.3. *For every sequence $\mathcal{A} \subseteq \mathbb{N}$,*

$$\mathcal{M}_*(A) = \frac{1}{\mathcal{M}^*(a_{[\cdot]})}, \quad \mathcal{M}^*(A) = \frac{1}{\mathcal{M}_*(a_{[\cdot]})};$$

adopting the conventions “ $1/0 = +\infty$ ” and “ $1/(+\infty) = 0$ ”.

¹⁸Theorem 2.2.2, p. 72 of Bingham et al. [2].

¹⁹Theorem 2.1.7, p. 71 of Bingham et al. [2].

Proof. We will only show that $\mathcal{M}_*(A) = 1/\mathcal{M}^*(a_{[\cdot]})$. The argument for the other equation is entirely analogous, but with opposite inequalities. First, suppose $\mathcal{M}_*(A) > 0$. Taking $0 < \gamma < \mathcal{M}_*(A)$, there exists $m > 0$ and some $x_0 \in \mathbb{R}_+$ for which

$$\frac{A(y)}{y^\gamma} \geq m \frac{A(x)}{x^\gamma}, \quad \forall y \geq x \geq x_0.$$

Hence, there must be some $n_0 \in \mathbb{N}$ such that

$$\frac{A(a_N)}{a_N^\gamma} \geq m \frac{A(a_n)}{a_n^\gamma}, \quad \forall N \geq n \geq n_0;$$

but since $A(a_k) = k + 1$,

$$\frac{a_N}{(N+1)^{1/\gamma}} \leq \frac{1}{m^{1/\gamma}} \frac{a_n}{(n+1)^{1/\gamma}}, \quad \forall N \geq n \geq n_0;$$

which means $a_n/n^{1/\gamma}$ is almost decreasing, thus $\mathcal{M}^*(a_{[\cdot]}) \geq 1/\mathcal{M}_*(A)$.

To see that strict inequality cannot hold, we now take $\gamma > \mathcal{M}_*(A)$. Note that now we are considering the possibility of having $\mathcal{M}_*(A) = 0$, for the next argument will also apply to the “0, $+\infty$ ” case. For this chosen γ , the function $A(x)/x^\gamma$ will not be almost increasing; that is, $\forall \varepsilon > 0, \forall M \in \mathbb{R}_+$ there are $y_\varepsilon > x_\varepsilon \geq M$ such that

$$\frac{A(y_\varepsilon)}{y_\varepsilon^\gamma} \leq \varepsilon \frac{A(x_\varepsilon)}{x_\varepsilon^\gamma}.$$

Letting $n_\varepsilon := A(x_\varepsilon)$ and $N_\varepsilon := A(y_\varepsilon) + 1$, the above inequality implies

$$\frac{a_{N_\varepsilon}}{(N_\varepsilon - 1)^{1/\gamma}} \geq \frac{1}{\varepsilon^{1/\gamma}} \frac{a_{n_\varepsilon}}{n_\varepsilon^{1/\gamma}},$$

which means $a_n/n^{1/\gamma}$ cannot be almost decreasing, thus $\mathcal{M}^*(a_{[\cdot]}) < 1/\gamma$. Hence equality must hold when $\mathcal{M}_*(A) > 0$, and $\mathcal{M}_*(A)$ vanishes iff $\mathcal{M}^*(a_{[\cdot]}) = +\infty$. \square

Lemma A.4. *For every sequence $\mathcal{A} \subseteq \mathbb{N}$ and every $h \geq 1$,*

$$\mathcal{M}_*(s_{\mathcal{A},h}) = h \cdot \mathcal{M}_*(A), \quad \mathcal{M}^*(s_{\mathcal{A},h}) = h \cdot \mathcal{M}^*(A).$$

Proof. As in Lemma A.3, we will only show the \mathcal{M}_* case, for the other is analogous but with opposite inequalities and “ $+\infty$ ” instead of “0”. First, suppose $\mathcal{M}_*(A) > 0$. Assume for the sake of contradiction that $\mathcal{M}_*(s_{\mathcal{A},h}) < h \cdot \mathcal{M}_*(A)$. This means that there is $\mathcal{M}_*(s_{\mathcal{A},h}) < \gamma < h \cdot \mathcal{M}_*(A)$ for which there is $(x_k)_{k \in \mathbb{N}}, (y_k)_{k \in \mathbb{N}}$ with $y_k > x_k \rightarrow +\infty$ as $k \rightarrow +\infty$ such that

$$\frac{s_{\mathcal{A},h}(y_k)}{s_{\mathcal{A},h}(x_k)} < \frac{1}{k} \left(\frac{y_k}{x_k} \right)^\gamma.$$

From Lemma A.1 we have

$$\frac{A(y_k/2)^h}{A(x_k)^h} \leq \frac{s_{\mathcal{A},h}(y_k)}{s_{\mathcal{A},h}(x_k)},$$

but, since $\gamma/h < \mathcal{M}_*(A)$, we conclude that $y_k/2 < x_k$ for all large k . Hence

$$\frac{s_{\mathcal{A},h}(y_k)}{s_{\mathcal{A},h}(x_k)} < \frac{2^\gamma}{k} \xrightarrow{k \rightarrow +\infty} 0,$$

a contradiction. Therefore $\mathcal{M}_*(A) > 0$ implies $\mathcal{M}_*(s_{\mathcal{A},h}) \geq h \cdot \mathcal{M}_*(A)$.

To see that strict inequality cannot hold, notice that $A(x)$ is always an increasing function and $A(x) \leq x + 1$, hence $0 \leq \mathcal{M}_*(A) \leq 1$. Take $\gamma > h \cdot \mathcal{M}_*(A)$, considering the possibility of having $\mathcal{M}_*(A) = 0$. Similarly to the previous case, there is $(x_k)_{k \in \mathbb{N}}$, $(y_k)_{k \in \mathbb{N}}$ with $y_k > x_k \rightarrow +\infty$ as $k \rightarrow +\infty$ for which

$$\frac{A(y_k)}{A(x_k)} < \frac{1}{k^{1/h}} \left(\frac{y_k}{x_k} \right)^{\gamma/h}.$$

From Lemma A.1 we have

$$\frac{s_{\mathcal{A},h}(y_k)}{s_{\mathcal{A},h}(x_k/2)} \leq \frac{A(y_k)^h}{A(x_k)^h},$$

therefore we may deduce that $\mathcal{M}_*(s_{\mathcal{A},h}) \leq \gamma$, as required. \square

Thus, we define a *PI sequence* to be a sequence satisfying any of the following equivalent conditions.

Proposition A.5 (PI sequences). *Let $\mathcal{A} \subseteq \mathbb{N}$ be a sequence. The following are equivalent:*

- (i) A has positive increase;
- (ii) $a_{2n} = O(a_n)$;
- (iii) $\liminf_{x \rightarrow +\infty} A(\lambda x)/A(x) > 1$ for some $\lambda > 1$;
- (iv) $s_{\mathcal{A},h}$ has positive increase for all or at least some $h \geq 1$.

Proof. (i) \iff (ii) follows from Lemma A.3 and the observation that a function is OR exactly when both its Matuszewska indices are finite, whereas (i) \iff (iv) follows from Lemma A.4.

For (iv) \implies (iii), notice that if $\mathcal{M}_*(s_{\mathcal{A},h}) > 0$ for some $h \geq 1$, then for $0 < \gamma < \mathcal{M}_*(s_{\mathcal{A},h})$ there is $m > 0$ and $x_0 \in \mathbb{R}_+$ such that, for all $\mu > 1$,

$$\frac{s_{\mathcal{A},h}(\mu x)}{s_{\mathcal{A},h}(x)} \geq m\mu^\gamma, \quad \forall x \geq x_0.$$

From Lemma A.1, we know that

$$\frac{A(\mu x)^h}{A(x/2)^h} \geq \frac{s_{\mathcal{A},h}(\mu x)}{s_{\mathcal{A},h}(x)},$$

hence

$$\frac{A(2\mu x)}{A(x)} \geq m^{1/h} \mu^{\gamma/h}, \quad \forall x \geq x_0/2.$$

Thus, if μ is large enough, we have $\liminf_{x \rightarrow +\infty} A(\lambda x)/A(x) > 1$ for $\lambda = 2\mu$.

Finally, to show (iii) \implies (i), assume that (i) is false. Then for every $\varepsilon > 0$ there is a sequence $(n_k)_{k \in \mathbb{N}}$ with $n_k = n_k(\varepsilon) \rightarrow +\infty$ as $k \rightarrow +\infty$ for which

$$\frac{a_{\lfloor (1+\varepsilon)n_k \rfloor}}{a_{n_k}} > k.$$

This is the same as $ka_{n_k} < a_{\lfloor (1+\varepsilon)n_k \rfloor}$. Since $A(a_k) = k + 1$, it follows that

$$\frac{A(ka_{n_k})}{A(n_k)} < \frac{\lfloor (1+\varepsilon)n_k \rfloor + 1}{n_k + 1} \leq 1 + \varepsilon + \frac{1 - \varepsilon}{n_k + 1};$$

that is, for every real number $\lambda > 1$,

$$\frac{A(\lambda n_k)}{A(n_k)} < 1 + \varepsilon + o(1).$$

as $k \rightarrow +\infty$. Since for every $\varepsilon > 0$ there is such $(n_k)_{k \in \mathbb{N}}$, we conclude that for every $\lambda > 1$ it holds that $\liminf_{x \rightarrow +\infty} A(\lambda x)/A(x) = 1$, the negation of which is exactly item (iii). \square

Remark A.6 (Criteria for non-bases). We describe shortly another motivation for PI sequences. In Section 7 of Stöhr [17] it is presented a list of criteria for non-bases, i.e. criteria which if a sequence $\mathcal{A} \subseteq \mathbb{N}$ satisfies it, then it cannot be an additive basis. Among these, we highlight the following two:

- *Criterion #1*: $A(x) \not\gg x^\varepsilon$ for every $\varepsilon > 0$.
- *Criterion #2*: $\limsup_{n \rightarrow +\infty} a_{n+k}/a_n = +\infty$ for some $k \geq 1$.

When studying bases, it is then natural to focus only on sequences that avoid such criteria. A simple way of doing this would be to consider a condition that simultaneously avoid the above two statements. We may then consider the following questions:

- (1) What is the smallest f for which if $A(n) \gg f(n)$ then \mathcal{A} does not satisfy criterion #2?
- (2) What is the smallest f for which if $\limsup_{n \rightarrow +\infty} a_{n+f(n)}/a_n < +\infty$ then \mathcal{A} does not satisfy criterion #1?

It is an exercise to show that, surprisingly enough, the answer to both questions is $f(n) = \Theta(n)$. For question (1), it has been long known that a sequence $\mathcal{A} \subseteq \mathbb{N}$ with $A(n) = \Theta(n)$ is an additive bases if and only if it satisfies the trivial requirement $\{0, 1\} \subseteq \mathcal{A}$; this is indeed equivalent to a celebrated theorem by Schnirelmann (cf. Subsection A.5).

For question (2), on the other hand, we may see that taking $f(n) = \Theta(n)$ implies item (ii) of Proposition A.5, therefore we arrive back at PI sequences. This is a much more general condition than $A(n) = \Theta(n)$, but not all bases are necessarily PI (cf. Remark A.9). An interesting line of research would be to identify which PI sequences are not bases.

A.4. OR_+ sequences. Finally we arrive at the central point of this appendix, which are sequences that are OR and PI at the same time. The reason behind considering these two conditions simultaneously comes from the following lemma, which is of central importance to this paper.

Lemma A.7 (OR-PI lemma). *Let $f : [x_0, +\infty) \rightarrow \mathbb{R}_+$ be a positive and locally integrable real function. The following are equivalent:*

- (i) $\int_{x_0}^x \frac{f(t)}{t} dt = \Theta(f(x))$;
- (ii) $0 < \mathcal{M}_*(f) \leq \mathcal{M}^*(f) < +\infty$;
- (iii) f is O -regularly varying and has positive increase.

Proof. Corollary 2.6.2, p. 96 of Bingham et al. [2]. \square

We define an OR_+ sequence to be a sequence satisfying any of the equivalent conditions of the following proposition. The “+” in OR_+ is used to indicate the positivity of both Matuszewska indices.

Proposition A.8 (OR_+ sequences). *Let $\mathcal{A} \subseteq \mathbb{N}$ be a sequence. The following are equivalent:*

- (i) A is O -regularly varying and has positive increase;
- (ii) $\int_1^x \frac{A(t)}{t} dt = \Theta(A(x))$;
- (iii) Both $A(2x) = O(A(x))$ and $a_{2n} = O(a_n)$;
- (iv) $\int_1^x \frac{s_{\mathcal{A},h}(t)}{t} dt = \Theta(s_{\mathcal{A},h}(x))$ for all or at least some $h \geq 1$.

The proof is just a straightforward combination of Proposition A.2, Proposition A.5 and Lemma A.7.

Remark A.9 (Counterexamples). Almost all common examples of additive bases are OR_+ , such as the sequence of primes and sequences generated by polynomials, which even have regularly varying counting functions. Nonetheless, not all bases are common! For instance, the sequence

$$\mathcal{C} := [0, 2^{2^{17}}] \cup \left(\bigcup_{k \geq 17} \left\{ 2^{2^k} + t(k-1) : 0 \leq t \leq 2^{2^k} \right\} \cup [k \cdot 2^{2^k}, 2^{2^{k+1}}] \right)$$

is an example of a 2-basis, i.e. a basis of order 2, which is not PI. It is a relatively simple exercise to show that $2\mathcal{C} = \mathbb{N}$; to see that it is not PI, let n_k be such that $c_{n_k} = 2^{2^k}$. One can then show that $n_k \sim 2^{2^k}$ as $k \rightarrow +\infty$, and thus $c_{2n_k}/c_{n_k} = k - o(k)$, implying $c_{2n} \neq O(c_n)$, which by the last lemma yields $\mathcal{M}_*(C) = 0$.

On the other hand, the sequence

$$\mathcal{D} := \mathcal{P}_3 \cup \left\{ (2^{2^k})^3 + t : 0 \leq t \leq (2^{2^k})^2, k \in \mathbb{N} \right\},$$

where \mathcal{P}_3 is the sequence of cubes, is not OR . It is a basis, for it contains \mathcal{P}_3 , which is well-known to be a 9-basis;²⁰ to see it is not OR , let $n_k = (2^{2^k})^3$. It is not hard to show that $D(n_k) = o(n_k^{2/3})$ as $k \rightarrow +\infty$, and then

$$\frac{D(2n_k)}{D(n_k)} \geq \frac{D(n_k + n_k^{2/3})}{D(n_k)} = 1 + \frac{n_k^{2/3}}{D(n_k)},$$

which diverges, implying $D(2x) \neq O(D(x))$ and yielding $\mathcal{M}^*(D) = +\infty$.

A.5. An application to additive bases. We finish this appendix by showing a relatively general version of what is sometimes loosely referred to in the literature as *Schnirelmann’s method*. This refers to the common element between the elementary proofs of the sequence of primes with $\{0, 1\}^{21}$ and the sequence of k -powers for all $k \geq$

²⁰Wieferich–Kempner theorem (cf. Chapter 2 of Nathanson [13]).

²¹Schnirelmann–Goldbach theorem (cf. Chapter 7 of Nathanson [13]).

¹²² constituting additive bases, with “elementary” meaning avoiding the complex analytical approach of the Hardy–Littlewood circle method. Although seemingly paradoxically, we believe that our idea is most concisely stated on the context of Vinogradov’s form of the circle method. This part is of complete independent interest from the main subject of this paper, although some analogies can certainly be drawn as we shall point out. We now introduce the necessary concepts and notation.

Fix $\mathcal{A} \subseteq \mathbb{N}$ a sequence. For $\alpha \in \mathbb{R}/\mathbb{Z}$, let $e(\alpha) := e^{2\pi i \alpha}$. This is well-defined, for the value of $e^{2\pi i x}$ for $x \in \mathbb{R}$ depends only on the fractional part of x . Consider then the following truncated generating-type function:

$$(A.3) \quad \mathcal{G}_{\mathcal{A}}^{(n)}(\alpha) := \sum_{k \leq n} \mathbb{1}_{\mathcal{A}}(k) e(k\alpha).$$

The functions $\{e(k\alpha)\}_{k \in \mathbb{Z}}$ form an orthonormal basis of $L^2(\mathbb{R}/\mathbb{Z})$. From this orthonormality relation it follows, similarly to the formal series in (1.1),

$$(A.4) \quad r_{\mathcal{A},h}(n) = \int_{\mathbb{R}/\mathbb{Z}} \mathcal{G}_{\mathcal{A}}^{(n)}(\alpha)^h e(-n\alpha) d\alpha.$$

This is the starting point of Vinogradov’s form of the circle method. Following mainly Chapters 5 and 8 of Nathanson [13], we describe something which may be thought of as a general framework of the method. In both Waring’s and Goldbach’s ternary representation problem, a similar sort of *ansatz* is employed. The idea is to first consider the following “averaged” form of (A.3):

$$(A.5) \quad \tilde{\mathcal{G}}_{\mathcal{A}}^{(n)}(\alpha) := \sum_{k \leq n} \frac{A(k)}{k+1} e(k\alpha).$$

This is similar to what we did in the construction of $\mathcal{S}_{\mathcal{A}}$. Even though we are not necessarily requiring \mathcal{A} to be OR_+ , our goal is to show that the OR_+ condition fits naturally into this framework. Indeed, the next step is to approximate (A.4) by

$$(A.6) \quad J_{\mathcal{A},h}(n) := \int_{\mathbb{R}/\mathbb{Z}} \tilde{\mathcal{G}}_{\mathcal{A}}^{(n)}(\alpha)^h e(-n\alpha) d\alpha,$$

the so-called *singular integral* in the context of the usual circle method. Again, the orthonormality of $\{e(k\alpha)\}_{k \in \mathbb{Z}}$ implies that

$$(A.7) \quad J_{\mathcal{A},h}(n) = \sum_{\substack{k_1 + \dots + k_h = n \\ k_1, \dots, k_h \geq 0}} \frac{A(k_1)}{k_1 + 1} \dots \frac{A(k_h)}{k_h + 1}.$$

When \mathcal{A} is OR_+ one can show that $J_{\mathcal{A},h}(n) = \Theta(A(n)^h/n)$ for all $h \geq 1$; this is exactly equivalent to Lemma 4.2 in the main text, which we called the *fundamental lemma*. Finally, one aims to arrive at

$$(A.8) \quad r_{\mathcal{A},h}(n) = \mathfrak{S}_{\mathcal{A},h}(n) J_{\mathcal{A},h}(n) + \text{error},$$

where $\mathfrak{S}_{\mathcal{A},h}$ is some arithmetic function and the error term is at most $o(r_{\mathcal{A},h}(n))$. It is worth noting that *every sequence* $\mathcal{A} \subseteq \mathbb{N}$ has such a representation for $r_{\mathcal{A},h}$ for every $h \geq 1$; i.e. we can always choose $\mathfrak{S}_{\mathcal{A},h}$ such that $r_{\mathcal{A},h}$ satisfies (A.8), the trivial choice being $\mathfrak{S}_{\mathcal{A},h}(n) = r_{\mathcal{A},h}(n)/J_{\mathcal{A},h}(n)$ when the denominator is greater than 0,

²²Linnik’s solution to Waring’s problem (cf. Chapter 2 of Gel’fond & Linnik [9]).

and $\mathfrak{S}_{\mathcal{A},h}(n) = 1$ otherwise. Furthermore, by the way we defined the error term, it is always the case that

$$(A.9) \quad \mathfrak{S}_{\mathcal{A},h}(n) \sim \frac{r_{\mathcal{A},h}(n)}{J_{\mathcal{A},h}(n)} \text{ as } n \rightarrow +\infty \text{ through } h\mathcal{A}.$$

Remark A.10 (Relation to Main Theorem). One way to interpret the conclusion after Theorem 5.6 is that, for suitable \mathcal{A} , in $\mathcal{S}_{\mathcal{A}}$ one has $\mathfrak{S}_{\omega,h}(n) \stackrel{\text{a.s.}}{=} \Theta(1)$.

In both Waring’s and Goldbach’s problem, the idea of the method centers around finding a good description to $\mathfrak{S}_{\mathcal{A},h}$ (the so-called *singular series*) and controlling the error term. For that aim, one partitions \mathbb{R}/\mathbb{Z} into *major* and *minor arcs* for each sufficiently large n , usually denoted by $\mathfrak{M} = \mathfrak{M}(n)$ and $\mathfrak{m} = \mathfrak{m}(n)$ resp.. This partition is constructed in a way such that the integral in (A.4) over the minor arcs constitutes part of the error term, while over the major arcs one seeks for a representation of $r_{\mathcal{A},h}$ as in (A.8).

The method as generally employed, however, draws heavily upon the number-theoretical properties of the sequences being considered, demonstrated by the use of Hua’s lemma and Weyl’s inequality in the case of Waring, and specific estimates of exponential sums over primes together with non-trivial estimates for the error term in Dirichlet’s prime number theorem in the case of ternary Goldbach.

If on the other hand we dispense with the precision achieved by the method in the case of highly number-theoretical sequences and focus only on the more relaxed property of just being an additive basis, we find that through OR_+ sequences it is possible to marry our description of the circle method with the elementary approach of Schnirelmann’s method. To properly state this result, we only need two more things. The first one was already mentioned in Remark A.6.

Schnirelmann’s theorem (Theorem 4, p. 8 of Halberstam & Roth [10]). *If a sequence $\mathcal{A} \subseteq \mathbb{N}$ contains $\{0, 1\}$ and satisfies $A(x) = \Theta(x)$, then \mathcal{A} is an additive basis.*

The second one is the notion of *stability* in additive bases as described in Section 11.2 of Nathanson [14]. A sequence $\mathcal{A} \subseteq \mathbb{N}$ is said to be a *stable basis* when \mathcal{A} is a basis and every subsequence $\mathcal{B} \subseteq \mathcal{A}$ with $\{0, 1\} \subseteq \mathcal{B}$ and $B(x) = \Theta(A(x))$ is also a basis. Without further ado:

Theorem A.11. *Let $\mathcal{A} \subseteq \mathbb{N}$ be an OR_+ sequence with $\{0, 1\} \subseteq \mathcal{A}$. If there is $\varepsilon > 0$ such that*

$$(A.10) \quad \frac{1}{x} \sum_{n \leq x} \mathfrak{S}_{\mathcal{A},h}(n)^{1+\varepsilon} = O(1)$$

in (A.8) for some $h \geq 1$, then \mathcal{A} is a stable additive basis.

Proof. From equations (A.7) and (A.9), the OR_+ condition implies that (A.10) is equivalent to $\sum_{n \leq x} (n \cdot r_{\mathcal{A},h}(n)/A(n)^h)^{1+\varepsilon} = O(x)$. Recall that \mathcal{A} is, in particular, PI. By item (iii) of Proposition A.5, there is $\lambda > 1$ and $\delta > 0$ for which $A(\lambda x)/A(x) > 1 + \delta$ for all large x , therefore

$$A(x)^{(1+\varepsilon)h/\varepsilon} < \delta^{-(1+\varepsilon)h/\varepsilon} (A(\lambda x) - A(x))^{(1+\varepsilon)h/\varepsilon}$$

$$\begin{aligned}
&= \delta^{-(1+\varepsilon)h/\varepsilon} \left(\sum_{x < n \leq \lambda x} \mathbb{1}_{\mathcal{A}}(n) \right)^{(1+\varepsilon)h/\varepsilon} \\
&\leq \delta^{-(1+\varepsilon)h/\varepsilon} \left(\sum_{hx < n \leq h\lambda x} r_{\mathcal{A},h}(n) \right)^{(1+\varepsilon)/\varepsilon}.
\end{aligned}$$

We know that the sequence \mathcal{A} is also, in particular, OR. Therefore, by Hölder's inequality,²³ since $1 = 1/(1+\varepsilon) + \varepsilon/(1+\varepsilon)$,

$$\begin{aligned}
\left(\sum_{hx < n \leq h\lambda x} r_{\mathcal{A},h}(n) \right)^{(1+\varepsilon)/\varepsilon} &\leq |h\mathcal{A} \cap [hx, h\lambda x]| \left(\sum_{hx < n \leq h\lambda x} r_{\mathcal{A},h}(n)^{1+\varepsilon} \right)^{1/\varepsilon} \\
&\leq |h\mathcal{A} \cap [hx, h\lambda x]| \left(\frac{A(h\lambda x)^h}{hx} \right)^{(1+\varepsilon)/\varepsilon} \left(\sum_{hx < n \leq h\lambda x} \left(n \frac{r_{\mathcal{A},h}(n)}{A(n)^h} \right)^{1+\varepsilon} \right)^{1/\varepsilon} \\
&\ll |h\mathcal{A} \cap [hx, h\lambda x]| \frac{A(x)^{(1+\varepsilon)h/\varepsilon}}{x},
\end{aligned}$$

hence $|h\mathcal{A} \cap [0, x]| = \Theta(x)$. By Schnirelmann's theorem, we deduce that \mathcal{A} is a basis. To see that \mathcal{A} is a stable basis, notice that if $\mathcal{B} \subseteq \mathcal{A}$ then $r_{\mathcal{B},h}(n) \leq r_{\mathcal{A},h}(n)$, thus $B(x) \asymp A(x)$ implies $\mathfrak{S}_{\mathcal{B},h}(n) \ll \mathfrak{S}_{\mathcal{A},h}(n)$. Hence, repeating the same argument for \mathcal{B} , since $\{0, 1\} \subseteq \mathcal{B}$, we apply Schnirelmann's theorem again. \square

A similar version for asymptotic bases can be deduced from the appropriate variant of Schnirelmann's theorem. A sequence $\mathcal{A} \subseteq \mathbb{N}$ is called an *asymptotic basis* when there is $h \geq 2$ such that $\mathbb{N} \setminus h\mathcal{A}$ is finite, and from Theorem 11.6, p. 366 of Nathanson [14] one knows that, if $A(n) = \Theta(n)$, then not being contained in some non-trivial arithmetic progression ensures that \mathcal{A} is an asymptotic basis.

Remark A.12 (Applications of Theorem A.11). The connection with elementary methods comes from (A.9), as both the proof of Schnirelmann–Goldbach's theorem and Linnik's elementary solution to Waring's problem is achieved by giving non-trivial upper bounds to $r_{\mathcal{A},h}(n)/s_{\mathcal{A},h}(n)$ for some large enough $h \geq 2$.

Following Chapter 11 of Nathanson [14], let $f \in \mathbb{Q}[x]$ be an integer-valued polynomial (i.e. $f(k) \in \mathbb{Z}$ when $k \in \mathbb{Z}$) with positive leading coefficient. Define $\mathcal{P}_f := \{f(n) : n \in \mathbb{N}\} \cap \mathbb{N}$, the sequence generated by f . Since $|\mathcal{P}_f \cap [0, x]| = \Theta(x^{\deg(f)})$, we know that \mathcal{P}_f is OR_+ . To show that it is a stable basis, one proves that there is $h = h_f$ for which $\mathfrak{S}_{\mathcal{P}_f,h}(n) = O(1)$.²⁴

In the case of primes, the weak form of the Prime Number Theorem says that $\pi(x) = \Theta(x/\log(x))$, where $\pi(x) := |\mathbb{P} \cap [0, x]|$, the sequence of primes; hence \mathbb{P} is OR_+ . At the heart of Schnirelmann's proof that $\mathbb{P} \cup \{0, 1\}$ forms a stable basis is the fact that $\mathfrak{S}_{\mathbb{P},2}(n) = O(\prod_{p|n} (1 + p^{-1}))$,²⁵ and one can then show that \mathbb{P} satisfies (A.10) with $\varepsilon = 1$. For the reader familiar with algebraic number theory, one can take \mathbb{P}_K to be the set of primes that split completely in a number field K/\mathbb{Q} . From

²³Theorem A in Section 42, p. 175 of Halmos [11].

²⁴Theorem 11.8, p. 370 of Nathanson [14].

²⁵Theorem 7.2, p. 186 of Nathanson [13].

Chebotarev’s density theorem²⁶ one has $\pi_K(x) \asymp \pi(x)$, where $\pi_K(x) := |\mathbb{P}_K \cap [0, x]|$, therefore $\mathbb{P}_K \cup \{0, 1\}$ is also a basis.

Remark A.13. As an interesting side note, it is possible to show that when \mathcal{A} is OR_+ , we have $\sum_{n \leq x} \mathfrak{S}_{\mathcal{A}, h}(n) = O(x)$ for all $h \geq 1$, which is the same of having $\varepsilon = 0$ in (A.10). Since this fact is not essential to Theorem A.11, we chose to omit the proof, which is essentially a lengthy calculation.

REFERENCES

1. N. Alon and J. H. Spencer, *The probabilistic method*, 4th ed., John Wiley & Sons, New Jersey, 2016.
2. N. H. Bingham, C. M. Goldie, and J. L. Teugels, *Regular variation*, Cambridge Univ. Press, 1989.
3. R. B. Boppana and J. H. Spencer, *A useful elementary correlation inequality*, J. Comb. Theory A **50** (1989), 305–307.
4. P. Erdős, *Problems and results in additive number theory*, Colloque sur la Theorie des Nombres (CBRM) (Bruxelles), 1956, pp. 127–137.
5. P. Erdős and A. Rényi, *Additive properties of random sequences of positive integers*, Acta Arith. **6** (1960), 83–110.
6. P. Erdős and P. Tetali, *Representations of integers as the sum of k terms*, Random Struct. Algor. **1** (1990), 245–261.
7. W. Feller, *An introduction to probability theory and its applications*, 3rd ed., vol. 1, John Wiley & Sons Inc., New York, 1967.
8. ———, *An introduction to probability theory and its applications*, 2nd ed., vol. 2, John Wiley & Sons Inc., New York, 1971.
9. A. O. Gel’fond and Yu. V. Linnik, *Elementary methods in the analytic theory of numbers*, Pergamon Press, Oxford, 1966, D. E. Brown translated version.
10. H. Halberstam and K. F. Roth, *Sequences*, revised ed., Springer, 1983.
11. P. R. Halmos, *Measure theory*, Graduate Texts in Mathematics, vol. 18, Springer, 1974.
12. M. N. Kolountzakis, *An effective additive basis for the integers*, Discrete Math. **145** (1995), 307–313.
13. M. B. Nathanson, *Additive number theory: The classical bases*, 2nd ed., Graduate Texts in Mathematics, vol. 164, Springer, 1996.
14. ———, *Elementary methods in number theory*, Graduate Texts in Mathematics, vol. 195, Springer, 1999.
15. ———, *Thin bases in additive number theory*, Discrete Math. **312** (2012), 2069–2075.
16. J. Neukirch, *Algebraic number theory*, Grundlehren der mathematischen Wissenschaften, vol. 322, Springer, 1999.
17. A. Stöhr, *Gelöste und ungelöste Fragen über Basen der natürlichen Zahlenreihe i* , J. reine angew. Math. **194** (1955), 40–65.
18. T. Tao and V. H. Vu, *Additive combinatorics*, Cambridge Stud. Adv. Math., vol. 105, Cambridge Univ. Press, 2006.
19. V. H. Vu, *On a refinement of Waring’s problem*, Duke Math. J. **105** (2000), 107–134.
20. ———, *On the concentration of multivariate polynomials with small expectation*, Random Struct. Algor. **16** (2000), 344–363.
21. L. Warnke, *Upper tails for arithmetic progressions in random subsets*, Israel Journal of Mathematics **221** (2017), 317–365.

RIMS, KYOTO UNIVERSITY, 606-8502 KYOTO, JAPAN
E-mail address: tafula@kurims.kyoto-u.ac.jp

²⁶Theorem 13.4 in Chapter VII, p. 545 of Neukirch [16].