NOTE ON THE NUMBER OF ANTICHAINS IN GENERALIZATIONS OF THE BOOLEAN LATTICE

Jinyoung Park*[1], Michail Sarantis†[2], and Prasad Tetali‡[3]

[1] *Department of Mathematics, Courant Institute of Mathematical Sciences, New York University, U.S.A.*
jinyoungpark@nyu.edu
[2] *Institute of Discrete Mathematics, Technical University of Graz, Austria*
msarantis@tugraz.at
[3] *School of Mathematical Sciences, Carnegie Mellon University, Pittsburgh, U.S.A.*
ptetali@cmu.edu

Submitted: May 25, 2023; Accepted: Sep 20, 2024; Published: Mar 15, 2025

Abstract. We give a short and self-contained argument that shows that, for any positive integers t and n with $t = O\left(\frac{n}{\log n}\right)$, the number $\alpha([t]^n)$ of antichains of the poset $[t]^n$ is at most

$$\exp_2\left[\left(1 + O\left(\left(\frac{t \log^3 n}{n}\right)^{1/2}\right)\right) N(t,n)\right],$$

where $N(t,n)$ is the size of a largest level of $[t]^n$. This, in particular, says that if $t \ll n/\log^3 n$ as $n \to \infty$, then $\log \alpha([t]^n) = (1 + o(1))N(t,n)$, giving a (partially) positive answer to a question of Moshkovitz and Shapira for t, n in this range.

Particularly for $t = 3$, we prove a better upper bound:

$$\log \alpha([3]^n) \leqslant (1 + 4\log 3/n)N(3,n),$$

which is the best known upper bound on the number of antichains of $[3]^n$.

Keywords. Boolean lattice, antichains, entropy method

Mathematics Subject Classifications. 05A16, 06A07

---

*Supported by NSF grant DMS-1926686, CCF-1900460, DMS-2153844 and DMS-2324978.

†Research conducted while affiliated with Carnegie Mellon University. Supported in part by the NSF grant DMS-2151283 and the Onassis Foundation - Scholarship F ZP 051-1/2019-2020.

‡Research supported in part by the NSF grant DMS-2151283 and Alexander M. Knaster Professorship.

1. Introduction

An *antichain* of a poset P is a set of elements of P, any two of which are incomparable in the partial order. We denote by $\alpha(P)$ the number of antichains of P. "Dedekind's Problem" of 1897 [Ded97] asks for the number of antichains of the Boolean lattice $B_n = \{0,1\}^n$. Since any subsets of a middle layer of B_n is an antichain, a trivial lower bound is $\alpha(B_n) \geqslant 2^{\binom{n}{\lfloor n/2 \rfloor}}$. Kleitman [Kle69] and subsequently Kleitman and Markowsky [KM75] proved that this trivial lower bound is optimal in the logarithmic scale, namely,

$$\log \alpha(B_n) \leqslant \left(1 + O\left(\frac{\log n}{n}\right)\right)\binom{n}{\lfloor n/2 \rfloor}. \tag{1.1}$$

(In this paper, $\log$ always means $\log_2$.) Asymptotics for $\alpha(B_n)$ itself were obtained by Korshunov [Kor80] via a very involved argument. Later, Sapozhenko [Sap89] gave a simpler (yet still difficult) proof for the asymptotics. The main tool that Sapozhenko used in this work is now called *(Sapozhenko's) graph container method*, and has been very influential; see e.g. [Gal19] for an excellent exposition.

Kahn [Kah01] used entropy methods to give an optimal upper bound on the number of independent sets of a regular bipartite graph. In [Kah02], he extended this idea to the layers of a graded poset to recover the result of Kleitman and Markowsky in (1.1). Independently, Pippenger [Pip99] also gave a slightly weaker bound $\log \alpha(B_n) \leqslant \left(1 + O\left(\frac{\log^{3/2} n}{n^{1/4}}\right)\right)\binom{n}{\lfloor n/2 \rfloor}$. Pippenger's work also uses entropy functions, but his approach was more akin to that of [Kle69]. A shorter proof of a version of (1.1) was also given by Balogh, Treglown, and Wagner [BTW16] with a weaker error term $O\left(\frac{\log n}{\sqrt{n}}\right)$, using the graph container method.

Carroll, Cooper and Tetali [CCT12] considered the question of counting antichains for the following natural generalizations of the Boolean lattice: for an integer $t \geqslant 2$, let $[t]^n$ be the poset consisting of all n-tuples $(x_1, \ldots, x_n)$ of integers in $\{0, 1, \ldots, t-1\}$ with the partial order $\preceq$ defined by $x \preceq y \Leftrightarrow x_i \leqslant y_i$, for all $1 \leqslant i \leqslant n$. Following Pippenger's approach, they proved an analogous result for this generalized Boolean lattice:

Theorem 1.1 ([CCT12]). *For integers t, n such that $1 < t < n$,*

$$\log \alpha([t]^n) \leqslant \left(1 + \frac{11t^2 \log t (\log n)^{3/2}}{n^{1/4}}\right) N(t,n), \tag{1.2}$$

where $N(t,n)$ is the size of (one of) the middle layer(s) of $[t]^n$.

We note that the special case of $t = 3$ was independently considered in [NSS18], in which it was proved that $\log \alpha([3]^n) \leqslant \left(1 + O\left(\sqrt{\log n/n}\right)\right) N(3,n)$. The asymptotics for $N(t,n)$ on the right-hand side of (1.2) was obtained in 1960's: in [And67] (and, more recently, in [MR08]), it was proved that for $n \to \infty$ and every t,

$$N(t,n) = t^n \sqrt{\frac{6}{\pi(t^2-1)n}}\,(1 + o_n(1)). \tag{1.3}$$

Similarly to the Boolean lattice, $N(t,n)$ is a trivial lower bound on $\log \alpha([t]^n)$. So Theorem 1.1 tells us that even in this generalized setting of $[t]^n$, it still holds that $\log \alpha([t]^n)$ is asymptotically bounded *above* by the trivial lower bound, as long as t is much smaller than around $n^{1/8}$ (up to some poly-log terms).

We note that getting the right dependency of $\alpha([t]^n)$ on each of t and n has an interesting connection to some Ramsey-type problems, which we will briefly discuss in Section 1.2. Tsai [Tsa19] gave the following bound which is better than the bound in (1.2) for $t \gg \frac{n^{1/8}}{(\log n)^{3/4}}$.

Theorem 1.2 ([Tsa19]). *For any integers $t, n \geqslant 1$,*

$$\log \alpha([t]^n) \leqslant N(t,n)\log(t+1).$$

Recently, Pohoata and Zakharov [PZ24] used the graph container method to treat the case of constant t, which yields better dependency on n (but dependency unspecified on t).

Theorem 1.3 ([PZ24]). *For any integer $n \geqslant 2$,*

$$\log \alpha([t]^n) \leqslant \left(1 + C_t \frac{\log n}{\sqrt{n}}\right) N(t,n),$$

where $C_t > 0$ is a constant depending solely on t.

1.1. Our results

Our first result is an upper bound on $\log \alpha([3]^n)$ with a better dependency on n.

Theorem 1.4. *For any integer $n \geqslant 1$,*

$$\log \alpha([3]^n) \leqslant \left(1 + \frac{4\log 3}{n}\right) N(3,n).$$

We remark that our proof almost identically applies to $\alpha(B_n)$ to improve the right-hand side of (1.1) to $(1 + O(1/n))N(2,n)$. (Of course, even with this improvement, this bound is much looser than the actual asymptotics proved by Korshunov and Shapozhenko.) The proof of Theorem 1.4 was inspired by Kahn's entropy approach in [Kah02], but Kahn's result applies under quite a strict condition on the degrees of the vertices in the bipartite graph induced by two consecutive layers of B_n. More precisely, any two consecutive layers of B_n form a bi-regular bipartite graph, but for any $t > 2$, two consecutive layers of $[t]^n$ do not satisfy this condition.

Our approach to overcome this difficulty is roughly as the following: we consider weighted antichains in a manner that takes into account the degrees of the vertices of the poset. It turned out that much more careful and delicate analysis were required to handle the weighted antichains, and to that end, we prove a refined version of Kahn's inequality for the independence polynomial of a bipartite graph (Theorem 3.1). We then use it to upper bound the weighted sum of antichains of a graded poset (Theorem 3.6). The seemingly complicated inequality becomes easier to handle in the case of equal weights, and analyzing it more carefully on $[3]^n$ allows us to obtain Theorem 1.4.

Unfortunately, the proof of Theorem 1.4 does not extend to $t > 3$ because of the structural difference of the bipartite graphs formed by consecutive layers of $[t]^n$. However, it is not clear whether the inequality obtained from Theorem 3.6 itself fails.

Our second result concerns any t and n, whenever n is large enough compared to t.

Theorem 1.5. *There exists an absolute constant C such that for integers t, n such that $1 \leqslant t < n/(100\log n)$,*

$$\log \alpha([t]^n) \leqslant \left(1 + C\left(\frac{t\log^3 n}{n}\right)^{1/2}\right) N(t,n). \tag{1.4}$$

Theorem 1.5 confirms the very interesting phenomenon that $\log \alpha([t]^n)$ is asymptotically equal to its trivial lower bound, $N(t,n)$, for the range of pairs (t,n) even further than what is confirmed in Theorem 1.1. We note that the constant 100 in the restriction on t is not important – we can make it any constant by choosing C appropriately. In the proof, for example, we will see that the choice of 100 suggests $C \approx 15$.

The proof of Theorem 1.5 adapts the beautiful entropy approach in [Pip99]: note that antichains are in 1-1 correspondence with monotone Boolean functions, by mapping each antichain to the function which equals 0 below each element of the antichain and 1 everywhere else. With $\mathbf{f}$ a random monotone Boolean function on $[t]^n$ chosen uniformly at random, we have

$$\log \alpha([t]^n) = H(\mathbf{f}),$$

where $H(\cdot)$ is the binary entropy function. (See Section 2 for entropy basics.) To bound $H(\mathbf{f})$, we define a series of auxiliary random variables that determine $\mathbf{f}$, and will bound the entropy of those auxiliary random variables instead. At the high-level, this approach is similar to that of [Pip99, CCT12], but we had to input extra insights to the choice of the auxiliary random variables to obtain the improvements of the error terms both in [Pip99, CCT12].

As a final remark, we do not expect matching lower bounds to our results: in the Boolean case, $\alpha([2]^n)$ is actually much closer to $\binom{n}{\lfloor n/2 \rfloor}$ than the right-hand side of (1.1) (as proved in [Kor80], [Sap89]), and there is no reason to believe a similar behavior should not manifest in $[t]^n$.

1.2. Connection to a Ramsey-type problem

Antichains of $[t]^n$ are closely related to another interesting combinatorial problem that we briefly discuss in this section. Following [FPSS12], for any sequence of positive integers $j_1 < j_2 < \ldots < j_l$, we say that the k-tuples $(j_i, j_{i+1}, \ldots, j_{i+k-1})$ $(i = 1, 2, \ldots, l-k+1)$ form a *monotone path* of length l. Let $M_k(t,n)$[1] be the smallest integer M with the property that no matter how we color all k-element subsets of $\{1, 2, \ldots, M\}$ with n colors, we can always find a monochromatic monotone path of length t. In this language, the celebrated results of Erdős and Szekeres [ES35] can be written as $M_2(t,n) = (t-1)^n + 1$ and $M_3(t,2) = \binom{2t-4}{t-2} + 1$. Fox, Pach, Sudakov, and Suk [FPSS12] showed that

$$2^{(t/n)^{n-1}} \leqslant M_3(t,n) \leqslant 2^{t^{n-1}\log t} \quad \text{for } n \geqslant 2 \text{ and } t \geqslant n+2,$$

and suggested closing the gap between the lower and upper bound as an interesting question. Answering this question, Moshkovitz and Shapira [MS14] proved the following result.

[1] We use $M_k(t,n)$ for the notation $N_k(q,n)$ in [FPSS12] to make it consistent with our main theorems.

Theorem 1.6. *[MS14] For every $t, n \geqslant 2$,*

$$2^{\frac{2}{3}t^{n-1}/\sqrt{n}} \leqslant M_3(t,n) \leqslant 2^{2t^{n-1}}.$$

The proof of the above theorem uses the fascinating relationship

$$\alpha([t]^n) = M_3(t,n) - 1 \quad \text{for all } t, n \geqslant 2$$

that is also proved in [MS14] and independently in [MSW15]. Moshkovitz and Shapira conjectured that

$$M_3(t,n) = 2^{\Theta(t^{n-1}/\sqrt{n})},$$

and more boldly, asked whether (as $n \to \infty$)

$$M_3(t,n) = 2^{(1+o_n(1))N(t,n)}. \tag{1.5}$$

Of course, the motivation of this bold question is the fascinating phenomenon proved by Kleitman and Markowsky that almost all antichains in the Boolean lattice are subsets of the middle layer(s). Our second main result, Theorem 1.5, confirms that (1.5) holds for t, n with $t \ll \frac{n}{\log^3 n}$ as $n \to \infty$ (t is not necessarily fixed), improving on the previous best range $t \ll n^{1/8}$ of [CCT12].

We note that the case where t is much larger than n as $n \to \infty$ remains interesting for the verification of (1.5) in all ranges, as well as the Ramsey-type problem of Fox et al. when the length of the path t is much larger than the number of colors. It was recently announced by Falgas-Ravry, Räty and Tomon [FRRT23] (after we submitted this paper) that this holds, indeed, for all t, n when $n \to \infty$.

Organization. Section 2 collects basic properties of entropy. The two main theorems are proved in Sections 3 and 4, respectively.

2. Entropy Basics

For $p \in [0,1]$, we denote by $H(p) = -p\log p - (1-p)\log(1-p)$ (where $0 \log 0 := 0$) the *binary entropy function*. For a discrete random variable $\mathbf{X}$, we define its *entropy* as

$$H(\mathbf{X}) = \sum_x -p(x)\log p(x),$$

where $p(x) = \mathbb{P}(\mathbf{X} = x)$. Note that by Jensen's inequality we get

$$H(\mathbf{X}) \leqslant \log|\mathrm{range}(\mathbf{X})| \quad \text{(equality holds iff } \mathbf{X} \text{ is uniform on } \mathrm{range}(\mathbf{X})). \tag{2.1}$$

For any event T, we define the entropy of $\mathbf{X}$ given T as

$$H(\mathbf{X}|T) = \sum_x -p(x|T)\log p(x|T),$$

where $p(x|T) = \mathbb{P}(\mathbf{X} = x|T)$. The *conditional entropy* of $\mathbf{X}$ with respect to a discrete random variable $\mathbf{Y}$ is

$$H(\mathbf{X}|\mathbf{Y}) = \mathbb{E}(H(\mathbf{X}|\mathbf{Y} = y)) = \sum_{y} p(y) \sum_{x} -p(x|y) \log p(x|y). \tag{2.2}$$

Below are some standard facts about entropy (for their proofs, see, for example, [Gal14]); for any discrete random variables $\mathbf{X}$ and $\mathbf{Y}$,

$$H(\mathbf{X}, \mathbf{Y}) = H(\mathbf{X}) + H(\mathbf{Y}|\mathbf{X}), \tag{2.3}$$

and

$$\text{if } \mathbf{Y} \text{ determines } \mathbf{X}, \text{ then } H(\mathbf{Z}|\mathbf{Y}) \leqslant H(\mathbf{Z}|\mathbf{X}). \tag{2.4}$$

Fact 2.1 (Subadditivity of entropy)**.** *For any random vector* $\mathbf{X} = (\mathbf{X}_1, \ldots, \mathbf{X}_n)$ *we have*

$$H(\mathbf{X}) \leqslant H(\mathbf{X}_1) + \cdots + H(\mathbf{X}_n).$$

We will also need a celebrated inequality due to Shearer, which generalizes the subadditivity property of entropy.

Lemma 2.2 ([GGFS86])**.** *Let* $\mathbf{X} = (\mathbf{X}_1, \ldots, \mathbf{X}_k)$ *be a random vector and for every* $A \subseteq [k]$ *let* $\alpha_A \in \mathbb{R}^+$*. If* $\sum_{A \ni i} \alpha_A \geqslant 1$ *for all* $i \in [k]$*, then*

$$H(\mathbf{X}) \leqslant \sum_{A \subseteq [k]} \alpha_A H(\mathbf{X}_A),$$

where $\mathbf{X}_A = (\mathbf{X}_i : i \in A)$*.*

A derivation of the following elementary fact can be found in [Kah02].

Fact 2.3. *Let* $\mathbf{X}$ *be a discrete random variable. Then,*

$$H(\mathbf{X}|\mathbf{X} \neq 0) = \frac{H(\mathbf{X}) - H(\mathbb{P}(\mathbf{X} = 0))}{1 - \mathbb{P}(\mathbf{X} = 0)}.$$

3. Proof of Theorem 1.4

As sketched in Section 1.1, our goal is to bound the weighted sum of antichains of a graded poset where the bipartite graphs induced by two consecutive layers of the poset are irregular. We will proceed with induction, and to that end, we first extend [Kah02, Theorem 1.2] to bound the weighted sum of independent sets in an irregular graph. In this setting, we allow the vertices of one of the parts to have different weights (activities).

For any graph G and any vertex $v \in V(G)$, we denote by $d(v)$ the degree of v and by $N(v)$ the set of neighbors of v. An *independent set* in a graph is a set of vertices in which no pairs are adjacent, and $\mathcal{I}(G)$ is the collection of independent sets of G.

Theorem 3.1. *Let G be a bipartite graph on $A \cup B$ with the weight for each vertex $x \in A \cup B$ defined to be $\lambda_x \geqslant 1$. For $v \in A$, let*

$$d_{N(v)} = \min\{d(u) : u \in N(v)\}$$

and assume that the weight on each $v \in A$ is

$$\lambda_v = \mu\,,$$

for some μ. Then

$$\sum_{I \in \mathcal{I}(G)} \prod_{x \in I} \lambda_x \leqslant \prod_{v \in A} \Big[(1+\mu)^{d_{N(v)}} + \prod_{u \in N(v)} (1+\lambda_u) - 1\Big]^{1/d_{N(v)}}.$$

We use the setup in [Kah02] for our proof. For the time being, $G = (V, E)$ is an arbitrary graph and $\lambda : V \to [1, \infty)$ is an arbitrary assignment of weights to the vertices. Set

$$Z = Z(G, \lambda) = \sum_{I \in \mathcal{I}(G)} \prod_{x \in I} \lambda_x.$$

With each $v \in V$ we associate a set $S_v \ni 0$ and nonnegative weights $\alpha_v(s)$, $s \in S_v$, such that

$$\alpha_v(0) = 1, \sum_{s \neq 0} \alpha_v(s) = \lambda_v,$$

and the r.v. $\mathbf{X}_v$ given by $\mathbb{P}(\mathbf{X}_v = s) = \frac{\alpha_v(s)}{1+\lambda_v}$ satisfies

$$H(\mathbf{X}_v) = \log(1+\lambda_v). \tag{3.1}$$

(As noted in [Kah02], this is possible iff $\lambda_v \geqslant 1$.) We say that a vector $(s_v : v \in V) \in \prod S_v$ is *independent* if $\{v : s_v \neq 0\} \in \mathcal{I}(G)$. Finally, let $\mathbf{Y} = (\mathbf{Y}_v : v \in V)$ be chosen from the independent vectors in $\prod S_v$ so that $\mathbb{P}(\mathbf{Y} = (s_v))$ is proportional to $\prod \alpha_v(s_v)$.

Remark 3.2. If we define a random independent set $\mathbf{I} = \{v \in V : \mathbf{Y}_v \neq 0\}$ with $\mathbf{Y}$ as above, then $\mathbb{P}(\mathbf{I} = I) = \prod_{v \in I} \lambda_v / Z$ for any $I \in \mathcal{I}(G)$ (because $\mathbb{P}(\mathbf{I} = I)$ is equal to the probability for the event that $\mathbf{Y}_v \neq 0$ if and only if $v \in I$).

It was proved in [Kah02] that

$$H(\mathbf{Y}) = \log Z. \tag{3.2}$$

Proof of Theorem 3.1. Let G and the values λ_x be as in the statement of the theorem, and $\mathbf{Y}$ be the random independent vector defined as above. Our goal is to show that

$$H(\mathbf{Y}) \leqslant \sum_{v \in A} \frac{1}{d_{N(v)}} \log\Big[(1+\mu)^{d_{N(v)}} + \prod_{u \in N(v)} (1+\lambda_u) - 1\Big], \tag{3.3}$$

from which the conclusion of Theorem 3.1 will follow using (3.2). Denote by Q_v the event $\{\mathbf{Y}_w = 0 \ \forall w \in N(v)\}$, and set $q_v = \mathbb{P}(Q_v)$. Note that, by the definition of $\mathbf{Y}$, we have

$$H(\mathbf{Y}_v | Q_v) = H(\mathbf{X}_v). \tag{3.4}$$

Let $\mathbf{Y}_W = (\mathbf{Y}_w : w \in W)$. Then,

$$\begin{aligned} H(\mathbf{Y}) = H(\mathbf{Y}_A, \mathbf{Y}_B) &\overset{(2.3)}{=} H(\mathbf{Y}_A|\mathbf{Y}_B) + H(\mathbf{Y}_B) \\ &\leqslant \sum_{v \in A} \left[H(\mathbf{Y}_v|\mathbf{Y}_{N(v)}) + \frac{1}{d_{N(v)}} H(\mathbf{Y}_{N(v)}) \right], \end{aligned} \tag{3.5}$$

where the second inequality uses subadditivity for conditional entropy, (2.4), and Lemma 2.2. The first term in the above sum is

$$\begin{aligned} H(\mathbf{Y}_v|\mathbf{Y}_{N(v)}) &\overset{(2.2)}{=} \sum_{\xi} \mathbb{P}(\mathbf{Y}_{N(v)} = \xi) H(\mathbf{Y}_v|\mathbf{Y}_{N(v)} = \xi) \\ &\overset{(\dagger)}{=} q_v H(\mathbf{Y}_v|Q_v) \\ &\overset{(3.1),(3.4)}{=} q_v \log(1+\mu), \end{aligned} \tag{3.6}$$

where ξ ranges over all possible choices for $\mathbf{Y}_{N(v)}$ and $(\dagger)$ holds because $H(\mathbf{Y}_v|\mathbf{Y}_{N(v)} = \xi) = 0$ unless ξ is the zero vector. The second term in the sum in (3.5) is, with $\mathbf{1}_A$ the indicator of A,

$$\begin{aligned} H(\mathbf{Y}_{N(v)}) = H(\mathbf{Y}_{N(v)}, \mathbf{1}_{Q_v}) &\overset{(2.3)}{=} H(\mathbf{1}_{Q_v}) + H(\mathbf{Y}_{N(v)}|\mathbf{1}_{Q_v}) \\ &= H(q_v) + q_v H(\mathbf{Y}_{N(v)}|Q_v) + (1-q_v) H(\mathbf{Y}_{N(v)}|\overline{Q_v}) \\ &= H(q_v) + (1-q_v) H(\mathbf{Y}_{N(v)}|\overline{Q_v}). \end{aligned} \tag{3.7}$$

Claim 3.3. *For any $v \in A$,*

$$H(\mathbf{Y}_{N(v)}|\overline{Q_v}) \leqslant \log\left(\prod_{u \in N(v)} (1+\lambda_u) - 1 \right). \tag{3.8}$$

We now state an easy proposition that we will use to prove the above claim:

Proposition 3.4. *For any set B, $\lambda : B \to [0,\infty)$, $\mathcal{S} \subseteq 2^B$, and probability distribution p on $\mathcal{S}$,*

$$\sum_{S \in \mathcal{S}} p(S) \Big(\log(1/p(S)) + \sum_{x \in S} \log \lambda_x \Big) \leqslant \log \left(\sum_{S \in \mathcal{S}} \prod_{x \in S} \lambda_x \right). \tag{3.9}$$

Proof. Set $W = \sum_{S \in \mathcal{S}} \prod_{x \in S} \lambda_x$ and $q(S) = \prod_{x \in S} \lambda_x / W$. Then the left-hand side of (3.9) is $\sum_{S \in \mathcal{S}} p(S) \log(q(S)/p(S)) + \log W$. By Jensen's inequality, $\sum_{S \in \mathcal{S}} p(S) \log(q(S)/p(S)) \leqslant \log \left(\sum_{S \in \mathcal{S}} q(S) \right) = 0$, from which the conclusion follows. $\square$

Proof of Claim 3.3. Let $\mathbf{I}$ be the random independent set defined in Remark 3.2, and $\mathbf{T} = \mathbf{I} \cap N(v)$. Observe that $H(\mathbf{Y}_{N(v)}|\mathbf{T} = T) = \sum_{u \in T} H(\mathbf{X}_u|\mathbf{X}_u \neq 0) \overset{\text{Fact 2.3}}{=} \sum_{u \in T} \log \lambda_u$. Thus, setting $p_T := \mathbb{P}(\mathbf{T} = T|\overline{Q_v})$,

$$\begin{aligned} H(\mathbf{Y}_{N(v)}|\overline{Q_v}) = H(\mathbf{Y}_{N(v)}, \mathbf{T}|\overline{Q_v}) &= H(\mathbf{T}|\overline{Q_v}) + H(\mathbf{Y}_{N(v)}|\mathbf{T}, \overline{Q_v}) \\ &= \sum_{T \neq \varnothing} p_T \Big(\log(1/p_T) + \sum_{u \in T} \log \lambda_u \Big), \end{aligned}$$

where the last equality uses

$$\begin{aligned} H(\mathbf{Y}_{N(v)}|\mathbf{T},\overline{Q_v}) &\overset{(2.3)}{=} \sum_{T} \mathbb{P}(\mathbf{T}=T|\overline{Q_v}) H(\mathbf{Y}_{N(v)}|\mathbf{T}=T,\overline{Q_v}) \\ &= \sum_{T\neq\varnothing} p_T H(\mathbf{Y}_{N(v)}|\mathbf{T}=T) \\ &= \sum_{T\neq\varnothing} p_T \sum_{u\in T} \log\lambda_u. \end{aligned}$$

By Proposition 3.4, this is at most

$$\log\left(\sum_{\varnothing\neq T\subseteq N(v)} \prod_{u\in T}\lambda_u\right) = \log\left(\prod_{u\in N(v)}(1+\lambda_u)-1\right). \qquad \square$$

The combination of (3.5)-(3.8) gives

$$\begin{aligned} H(\mathbf{Y}) \leqslant \sum_{v\in A} \Bigg[& q_v \log(1+\mu) \\ & + \frac{1}{d_{N(v)}}\left\{H(q_v) + (1-q_v)\log\left(\prod_{u\in N(v)}(1+\lambda_u)-1\right)\right\}\Bigg]. \end{aligned} \tag{3.10}$$

The contribution of v is

$$\begin{aligned} &\frac{1}{d_{N(v)}}\log\left(\prod_{u\in N(v)}(1+\lambda_u)-1\right) \\ &\quad + \frac{1}{d_{N(v)}}\left[H(q_v) + q_v\left\{d_{N(v)}\log(1+\mu) - \log\left(\prod_{u\in N(v)}(1+\lambda_u)-1\right)\right\}\right]. \end{aligned}$$

Note that only the term in the brackets depends on q_v, and the function $H(q)+q\cdot R$ is maximized at $q=\frac{2^R}{1+2^R}$. Therefore, the contribution of v is maximized at

$$q_v = \frac{2^R}{2^R+1} = \frac{(1+\mu)^{d_{N(v)}}}{(1+\mu)^{d_{N(v)}} + \prod_{u\in N(v)}(1+\lambda_u)-1},$$

where $R = d_{N(v)}\log(1+\mu) - \log[\prod_{u\in N(v)}(1+\lambda_u)-1]$.

Inserting this value of q_v in (3.10) gives

$$\log Z = H(\mathbf{Y}) \leqslant \sum_{v\in A}\frac{1}{d_{N(v)}}\log\left[(1+\mu)^{d_{N(v)}} + \prod_{u\in N(v)}(1+\lambda_u)-1\right]. \qquad \square$$

Let P be a graded poset with levels $P_1, P_2, \ldots, P_k$. For $X \subseteq P_j$ for some j, write

$$P_{<X} = \{z \in P : z < v \text{ for some } v \in X\}, \qquad M(X) = P_j \setminus X,$$
$$P_X = \{z \in P_1 \cup \cdots \cup P_{j-1} : z \not< x \ \forall x \in X\}, \qquad \text{and} \qquad \overline{P}_X = P_X \cup M(X).$$

If X is a singleton $\{v\}$, then we simply write $P_{<v}$ for $P_{<X}$. Finally, let

$$N^i(v) = P_i \cap \{y : y > v \text{ or } y < v\}, \qquad d^i(v) = |N^i(v)|$$
$$\text{and} \qquad d_{N^i(v)} = \min\{d^{i+1}(w) : w \in N^i(v)\}.$$

When not working with the main poset P, we will be adding subscripts to the notation to keep track of the poset we are considering.

For a poset P with k levels and $\lambda_1, \ldots, \lambda_k \geqslant 0$, define $f_P(\lambda_1, \ldots, \lambda_k)$ recursively as follows: for any poset Q with one level, $f_Q(\lambda_1) := (1 + \lambda_1)^{|Q|}$. If P has $k \geqslant 2$ layers, then

$$f_P(\lambda_1, \ldots, \lambda_k) := \prod_{v \in P_k} \left((1 + \lambda_k)^{d_{N^{k-1}(v)}} + f_{P_{<v}}(\lambda_1, \ldots, \lambda_{k-1}) - 1\right)^{1/d_{N^{k-1}(v)}}.$$

For example,

$$\text{if } P = P_1 \cup P_2, \text{ then } f_P(\lambda_1, \lambda_2) = \textstyle\prod_{v \in P_2} \left((1 + \lambda_2)^{d_{N^1(v)}} + (1 + \lambda_1)^{d^1(v)} - 1\right)^{1/d_{N^1(v)}}. \tag{3.11}$$

Observe that the assumption $\lambda_i \geqslant 0$ easily yields

$$f_P \geqslant 1\,, \quad \text{for any poset } P. \tag{3.12}$$

In Theorem 3.6, we prove that f_P is an upper bound on the weighted number of antichains in a graded poset, by applying Theorem 3.1 inductively. One might wish to avoid the nested products in the definition of f_P for the sake of simplicity, but then one would lose track of the structure of the previous layers. The recursive definition of f_P is, thus, crucial for obtaining Theorem 1.4, and more specifically for the proof of Lemma 3.7.

The following proposition is a key ingredient for the proof of Theorem 3.6.

Proposition 3.5. *Let P be a graded poset with k levels, and $\lambda_1, \ldots, \lambda_k \geqslant 0$. For any $Y \subseteq P_k$, $f_{\overline{P}_Y}(\lambda_1, \ldots, \lambda_k) \leqslant f_P(\lambda_1, \ldots, \lambda_k)$.*

Proof. Let $Q = \overline{P}_Y$. We use induction on the number of levels k. The assertion trivially holds for $k = 1$. For $k \geqslant 2$,

$$\begin{aligned} f_Q(\lambda_1, \ldots, \lambda_k) &= \prod_{v \in Q_k} \left((1 + \lambda_k)^{d_{N_Q^{k-1}(v)}} + f_{Q_{<v}}(\lambda_1, \ldots, \lambda_{k-1}) - 1\right)^{1/d_{N_Q^{k-1}(v)}} \\ &\leqslant \prod_{v \in Q_k} \left((1 + \lambda_k)^{d_{N^{k-1}(v)}} + f_{Q_{<v}}(\lambda_1, \ldots, \lambda_{k-1}) - 1\right)^{1/d_{N^{k-1}(v)}}, \end{aligned}$$

where the inequality uses the facts that $\frac{1}{x}\log(A^x + B)$ $(A, B \geqslant 0)$ is decreasing for $x > 0$ and that $d_{N_Q^{k-1}}(v) \geqslant d_{N^{k-1}(v)}$ for any $v \in Q_k$: Q is upwards closed in P, hence $d_Q^k(u) = d^k(u)$ for all $u \in N_Q^{k-1}(v)$ and $d_{N_Q^{k-1}}(v) = \min\{d_k(u) : u \in N_Q^{k-1}\} \geqslant \min\{d_k(u) : u \in N_Q^{k-1}\} = d_{N^{k-1}(v)}$. Furthermore, the induction hypothesis yields $f_{Q_{<v}}(\lambda_1, \ldots, \lambda_{k-1}) \leqslant f_{P_{<v}}(\lambda_1, \ldots, \lambda_{k-1})$ for any $v \in Q_k$ (because, with $Z := N^{k-1}(v) \cap P_{<Y}$, we have $Q_{<v} = \overline{(P_{<v})_Z}$). Therefore,

$$
\begin{aligned}
f_Q(\lambda_1, \ldots, \lambda_k) &\leqslant \prod_{v \in Q_k} \left((1+\lambda_k)^{d_{N^{k-1}(v)}} + f_{P_{<v}}(\lambda_1, \ldots, \lambda_{k-1}) - 1\right)^{1/d_{N^{k-1}(v)}} \\
&\leqslant \prod_{v \in P_k} \left((1+\lambda_k)^{d_{N^{k-1}(v)}} + f_{P_{<v}}(\lambda_1, \ldots, \lambda_{k-1}) - 1\right)^{1/d_{N^{k-1}(v)}} \\
&= f_P(\lambda_1, \ldots, \lambda_k)
\end{aligned}
$$

(the last inequality holds because each extra term is $\geqslant 1$). □

We are now ready to prove the main theorem of this section.

Theorem 3.6. *Let P be a graded poset with levels $P_1, P_2, \ldots, P_k$, and $\mathcal{A}(P)$ be the collection of antichains of P. For each $x \in P_i$, define $\lambda_x \equiv \lambda_i$ where $\lambda_j \geqslant 1$ for all $1 \leqslant j \leqslant k$. Then,*

$$\sum_{I \in \mathcal{A}(P)} \prod_{x \in I} \lambda_x \leqslant f_P(\lambda_1, \ldots, \lambda_k). \tag{3.13}$$

Proof. We proceed by induction on k. The base case ($k = 2$) follows from Theorem 3.1 (see (3.11)). Assume the theorem is true for any poset with $k-1$ levels and let P be a poset with k levels. Then for any $X \subseteq P_k$, we have

$$\sum_{\substack{I \in \mathcal{A}(P) \\ I \cap P_k = X}} \prod_{x \in I} \lambda_x = \lambda_k^{|X|} \sum_{I \in \mathcal{A}(P_X)} \prod_{x \in I} \lambda_x \leqslant \lambda_k^{|X|} f_{P_X}(\lambda_1, \ldots, \lambda_{k-1}).$$

Therefore,

$$
\begin{aligned}
\sum_{I \in \mathcal{A}(P)} \prod_{x \in I} \lambda_x &\leqslant \sum_{X \subseteq P_k} \lambda_k^{|X|} f_{P_X}(\lambda_1, \ldots, \lambda_{k-1}) \\
&= \sum_{X \subseteq P_k} \lambda_k^{|X|} \prod_{v \in (P_X)_{k-1}} \left((1+\lambda_{k-1})^{d_{N_{P_X}^{k-2}(v)}} + f_{(P_X)_{<v}}(\lambda_1, \ldots, \lambda_{k-2}) - 1\right)^{1/d_{N_{P_X}^{k-2}(v)}} \\
&\overset{(*)}{\leqslant} \sum_{X \subseteq P_k} \lambda_k^{|X|} \prod_{v \in (P_X)_{k-1}} \left((1+\lambda_{k-1})^{d_{N^{k-2}(v)}} + f_{(P_X)_{<v}}(\lambda_1, \ldots, \lambda_{k-2}) - 1\right)^{1/d_{N^{k-2}(v)}} \\
&\leqslant \sum_{X \subseteq P_k} \lambda_k^{|X|} \prod_{v \in (P_X)_{k-1}} \left((1+\lambda_{k-1})^{d_{N^{k-2}(v)}} + f_{P_{<v}}(\lambda_1, \ldots, \lambda_{k-2}) - 1\right)^{1/d_{N^{k-2}(v)}}
\end{aligned}
$$

where $(*)$ uses the facts that $\frac{1}{x}\log(A^x + B)$ $(A, B \geqslant 0)$ is decreasing for $x > 0$ and that $d_{N^{k-2}_{P_X}}(v) \geqslant d_{N^{k-2}}(v)$ for any $v \in (P_X)_{k-1}$; the last inequality follows from Proposition 3.5 (applied to $P_{<v}$ with $Y = (P_{<v})_{k-2} \cap P_{<X}$). For each $v \in P_{k-1}$, define

$$\mu_v = \left((1+\lambda_{k-1})^{d_{N^{k-2}(v)}} + f_{P_{<v}}(\lambda_1, \dots, \lambda_{k-2}) - 1\right)^{1/d_{N^{k-2}(v)}} - 1. \tag{3.14}$$

We have shown that

$$\sum_{I \in \mathcal{A}(P)} \prod_{x \in I} \lambda_x \leqslant \sum_{X \subseteq P_k} \lambda_k^{|X|} \prod_{v \in (P_X)_{k-1}} (1+\mu_v).$$

Recall from (3.12) that $f_Q \geqslant 1$ for any graded poset Q, so

$$\mu_v \geqslant \left((1+\lambda_{k-1})^{d_{N^{k-2}(v)}}\right)^{1/d_{N^{k-2}(v)}} - 1 = \lambda_{k-1} \geqslant 1.$$

Now, by applying Theorem 3.1 with $A = P_k,\ B = P_{k-1}$, $\mu = \lambda_k$, and μ_v in (3.14) the weight for each $v \in B$, we get

$$\begin{aligned}
\sum_{X \subseteq P_k} \lambda_k^{|X|} \prod_{v \in (P_X)_{k-1}} (1+\mu_v) &\leqslant \prod_{v \in P_k} \left[(1+\lambda_k)^{d_{N^{k-1}(v)}} + \prod_{u \in N^{k-1}(v)} (1+\mu_u) - 1\right]^{1/d_{N^{k-1}(v)}} \\
&\overset{(3.14)}{=} \prod_{v \in P_k} \Bigg[(1+\lambda_k)^{d_{N^{k-1}(v)}} \\
&\quad + \prod_{u \in N^{k-1}(v)} \left((1+\lambda_{k-1})^{d_{N^{k-2}(u)}} + f_{P_{<u}}(\lambda_1, \dots, \lambda_{k-2}) - 1\right)^{1/d_{N^{k-2}(u)}} - 1\Bigg]^{1/d_{N^{k-1}(v)}} \\
&= \prod_{v \in P_k} \left((1+\lambda_k)^{d_{N^{k-1}(v)}} + f_{P_{<v}}(\lambda_1, \dots, \lambda_{k-1}) - 1\right)^{1/d_{N^{k-1}(v)}} \\
&= f_P(\lambda_1, \dots, \lambda_k),
\end{aligned}$$

which completes the induction. □

For the rest of this section, we work specifically on $[3]^n$. Write $P_0, P_1, \dots, P_{2n}$ for the levels of $[3]^n$. Note that $[3]^n$ enjoys the special property that

$$d_{N^{i-1}(x)} - d^{i-1}(x) \geqslant n - i \quad \forall x \in P_i. \tag{3.15}$$

Indeed, assume x has a zeroes, b ones and c twos. The up-degree of any neighbor of x in P_{i-1} is the number of zeroes and ones it has, which can't be less than the same number for x; that is $d_{N^{i-1}(x)} \geqslant a + b$. On the other hand, it is clear that $d^{i-1}(x) = b + c$. Thus $d_{N^{i-1}(x)} - d^{i-1}(x) \geqslant a - c = (a+b+c) - (b+2c) = n - i$.

Let $P = P_0 \cup P_1 \cup \dots \cup P_n$ be the bottom half. Recall that $\alpha(P)$ denotes the number of antichains in P.

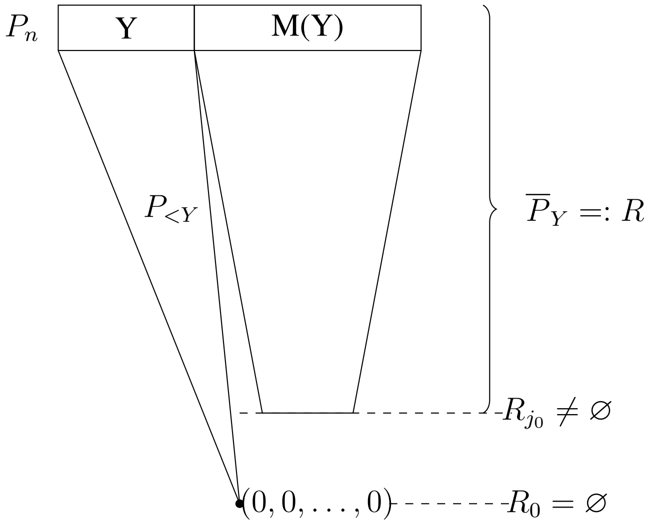

Figure 3.1: The poset $\overline{P}_Y$.

Lemma 3.7. *For any* $Y \subseteq P_n$*,*

$$\alpha\left(\overline{P}_Y\right) \leqslant 2^{|M(Y)|\left(1+\frac{2\log 3}{n}\right)}. \tag{3.16}$$

Proof. Let $R := \overline{P}_Y$ and $R_j = R \cap P_j$ for $0 \leqslant j \leqslant n$. Note that the R_j's are the levels of R, with possibly some empty sets in the beginning. Let j_0 be minimal such that $R_{j_0} \neq \varnothing$. See Figure 3.1 for the schematic presentation of R. We will prove that for any $j \geqslant j_0 + 1$,

$$f_{R_{<v}}(1, \ldots, 1) \leqslant 2^{d^{j-1}(v)\left(1+1/d_{N^{j-1}(v)}\right)} \quad \forall v \in R_j. \tag{3.17}$$

We first show that (3.17) implies the lemma. We have

$$\begin{aligned} \alpha(R) \overset{(3.13)}{\leqslant} f_R(1, \ldots, 1) &= \prod_{v \in R_n} \left(2^{d_{N_R^{n-1}(v)}} + f_{R_{<v}}(1, \ldots, 1) - 1\right)^{1/d_{N_R^{n-1}(v)}} \\ &\leqslant \prod_{v \in R_n} \left(2^{d_{N^{n-1}(v)}} + f_{R_{<v}}(1, \ldots, 1)\right)^{1/d_{N^{n-1}(v)}} \\ &\overset{(3.17)}{\leqslant} \prod_{v \in R_n} \left(2^{d_{N^{n-1}(v)}} + 2^{d^{n-1}(v)\left(1+1/d_{N^{n-1}(v)}\right)}\right)^{1/d_{N^{n-1}(v)}} \\ &\leqslant \prod_{v \in R_n} 2 \cdot 3^{1/d_{N^{n-1}(v)}} = \prod_{v \in R_n} 2^{1+\log 3/d_{N^{n-1}(v)}}. \end{aligned}$$

The last inequality is obtained using (3.15): Letting $a = d_{N^{n-1}(v)}$ and $b = d^{n-1}(v)$, we have $a - b \geqslant 0$. Thus

$$\left(2^a + 2^{b(1+1/a)}\right)^{1/a} = 2\left(1 + 2^{b-a+\frac{b}{a}}\right)^{1/a} \leqslant 2\left(1 + 2^{\frac{b}{a}}\right)^{1/a} \leqslant 2 \cdot 3^{1/a}.$$

Finally, the lemma follows by noticing that $d_{N^{n-1}(v)} \geqslant n/2$ for any $v \in R_n$. □

Proof of (3.17). We proceed by induction. For the base case, for any $v \in R_{j_0+1}$, we trivially have

$$f_{R_{<v}}(1) = 2^{d_R^{j_0}(v)} \leqslant 2^{d^{j_0}(v)} < 2^{d^0(v)\left(1+1/d_{N^0(v)}\right)}.$$

Next, assume the statement is true for any vertex up to layer $j-1$. For any $v \in R_j$,

$$\begin{aligned}
f_{R_{<v}}(1,\ldots,1) &= \prod_{u\in N^{j-1}(v)} \left(2^{d_{N_R^{j-2}(u)}} + f_{R_{<u}}(1,\ldots,1) - 1\right)^{1/d_{N_R^{j-2}(u)}} \\
&\leqslant \prod_{u\in N^{j-1}(v)} \left(2^{d_{N^{j-2}(u)}} + f_{R_{<u}}(1,\ldots,1) - 1\right)^{1/d_{N^{j-2}(u)}} \\
&\overset{(\dagger)}{\leqslant} \prod_{u\in N^{j-1}(v)} \left(2^{d_{N^{j-2}(u)}} + 2^{d^{j-2}(u)\left(1+1/d_{N^{j-2}(u)}\right)}\right)^{1/d_{N^{j-2}(u)}} \\
&\leqslant \prod_{u\in N^{j-1}(v)} 2\left(1 + 2^{d^{j-2}(u)-d_{N^{j-2}(u)}+d^{j-2}(u)/d_{N^{j-2}(u)}}\right)^{1/d_{N^{j-2}(u)}} \\
&\overset{(3.15)}{\leqslant} \prod_{u\in N^{j-1}(v)} 2^{1+1/d_{N^{j-2}(u)}},
\end{aligned}$$

where $(\dagger)$ uses the induction hypothesis. But for every $u \in N^{j-1}(v)$, we have $d_{N^{j-2}(u)} \geqslant d_{N^{j-1}(v)}$: as the up-degree increases when we traverse a chain towards lower layers, $d^{j-1}(w) \geqslant d^j(u)$ for all $w \in N^{j-2}(u)$, which means $d_{N^{j-2}(u)} \geqslant d^j(u) \geqslant \min\{d^j(u) : u \in N^{j-1}(v)\} = d_{N^{j-1}(v)}$. Hence, the above expression is at most $2^{d^{j-1}(v)\left(1+1/d_{N^{j-1}(v)}\right)}$. □

Proof of Theorem 1.4. For $I \in \mathcal{A}(P_{n+1}\cup\ldots\cup P_{2n})$, let $X = X(I) \subseteq P_n$ be the "lower shadow" of I on P_n, namely, $X = \{v \in P_n : \exists\, w \in I : v < w\}$. Then

$$\begin{aligned}
\alpha([3]^n) &= \sum_{Y\subseteq P_n} |\{I \in \mathcal{A}(P_{n+1}\cup\ldots\cup P_{2n}) : X(I) = Y\}| \cdot |\mathcal{A}(\bar{P}_Y)| \\
&\overset{(3.16)}{\leqslant} \sum_{Y\subseteq P_n} 2^{(1+2\log 3/n)|M(Y)|} |\{I \in \mathcal{A}(P_{n+1}\cup\ldots\cup P_{2n}) : X(I) = Y\}| \\
&\leqslant 2^{2\log 3|P_n|/n} \sum_{Y\subseteq P_n} 2^{|M(Y)|} |\{I \in \mathcal{A}(P_{n+1}\cup\ldots\cup P_{2n}) : X(I) = Y\}| \\
&= 2^{2\log 3|P_n|/n} |\mathcal{A}(P_n\cup\ldots\cup P_{2n})| \\
&\overset{(3.16)}{\leqslant} 2^{2\log 3|P_n|/n} 2^{(1+2\log 3/n)|P_n|} = 2^{(1+4\log 3/n)|P_n|},
\end{aligned}$$

where in the last step, (3.16) is applied for $Y = \varnothing$ in the dual poset of $P_n \cup \ldots \cup P_{2n}$, which is simply $P_0 \cup \ldots \cup P_n$. □

4. Proof of Theorem 1.5

The proposition below immediately follows from the proof of [dBvETK51, Theorem 2].

Proposition 4.1. *For any $t, n \geqslant 1$, the poset $[t]^n$ admits a chain partition $\mathcal{C} = \{C_1, C_2, \ldots, C_N\}$ of size $N = N(t,n)$ that satisfies the following property:*

$$\text{if } y \text{ is immediately below } x \text{ in a chain } C \in \mathcal{C}, \text{ then } y \lessdot x \text{ in } [t]^n. \tag{4.1}$$

We also recall from [MS14] the following easy lower bound on $N(t,n)$ that works for all t and n.

Lemma 4.2 (Lemma 2.6, [MS14])**.** *For all $t, n \geqslant 1$,*

$$N(t,n) \geqslant \frac{2t^{n-1}}{3\sqrt{n}}.$$

An element $x \in [t]^n = \{0, 1, \ldots, t-1\}^n$ is called a *point*. For $l \in \{0, 1, \ldots, t-1\}$, let $d_l(x)$ be the number of coordinates of x equal to l. We say x is *low* if $d_l(x) < \frac{n}{2t}$ for some $1 \leqslant l \leqslant t-1$ and *high* otherwise. We also say a chain $C_j \in \mathcal{C}$ is *low* if it contains a low point and *high* otherwise.

Let $(\mathcal{F}(t,n), \prec)$ be the poset on the family of monotone Boolean functions on $[t]^n$ where $f \prec g$ iff $f(x) \leqslant g(x)$ $\forall x \in [t]^n$. Observe that $|\mathcal{F}(t,n)| = \alpha([t]^n)$. Thus, with $\mathbf{f}$ a uniformly chosen element of $\mathcal{F}(t,n)$, we have

$$\log \alpha([t]^n) \overset{(2.1)}{=} H(\mathbf{f}).$$

Following the approaches in [Pip99] and [CCT12], we will define a series of random variables that determine $\mathbf{f}$ in order to bound $H(\mathbf{f})$. In what follows, we use bold-face letters $(\mathbf{f}, \mathbf{y}, \mathbf{Y}, \ldots)$ for random variables, while plain letters $(f, y, Y, \ldots)$ represent values that the corresponding random variable takes.

For $f \in \mathcal{F} = \mathcal{F}(t,n)$ and $j \in [N]$, let $\gamma_j(f) = |\{x \in C_j : f(x) = 1\}|$. Note that $(\gamma_j(f))_j$ determines f. But exposing $\gamma_j(f)$ for all j is too expensive, so we will make a random choice on which chains to expose. To that end, first define $\mathbf{y}_j$ as follows. Let $p = [t \log((t-1)n)/n]^{1/2}$. If the chain C_j is high, then $\mathbb{P}(\mathbf{y}_j = 1) = p = 1 - \mathbb{P}(\mathbf{y}_j = 0)$; if C_j is low, then $\mathbf{y}_j \equiv 1$. Having $\mathbf{y}_j = 1$ means we will expose $\gamma_j(f)$. Define $\tilde{\mathbf{Y}}_j(f) = \mathbf{y}_j \gamma_j(f)$.

As in [Pip99, CCT12], in order to complement $(\tilde{\mathbf{Y}}_j(f))_j$, we introduce another random variable. Write $\tilde{\mathbf{Y}} = (\tilde{\mathbf{Y}}_j)_j$ (and similarly for $\hat{\mathbf{Y}}$ later). Given $\tilde{\mathbf{Y}}(f) = \tilde{Y}$, let $\tilde{f}$ be the smallest in $\mathcal{F}$ that satisfies $\gamma_j(\tilde{f}) \geqslant \tilde{Y}_j$ for all $1 \leqslant j \leqslant N$, and set $\hat{Y}_j = \gamma_j(f) - \gamma_j(\tilde{f})$ $(\geqslant 0)$. Observe that f is determined by the pair $(\tilde{Y}(f), \hat{Y}(f))$, so in particular,

$$H(\mathbf{f}) \leqslant H(\tilde{\mathbf{Y}}(\mathbf{f}), \hat{\mathbf{Y}}(\mathbf{f})) \leqslant H(\tilde{\mathbf{Y}}(\mathbf{f})) + H(\hat{\mathbf{Y}}(\mathbf{f})).$$

(Note that, given f, the randomness of $\tilde{\mathbf{Y}}(f)$ and $\hat{\mathbf{Y}}(f)$ is inherited from $\mathbf{y}$.) Therefore, the next two assertions complete the proof of Theorem 1.5. Here and for the rest of the section, we will use $\epsilon, \epsilon', \ldots$ for small (absolute) constants.

$$H(\tilde{\mathbf{Y}}(\mathbf{f})) \leqslant (2+\epsilon) N \frac{t^{1/2}(\log((t-1)n))^{3/2}}{n^{1/2}}, \tag{4.2}$$

and

$$H(\hat{\mathbf{Y}}(\mathbf{f})) \leqslant N\left(1+\frac{(4+2\epsilon')t^{1/2}(\log(t^{1/2}(t-1)n))^{3/2}}{n^{1/2}}\right). \tag{4.3}$$

The following handy lemma is given in [Pip99].

Lemma 4.3 (Pippenger [Pip99])**.** *Suppose the random variable* $\mathbf{K}$ *takes values in* $\{0,1,\dots,n\}$*, and* $P(\mathbf{K}\geqslant k)\leqslant q$ *for some* $k\geqslant 1$ *and* $0\leqslant q\leqslant 1$*. Then*

$$H(\mathbf{K}) \leqslant h_1(q)+\log k+q\log n,$$

where

$$h_1(q)=\begin{cases}-q\log q-(1-q)\log(1-q) & \text{if } \ q\in[0,1/2];\\ 1 & \text{if } \ q\in[1/2,1].\end{cases}$$

For future reference, we remark that

$$h_1(q)\leqslant -2q\log q \quad \text{for } q\leqslant 1/2. \tag{4.4}$$

Proof of (4.2). If C_j is low, then we apply the naive bound $H(\tilde{\mathbf{Y}}_j(\mathbf{f}))\leqslant \log((t-1)n+2)$. If C_j is high, then noticing that $\tilde{\mathbf{Y}}_j(\mathbf{f})\geqslant 1$ implies $\mathbf{y}_j=1$, we apply Lemma 4.3 with $k=1$ and $q=p$ to obtain

$$H(\tilde{\mathbf{Y}}_j(\mathbf{f}))\leqslant h_1(p)+p\log((t-1)n+1).$$

Therefore, with M the number of low chains in $\mathcal{C}$,

$$H(\tilde{\mathbf{Y}}(\mathbf{f}))\leqslant \sum_{j=1}^{N} H(\tilde{\mathbf{Y}}_j(\mathbf{f}))\leqslant M\log((t-1)n+2)+(N-M)(h_1(p)+p\log((t-1)n+1)). \tag{4.5}$$

Since each low chain contains a low point, we may bound M by the number of low points. Any such point satisfies $d_l(x)<\frac{n}{2t}$ for some $l\in\{1,\dots,t-1\}$. Let $\mathbf{x}$ be a uniformly random point in $[t]^n$; equivalently, each coordinate of $\mathbf{x}$ is chosen uniformly and independently from $\{0,\dots,t-1\}$. Then for each $l\in[t-1]$, the Chernoff bound yields that $\mathbb{P}(d_l(\mathbf{x})<n/(2t))\leqslant e^{-n/(8t)}$, hence by the union bound,

$$M\leqslant (t-1)t^n e^{-n/(8t)}\leqslant t^{n+1}e^{-n/(8t)}. \tag{4.6}$$

By combining (4.4) (note $p\leqslant 1/2$ since $t\leqslant n/(100\log n)$), (4.5) and (4.6), we have

$$\begin{aligned} H(\tilde{\mathbf{Y}}(\mathbf{f})) &\leqslant t^{n+1}e^{-n/(8t)}\log((t-1)n+2)+N(-2p\log p+p\log((t-1)n+1))\\ &\leqslant (2+\epsilon)N\frac{t^{1/2}(\log((t-1)n))^{3/2}}{n^{1/2}}, \end{aligned}$$

where the last inequality uses Lemma 4.2 (and that $t\leqslant n/(100\log n)$). □

Proof of (4.3). For each $j \in [N]$, let $q_j = \mathbb{P}(\hat{\mathbf{Y}}_j(\mathbf{f}) \geqslant 2)$. By Lemma 4.3 with $k = 2$ and $q = q_j$, $H(\hat{\mathbf{Y}}_j(\mathbf{f})) \leqslant h_1(q_j) + 1 + q_j \log((t-1)n)$. Therefore, with $Q := \sum_{j=1}^{N} q_j$,

$$H(\hat{\mathbf{Y}}(\mathbf{f})) \leqslant \sum_j H(\hat{\mathbf{Y}}_j(\mathbf{f})) \leqslant \sum_j [h_1(q_j) + 1 + q_j \log((t-1)n)]$$
$$\overset{(*)}{\leqslant} N h_1(Q/N) + N + Q \log((t-1)n)\,,$$

where $(*)$ follows from the concavity of h_1 (so $\frac{1}{N}\sum_j h_1(q_j) \leqslant h_1\big(\sum_j q_j/N\big)$). We will prove that

$$Q \leqslant \frac{(2+\epsilon')t^{1/2}(\log(t^{1/2}(t-1)n))^{1/2}}{n^{1/2}} N. \tag{4.7}$$

Let's first show that (4.7) implies (4.3). By (4.4) (and $Q/N \leqslant 1/2$),

$$h_1\left(\frac{Q}{N}\right) \leqslant \frac{(2+\epsilon')t^{1/2}(\log(t^{1/2}(t-1)n))^{1/2}}{n^{1/2}} \log n,$$

thus

$$H(\hat{\mathbf{Y}}(\mathbf{f})) \leqslant N\left(1 + \frac{(4+2\epsilon')t^{1/2}(\log(t^{1/2}(t-1)n))^{3/2}}{n^{1/2}}\right).$$

The rest of this section is devoted to proving (4.7). To this end, we first define an event that is implied by $\hat{Y}_j(f) \geqslant 2$. Given f and $\tilde{Y}(f)$, say a point $x \in C_j$ is *bad* if

1. x is high (that is, $d_l(x) \geqslant \frac{n}{2t}$ for all $1 \leqslant l \leqslant t-1$); and
2. $\tilde{f}(x) = 0$; and
3. $f(y) = 1$ for $y \in C_j$ that is immediately below x.

Define r_x to be the probability (w.r.t. $\mathbf{f}$ and $\mathbf{y}$) that x is bad. Notice that if $\hat{Y}_j(f) \geqslant 2$, then the chain C_j must contain a bad vertex. Therefore,

$$Q = \sum_j q_j \leqslant \sum_j \mathbb{P}(\text{some } x \in C_j \text{ is bad}) \leqslant \sum_x r_x.$$

Set $s = p^{-1}\log(t^{1/2}(t-1)n) = (n/t)^{1/2}(\log(t^{1/2}(t-1)n))^{1/2}$. Say y is a *k-child* ($k \in \{1,\dots,t-1\}$) of x if $y \lessdot x$ and y differs from x in a coordinate that has the value k in x. Given f, define a point x to be *heavy* if there exists some $k \in \{1,\dots,t-1\}$ for which x has at least s of k-children y with $f(y) = 1$, and *light* otherwise. Say x is *occupied* if $f(x) = 1$.

Given $x \in [t]^n$, let $C_x \in \mathcal{C}$ be the chain containing x, and $y = y(x)$ be $y \lessdot x$ and $y \in C_x$. Such y exists by Proposition 4.1 unless x is the smallest in C_x. If x is the smallest in C_x, then write $y(x) = \varnothing$. Note that

$$\text{if } x \text{ is bad, then } x \text{ is high and occupied, } y(x) \neq \varnothing, \text{ and } f(y) = 1. \tag{4.8}$$

For convenience, denote by Q_x the event that x is high, occupied, and $y(x) \neq \varnothing$. Then by (4.8),

$$\begin{aligned} r_x &= \mathbb{P}(x \text{ heavy and bad}) + \mathbb{P}(x \text{ light and bad}) \\ &\leqslant \mathbb{P}(x \text{ heavy and bad}) + \mathbb{P}(x \text{ light}, Q_x) \cdot \mathbb{P}(\mathbf{f}(y) = 1 | x \text{ light}, Q_x). \end{aligned} \tag{4.9}$$

We will bound each term in the expression above.

Bounding $\mathbb{P}(x$ heavy and bad$)$. Say x is k-*heavy* for $k \in \{1, 2, \ldots, t-1\}$ if x has at least s of k-children y with $f(y) = 1$. In order for a k-heavy x to be bad, each of the chains C_i containing the k-children y with $f(y) = 1$ must have $\mathbf{y}_i = 0$, which happens with probability at most $1-p$. Therefore,

$$\begin{aligned}\mathbb{P}(x \text{ heavy and bad}) &\leqslant \sum_k \mathbb{P}(x\ k\text{-heavy and bad}) \\ &\leqslant \sum_k \mathbb{P}(x \text{ bad}|\ x\ k\text{-heavy}) \leqslant (t-1)(1-p)^s \leqslant (t-1)e^{-ps} = \frac{1}{t^{1/2}n}.\end{aligned} \tag{4.10}$$

Bounding $\mathbb{P}(\mathbf{f}(y) = 1|x$ light, $Q_x)$. Suppose y is a k-child of x for some k, and let $N_k(x)$ be the set of k-children of x. Consider S_n, the group of permutations on $[n]$. For $\pi \in S_n$, π acts on $x \in [t]^n$ by $\pi(x) = (x_{\pi(1)}, \ldots, x_{\pi(n)})$ and on $f \in \mathcal{F}$ by $\pi(f) = f(\pi(x))$. Let $\mathrm{Stab}(x) = \{\pi \in S_n : \pi(x) = x\}$, noticing that $\mathrm{Stab}(x)$ acts transitively on $N_k(x)$. Therefore, for each $f \in \mathcal{F}$, the fraction of elements g in $\mathrm{Orb}(f) := \{\pi(f) : \pi \in \mathrm{Stab}(x)\}$ with $g(y) = 1$ is precisely $|f^{-1}(1) \cap N_k(x)|/|N_k(x)|$. Given that x is high, light and occupied, $|N_k(x)| \geqslant n/(2t)$ and $|f^{-1}(1) \cap N_k(x)| \leqslant s$. Finally, $\{\mathrm{Orb}(f) : f \in \mathcal{F}\}$ partitions $\mathcal{F}$. Therefore,

$$\mathbb{P}(\mathbf{f}(y) = 1|x \text{ light}, Q_x) \leqslant \frac{s}{\frac{n}{2t}} = \frac{2t^{1/2}(\log(t^{1/2}(t-1)n))^{1/2}}{n^{1/2}}. \tag{4.11}$$

Bounding $\sum_x \mathbb{P}(x$ light, $Q_x)$. Given f, write $R(f)$ for the set of points that are light and satisfy Q_x. Say x is *marginal* if x is occupied, $y(x) \neq \varnothing$, and $f(y) = 0$. Then

$$\begin{aligned}\mathbb{P}(x \text{ non-marginal}|x \text{ light}, Q_x) &= \mathbb{P}(\mathbf{f}(y) = 1|x \text{ light}, Q_x) \\ &\overset{(4.11)}{\leqslant} \frac{2t^{1/2}(\log(t^{1/2}(t-1)n))^{1/2}}{n^{1/2}},\end{aligned}$$

so,

$$\mathbb{E}[|\{x : x \text{ marginal, light}, Q_x\}|] \geqslant \left(1 - \frac{2t^{1/2}(\log(t^{1/2}(t-1)n))^{1/2}}{n^{1/2}}\right)\mathbb{E}[|R(\mathbf{f})|].$$

However, for any f, the number of marginal points is at most N since each chain has at most one marginal point. Therefore, (using the fact that $1/(1-x) \leqslant 1 + 2x$ for $x \in [0, 1/2]$,)

$$\sum_x \mathbb{P}(x \text{ light}, Q_x) = \mathbb{E}[|R(\mathbf{f})|] \leqslant \left(1 + \frac{4t^{1/2}(\log(t^{1/2}(t-1)n))^{1/2}}{n^{1/2}}\right)N. \tag{4.12}$$

Now, by combining (4.9), (4.10), (4.11), and (4.12), we obtain that

$$Q \leqslant \frac{(2+\epsilon')t^{1/2}(\log(t^{1/2}(t-1)n))^{1/2}}{n^{1/2}}N. \qquad \square$$

Acknowledgements

The authors would like to thank the referees for carefully reading the manuscript and proposing various improvements in the presentation of the paper.

References

[And67] I. Anderson. On primitive sequences. *Journal of the London Mathematical Society*, s1-42(1):137–148, 1967. doi:10.1112/jlms/s1-42.1.137.

[BTW16] József Balogh, Andrew Treglown, and Zsolt Adam Wagner. Applications of graph containers in the boolean lattice. *Random Structures & Algorithms*, 49(4), 2016. doi:10.1002/rsa.20666.

[CCT12] Teena Carroll, Joshua N. Cooper, and Prasad Tetali. Counting antichains and linear extensions in generalizations of the Boolean lattice. 2012. URL: https://tetali.math.gatech.edu/PUBLIS/CCT.pdf.

[dBvETK51] Nicolaas G de Bruijn, Cornelia van Ebbenhorst Tengbergen, and Dirk Kruyswijk. On the set of divisors of a number. *Nieuw Arch. Wiskunde (2)*, 23:191–193, 1951.

[Ded97] R. Dedekind. *Über Zerlegungen von Zahlen Durch Ihre Grössten Gemeinsamen Theiler*, pages 1–40. Vieweg+Teubner Verlag, Wiesbaden, 1897. doi:10.1007/978-3-663-07224-9_1.

[ES35] P. Erdős and G. Szekeres. A combinatorial problem in geometry. *Compositio Math.*, 2:463–470, 1935.

[FPSS12] Jacob Fox, János Pach, Benny Sudakov, and Andrew Suk. Erdős–Szekeres-type theorems for monotone paths and convex bodies. *Proceedings of the London Mathematical Society*, 105(5), 2012. doi:10.1112/plms/pds018.

[FRRT23] Victor Falgas-Ravry, Eero Räty, and István Tomon. Dedekind's problem in the hypergrid. 2023. arXiv:2310.12946.

[Gal14] David J. Galvin. Three tutorial lectures on entropy and counting. 2014. arXiv:1406.7872.

[Gal19] David Galvin. Independent sets in the discrete hypercube. 2019. arXiv:1901.01991.

[GGFS86] Fan Chung Graham, Ronald L. Graham, Peter Frankl, and James B. Shearer. Some intersection theorems for ordered sets and graphs. *Journal of Combinatorial Theory, Series A*, 43(1):23–37, 1986. doi:10.1016/0097-3165(86)90019-1.

[Kah01] Jeff Kahn. An entropy approach to the hard-core model on bipartite graphs. *Combinatorics, Probability, & Computing*, 10(3):219–237, 2001. doi:10.1017/s0963548301004631.

[Kah02] Jeff Kahn. Entropy, independent sets and antichains: A new approach to Dedekind's problem. *Proceedings of the American Mathematical Society*, 130(2):371–378, 2002. doi:10.1090/s0002-9939-01-06058-0.

[Kle69] Daniel J. Kleitman. On Dedekind's problem: The number of monotone Boolean functions. *Proceedings of the American Mathematical Society*, 21(3):677–682, 1969. doi:10.1090/s0002-9939-1969-0241334-6.

[KM75] Daniel J. Kleitman and George Markowsky. On Dedekind's problem: the number of isotone Boolean functions. II. *Transactions of the American Mathematical Society*, 213:373–390, 1975. doi:10.1090/s0002-9947-1975-0382107-0.

[Kor80] Aleksej Dmitrievich Korshunov. The number of monotone boolean functions (russian). *Problemy Kibernet*, 38:5–108, 1980.

[MR08] Lutz Mattner and Bero Roos. Maximal probabilities of convolution powers of discrete uniform distributions. *Statistics & Probability Letters*, 78(17):2992–2996, 2008. doi:10.1016/j.spl.2008.05.005.

[MS14] Guy Moshkovitz and Asaf Shapira. Ramsey theory, integer partitions and a new proof of the Erdős–Szekeres theorem. *Advances in Mathematics*, 262:1107–1129, 2014. doi:10.1016/j.aim.2014.06.008.

[MSW15] Kevin Milans, Derrick Stolee, and Douglas West. Ordered Ramsey theory and track representations of graphs. *Journal of Combinatorics*, 6(4):445–456, 2015. doi:10.4310/joc.2015.v6.n4.a3.

[NSS18] Jonathan Noel, Alex Scott, and Benny Sudakov. Supersaturation in posets and applications involving the container method. *Journal of Combinatorial Theory, Series A*, 154:247–284, 2018. doi:10.1016/j.jcta.2017.08.019.

[Pip99] Nicholas Pippenger. Entropy and enumeration of Boolean functions. *IEEE Trans. Inf. Theory*, 45(6):2096–2100, 1999. doi:10.1109/18.782146.

[PZ24] Cosmin Pohoata and Dmitrii Zakharov. On the number of high-dimensional partitions. *Proceedings of the London Mathematical Society*, 128(2), February 2024. doi:10.1112/plms.12586.

[Sap89] Alexander Antonovich Sapozhenko. The number of antichains in multilayered ranked sets (russian). *Diskret. Mat.*, 1(1):110–128, 1989.

[Tsa19] Shen-Fu Tsai. A simple upper bound on the number of antichains in $[t]^n$. *Order*, 36:507–510, 2019. doi:10.1007/s11083-018-9480-5.