

The Number of Linear Extensions of the Boolean Lattice

Graham R. Brightwell *
Department of Mathematics
London School of Economics
Houghton St.
London WC2A 2AE
England

Prasad Tetali †
School of Mathematics & College of Computing
Georgia Tech.
Atlanta
GA 30332-0160
U.S.A.

28 October 2003

CDAM Research Report LSE-CDAM-2003-19

Abstract

Let $L(Q^t)$ denote the number of linear extensions of the t -dimensional Boolean lattice Q^t . We use the entropy method of Kahn to show that

$$\frac{\log(L(Q^t))}{2^t} = \log \binom{t}{\lfloor t/2 \rfloor} - \frac{3}{2} \log e + o(1),$$

where the logarithms are base 2. We also find the exact maximum number of linear extensions of a d -regular bipartite order on n elements, in the case when n is a multiple of $2d$.

*Email: g.r.brightwell@lse.ac.uk

†Research supported in part by NSF grant DMS-0100289; email: tetali@math.gatech.edu

1 Introduction

Let $L(P)$ denote the number of linear extensions of a poset P . A natural problem is to estimate $L(P)$ when P is the Boolean lattice Q^t , consisting of the subsets of $[t] = \{1, 2, \dots, t\}$, ordered by inclusion. This problem was apparently first posed by Richard Stanley (see Sha and Kleitman [14]), although it has also been raised by several others independently.

The Boolean lattice Q^t is naturally partitioned into its *level sets* (or simply *levels*) A_j , where A_j consists of those subsets of $[t]$ of size exactly j . A trivial lower bound on $L(Q^t)$ is $\prod_{j=0}^t \binom{t}{j}!$, since this counts exactly those linear extensions of Q^t in which all of the j th level comes below the $(j+1)$ st, for each j .

A trivial upper bound is $\binom{t}{\lfloor t/2 \rfloor}^{2^t}$; to see this, consider writing down a linear extension from the bottom up – at each stage the set of possible next elements is an antichain in Q^t , and so has size at most $\binom{t}{\lfloor t/2 \rfloor}$. For many purposes, these bounds are sufficiently close: they can be written as

$$\log \binom{t}{\lfloor t/2 \rfloor} - \frac{3}{2} \log e + o(1) \leq \frac{\log(L(Q^t))}{2^t} \leq \log \binom{t}{\lfloor t/2 \rfloor}.$$

Here and throughout, \log denotes the base 2 logarithm.

The only previous improvement on these trivial bounds was made by Sha and Kleitman [14], who improved the upper bound to

$$L(Q^t) \leq \prod_{j=0}^t \binom{t}{j}^{\binom{t}{j}} \leq \prod_{j=0}^t \binom{t}{j}! \exp(2^t),$$

yielding

$$\frac{\log(L(Q^t))}{2^t} \leq \log \binom{t}{\lfloor t/2 \rfloor} - \frac{1}{2} \log e + o(1).$$

Sha and Kleitman suggested that the simple lower bound in fact gives the right value of the constant term in the asymptotic expansion for $\log(L(Q^t))/2^t$, namely $\frac{3}{2} \log e$. They also remarked that the lower bound can be improved slightly to

$$\frac{\log L(Q^t)}{2^t} \geq \log \binom{t}{\lfloor t/2 \rfloor} - \frac{3}{2} \log e + \frac{C}{t},$$

for some positive constant C . Our primary purpose in this paper is to prove the following result, which indeed shows that the constant term in the asymptotic expansion is $\frac{3}{2} \log e$, and gets reasonably close to the error term.

Theorem 1.1 *For any positive integer t ,*

$$L(Q^t) \leq \exp \left(2^t \frac{6 \ln t}{t} \right) \prod_{j=0}^t \binom{t}{j}!.$$

Corollary 1.2

$$\frac{\log(L(Q^t))}{2^t} = \log \binom{t}{\lfloor t/2 \rfloor} - \frac{3}{2} \log e + O\left(\frac{\ln t}{t}\right).$$

Theorem 1.1 supports the intuition that, in a typical linear extension of Q^t , most of the j th level appears below most of the $(j + 1)$ st, for each j .

Theorem 1.1 is a special case of a more general result. A *ranked poset* P is a poset in which every maximal chain has the same length. A ranked poset is naturally partitioned into *ranks* A_1, \dots, A_k , where A_j contains those elements x such that the longest chain with top element x has exactly j elements. In a ranked poset, if $x \in A_j$ is covered by an element y , then $y \in A_{j+1}$. Also, every element in A_j is above some element of A_{j-1} (provided $j > 1$) and below some element of A_{j+1} (provided $j < k$).

For a positive integer $k \geq 2$, and for $j = 1, 2, \dots, k$, let n_j , u_j , and d_j be arbitrary positive integers. Now let P be an n -element ranked poset with ranks A_1, A_2, \dots, A_k , of sizes n_1, n_2, \dots, n_k respectively, such that every element of A_j is above exactly d_j elements of A_{j-1} ($j \geq 2$), and every element of A_j is below exactly u_j elements of A_{j+1} ($j \leq k - 1$). Finally let $r = \sum_{j=2}^k (n_{j-1}/d_j)$.

Theorem 1.3 *For P a poset as above, the number $L(P)$ of linear extensions of P satisfies*

$$L(P) \leq \left(\frac{2e(k-1)n}{r}\right)^r n_1! \cdots n_k! . \tag{1}$$

Of course, $n_1! \cdots n_k!$ is a lower bound on $L(P)$ for such a poset P . We get a useful bound whenever $r = o(n/\ln k)$; if we think of r as approximately n/\tilde{d} , where \tilde{d} is the harmonic mean of the “down-degrees” of elements of P , then this corresponds to the mild condition $\tilde{d} = \omega(\ln k)$. In this case, the multiplicative “error term” $\left(\frac{2e(k-1)n}{r}\right)^r$ is $2^{o(n)}$; in the context of counting linear extensions, this means that $n_1! \cdots n_k!$ is a good estimate of $L(P)$.

Our proof of this theorem is based on what seems to be emerging as an “entropy method,” stemming from Kahn’s proof [10] of a result of Kleitman and Markowsky on Dedekind’s problem concerning the number of antichains in the Boolean lattice. Other applications of this method can be found in [2], [8] and [9].

In the case where there are just two ranks, i.e., the poset P is *bipartite*, we are able to prove a more precise result.

Theorem 1.4 *Let P be a ranked poset on n elements, with two ranks A and B , such that every element of A is below exactly u elements of B , and every element of B is above exactly d elements of A . Then*

$$L(P) \leq n! \binom{d+u}{u}^{-n/(d+u)} .$$

This result is best possible for n a multiple of $d + u$. Indeed the expression given is the number of linear extensions of the disjoint union of $n/(d + u)$ copies of the complete bipartite poset $K_{d,u}$ with d minimal and u maximal elements.

Perhaps curiously, we cannot replace “exactly” by “at least” in the degree conditions of Theorem 1.4. Indeed, for n large, form a bipartite poset P with some d elements of A below all elements of B , some u elements of B above all elements of A , and no other relations. Then $L(P) \geq (n - d - u)!$, which is much larger than the bound in Theorem 1.4.

Fishburn and Trotter [7] studied a seemingly related problem, where they fixed the number n of elements and the number m of comparable pairs, and sought to maximize the number of linear extensions. They proved that the extremal poset for their problem is always a semiorder; our extremal examples are not.

In Section 2, we provide the necessary background on entropy. Then we prove Theorem 1.4 in Section 3; this proof is particularly close to one given by Kahn in [9].

In Section 4, we give the proof of Theorem 1.3, which requires overcoming various technical difficulties. Theorem 1.1 follows almost immediately from Theorem 1.3.

The result, and the method, of Sha and Kleitman was extended by Shastri [15] to cover any ranked poset satisfying the LYM condition, the sizes of whose ranks form a symmetric unimodal sequence. In Section 5, we show that the Sha-Kleitman bound $L(P) \leq \prod_{j=1}^k n_j^{n_j}$ holds for any ranked poset satisfying the LYM condition, regardless of the degrees and rank sizes n_1, \dots, n_k : in fact this is an immediate consequence of a result of Kahn and Kim [11] – for which we supply an alternative proof that might be of some interest.

2 Basics on Entropy

In this section we recall various basic facts on entropy, including a powerful lemma of Shearer. Our presentation follows that in [9] very closely. Some standard references for material on entropy include [5], [6], and [12].

Let \mathbf{X} , \mathbf{Y} , and \mathbf{Z} denote arbitrary discrete random variables. The *entropy* of \mathbf{X} is defined to be

$$H(\mathbf{X}) = \sum_x p(x) \log \frac{1}{p(x)},$$

where $p(x) = \mathbf{P}(\mathbf{X} = x)$. The *conditional entropy* of \mathbf{X} given \mathbf{Y} is

$$H(\mathbf{X} \mid \mathbf{Y}) = \mathbf{E}H(\mathbf{X} \mid \mathbf{Y} = y) = \sum_y p(y) \sum_x p(x|y) \log \frac{1}{p(x|y)},$$

where $p(x|y) = \mathbf{P}(\mathbf{X} = x \mid \mathbf{Y} = y)$.

For a random vector $\mathbf{X} = (\mathbf{X}_1, \dots, \mathbf{X}_n)$, viewed as a random variable, the *chain rule* of entropy tells us that

$$H(\mathbf{X}) = H(\mathbf{X}_1) + H(\mathbf{X}_2 \mid \mathbf{X}_1) + \dots + H(\mathbf{X}_n \mid \mathbf{X}_1, \dots, \mathbf{X}_{n-1}). \quad (2)$$

We will also make use of the inequalities

$$H(\mathbf{X}) \leq \log |\text{range}(\mathbf{X})|, \quad \text{with equality if } \mathbf{X} \text{ is uniform,} \quad (3)$$

$$H(\mathbf{X} \mid \mathbf{Y}) \leq H(\mathbf{X}),$$

and more generally,

$$\text{if } \mathbf{Y} \text{ determines } \mathbf{Z} \text{ then } H(\mathbf{X} \mid \mathbf{Y}) \leq H(\mathbf{X} \mid \mathbf{Z}), \quad (4)$$

where “ \mathbf{Y} determines \mathbf{Z} ” means formally that $H(\mathbf{Z} \mid \mathbf{Y}) = 0$. Note that (2) and (4) imply

$$H(\mathbf{X}) \leq H(\mathbf{Y}) + H(\mathbf{X} \mid \mathbf{Y}) \quad (5)$$

and also the sub-additive property of entropy:

$$H(\mathbf{X}_1, \dots, \mathbf{X}_n) \leq \sum H(\mathbf{X}_i). \quad (6)$$

We also have a conditional version of (6):

$$H(\mathbf{X}_1, \dots, \mathbf{X}_n \mid \mathbf{Y}) \leq \sum H(\mathbf{X}_i \mid \mathbf{Y}).$$

Finally we need the following lemma of Shearer (see page 33 in [4]), which refines the sub-additivity property (6). A particularly elegant and short proof of this useful lemma is given in [13]. For a random vector $\mathbf{X} = (\mathbf{X}_1, \dots, \mathbf{X}_m)$ and $A \subseteq [m]$, set $\mathbf{X}_A = (\mathbf{X}_i : i \in A)$.

Lemma 2.1 *Let $\mathbf{X} = (\mathbf{X}_1, \dots, \mathbf{X}_m)$ be a random vector and \mathcal{A} a collection of (not necessarily distinct) subsets of $[m]$, with each element of $[m]$ contained in at least d members of \mathcal{A} . Then*

$$H(\mathbf{X}) \leq \frac{1}{d} \sum_{A \in \mathcal{A}} H(\mathbf{X}_A).$$

The conditional analog of Shearer’s lemma, where all entropies are taken conditional on the same random variable \mathbf{Y} , is true as well, and we make use of it in our proof.

3 Bipartite posets

In this section, we prove Theorem 1.4.

Let C_M denote the M -element chain with groundset $[M] = \{1, \dots, M\}$, with the standard order. We make use of an idea of Shepp [16, 17]: rather than work directly with linear extensions, we count the number $O_M(P)$ of *order-preserving maps* from the n -element poset P to C_M , i.e., functions f from P to C_M such that $x < y$ implies $f(x) \leq f(y)$. The connection with linear extensions is given by the following easy lemma, first observed by Shepp.

Lemma 3.1 For any n -element poset P ,

$$\lim_{M \rightarrow \infty} \frac{O_M(P)}{M^n} = \frac{L(P)}{n!}.$$

Thus Theorem 1.4 will be an easy consequence of the following result.

Theorem 3.2 Let P be an n -element ranked poset with two ranks A and B , such that each element of A is below exactly u elements of B , and each element of B is above exactly d elements of A , and let M be any natural number. Then

$$O_M(P) \leq \left(\sum_{j=1}^M [(M-j+1)^u - (M-j)^u] j^d \right)^{n/(d+u)} = O_M(K_{d,u})^{n/(d+u)}.$$

To see the final identity, note that the j -term in the sum counts order-preserving maps f of $K_{d,u}$ in which $\min_{y \in B} f(y)$ is equal to j . Of course, here too we have equality when P is the disjoint union of copies of $K_{d,u}$.

For $M = 2$, order-preserving maps from P to C_2 are in 1-1 correspondence with antichains in P , which in turn are exactly independent sets in the (bipartite) comparability graph. So the number of independent sets in a “ (u, d) -biregular” bipartite graph on n vertices is at most $(2^u + 2^d - 1)^{n/(d+u)}$, with equality if the graph is a disjoint union of $K_{d,u}$ s. This result is in fact Theorem 4.3 of Kahn [9], and in this case our proof is nothing more than a direct translation of Kahn’s. Indeed, our proof of Theorem 3.2 follows Kahn’s proof very closely.

Proof. Let P be a poset with ranks A and B satisfying the conditions of the theorem, and let \mathbf{f} be a uniform random variable over all elements of $O_M(P)$. Set $a = |A|$ and $b = |B|$, so that $n = a + b$.

For X a subset of the ground-set of P , let \mathbf{f}_X be the restriction of \mathbf{f} to X . For $x \in A$, let $U(x)$ be the set of elements in B covering x , and define the random variable $\mathbf{Y}_x = \min \mathbf{f}_{U(x)}$. Finally, for $x \in A$ and $j \in [M]$, let $p_x(j)$ be the probability that \mathbf{Y}_x is equal to j .

Now we have

$$\log O_M(P) = H(\mathbf{f}) = H(\mathbf{f}_B) + H(\mathbf{f}_A \mid \mathbf{f}_B).$$

If $B = \{y_1, \dots, y_b\}$, set $\mathbf{f}_i = \mathbf{f}(y_i)$, and think of \mathbf{f}_B as the vector $(\mathbf{f}_1, \dots, \mathbf{f}_b)$. We apply Shearer’s lemma to this vector, with \mathcal{A} the collection of sets $\{i : y_i \in U(x)\}$, for $x \in A$. As each y_i is in d of these sets, we get

$$H(\mathbf{f}_B) \leq \frac{1}{d} \sum_{x \in A} H(\mathbf{f}_{U(x)}).$$

Furthermore, noting that $H(\mathbf{Y}_x \mid \mathbf{f}_{U(x)}) = 0$, we have

$$H(\mathbf{f}_{U(x)}) = H(\mathbf{Y}_x) + H(\mathbf{f}_{U(x)} \mid \mathbf{Y}_x) = \sum_{j=1}^M p_x(j) \log(1/p_x(j)) + \sum_{j=1}^M p_x(j) H(\mathbf{f}_{U(x)} \mid \mathbf{Y}_x = j).$$

As $|U(x)| = u$ for every x , there are exactly $(M - j + 1)^u - (M - j)^u$ values that $\mathbf{f}_{U(x)}$ can take consistent with $\min \mathbf{f}_{U(x)} = \mathbf{Y}_x = j$, so

$$H(\mathbf{f}_{U(x)} \mid \mathbf{Y}_x = j) \leq \log [(M - j + 1)^u - (M - j)^u].$$

The term $H(\mathbf{f}_A \mid \mathbf{f}_B)$ is at most

$$\begin{aligned} \sum_{x \in A} H(\mathbf{f}(x) \mid \mathbf{f}_B) &\leq \sum_{x \in A} H(\mathbf{f}(x) \mid \mathbf{Y}_x), \quad \text{using (4)} \\ &= \sum_{x \in A} \sum_{j=1}^M p_x(j) H(\mathbf{f}(x) \mid \mathbf{Y}_x = j) \leq \sum_{x \in A} \sum_{j=1}^M p_x(j) \log j, \end{aligned}$$

as, if $\min \mathbf{f}_{U(x)} = \mathbf{Y}_x = j$, then $\mathbf{f}(x)$ takes one of the values $1, \dots, j$.

Combining all these estimates, we have:

$$\log O_M(P) \leq \frac{1}{d} \sum_{x \in A} \sum_{j=1}^M p_x(j) \log \left(\frac{[(M - j + 1)^u - (M - j)^u] j^d}{p_x(j)} \right).$$

Finally, we apply Jensen's inequality to the sum, and find that

$$\log O_M(P) \leq \frac{1}{d} \sum_{x \in A} \log \left(\sum_{j=1}^M [(M - j + 1)^u - (M - j)^u] j^d \right),$$

so

$$O_M(P) \leq \left(\sum_{j=1}^M [(M - j + 1)^u - (M - j)^u] j^d \right)^{a/d}.$$

This is the desired result, as $au = bd$, so $a/d = b/u = n/(d + u)$. ■

To deduce Theorem 1.4, we could estimate the sum in Theorem 3.2, but it is easier and more informative to use the result in the form $O_M(P) \leq O_M(K_{d,u})^{n/(d+u)}$.

Proof of Theorem 1.4. For P a poset satisfying the conditions of the theorem, we have, using Theorem 3.2 and Lemma 3.1,

$$\begin{aligned} \frac{L(P)}{n!} &= \lim_{M \rightarrow \infty} \frac{O_M(P)}{M^n} \leq \lim_{M \rightarrow \infty} \frac{O_M(K_{d,u})^{n/(d+u)}}{M^n} = \left(\lim_{M \rightarrow \infty} \frac{O_M(K_{d,u})}{M^{d+u}} \right)^{n/(d+u)} \\ &= \left(\frac{L(K_{d,u})}{(d+u)!} \right)^{n/(d+u)} = \left(\frac{d! u!}{(d+u)!} \right)^{n/(d+u)}, \end{aligned}$$

which is the required result. ■

As an example of an application, consider the poset consisting of the middle two levels of the Boolean cube Q^t , with $t = 2s - 1$, so the number of elements is $n = 2\binom{2s-1}{s}$. This is a regular bipartite poset with up- and down-degree both equal to s , so the number of linear extensions is at most

$$n! \binom{2s}{s}^{-n/2s} \leq (n/2)!^2 \exp((1 + o(1))n \log s/4s).$$

We do not know of any other way to obtain this good an estimate.

4 Proof of Main Theorem

As in the introduction, let P be a ranked poset with ranks A_1, A_2, \dots, A_k , of sizes n_1, n_2, \dots, n_k respectively, such that every element of A_j is below exactly u_j elements of A_{j+1} ($j \leq k-1$), and every element of A_j is above exactly d_j elements of A_{j-1} ($j \geq 2$).

While our basic approach will be the same as in the previous section, we have a number of additional difficulties to overcome in this more general case.

Proof of Theorem 1.3. As in the bipartite case, our plan is to bound the number $O_M(P)$ of order-preserving maps from P to the chain C_M , for M suitably larger than n , and then to deduce the desired bound on $L(P)$ from it.

Again, let \mathbf{f} be a uniform random variable on the set of all order-preserving maps from P to C_M , and let \mathbf{f}_X denote the restriction of the map \mathbf{f} to the subset X . For $x \in A_i$ with $i \leq k-1$, set $U(x) = \{y \in A_{i+1} : y > x\}$ (the set of u_i elements covering x), and let the random variable \mathbf{B}_x be $\left\lceil \frac{s}{M} \min \mathbf{f}_{U(x)} \right\rceil$, where $s \leq M$ is fixed and will be specified later. Thus \mathbf{B}_x takes an integer value between 1 and s inclusive; we think of C_M as divided into blocks of consecutive integers of size M/s : \mathbf{B}_x tells us which is the lowest block an element of $U(x)$ is mapped to. Suppose that $x > y$ in P , with $\mathbf{B}_x = b_x$ and $\mathbf{B}_y = b_y$; then $\mathbf{f}(x) \leq \min \mathbf{f}_{U(x)} \leq \frac{M}{s} b_x$, while $\mathbf{f}(x) \geq \min \mathbf{f}_{U(y)} > \frac{M}{s} (b_y - 1)$, so $\mathbf{f}(x)$ takes one of at most $\frac{M}{s} (b_x - b_y + 1)$ values. Therefore,

$$H(\mathbf{f}_x \mid \mathbf{B}_x = b_x, \mathbf{B}_y = b_y) \leq \log((M/s)(b_x - b_y + 1)), \quad (7)$$

which is a crucial bound in what follows.

For $x \in P$ and $b \in [s]$, let $p_x(b)$ be the probability that $\mathbf{B}_x = b$. For $y < x$ in P , and $c \leq b$ in $[s]$, let $p_{y|x}(c \mid b)$ be the probability that $\mathbf{B}_y = c$ conditioned on $\mathbf{B}_x = b$.

We have

$$\log O_M(P) = H(\mathbf{f}) = H(\mathbf{f}_{A_k}) + H(\mathbf{f}_{A_{k-1}} \mid \mathbf{f}_{A_k}) + \dots + H(\mathbf{f}_{A_1} \mid \mathbf{f}_{A_2 \cup \dots \cup A_k}).$$

We claim that, for $j = 1, \dots, k$, the sum of the final j terms above, i.e.,

$$H(\mathbf{f}_{A_1 \cup \dots \cup A_j} \mid \mathbf{f}_{A_{j+1} \cup \dots \cup A_k}),$$

is bounded above by

$$\frac{1}{n_j} \sum_{x \in A_j} \sum_{b_j=1}^s p_x(b_j) \log \left(s^{r_j} (M/s)^{n_1 + \dots + n_j} \max \left[b_1^{n_1} \prod_{i=2}^j (b_i - b_{i-1} + 1)^{n_i} \right] \right),$$

where $r_j = \sum_{i=2}^j (n_{i-1}/d_i)$ (so $r_k = r$) and the maximum is taken over all (b_1, \dots, b_{j-1}) with $1 \leq b_1 \leq b_2 \leq \dots \leq b_j$: it is well-known and easy to check (using Lagrange multipliers) that this maximum is attained by taking each $b_i - b_{i-1} + 1$ proportional to n_i (where $b_0 := 1$).

We prove the above claim by induction on j . For $j = 1$, we argue that:

$$\begin{aligned} H(\mathbf{f}_{A_1} \mid \mathbf{f}_{A_2 \cup \dots \cup A_k}) &\leq \sum_{x \in A_1} H(\mathbf{f}(x) \mid \mathbf{f}_{A_2 \cup \dots \cup A_k}) \leq \sum_{x \in A_1} H(\mathbf{f}(x) \mid \mathbf{B}_x) \\ &= \sum_{x \in A_1} \sum_{b_1=1}^s p_x(b_1) H(\mathbf{f}_x \mid \mathbf{B}_x = b_1) \leq \sum_{x \in A_1} \sum_{b_1=1}^s p_x(b_1) \log[(M/s)b_1] \\ &= \frac{1}{n_1} \sum_{x \in A_1} \sum_{b_1=1}^s p_x(b_1) \log[(M/s)^{n_1} b_1^{n_1}], \end{aligned}$$

which is as claimed. (Note that $r_1 = 0$.)

For the induction step, we assume the result for $j - 1$, and consider the new term $H(\mathbf{f}_{A_j} \mid \mathbf{f}_{A_{j+1} \cup \dots \cup A_k})$. This is at most

$$\begin{aligned} &\frac{1}{d_j} \sum_{y \in A_{j-1}} H(\mathbf{f}_{U(y)} \mid \mathbf{f}_{A_{j+1} \cup \dots \cup A_k}), \quad \text{by Shearer's lemma} \\ &\leq \frac{1}{d_j} \sum_{y \in A_{j-1}} (H(\mathbf{B}_y \mid \mathbf{f}_{A_{j+1} \cup \dots \cup A_k}) + H(\mathbf{f}_{U(y)} \mid \mathbf{B}_y, \mathbf{f}_{A_{j+1} \cup \dots \cup A_k})), \quad \text{using the conditional version of (5)} \\ &\leq \frac{1}{d_j u_{j-1}} \sum_{y \in A_{j-1}} \sum_{x \in U(y)} H(\mathbf{B}_y \mid \mathbf{B}_x) + \frac{1}{d_j} \sum_{y \in A_{j-1}} \sum_{x \in U(y)} H(\mathbf{f}_x \mid \mathbf{B}_y, \mathbf{B}_x), \quad \text{using (4)} \\ &= \frac{1}{d_j u_{j-1}} \sum_{y \in A_{j-1}} \sum_{x \in U(y)} \sum_{b_j=1}^s p_x(b_j) \sum_{b_{j-1}=1}^{b_j} p_{y|x}(b_{j-1} \mid b_j) \log(1/p_{y|x}(b_{j-1} \mid b_j)) + \\ &\quad + \frac{1}{d_j} \sum_{y \in A_{j-1}} \sum_{x \in U(y)} \sum_{b_j=1}^s p_x(b_j) \sum_{b_{j-1}=1}^{b_j} p_{y|x}(b_{j-1} \mid b_j) H(\mathbf{f}_x \mid \mathbf{B}_y = b_{j-1}, \mathbf{B}_x = b_j) \\ &\leq \frac{1}{d_j u_{j-1}} \sum_{y \in A_{j-1}} \sum_{x \in U(y)} \sum_{b_j=1}^s p_x(b_j) \sum_{b_{j-1}=1}^{b_j} p_{y|x}(b_{j-1} \mid b_j) \log \left(\frac{(M/s)^{u_{j-1}} (b_j - b_{j-1} + 1)^{u_{j-1}}}{p_{y|x}(b_{j-1} \mid b_j)} \right), \end{aligned}$$

using (7).

By the induction hypothesis, the sum of the final $j - 1$ terms is at most

$$\frac{1}{n_{j-1}} \sum_{y \in A_{j-1}} \sum_{b_{j-1}=1}^s p_y(b_{j-1}) \log X(b_{j-1}),$$

where

$$X(b_{j-1}) = s^{r_{j-1}}(M/s)^{n_1+\dots+n_{j-1}} \max [b_1^{n_1} \dots (b_{j-1} - b_{j-2} + 1)^{n_{j-1}}],$$

and this sum can be rewritten as

$$\frac{1}{n_{j-1}u_{j-1}} \sum_{y \in A_{j-1}} \sum_{x \in U(y)} \sum_{b_{j-1}=1}^s \sum_{b_j=b_{j-1}}^s p_x(b_j) p_{y|x}(b_{j-1} | b_j) \log X(b_{j-1}).$$

Therefore the sum of all the relevant terms is at most

$$\begin{aligned} & \frac{1}{d_j u_{j-1}} \sum_{y \in A_{j-1}} \sum_{x \in U(y)} \sum_{b_j=1}^s p_x(b_j) \sum_{b_{j-1}=1}^{b_j} p_{y|x}(b_{j-1} | b_j) \log \left(\frac{(M/s)^{u_{j-1}} (b_j - b_{j-1} + 1)^{u_{j-1}} (X(b_{j-1}))^{d_j/n_{j-1}}}{p_{y|x}(b_{j-1} | b_j)} \right) \\ & \leq \frac{1}{d_j u_{j-1}} \sum_{x \in A_j} \sum_{\substack{y \in A_{j-1} \\ y < x}} \sum_{b_j=1}^s p_x(b_j) \log \left(\sum_{b_{j-1}=1}^{b_j} (M/s)^{u_{j-1}} (b_j - b_{j-1} + 1)^{u_{j-1}} (X(b_{j-1}))^{d_j/n_{j-1}} \right), \end{aligned}$$

by Jensen's inequality, and the above is in turn at most

$$\leq \frac{1}{u_{j-1}} \sum_{x \in A_j} \sum_{b_j=1}^s p_x(b_j) \log \left(s (M/s)^{u_{j-1}} \max_{b_{j-1}} [(b_j - b_{j-1} + 1)^{u_{j-1}} X(b_{j-1})^{d_j/n_{j-1}}] \right)$$

(now we use that $1/u_{j-1} = (n_{j-1}/d_j)(1/n_j)$)

$$\begin{aligned} & = \frac{1}{n_j} \sum_{x \in A_j} \sum_{b_j=1}^s p_x(b_j) \log \left(s^{n_{j-1}/d_j} (M/s)^{n_j} \max_{b_{j-1}} [(b_j - b_{j-1} + 1)^{n_j} X(b_{j-1})] \right) \\ & = \frac{1}{n_j} \sum_{x \in A_j} \sum_{b_j=1}^s p_x(b_j) \log \left(s^{r_j} (M/s)^{n_1+\dots+n_j} \max \left[b_1^{n_1} \prod_{i=2}^j (b_i - b_{i-1} + 1)^{n_i} \right] \right), \end{aligned}$$

which completes the induction step.

Carrying it through to the end, we can treat \mathbf{B}_x as identically equal to s for x in the top class A_k , and we emerge with

$$O_M(P) \leq s^r (M/s)^{n_1+\dots+n_k} \max b_1^{n_1} (b_2 - b_1 + 1)^{n_2} \dots (b_{k-1} - b_{k-2} + 1)^{n_{k-1}} (s - b_{k-1} + 1)^{n_k}.$$

Now we do the maximization and, setting $n = n_1 + \dots + n_k$, we get:

$$\begin{aligned} O_M(P) & \leq s^r (M/s)^n \left(\frac{n_1}{n} (s + k - 1) \right)^{n_1} \dots \left(\frac{n_k}{n} (s + k - 1) \right)^{n_k} \\ & = s^r (M/s)^n (s + k - 1)^n \frac{n_1^{n_1} \dots n_k^{n_k}}{n^n}. \end{aligned}$$

For any M , we certainly have $O_M(P) \geq L(P) \binom{M}{n}$. Therefore

$$\begin{aligned} L(P) &\leq s^r M^n \left(1 + \frac{k-1}{s}\right)^n \frac{n_1^{n_1} \cdots n_k^{n_k}}{n^n} \frac{n!}{(M-n)^n} \\ &\leq s^r \left(1 + \frac{n}{M-n}\right)^n \left(1 + \frac{k-1}{s}\right)^n n_1! \cdots n_k!. \end{aligned}$$

Taking the limit as $M \rightarrow \infty$ now yields, for any choice of s :

$$L(P) \leq s^r e^{(k-1)n/s} n_1! \cdots n_k! .$$

Finally we optimize this bound by setting $s = \lceil (k-1)n/r \rceil$; as $r \leq n$ and $k-1 \geq 1$, we have $(k-1)n/r \leq s \leq 2(k-1)n/r$. Hence

$$L(P) \leq \left(\frac{2(k-1)n}{r}\right)^r e^r n_1! \cdots n_k! ,$$

proving the theorem. ■

Proof of Theorem 1.1. Recall that $r = \sum_{j=2}^k (n_{j-1}/d_j)$. For the Boolean cube $P = Q^t$, we have $n = 2^t$, $k = t+1$, $n_j = \binom{t}{j-1}$ and $d_j = j-1$, so

$$r = \sum_{i=0}^{t-1} \binom{t}{i} \frac{1}{i+1} = \frac{1}{t+1} (2^{t+1} - 2) < 2n/t.$$

Therefore, by Theorem 1.3, we have, for $t \geq 3$,

$$L(Q^t) \leq (e^{t^2})^{2n/t} \prod_{j=0}^t \binom{t}{j}! \leq e^{6n \ln t/t} \prod_{j=0}^t \binom{t}{j}! .$$

Of course, we have $L(Q^t) = \prod \binom{t}{j}!$ for $t < 3$, so the result follows. ■

5 LYM posets

Let P be a ranked poset on $[n]$ with ranks A_1, A_2, \dots, A_k of sizes n_1, n_2, \dots, n_k respectively. The *weight* $w(x)$ of an element $x \in A_j$ is defined to be $1/n_j$. The *LYM condition* is that, for any antichain A of P , the sum $w(A)$ of the weights of elements of A is at most 1. (Note that equality is achieved for each of the ranks A_j .)

Theorem 5.1 *Let P be a ranked poset satisfying the LYM condition, with rank sizes n_1, \dots, n_k . Then $L(P) \leq \prod_{j=1}^k n_j^{n_j}$.*

The well-known *LYM theorem* states that the Boolean lattice Q^t satisfies the LYM condition (see for instance Bollobás [1]), so Theorem 5.1 extends the result of Sha and Kleitman [14], as well as that of Shastri [15], who proved it under the additional assumption that (n_i) is a symmetric unimodal sequence.

Theorem 5.1 is a special case of the following result from Kahn and Kim [11].

Theorem 5.2 *Let P be any n -element poset, and let $w(x)$ be any non-negative function on P satisfying $w(A) = \sum_{x \in A} w(x) \leq 1$ for all antichains A of P . Then $e(P) \leq \prod_{x \in P} 1/w(x)$.*

Theorem 5.1 follows immediately from Theorem 5.2, as the function given by $w(x) = 1/n_j$ for $x \in A_j$ satisfies the hypotheses, so that each of the n_j elements of A_j contributes a factor of n_j to the product $\prod_{x \in P} 1/w(x)$.

Kahn and Kim's proof of Theorem 5.2 uses polyhedral methods, and is based on a theorem of Stanley. (They go on to use these techniques to prove important results on comparison sorting algorithms.) A research report [3] of the first author contains further information on this topic, and also gives an alternative proof of Theorem 5.2, which we include here as it is very short and natural, and may be of some interest.

Proof. Consider the following random procedure for building a linear extension of P from the bottom up. At each stage, from the set of available elements, choose x as the next element with probability proportional to $w(x)$.

Now consider any single linear extension \prec of P , say $x_1 \prec x_2 \prec \dots \prec x_n$. The probability that our random procedure results in \prec is exactly $\prod_{i=1}^n w(x_i)/w(B_i)$, where B_i is the set of elements minimal among $\{x_i, \dots, x_n\}$. Since the B_i are antichains, each $w(B_i)$ is at most 1, and the probability of \prec is at least $\prod_{i=1}^n w(x_i)$. This is independent of the particular linear extension \prec , so we conclude that the total number of possible linear extensions is at most $1/\prod w(x_i)$, as required. ■

The disjoint union P of n/k chains, each of k elements, is an LYM poset, and

$$L(P) = n!/k!^{n/k} = \left(\frac{n}{k}(1 - o(1))\right)^n,$$

in close agreement with the upper bound in Theorem 5.1. What this means is that Theorem 5.1 gives (in some sense) close to the best possible result for LYM posets in general, so that information on the degrees was essential for our improved upper bound in the case of the Boolean lattice.

Acknowledgments. We thank the Institute for Mathematics and Applications (IMA) for funding the summer school in Combinatorics and Applications at Georgia Tech. during July 2003. This was pivotal to our collaboration. We also thank Tom Trotter and the staff of the School of Mathematics (Georgia Tech.) for organizing and hosting the summer school. Finally, we thank the anonymous referee for a very thorough checking of our manuscript.

References

- [1] B. Bollobás, *Combinatorics*, Cambridge University Press 1986.
- [2] B. Bollobás and G.R. Brightwell, The number of k-SAT functions, *Random Structures and Algorithms* **22** (2003) 227–247.
- [3] G.R. Brightwell, The number of linear extensions of ranked posets, LSE-CDAM Research Report LSE-CDAM-2003-18 (2003).
- [4] F.R.K. Chung, P. Frankl, R. Graham and J.B. Shearer, Some intersection theorems for ordered sets and graphs, *J. Combin. Theory Ser. A.* **48** (1986), 23–37.
- [5] T. Cover and J. Thomas, *Elements of Information Theory*, John Wiley, New York (1991).
- [6] I. Csiszár and J. Körner, *Information Theory: Coding Theorems for discrete memoryless systems*, Academic Press, New York (1981).
- [7] P.C. Fishburn and W.T. Trotter, Linear extensions of semiorders: a maximization problem, *Discrete Math.* **103** (1992) 25–40.
- [8] D. Galvin and P. Tetali, On weighted graph homomorphisms, Special DIMACS-AMS volume on *Graph Homomorphisms and Statistical Physics Models*, to appear.
- [9] J. Kahn, An entropy approach to the hard-core model on bipartite graphs, *Combin. Prob. Comp.* **10** (2001), 219–237.
- [10] J. Kahn, Entropy, independent sets and antichains: a new approach to Dedekind’s problem, *Proc. Amer. Math. Soc.* **130** (2002), 371–378.
- [11] J. Kahn and J.H. Kim, Entropy and sorting, *J. Comput. System Sci.* **51** (1995) 390–399
- [12] R.J. McEliece, *The Theory of Information and Coding*, Addison-Wesley, London, 1977.
- [13] J. Radhakrishnan, Entropy and Counting, Preprint of a survey article (November 2001).
- [14] J. Sha and D. J. Kleitman, The number of linear extensions of subset ordering, *Discrete Math.*, **63** (1987), 271–279.

- [15] A. Shastri, Number of linear extensions of symmetric KLYM posets, *Util. Math.* **53** (1998) 25–35.
- [16] L.A. Shepp, The FKG inequality and some monotonicity properties of partial orders, *SIAM J. Algebraic Discrete Methods* **1** (1980) 295–299.
- [17] L. Shepp, The *XYZ* conjecture and the FKG inequality, *Ann. Probab.* **10** (1982) 824–827.