$\Sigma\Pi\Sigma$ Threshold Formulas

Jaikumar Radhakrishnan*
Department of Computer Science
Rutgers University
New Brunswick, NJ 08903, USA

Abstract

A $\Sigma\Pi\Sigma$ formula has the form $\bigvee_u \bigwedge_v \bigvee_w L_{uvw}$, where each L is either a variable or a negated variable. In this paper we study the computation of threshold functions by $\Sigma\Pi\Sigma$ formulas. By combining the proof of the Fredman-Komlós bound [5, 10] and a counting argument, we show that for k and n large and $k \leq n/2$, every $\Sigma\Pi\Sigma$ formula computing the threshold function T_k^n has size at least $\exp(\Omega(\sqrt{k/\ln k}))n \log n$. For k and n large and $k \leq n^{2/3}$, we show that there exist $\Sigma\Pi\Sigma$ formulas for computing T_k^n with size at most $\exp(2\sqrt{k}\ln k)n \log n$.

1 Introduction

The kth threshold function, T_k^n, is the Boolean function on n variables that takes the value 1 precisely when there are at least k 1's in the input. Threshold functions play a central role in the investigation of the complexity of Boolean functions. Their complexity has been studied in various circuit models (see Boppana and Sipser [3], Wegener [18]). In this paper, upper and lower bounds are shown for computing T_k^n using $\Sigma\Pi\Sigma$ formulas. A $\Sigma\Pi\Sigma$ formula has the form $\bigvee_{i=1}^{p} \bigwedge_{j=1}^{t_i} \bigvee_{q \in S_{ij}} q$, where each S_{ij} is a subset of variables and their negations.

The complexity of computing the *majority* function, $T_{\lceil n/2 \rceil}^n$, using constant depth circuits has been well studied [3]. Hastad [6] obtained a nearly optimal lower bound on the size of such circuits. His result implies that any depth d circuit computing T_k^n, $k \leq n/2$, has size $\exp(\Omega(k^{1/(d-1)}))$. Note that for small values of k Hastad's results do not give superlinear lower bounds. Indeed, it has been shown by Newman, Ragde, and Wigderson [12] that for small values of k (bounded by a function of the form $(\log n)^r$, for some constant r), there do exist linear size constant depth circuits computing T_k^n.

The complexity of computing T_k^n using formulas over the basis {AND, OR, NOT} has also been studied. Paterson, Pippenger, and Zwick [13] showed that all threshold functions can be computed by formulas of size $O(n^{4.57})$. Khrapchenko [9] showed that any such formula must have size at least $k(n-k+1)$. Hansel [7] and Krichevskii [11] showed that any formula computing T_k^n, $2 \leq k \leq n-1$, has size $\Omega(n \log n)$. In the monotone case, when only AND and OR gates are allowed, Valiant showed that the majority function can be computed by formulas of size $O(n^{5.3})$. Boppana [2] showed that T_k^n can be computed by monotone formulas of size $O(k^{4.3} n \log n)$. Radhakrishnan [14] showed that any monotone formula computing T_k^n, $2 \leq k \leq \frac{n}{2}$, has size at least $\left\lfloor \frac{k}{2} \right\rfloor n \log(\frac{n}{k-1})$.

For large values of k, the results for constant depth circuits mentioned above provide nearly optimal bounds for constant depth formulas as well. However, the situation is different for small

*Present address: Theoretical Computer Science Group, Tata Institute of Fundamental Research, Bombay, INDIA 400 005.

thresholds. While the $\Omega(n \log n)$ lower bound for T_2^n, due to Hansel and Krichevskii, is tight for $\Sigma\Pi\Sigma$ formulas, for larger thresholds such tight bounds are not known. To better understand the computation of T_k^n by constant depth formulas, Newman, Ragde, and Wigderson [12] considered $\Sigma\Pi\Sigma$ formulas computing T_k^n for small values of k. They showed, under the assumption that each $t_i = k$ (fanin of the AND gates is k), that every $\Sigma\Pi\Sigma$ formula computing T_k^n has size at least $\Omega(kn \log n)$. Under their assumption the problem is equivalent to the problem of covering the complete uniform hypergraph using k-partite hypergraphs. In this setting the problem was studied earlier by Snir [16], who obtained the same lower bounds. It was shown by Radhakrishnan [15] that the results of Snir can be improved using the techniques of Körner [10] and Fredman and Komlós [5]. The result of [15] implies that every $\Sigma\Pi\Sigma$ formula computing T_k^n, with the restriction that the fanin for the AND gates be k, has size $\Omega(\frac{\exp(k)}{k\sqrt{k}} n \log n)$. Using a random family of k-partite hypergraphs one may obtain $\Sigma\Pi\Sigma$ formulas of size $O(\sqrt{k} \exp(k) n \log n)$ [8, 5]. Thus, there exist almost tight bounds on the size of such restricted $\Sigma\Pi\Sigma$ formulas computing T_k^n.

In this paper, we consider $\Sigma\Pi\Sigma$ formulas computing T_k^n, $k \leq \frac{n}{2}$, with no restriction. That is, the t_i need not be k and the formula is permitted to contain negations. We obtain the following results. Suppose that k and n are large numbers.

Result 1. If $k \leq n/2$, then every $\Sigma\Pi\Sigma$ formula computing T_k^n has size $\exp(\sqrt{k}/3)n$.

Result 2. If $k < (\log \log n)^2$, then every $\Sigma\Pi\Sigma$ formula computing T_k^n has size at least $\exp(\delta(k))n \log n$, where $\delta(k) = \frac{1}{50}\sqrt{\frac{k}{\ln k}}$.

Result 3. If $k^{3/2}$ is an integer that divides n, then there exist $\Sigma\Pi\Sigma$ formulas computing T_k^n with size at most $\exp(2\sqrt{k} \log k)n \log n$. These formulas are monotone.

Note that for $k \geq (\log \log n)^2$ the lower bound claimed in the abstract is implied by Result 1. The main contribution of this work is Result 2, which combines an exponential dependence on k, suggested by the small depth circuit lower bounds for the majority function, with the $\Omega(n \log n)$ lower bound of Hansel and Krichevskii. The proof is based on the proof of the Fredman-Komlós bound presented by Körner [10]. Our proof, like Körner's proof, makes use of the notion of *graph entropy*. The idea is to associate graphs with formulas in such a way that graphs of small formulas have low entropy, while a formula computing T_k^n has graph of high entropy.

1.1 Overview

The rest of the paper is organized as follows. In section 2, we recall definitions and facts about *graph entropy*. In section 3, we describe the lower bound results. The lower bound stated as Result 1 above is shown in subsection 3.2. Our main result, Result 2, is derived in subsection 3.3. Assuming a combinatorial lemma, the main argument is presented in subsections 3.3.1 and 3.3.2. The proof of the combinatorial lemma is presented in subsection 3.3.3. Finally, the proof of Result 3 is presented in section 4.

We derive Result 1 using counting arguments that are related to those used later in the proof of the combinatorial lemma. Result 1 can also be obtained from Hastad's lower bound for constant depth circuits computing majority (see Hastad [6, p. 37]). However, we believe that familiarity with the arguments we present will help the reader follow the more involved proof of the combinatorial lemma.

2 Graph Entropy

We shall need the following standard definitions from information theory (see [4]).

Definition 2.1 (Entropy) For a random variable X with finite support, the entropy of X is given by

$$H(X) = -\sum_x \Pr[X = x] \log \Pr[X = x].$$

If X and Y are random variables then (X, Y) will be the random variable taking values in $support(X) \times support(Y)$ according to the joint distribution of X and Y.

Definition 2.2 (Mutual Information) If X and Y are random variables, then their mutual information is given by

$$I(X \wedge Y) = H(X) + H(Y) - H((X, Y)).$$

We shall need the following definitions and facts about graph entropy (see [10]).

Definition 2.3 (Graph Entropy) Let G be a graph. Let $\mathcal{A}(G)$ denote the set of independent sets of G. Let X and Y be random variables taking values in $V(G)$ and $\mathcal{A}(G)$ respectively. We shall say that the pair (X, Y) is *admissible* for G if

1. X takes values in $V(G)$ with uniform distribution; and
2. $\Pr[X = v \text{ and } Y = A] = 0$ whenever $v \notin A$.

The graph entropy $H(G)$ is defined by

$$H(G) = \min\{I(X \wedge Y) : (X, Y) \text{ is admissible for } G\}.$$

Lemma 2.4 (Subadditivity) If F and G are graphs such that $V(G) = V(F) = V$, then $H(F \cup G) \leq H(F) + H(G)$. Here $F \cup G$ denotes the graph on vertex set V with $E(F \cup G) = E(F) \cup E(G)$. □

Lemma 2.5 (Additivity) Let $G_1, G_2, \ldots, G_r$ be the connected components of the graph G. Then,

$$H(G) = \sum_{i=1}^{r} \frac{|V(G_i)|}{|V(G)|} H(G_i). \quad \square$$

The following lemma is a direct consequence of the definition of graph entropy.

Lemma 2.6 (Monotonicity) If F and G are graphs on the same vertex set and $E(F) \subseteq E(G)$, then $H(F) \leq H(G)$. □

Definition 2.7 (Coloring, Entropy of a Coloring) For a graph G, a function c with domain $V(G)$ is a coloring of G if $c(x) \neq c(y)$ whenever $(x, y) \in E(G)$. The entropy of a coloring c is given by $H(c) = H(c(X))$, where the random variable X takes values in $V(G)$ with uniform distribution.

Lemma 2.8 Let c be a coloring of the graph G. Then, $H(G) \leq H(c)$. □

Lemma 2.9 (a) $H(K_n) = \log n$; (b) If $E(G) = \emptyset$, then $H(G) = 0$. □

Let $L(x) = \log((x+1)e)$. We shall need the following lemma relating the entropy and the expectation of a random variable. It can be obtained from Corollary 3.2, Csiszár and Körner [4, page 56].

Lemma 2.10 If X is non-negative integer valued, then $H(X) \leq L(E(X))$. Here $E(X)$ denotes the expectation of X. □

The following lemma is due to Boppana [1].

Lemma 2.11 Every graph G with n vertices and m edges has a coloring with entropy at most $L(\frac{m}{n})$. □

Using Lemma 2.8 we obtain the following corollary to Lemma 2.11.

Corollary 2.12 Every graph with n vertices and m edges has entropy at most $L(\frac{m}{n})$. □

3 The Lower Bounds

In this section we describe the two lower bound results, Result 1 and Result 2.

3.1 Notation

We shall use the following notation and conventions.

A $\Sigma\Pi\Sigma$ formula has the form $\bigvee_{i=1}^{p} \bigwedge_{j=1}^{t_i} \bigvee_{q \in S_{ij}} q$, where each S_{ij} is a subset of variables and their negations. The size of a formula F, denoted by size(F), is the number of occurrences of literals in it. Thus, if F is the $\Sigma\Pi\Sigma$ formula $\bigvee_{i=1}^{p} \bigwedge_{j=1}^{t_i} \bigvee_{q \in S_{ij}} q$, then $\text{size}(F) = \sum_{i=1}^{p} \sum_{j=1}^{t_i} |S_{ij}|$. A $\Pi\Sigma$ formula has the form $\bigwedge_{j=1}^{t} \bigvee_{q \in S_j} q$. We use S_j^+ to denote the set of non-negated variables in S_j.

We use $[n]$ to denote the set $\{1, 2, \ldots, n\}$. For a set S, $\binom{S}{k}$ denotes the set of all k sized subsets of S. A set of size k is called a k-set. $[n]_k$ denotes the set of all sequences of length k of elements of $[n]$ where no element is repeated. We write $(n)_k$ for $|[n]_k|$. Thus, $(n)_k = n(n-1)\ldots(n-k+1)$.

Let f be a function with n variables $x_1, x_2, \ldots, x_n$. We say that f *accepts* $T \subseteq [n]$ if f evaluates to 1 when all the variables $x_i, i \in T$, are given the value 1 and the remaining variables are given the value 0. We say that f *accepts* a sequence $(i_1, i_2, \ldots, i_r)$ of elements of $[n]$ if f accepts the set $\{i_1, i_2, \ldots, i_r\}$. We identify a set of variables with the set of indices of those variables. Similarly, we identify a sequence of variables with the sequence formed by their indices. A function f is said to be *l-immune* if it accepts no set T with $|T| \leq l$. Thus the threshold function T_k^n is $(k-1)$-immune. We shall use this notation for formulas also. For example, a formula *accepts* a set T if the function it computes accepts T; and a formula is *l-immune* if the function it computes is l-immune.

When considering a formula with n variables, we shall normally assume that the n variables are $x_1, x_2, \ldots, x_n$. If F is a formula with n variables and $T \subseteq [n]$, then $F|_T$ denotes the formula obtained from F by fixing the variables appearing in T at 1. Unless it is indicated otherwise, we shall continue to think of $F|_T$ as a formula with n variables (only some of the variables do not appear explicitly). We shall also use the extension of this notation and refer to $F|_\sigma$, where σ is a sequence of elements of $[n]$. For a $\Pi\Sigma$ formula $F = \bigwedge_{j=1}^{t} \bigvee_{q \in S_j} q$ and $T \subseteq [n]$, $F|_T$ will denote the formula obtained from F as follows: if some S_j contains a variable in T, that S_j will not appear in $F|_T$; if some S_j contains the negation of a variable that appears in T, that variable will be deleted from S_j. The formula $F|_\sigma$, where σ is a sequence of elements of $[n]$, is obtained similarly.

3.2 $\Sigma\Pi\Sigma$ formulas of large thresholds

Suppose that k is a large positive number, $n \geq 2k$, and F is a $\Sigma\Pi\Sigma$ formula computing T_k^n. Let $F = \bigvee_{i=1}^{p} \bigwedge_{j=1}^{t_i} \bigvee_{q \in S_{ij}} q$. Let $A_i = \bigwedge_{j=1}^{t_i} \bigvee_{q \in S_{ij}} q$, for $i = 1, \ldots, p$. Since F is $(k-1)$-immune each A_i is $(k-1)$-immune. Further, every input accepted by F is accepted by at least one of the A_i, $i = 1, 2, \ldots, p$.

To show the lower bound on size(F) we proceed as follows. We first show that a $(k-1)$-immune $\Pi\Sigma$ formula of small size cannot accept many k-sets. More precisely, we show that a $(k-1)$-immune $\Pi\Sigma$ formula A accepts at most

$$(\frac{\text{size}(A)}{n}) \exp(-\frac{\sqrt{k}}{3}) \binom{n}{k} \tag{1}$$

of the k-sets. Since F accepts every k-set, it will follow that size$(F) \geq \exp(\frac{\sqrt{k}}{3})n$.

Instead of directly estimating the number of k-sets accepted by a $\Pi\Sigma$ formula A, we will find it more convenient to estimate the probability that a randomly chosen sequence $\sigma \in [n]_k$ is accepted by A. We shall divide such a sequence σ into subsequences σ_L and σ_R, so that $\sigma = \sigma_L \sigma_R$. The right subsequence σ_R will have length k' and the left subsequence σ_L will have length $k - k'$, where $k' = \left\lfloor \sqrt{k} \right\rfloor$.

Assume $A = \bigwedge_{j=1}^{t} \bigvee_{q \in S_j} q$ is a $(k-1)$-immune $\Pi\Sigma$ formula, and σ is chosen from $[n]_k$ with uniform distribution. We say that $\bigvee_{q \in S_j} q$ is a *big* OR if $|S_j^+| \geq \frac{n}{2\sqrt{k}}$. Lemma 3.1 shows that if σ is chosen at random then $A|_{\sigma_L}$ is unlikely to have a big OR. Lemma 3.2 shows that if $A|_{\sigma_L}$ has no big OR then A is unlikely to accept σ. These two facts, when combined, give the bound (1).

Lemma 3.1 $\Pr[A|_{\sigma_L} \text{ has a big OR}] \leq 2\sqrt{k}(\frac{\text{size}(A)}{n}) \exp(-\frac{k-k'}{2\sqrt{k}})$.

Proof: Let σ be chosen randomly from $[n]_k$. For any fixed big OR $\bigvee_{q \in S} q$ of A, the probability that σ_L contains no variables in S^+ is at most $(1 - \frac{1}{2\sqrt{k}})^{k-k'}$. Thus the probability that there is some big OR in $A|_{\sigma_L}$ is at most (*number of big ORs in* A)$(1 - \frac{1}{2\sqrt{k}})^{k-k'}$. It follows that

$$\Pr[A|_{\sigma_L} \textit{ has a big OR }] \leq 2\sqrt{k}(\frac{\text{size}(A)}{n})(1 - \frac{1}{2\sqrt{k}})^{k-k'} \leq 2\sqrt{k}(\frac{\text{size}(A)}{n}) \exp(-\frac{k-k'}{2\sqrt{k}}).$$

□

Lemma 3.2 $\Pr[A|_{\sigma_L \sigma_R} \equiv 1 \mid A|_{\sigma_L} \text{ has no big OR}] \leq k' \exp(-k')$.

Proof: Consider a σ_L such that $A|_{\sigma_L}$ has no big OR. We shall show that the number of extensions σ_R of σ_L such that $A|_{\sigma_L \sigma_R} \equiv 1$ is at most

$$\left(\frac{n}{2\sqrt{k}}\right)^{k'} k'!. \tag{2}$$

On the other hand, the number of all extensions is at least $(n - k + k')_{k'} \geq (\frac{n}{2})^{k'}$. We may then complete the proof of the lemma by noting that

$$\Pr[A|_{\sigma_L \sigma_R} \equiv 1 \mid A|_{\sigma_L} \textit{ has no big OR}] \leq \frac{\left(\frac{n}{2\sqrt{k}}\right)^{k'} k'!}{\left(\frac{n}{2}\right)^{k'}} \leq \left(\frac{1}{\sqrt{k}}\right)^{k'} \left(\frac{k'}{e}\right)^{k'} k' \leq k' \exp(-k').$$

We still have to show the bound (2) on the number of extensions σ_R. The following nondeterministic procedure generates such extensions.

1. Initially set $\sigma'_R \leftarrow$empty.

2. Repeat the following steps until some condition for stopping is met.

   (a) If $A|_{\sigma_L\sigma'_R}$ is identically 0, stop.

   (b) If $|\sigma'_R| = k'$, stop and output σ'_R.

   (c) Since A is a $(k-1)$-immune formula, there must be an OR of $A|_{\sigma_L\sigma'_R}$ that now contains only non-negated variables; for otherwise, A would accept the sequence $\sigma_L\sigma'_R$ of length at most $k-1$. Take the first such OR, say X_0, and extend σ'_R to $\sigma'_R v$, where v is a variable that appears in X_0.

We claim that for every extension σ_R of σ_L, such that A accepts $\sigma_L\sigma_R$, we can find a sequence σ'_R produced by the above procedure that is a reordering of σ_R. To justify this, we shall show that there is a sequence of choices in step (c) that produces a σ'_R that is a reordering of σ_R. Since initially σ'_R is empty, it is contained in σ_R. It will suffice to show that in each iteration we can extend σ'_R by a new choice v so that $\sigma'_R v$ is contained in σ_R. Suppose we are at the ith iteration for some $i \leq k'$ and σ'_R is contained in σ_R. Now A accepts $\sigma_L\sigma_R$. Hence when we are in step (c) of the ith iteration, one of the choices must be a variable in σ_R. Thus we may extend σ'_R with a variable v so that $\sigma'_R v$ is contained in σ_R.

We may, therefore, estimate the number of such extensions σ_R by multiplying by $k'!$ the number of σ'_R produced by the procedure. To bound the number number of σ'_R produced, we observe that the number of choices in each iteration of step (c) is at most $\frac{n}{2\sqrt{k}}$. Thus the number of of σ'_R generated by the above procedure is at most $(\frac{n}{2\sqrt{k}})^{k'}$. Hence the number of extensions σ_R for a fixed σ_L is at most $(\frac{n}{2\sqrt{k}})^{k'}k'!$, establishing the bound (2). This completes the proof of the lemma. □

Lemma 3.3 The number of k-sets accepted by A is at most

$$(\frac{\text{size}(A)}{n})\exp(-\frac{\sqrt{k}}{3})\binom{n}{k}.$$

Proof: We have two cases. If $\text{size}(A) < \frac{n}{2}$ then there are at most $\frac{n}{2}$ variables that appear in A. Since A is $(k-1)$-immune every k-set accepted by A must contain the variables that explicitly appear in A. The number of such sets is at most $\binom{\frac{n}{2}}{k}$. Since for k large this is less than $\exp(-\frac{\sqrt{k}}{3})\binom{n}{k}$, the claim of the lemma holds in this case.

For the second case we have $\text{size}(A) \geq \frac{n}{2}$. Let σ be chosen from $[n]_k$ with uniform distribution. Then

$$\Pr[A|_\sigma \equiv 1] \leq \Pr[A|_{\sigma_L} \textit{ has a big OR}] + \Pr[A|_{\sigma_L\sigma_R} \equiv 1 \mid A|_{\sigma_L} \textit{ has no big OR}].$$

Thus, using Lemma 3.1 and Lemma 3.2 we have

$$\begin{aligned}
\Pr[A|_\sigma \equiv 1] &\leq 2\sqrt{k}(\frac{\text{size}(A)}{n})\exp(-\frac{k-k'}{2\sqrt{k}}) + k'\exp(-k') \\
&\leq (\frac{\text{size}(A)}{n})[2\sqrt{k}\exp(-\frac{k-k'}{2\sqrt{k}}) + 2k'\exp(-k')] \\
&\leq (\frac{\text{size}(A)}{n})4\sqrt{k}\exp(-\frac{k-k'}{2\sqrt{k}}).
\end{aligned}$$

The last inequality holds because $\sqrt{k} \geq k'$ and $\frac{k-\lfloor\sqrt{k}\rfloor}{2\sqrt{k}} \leq \left\lfloor\sqrt{k}\right\rfloor$. For k large enough, we have that

$$4\sqrt{k}\exp(-\frac{k-k'}{2\sqrt{k}}) \leq \exp(-\frac{\sqrt{k}}{3}).$$

Since each k-set corresponds to precisely $k!$ sequences in $[n]_k$, and every sequence in $[n]_k$ corresponds to some k-set, we have that the number of k-sets accepted by A is at most

$$(\frac{\text{size}(A)}{n})\exp(-\frac{\sqrt{k}}{3})\binom{n}{k}.$$

This completes the proof of the lemma. □

Theorem 3.4 For all large k and $n \geq 2k$, every $\Sigma\Pi\Sigma$ formula computing T_k^n has size at least $\exp(\frac{\sqrt{k}}{3})n$.

Proof: Let $F = \bigvee_{i=1}^{p}\bigwedge_{j=1}^{t_i}\bigvee_{q\in S_{ij}} q$ be a $\Sigma\Pi\Sigma$ formula computing T_k^n. For $i = 1, \ldots, p$, let A_i denote the subformula $\bigwedge_{j=1}^{t_i}\bigvee_{q\in S_{ij}} q$. As discussed above each A_i is $(k-1)$-immune and every k-set is accepted by one of the A_i. Using Lemma 3.3 we have that

$$\sum_{i=1}^{p}(\frac{\text{size}(A_i)}{n})\exp(-\frac{\sqrt{k}}{3})\binom{n}{k} \geq \binom{n}{k}.$$

It follows that $\text{size}(F) = \sum_{i=1}^{p}\text{size}(A_i) \geq \exp(\frac{\sqrt{k}}{3})n$. The theorem follows from this. □

3.3 $\Sigma\Pi\Sigma$ formulas for small thresholds

In the previous section we showed a lower bound of $\exp(\Omega(\sqrt{k}))n$ on the size of any $\Sigma\Pi\Sigma$ formula computing T_k^n. However, for small values for k this bound is weak. For example, a lower bound of $n\log n$ is known on the size of any formula computing T_2^n, even when no restrictions are imposed on the depth. For constant k, the result of the previous section does not give any superlinear lower bound. In this section we shall show better lower bounds for small values of k. The main result of this section is the following. For k and n large enough, $k < (\log\log n)^2$, every $\Sigma\Pi\Sigma$ formula computing T_k^n has size at least

$$\exp(\delta(k))n\log n,$$

where $\delta(k) = \frac{1}{50}\sqrt{\frac{k}{\ln k}}$.

We shall show this bound by combining the counting argument of the last section and the entropy arguments used in the proof of the Fredman-Komlós bound [5, 10, 15]. However, as will become clear later, the counting argument needed is much more technical than in the last section.

3.3.1 Preliminaries

Let k and n be fixed. With each formula on n variables we associate a graph. Under this association the graph of a formula computing T_k^n will have high entropy.

Definition 3.5 (Fredman-Komlós graph) Let f be a formula on n variables. For $k \geq 2$, the graph $G(f,k)$ is defined by

$$\begin{aligned} V(G(f,k)) &= \{(C,x) : C \in \binom{[n]}{k-2} \text{ and } x \in [n]-C\}; \\ E(G(f,k)) &= \{((C,x),(D,y)) : C = D \text{ and } f \text{ accepts } C\cup\{x,y\}\}. \end{aligned}$$

In the special case of $k = 2$ we may think of $G(f, k)$ as a graph with vertex set $[n]$ where (i, j) is an edge if and only if $\{i, j\}$ is accepted by f. In our discussion, the parameter k in the above definition will often be clear from the context. For notational convenience, we will then write $G(f)$ instead of $G(f, k)$.

Let f be a formula computing T_k^n. Then $G(f, k)$ consists of $\binom{n}{k-2}$ components, where each component is a complete graph on $n-k+2$ vertices. The following lemma is a direct consequence of Lemma 2.5 and Lemma 2.9(a).

Lemma 3.6 If f is a formula computing T_k^n for $k \geq 2$, then $H(G(f, k)) = \log(n - k + 2)$. □

In general, for any formula f, the subgraph of $G(f, k)$ induced by those vertices (C, x) that have the same value for C will be called a *block* of $G(f, k)$. Thus, there are $\binom{n}{k-2}$ blocks, one for each $C \in \binom{[n]}{k-2}$.

A $\Pi\Sigma$ formula f is *k-optimal* if f is $(k-1)$-immune and no $(k-1)$-immune $\Pi\Sigma$ formula g (on the same set of variables as f) satisfies $\text{size}(g) < \text{size}(f)$ and $G(f, k)$ is a subgraph of $G(g, k)$.

Lemma 3.7 Let $k \geq 2$ and let $f = \bigwedge_{j=1}^{t} \bigvee_{q \in S_j} q$ be a k-optimal formula. Then no S_j has more than $k - 1$ negated variables.

Proof: Suppose some S_j, say S_{j_0}, has at least k negated variables. Consider the formula g obtained by omitting S_{j_0}. We claim that g is $(k - 1)$-immune. To justify this, first note that $\bigvee_{q \in S_{j_0}} q$ accepts all sets T with $|T| < k$. Then, since $f \equiv g \wedge \bigvee_{q \in S_{j_0}} q$ and f is $(k - 1)$-immune (because f is k-optimal), it follows that g is also $(k-1)$-immune. Clearly, $G(f, k)$ is a subgraph of $G(g, k)$ and $\text{size}(g) < \text{size}(f)$. But this contradicts the optimality of f. Hence, no S_j has more than $k - 1$ negated variables. □

Lemma 3.8 Let $f = \bigwedge_{j=1}^{t} \bigvee_{q \in S_j} q$ be a 2-optimal formula. Then no two S_j have the same negated variable.

Proof: Suppose S_i and S_j $(i \neq j)$ have the same negated variable, say $\overline{x}_1$. By the previous lemma they have no other negated variable. Let g be the formula obtained from f by omitting S_i. As before, $G(f, 2)$ is a subgraph of $G(g, 2)$ and $\text{size}(g) < \text{size}(f)$. We claim that g is 1-immune. First, note that if g accepts T and $1 \notin T$ then f accepts T. Also, g does not accept $\{1\}$ because $\bigvee_{q \in S_j} q$ evaluates to 0 on $\{1\}$ (Since f is optimal, S_j cannot contain x_1). Thus if g accepts T and $|T| < 2$ then so does f. Since f is 1-immune, it follows that g is 1-immune. But this contradicts the optimality of f. The lemma follows from this. □

3.3.2 The lower bound

Consider the $\Sigma\Pi\Sigma$ formula $F = \bigvee_{i=1}^{p} \bigwedge_{j=1}^{t_i} \bigvee_{q \in S_{ij}} q$. Suppose that F computes T_k^n. Then $G(F)$ consists of $\binom{n}{k-2}$ disjoint complete graphs of size $n - k + 2$ and has entropy $\log(n - k + 2)$. Let $A_i = \bigwedge_{j=1}^{t_i} \bigvee_{q \in S_{ij}} q$. Note that $G(F)$ is the union of the graphs $G(A_i)$, $i = 1, \ldots, p$. Roughly speaking, we shall show that if the size of A_i is small then $H(G(A_i))$ is also small. Thus we shall obtain a lower bound on $\sum_{i=1}^{p} \text{size}(A_i)$, since the subadditivity of graph entropy provides us the lower bound, $\sum_{i=1}^{p} H(G(A_i)) \geq H(G(F)) = \log(n - k + 2)$, on the sum of the entropies.

To relate $\text{size}(A_i)$ to $H(G(A_i))$, we need to show two results. The first is a combinatorial result which shows, roughly speaking, that if $\text{size}(A_i)$ is small then only a small number of blocks are nonempty in $G(A_i)$. Due to technical difficulties, introduced by the presence of negated variables, we actually show that if some edges are deleted from $G(A_i)$ then most of the blocks are empty. The edges deleted from the different $G(A_i)$ put together are so few that

they do not contribute significantly to the entropy of the final graph. This result, stated as the Combinatorial Lemma (Lemma 3.21), is discussed in detail in section 3.3.3.

However, this result in itself is not sufficient to complete the proof of the lower bound. Even if the number of nonempty blocks is small, each such block may be very dense and $G(A_i)$ could still have entropy which is not small enough for our purposes. The reader may recall that in the proof of the Fredman-Komlós bound as stated in Körner [10] (see also [15]), we were faced with a very similar situation. There, it turned out that the edges within one block were arranged as a bipartite graph and therefore had small entropy. In our case, we cannot make such a strong claim. Instead, to bound the entropy of a block, we observe that the edges contained in a block correspond to the edges accepted by a certain 1-immune $\Pi\Sigma$ formula. For example, the edges contained in the block corresponding to $C \in \binom{[n]}{k-2}$ are in direct correspondence with the edges of $G(A_i|_C, 2)$. Lemma 3.9 relates the size of a 1-immune $\Pi\Sigma$ formula with the entropy of its graph.

Lemma 3.9 Let $A = \bigwedge_{j=1}^{t} \bigvee_{q \in S_j} q$ be a 1-immune $\Pi\Sigma$ formula on n variables. Then

$$H(G(A,2)) \leq 2L(\frac{\text{size}(A)}{n}).$$

Proof: Let B be the smallest 1-immune $\Pi\Sigma$ formula (on the same variables as A) such that $G(A,2)$ is a subgraph of $G(B,2)$. Then B is 2-optimal and $\text{size}(B) \leq \text{size}(A)$. Now, if the statement of the lemma is true for all 2-optimal formulas, then we have, using Lemma 2.6, that

$$H(G(A,2)) \leq H(G(B,2)) \leq 2L(\frac{\text{size}(B)}{n}) \leq 2L(\frac{\text{size}(A)}{n}),$$

and the statement is true for A also. Hence it is sufficient to prove the lemma under the assumption that A is 2-optimal.

Assume that A is 2-optimal. By Lemma 3.7 an S_j may have at most one negated variable. Since A does not accept the empty set, not all S_j have a negated variable. By Lemma 3.8 no two S_j have the same negated variable. Let $S_1, S_2, \ldots, S_{t'}$ not have any negated variable and $S_{t'+1}, S_{t'+2}, \ldots, S_t$ have some negated variable. Further, let the negated variable in $S_{t'+j}$ be x_j for $1 \leq j \leq t - t'$.

Let G_1 be the subgraph of $G(A,2)$ with vertex set $[n]$ and consisting of all edges that have at least one end in $\{1, 2, \ldots, t - t'\}$. Let G_2 be the graph with vertex set $[n]$ consisting of the remaining edges of $G(A,2)$.

Now, for $1 \leq j \leq t - t'$, if $(i,j) \in E(G_1)$ then $x_i \in S_{t'+j}$. It follows that $|E(G_1)| \leq \sum_{j=1}^{t-t'} (|S_{t'+j}| - 1) \leq \text{size}(A)$. By Corollary 2.12 we have that $H(G_1) \leq L(\frac{\text{size}(A)}{n})$.

Next we consider the entropy of G_2. Let $\chi : [n] \to [t']$ be defined as follows.

$$\chi(j) = \begin{cases} 1 & \text{if } 1 \leq j \leq t - t'; \\ \min\{r : S_r \textit{ does not contain } x_j\} & \text{if } t - t' < j \leq n. \end{cases}$$

Since A is 1-immune, every variable x_j not appearing in the negated form in A satisfies $x_j \notin S_r$ for some r, $1 \leq r \leq t'$. Thus χ is well defined. We claim that χ is a coloring of G_2. Let $(i_1, i_2) \in E(G_2)$. Since vertices $1, \ldots, t - t'$ are isolated in G_2, $t - t' < i_1, i_2 \leq n$. Suppose χ colors both i_1 and i_2 by the same color, say r. Then $\bigvee_{q \in S_r} q$ evaluates to 0 on $\{i_1, i_2\}$ and hence (i_1, i_2) is not an edge of $G(A,2)$ and therefore not an edge in G_2. This contradicts our assumption. Hence χ is a coloring of G_2. By our definition $\sum_{j=1}^{n} (\chi(j) - 1) \leq \text{size}(A)$. Thus if the random variable X takes values in $[n]$ with uniform distribution then $E(\chi(X) - 1) \leq \text{size}(A)/n$. It follows from Lemma 2.8 and Lemma 2.10 that

$$H(G_2) \leq H(\chi(X)) = H(\chi(X) - 1) \leq L(\frac{\text{size}(A)}{n}).$$

Using Lemma 2.4 we have $H(G(A,2)) \leq H(G_1) + H(G_2) \leq 2L(\frac{\text{size}(A)}{n})$. □

We now state the combinatorial result to be proved in the next section.

Lemma 3.21 (Combinatorial Lemma) Let $A = \bigwedge_{j=1}^{t} \bigvee_{q \in S_j} q$ be a $(k-1)$-immune $\Pi\Sigma$ formula. Let $\Gamma = \{\gamma \in \binom{[n]}{k} : A \text{ accepts } \gamma\}$. Let $\alpha(k) = \frac{1}{6}\left\lfloor \sqrt{\frac{k}{e^4 \ln k}} \right\rfloor$.

(a) Suppose $\text{size}(A) \leq \frac{n}{2}$. Let

$$\Psi = \{\overline{a} \in \binom{[n]}{k-2} : \exists\, x, y \text{ such that } \overline{a} \cup \{x,y\} \in \Gamma\}.$$

Then $|\Psi| \leq \left(\frac{\text{size}(A)}{n}\right) e^{-\alpha(k)} \binom{n}{k-2}$.

(b) Suppose $\text{size}(A) > \frac{n}{2}$. Then there exists a set $\Delta \subseteq \Gamma$, $|\Delta| \leq n^{-\frac{1}{3}}\binom{n}{k}$, such that if

$$\Psi = \{\overline{a} \in \binom{[n]}{k-2} : \exists\, x, y \text{ such that } \overline{a} \cup \{x,y\} \in \Gamma - \Delta\},$$

then $|\Psi| \leq \left(\frac{\text{size}(A)}{n}\right) e^{-\alpha(k)} \binom{n}{k-2}$.

This result is useful for the following reason. Consider part (b) of the statement. Notice that a set in Δ contributes exactly $\binom{k}{2}$ edges to the graph $G(A,k)$. Hence, roughly speaking, we can infer that if we remove $\binom{k}{2} n^{-\frac{1}{3}} \binom{n}{k}$ edges from $G(A,k)$, then the number of nonempty blocks in the remaining graph is small.

Lemma 3.10 Let A be a $(k-1)$-immune $\Pi\Sigma$ formula on n variables. Let G' be a subgraph of $G(A,k)$ with at most $(\frac{\text{size}(A)}{n}) e^{-\alpha(k)} \binom{n}{k-2}$ nonempty blocks. Then

$$\begin{aligned} H(G') &\leq \frac{\text{size}(A)}{n} \exp(-\alpha(k)) 2L(\frac{\text{size}(A)}{n-k+2}) \quad \textit{if } \text{size}(A) \leq n\exp(\alpha(k)); \\ H(G') &\leq 2L(\frac{\text{size}(A)}{n-k+2}) \quad \textit{if } \text{size}(A) > n\exp(\alpha(k)). \end{aligned}$$

Proof: We think of G' as consisting of $\binom{n}{k-2}$ disjoint blocks G'_D, one for each $D \in \binom{[n]}{k-2}$. There is a natural correspondence between the vertex sets of G' and $G(A|_D, 2)$ and, with this correspondence, $E(G'_D) \subseteq E(G(A|_D, 2))$. (Here we think of $A|_D$ as a formula on $n-k+2$ variables). Since A is $(k-1)$-immune $A|_D$ is 1-immune. By Lemma 2.6 and Lemma 3.9 we have

$$H(G'_D) \leq 2L(\frac{\text{size}(A|_D)}{n-k+2}) \leq 2L(\frac{\text{size}(A)}{n-k+2}).$$

Since the number of nonempty blocks in G' is at most $\frac{\text{size}(A)}{n} e^{-\alpha(k)} \binom{n}{k-2}$, we can conclude from Lemma 2.4 that

$$H(G') \leq \frac{\text{size}(A)}{n} e^{-\alpha(k)} 2L(\frac{\text{size}(A)}{n-k+2}).$$

Since the number of nonempty blocks is at most $\binom{n}{k-2}$ we always have

$$H(G') \leq 2L(\frac{\text{size}(A)}{n-k+2}).$$

□

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

Theorem 3.11 Assume k and n are large numbers such that $k < (\log\log n)^2$. Suppose that $F = \bigvee_{i=1}^{p} \bigwedge_{j=1}^{t_i} \bigvee_{q\in S_{ij}} q$ computes T_k^n. Then

$$\text{size}(F) \geq \exp(\delta(k)) n \log n,$$

where $\delta(k) = \frac{1}{50}\sqrt{\frac{k}{\ln k}}$.

Proof: Let $A_i \equiv \bigwedge_{j=1}^{t_i} \bigvee_{q\in S_{ij}} q$. Let

$$\begin{aligned} I_1 &= \{i : \text{size}(A_i) \leq \frac{n}{2}\}; \\ I_2 &= \{i : \frac{n}{2} < \text{size}(A_i) \leq ne^{\alpha(k)}\}; \\ I_3 &= \{i : \text{size}(A_i) > ne^{\alpha(k)}\}. \end{aligned}$$

For $i \in I_2$, using the Combinatorial Lemma, we write $G(A_i) = G_1(A_i) \cup G_2(A_i)$, where $G_1(A_i)$ has at most $\frac{\text{size}(A_i)}{n} e^{-\alpha(k)} \binom{n}{k-2}$ nonempty blocks and $G_2(A_i)$ has at most $\binom{k}{2} n^{-\frac{1}{3}} \binom{n}{k}$ edges. Now

$$G(F) = \bigcup_i G(A_i) = \bigcup_{i\in I_1} G(A_i) \cup \bigcup_{i\in I_2} G_1(A_i) \cup \bigcup_{i\in I_2} G_2(A_i) \cup \bigcup_{i\in I_3} G(A_i).$$

Let $G' = \bigcup_{i\in I_2} G_2(A_i)$. Thus $|E(G')| \leq |I_2| \binom{k}{2} n^{-\frac{1}{3}} \binom{n}{k}$ and

$$\frac{|E(G')|}{|V(G')|} \leq \frac{|I_2| \binom{k}{2} n^{-\frac{1}{3}} \binom{n}{k}}{\binom{n}{k-2}(n-k+2)} = \frac{n^{-\frac{1}{3}} |I_2|(n-k+1)}{2} \leq |I_2| n^{\frac{2}{3}}.$$

By Lemma 2.4 and Lemma 3.6 we have that

$$\sum_{i\in I_1} H(G(A_i)) + \sum_{i\in I_2} H(G_1(A_i)) + H(G') + \sum_{i\in I_3} H(G(A_i)) \geq H(G(F)) = \log(n-k+2).$$

By Corollary 2.12, Lemma 3.10 and Lemma 3.21 we have

$$\sum_{i\in I_1\cup I_2} \frac{\text{size}(A_i)}{n} e^{-\alpha(k)} 2L(\frac{\text{size}(A_i)}{n-k+2}) + L(\frac{|E(G')|}{|V(G')|}) + \sum_{i\in I_3} 2L(\frac{\text{size}(A_i)}{n-k+2}) \geq \log(n-k+2).$$

Therefore, at least one of the following cases holds.

Case 1

$$\begin{aligned} \sum_{i\in I_1\cup I_2} \frac{\text{size}(A_i)}{n} e^{-\alpha(k)} 2L(\frac{\text{size}(A_i)}{n-k+2}) &\geq \frac{1}{8}\log(n-k+2) \\ i.e. \quad \sum_{i\in I_1\cup I_2} \text{size}(A_i) &\geq \frac{ne^{\alpha(k)}\log(n-k+2)}{16L(\frac{e^{\alpha(k)}n}{n-k+2})}. \end{aligned}$$

Case 2

$$\begin{aligned} L(\frac{|E(G')|}{|V(G')|}) &\geq \frac{3}{4}\log(n-k+2) \\ i.e. \quad L(|I_2| n^{\frac{2}{3}}) &\geq \frac{3}{4}\log(n-k+2) \\ i.e. \quad |I_2| &\geq \frac{(n-k+2)^{\frac{3}{4}} - e}{en^{\frac{2}{3}}} \end{aligned}$$

Since $k \leq (\log\log n)^2$, it follows that

$$\text{size}(F) \geq \sum_{i\in I_2} \text{size}(A_i) \geq \frac{n}{2}|I_2| \geq e^{\delta(k)} n \log n, \quad \text{for large } n.$$

Case 3

$$\sum_{i \in I_3} 2L(\frac{\text{size}(A_i)}{n-k+2}) \geq \frac{1}{8}\log(n-k+2).$$

For $i \in I_3$, let $r_i = \frac{\text{size}(A_i)}{n-k+2}e^{-\alpha(k)}$. It can be easily verified that $L(rx) \leq rL(x)$ for $r \geq 1$ and $x \geq 0$. Now for each $i \in I_3$, $r_i > 1$. Hence

$$L(\frac{\text{size}(A_i)}{n-k+2}) = L(r_i e^{\alpha(k)}) \leq r_i L(e^{\alpha(k)}).$$

Thus, $\sum_{i \in I_3} r_i \geq \frac{\log(n-k+2)}{16L(e^{\alpha(k)})}$ and

$$\sum_{i \in I_3} \text{size}(A_i) \geq \frac{(n-k+2)e^{\alpha(k)}\log(n-k+2)}{16L(e^{\alpha(k)})}.$$

Since k and n are large numbers and $k < (\log\log n)^2$, we have $\text{size}(F) \geq \exp(\delta(k))n\log n$ in each case. □

3.3.3 Proof of the combinatorial lemma

Consider a $(k-1)$-immune $\Pi\Sigma$ formula $A = \bigwedge_{j=1}^{t} \bigvee_{q \in S_j} q$. Our final goal, roughly speaking, is to show that if $\text{size}(A)$ is small then there are not many $(k-2)$-sets C that can be extended to a set $C \cup \{x, y\}$ accepted by A. As in the proof of Theorem 3.4, it will be more convenient to estimate the number of $\sigma \in [n]_{k-2}$ that can be extended to a sequence σxy accepted by A.

In the following we will set

$$\begin{aligned} k' &= \left\lfloor \sqrt{\frac{k}{e^4 \ln k}} \right\rfloor ; \\ \alpha(k) &= \frac{1}{6}k'. \end{aligned}$$

As before, we pick a $\sigma_L \in [n]_{k-k'-2}$ at random. Then we extend it to a sequence $\sigma = \sigma_L\sigma_R \in [n]_{k-2}$. We classify the S_j's of A into two kinds, *small* and *big*, based on the number of non-negated variables they contain. If the size of A is not very big then there cannot be too many big S_j's. If big and small are suitably defined, then we shall show, using arguments similar to the proof of Lemma 3.1, that for randomly chosen σ_L, $\Pr[A|_{\sigma_L}$ has a big $S_j]$ is very small. Now consider $A|_{\sigma_L}$. Clearly, $A|_{\sigma_L}$ is a $(k'+1)$-immune formula. We will show that for a $(k'+1)$-immune $\Pi\Sigma$ formula B with no big S_j there are very few $\sigma \in [n]_{k'}$ such that B accepts σxy for some x and y.

In the following B will denote a $\Pi\Sigma$ formula on n variables. We shall assume that B has the following properties. Let $B = \bigwedge_{j=1}^{t} \bigvee_{q \in S_j} q$.

(P1) B is $(k'+1)$-immune.

(P2) $\text{size}(B) \leq n\exp(\alpha(k))$.

(P3) $|S_j| \geq 1$ for $j = 1, \ldots, t$.

(P4) Every S_j has at most $\frac{nk'}{k}$ non-negated variables.

A sequence $\overline{b} \in [n]_{k'}$ is called *extendible* if B accepts $\overline{b}xy$ for some x, y. Our first goal is to show that there are very few extendible sequences. To clarify the main idea of the proof we first make some simplifying assumptions. Let us assume that there are no negations in B. Further, suppose a uniformity condition holds: all variables appearing in any sequence $\overline{b}xy$ accepted by B appear in the same number of S_j's.

Since the number of S_j's in B is t, the average number of occurrences of a variable is at most $\frac{t}{n}\frac{nk'}{k} = \frac{tk'}{k}$. On the other hand, if $\overline{b}xy \in [n]_{k'+2}$ is accepted by B, every S_j must include one of the variables in $\overline{b}xy$. Thus a variable in $\overline{b}xy$ appears in at least $\frac{t}{k'+2}$ of the S_j's. For our choice of k', $\frac{tk'}{k}$ is smaller than $\frac{t}{k'+2}$ by a factor of about $\ln k$. This suggests that the variables that appear in $\overline{b}xy$ are not typical. Since all the variables in $\overline{b}xy$ appear in the same number of S_j's, no variable in $\overline{b}$ is typical. The main idea of the proof is to exploit this fact and show that most sequences $\overline{b}$ are not extendible.

Even without the uniformity condition, reasoning as before, we can conclude that there is at least one extraordinary element in $\overline{b}xy$. But this is not enough to conclude that $\overline{b}$ is also atypical, because it may be that the extraordinary element is one of x and y. In our argument, this difficulty is surmounted by identifying such exceptional elements and eliminating them from the counting argument. The detailed argument, formalizing the rough sketch given above, is described below. Technical difficulties arise mainly because the uniformity condition need not hold and the formula may contain negations.

An S_j in the description of a $\Pi\Sigma$ formula will be called *positive* if it is not empty and it has no negated variables. If it contains a negated variable it will be called *negative*. Let $\overline{r}$ be a sequence of variables. We say that S_j *intersects* $\overline{r}$ if some non-negated variable of S_j appears in $\overline{r}$. Let $\overline{r}$ satisfy the following condition. (Note that B is $\Pi\Sigma$ formula and we assume the conventions described in section 3.1 when we refer $B|_{\overline{r}}$.)

$$B|_{\overline{r}} \text{ has no empty } S_j. \tag{3}$$

Definition 3.12 (Exception sequence) *For such a sequence $\overline{r}$, the* exception sequence *for B after $\overline{r}$, denoted by $Q_{\overline{r}}$, is given by the following procedure.*

1. Initially set $Q_{\overline{r}} \leftarrow$empty.
2. Repeat the following steps until some condition for stopping is met.
   (a) If B accepts $\overline{r}Q_{\overline{r}}$, then stop.
   Let $S_{av} = \frac{n}{l}$ be the average size of a positive S_j of $B|_{\overline{r}Q_{\overline{r}}}$. Let t^* be the number of positive S_j's in $B|_{\overline{r}Q_{\overline{r}}}$. Call a pair of variables $\{x, y\}$ *safe* for $B|_{\overline{r}Q_{\overline{r}}}$ if $\{x, y\} \subseteq [n] - \overline{r}Q_{\overline{r}}$ and $\{x, y\}$ does not include all the variables that appear negated in some negative S_j of $B|_{\overline{r}Q_{\overline{r}}}$.
   (b) If some safe pair of variables $\{x, y\}$ intersects more than $t^*(1 - \frac{k}{2lk'})$ of the positive S_j's of $B|_{\overline{r}Q_{\overline{r}}}$, then let $\{x_0, y_0\}$ be the lexicographically first such pair; set $Q_{\overline{r}} \leftarrow Q_{\overline{r}}x_0y_0$.
   (c) Otherwise, that is, if no such pair exists, stop.

Note that this procedure always terminates because eventually there will be no safe pair left.

Lemma 3.13 Let $\overline{r}$ be a sequence of variables satisfying condition (3). If $|\overline{r}| \leq \frac{k'}{2}$ then $|Q_{\overline{r}}| \leq \frac{k'}{2} + 1$ and $B|_{\overline{r}Q_{\overline{r}}}$ has a positive S_j.

Proof: By condition (3), $B|_{\overline{r}}$ has no empty S_j. Our definition of safe pair ensures that if some S_j of $B|_{\overline{r}}$ has negated variables, then not all of them are included in $Q|_{\overline{r}}$. It follows that $B|_{\overline{r}Q_{\overline{r}}}$

has no empty S_j. Thus, at any stage in the execution of the above procedure, if $B|_{\overline{r}Q_{\overline{r}}}$ has no positive S_j then B accepts $\overline{r}Q_{\overline{r}}$. We shall make use of this last observation.

Suppose, for contradiction, that $|\overline{r}| \leq \frac{k'}{2}$ but $|Q_{\overline{r}}| > \frac{k'}{2} + 1$. Suppose p iterations of step (b) of the above procedure were performed to produce $Q_{\overline{r}}$. In each iteration of step (b), exactly two variables are added to $Q_{\overline{r}}$. It follows that $p \geq \left\lfloor \frac{k'}{4} \right\rfloor + 1$

Let the value of t^* at the beginning of the ith iteration be t_i^*, for $i = 1, 2, \ldots, p$; similarly, let the value of S_{av} at the beginning of the ith iteration be $\frac{n}{l_i}$. Using the observation made above, we conclude that if at the beginning of any iteration $B|_{\overline{r}Q_{\overline{r}}}$ has no positive S_j, then B accepts $\overline{r}Q_{\overline{r}}$, and we stop in step (a). Since p iterations of step (b) were performed, $t_i^* > 1$ and l_i is well defined ($< \infty$), for $i = 1, 2, \ldots, p$.

The condition in step (b) implies that

$$t_{i+1}^* \leq t_i^* - t_i^*(1 - \frac{k}{2l_ik'}) = \frac{k}{2k'l_i}t_i^*,$$

for $i = 1, 2, \ldots, p-1$. Hence, for $i = 1, 2, \ldots, p-1$, we have

$$t_i^* \geq \frac{2k'l_i}{k}t_{i+1}^*.$$

Applying this inequality repeatedly and noting that $t_p^* \geq 1$, we have

$$t_1^* \geq \frac{2k'l_1}{k}t_2^* \geq \frac{2k'l_1}{k}\frac{2k'l_2}{k}t_3^* \geq \ldots \geq \left(\frac{2k'}{k}\right)^{p-1} l_1 \ldots l_{p-1}t_p^* \geq \left(\frac{2k'}{k}\right)^{p-1} l_1 l_2 \ldots l_{p-1}. \qquad (4)$$

By property (P4), $\frac{n}{l_i} \leq \frac{nk'}{k}$, that is, $l_i \geq \frac{k}{k'}$. It follows from the definition of t_1^* and $\frac{n}{l_1}$ that $\text{size}(B) \geq t_1^*(\frac{n}{l_1})$. Using (4) we get

$$\text{size}(B) \geq t_1^*(\frac{n}{l_1}) \geq 2^{p-1}\frac{k'l_1}{k}(\frac{n}{l_1}).$$

Since $p - 1 = \left\lfloor \frac{k'}{4} \right\rfloor$ and $4 \log e < 6$, we have for large enough k that

$$\text{size}(B) \geq n\frac{k'}{k}\exp(\frac{p-1}{\log e}) \geq n\frac{k'}{k}\exp(\frac{k'-4}{4\log e}) \geq n\exp(\frac{k'-4}{4\log e} - \ln k) > n\exp(\frac{k'}{6}).$$

But this contradicts (P2). Thus we have established that $|Q_{\overline{r}}| \leq \frac{k'}{2} + 1$.

Next, suppose that $|\overline{r}| \leq \frac{k'}{2}$ and $B|_{\overline{r}Q_{\overline{r}}}$ has no positive S_j. Then, by the observation above, B accepts $\overline{r}Q_{\overline{r}}$. But $|\overline{r}Q_{\overline{r}}| \leq k' + 1$, contradicting (P1). This contradiction establishes the second part of the lemma. □

For convenience we shall adopt the following notation.

$$\begin{aligned} w(\overline{r}, i) &= \textit{the number of times the variable } x_i \textit{ occurs in a positive } S_j \textit{ of } B|_{\overline{r}Q_{\overline{r}}}. \\ \frac{n}{l(\overline{r})} &= \textit{the average size of a positive } S_j \textit{ in } B|_{\overline{r}Q_{\overline{r}}}. \\ t^*(\overline{r}) &= \textit{number of positive } S_j\textit{'s in } B|_{\overline{r}Q_{\overline{r}}}. \end{aligned}$$

By Lemma 3.13 and (P3) all these are well defined, and $l(\overline{r}) < \infty$, if $|\overline{r}| \leq \frac{k'}{2}$. Note that the expected number of occurrences of a variable among the positive S_j's of $B|_{\overline{r}Q_{\overline{r}}}$ is precisely $E_v[w(\overline{r}, v)] = t^*(\overline{r})/l(\overline{r})$.

Our approach to showing that the number of extendible sequences is small is the following. We shall show that each extendible sequence can be reordered to form a special kind of sequence

called a π-sequence. A π-sequence $\overline{a}$ will consist of parts $\overline{a}_1, \overline{a}_2, \ldots, \overline{a}_{h+1}$. Let $\overline{r}_1 = \textit{empty}$ and $\overline{r}_{j+1} = \overline{r}_j \overline{a}_j$. It will be ensured while reordering an extendible sequence $\overline{b}$ into $\overline{a}$, that the elements in $\overline{a}_i$ are not typical in the following sense. For the indices j appearing in $\overline{a}_i$, the weight $w(\overline{r}_i, j)$ will be much higher than the expected value $E_v[w(\overline{r}_i, v)]$. We will then be able to conclude that there are only few π-sequences. It will follow, then, that even the extendible sequences that can be reordered to form π-sequences are few.

Below we first define a π-sequence precisely. Then we show that the number of π-sequences is small (Lemma 3.16) and conclude that the number of sequences that can be reordered to form π-sequences is also small (Lemma 3.17). These lemmas are direct consequences of our definitions and are based on straight forward counting arguments. To complete the proof, however, we still need to show that every extendible sequence can be reordered to form a π-sequence. To accomplish this we need a technical lemma (Lemma 3.18). Finally, using Lemma 3.18, we show in Lemma 3.19 that every extendible sequence can be reordered to form a π-sequence and hence the number of extendible sequences is small.

The outline of the proof given above would have been accurate had there been no negations. We, in reality, do not succeed in reordering all extendible sequences into π-sequences. Instead, we show that the ones we fail to reorder cannot be extended in many ways. This is made precise in the statement of Lemma 3.19.

Definition 3.14 (Good Partition) *A sequence of integers $\pi = (m_1, m_2, \ldots, m_h, m_{h+1})$ is a* good partition *if it satisfies the following conditions.*

$$\sum_{j=1}^{h+1} m_j = k'; \tag{5}$$

$$\sum_{j=1}^{h} m_j \geq \frac{k'}{2}; \tag{6}$$

$$m_j \geq 1, \quad \text{for } j = 1, \ldots, h; \tag{7}$$

$$m_{h+1} \geq 0. \tag{8}$$

Definition 3.15 (π-sequence) *Let $\pi = (m_1, m_2, \ldots, m_h, m_{h+1})$ be a good partition. Let $\overline{a} \in [n]_{k'}$. Let $\overline{a} = \overline{a}_1 \overline{a}_2 \ldots \overline{a}_{h+1}$, where $|\overline{a}_j| = m_j$, for $j = 1, \ldots, h+1$. Set $\overline{r}_1 = \textit{empty}$ and $\overline{r}_{j+1} = \overline{r}_j \overline{a}_j$, $j = 1, \ldots, h-1$. We say $\overline{a}$ is a π-sequence if*

$$w(\overline{r}_j, v) \geq \frac{t^*(\overline{r}_j)}{l(\overline{r}_j)} e^3 \binom{k'}{m_j}^{\frac{1}{m_j}} \text{ for } j = 1, \ldots, h \text{ and for all } v \in \overline{a}_j. \tag{9}$$

Lemma 3.16 For a good partition $\pi = (m_1, m_2, \ldots, m_h, m_{h+1})$, the number of π-sequences is at most

$$n^{k'} e^{-\frac{3k'}{2}} \prod_{j=1}^{h} \binom{k'}{m_j}^{-1}.$$

Proof: We shall show that only a few sequences satisfy (9), even if repetitions are permitted. Clearly, this would give us an upper bound on the number of π-sequences.

It is convenient to state the proof in terms of probability. We wish to estimate the probability that a sequence generated by randomly adding variables starting from an empty sequence is a π-sequence. Suppose we have generated $\overline{a}_1 \overline{a}_2 \ldots \overline{a}_{j-1}$, $1 \leq j \leq h$. We wish to estimate the probability that each of the next m_j random choices meets the condition (9) above. From the

definitions of $t^*(\overline{r}_j)$ and $\frac{n}{l(\overline{r}_j)}$ we have that the expected number of occurrences of a variable among the positive S_j's of $B_{\overline{r}_j Q_{\overline{r}_j}}$ is given by

$$E(w(\overline{r}_j, v)) = \frac{1}{n} t^*(\overline{r}_j) \frac{n}{l(\overline{r}_j)} = \frac{t^*(\overline{r}_j)}{l(\overline{r}_j)}.$$

Thus, by Markov's inequality,

$$\Pr[w(\overline{r}_j, v) \geq \frac{t^*(\overline{r}_j)}{l(\overline{r}_j)} e^3 \binom{k'}{m_j}^{\frac{1}{m_j}}] \leq e^{-3} \binom{k'}{m_j}^{-\frac{1}{m_j}}.$$

It follows that

$$\Pr[\overline{a}_j \textit{ meets condition (9)}] \leq e^{-3m_j} \binom{k'}{m_j}^{-1}.$$

Thus

$$\Pr[\overline{a}_1 \overline{a}_2 \ldots \overline{a}_{h+1} \textit{ is a } \pi\text{-sequence}] \leq e^{-3\sum_{j=1}^{h} m_j} \prod_{j=1}^{h} \binom{k'}{m_j}^{-1}.$$

By (6), $\sum_{j=1}^{h} m_j \geq \frac{k'}{2}$. The lemma follows from this. □

Let $\pi = (m_1, m_2, \ldots, m_{h+1})$ be a good partition and let $\overline{a} = \overline{a}_1 \overline{a}_2 \ldots \overline{a}_{h+1}$ be a π-sequence. A sequence $\overline{b}$ is said to be derived from $(\pi, \overline{a})$ if $\overline{b}$ is a reordering of the elements of $\overline{a}$ such that each $\overline{a}_i$ maintains its relative order. It is easy to see that the number of derived π-sequences that can be obtained from a fixed π-sequence $\overline{a}$ is at most

$$\prod_{j=1}^{h} \binom{k'}{m_j}.$$

Lemma 3.17 The number of sequences $\overline{b}$ that are derived π-sequences for some good partition π is at most $n^{k'} e^{-\frac{k'}{2}}$.

Proof: The number of choices for π is at most $2^{k'}$. To see this, consider the set of $\{0,1\}$-sequences of length k' with at least one 1. With each such sequence σ, we shall associate a partition of k' as follows. Suppose σ contains h 1's.

$$\begin{aligned}
m_1 &= (\textit{the number of } 0\text{'s } \textit{in } \sigma \textit{ before the 1st } 1) + 1; \\
m_2 &= (\textit{the number of } 0\text{'s } \textit{in } \sigma \textit{ between the 1st and 2nd } 1) + 1; \\
m_3 &= (\textit{the number of } 0\text{'s } \textit{in } \sigma \textit{ between the 2nd and 3rd } 1) + 1; \\
&\vdots \\
m_h &= (\textit{the number of } 0\text{'s } \textit{in } \sigma \textit{ between the } (h-1)\textit{st and } h\textit{th } 1) + 1; \\
m_{h+1} &= (\textit{the number of } 0\text{'s } \textit{in } \sigma \textit{ after the last } 1).
\end{aligned}$$

Clearly, $\sum_{j=1}^{h+1} m_j = k'$ and every good partition $\pi = (m_1, m_2, \ldots, m_h, m_{h+1})$ is associated with some $\{0,1\}$-sequence. Thus the number of good partitions is at most $2^{k'}$.

For each choice of π, by Lemma 3.16 the number of π-sequences is at most

$$n^{k'} e^{-\frac{3}{2}k'} \prod_{j=1}^{h} \binom{k'}{m_j}^{-1}.$$

As observed above, the number of derived π-sequences that can be obtained from any fixed π-sequence is at most $\prod_{j=1}^{h}\binom{k'}{m_j}$. Thus, the total number of derived π-sequences $\overline{b}$ is at most

$$2^{k'}n^{k'}e^{-\frac{3}{2}k'}\prod_{j=1}^{h}\binom{k'}{m_j}^{-1}\prod_{j=1}^{h}\binom{k'}{m_j} \leq 2^{k'}n^{k'}e^{-\frac{3}{2}k'} \leq n^{k'}e^{-\frac{k'}{2}}.$$

□

Lemma 3.18 (Technical Lemma) Let T be a nonempty subset of $[n]$ of size at most k'. Let δ be a positive real number. Let w be a weight function from T to the set of positive real numbers. Set $w(T) = \sum_{i\in T} w(i)$. Suppose $w(T) \geq \frac{\delta k}{2k'}$. Then $\exists\, R \subseteq T$, $R \neq \emptyset$, such that for each $i \in R$,

$$w(i) \geq \delta e^3 \binom{k'}{|R|}^{\frac{1}{|R|}}.$$

Proof: For $j = 1, \ldots, |T|$, let R_j be the set of j elements of T with the largest weights. We claim that one of the R_j meets the requirements of the lemma. Suppose for contradiction that none of the R_j meets the requirements of the lemma. Let w_j be the weight of the element of T with the jth largest weight. Then we have that

$$w_j < \delta e^3 \binom{k'}{j}^{\frac{1}{j}} \leq \delta e^3 \frac{ek'}{j}.$$

It follows that

$$w(T) = \sum_{j=1}^{|T|} w_j < \delta e^4 k' \sum_{j=1}^{|T|} \frac{1}{j} \leq \delta e^4 k'(1+\ln k') \leq \frac{\delta k}{2k'}[\frac{2e^4(k')^2(1+\ln k')}{k}] \leq \frac{\delta k}{2k'}.$$

The last inequality holds because $k' = \left\lfloor \sqrt{\frac{k}{e^4 \ln k}} \right\rfloor$. But this contradicts the condition that $w(T) \geq \frac{\delta k}{2k'}$ in the statement of the lemma, and the proof of the lemma is complete. □

The main part of our argument will appear in the proof of the following lemma. For this we will need the Lemma 3.13. Since that lemma was proved under the assumption that the formula B had properties (P1)–(P4), we need to ensure that these properties hold when we invoke that lemma. For easy reference, we state them again: **(P1)** B is $(k'+1)$-immune; **(P2)** size$(B) \leq n\exp(\alpha(k))$; **(P3)** $|S_j| \geq 1$ for $j = 1, \ldots, t$; **(P4)** Every S_j has at most $\frac{nk'}{k}$ non-negated variables. Also, recall that, for such a formula B and a sequence of variables $\overline{r}$, the *exception sequence* $Q_{\overline{r}}$ was defined only if $\overline{r}$ satisfied the condition (3). We reproduce this condition below for later reference.

$$B|_{\overline{r}} \text{ has no empty } S_j. \tag{3}$$

Lemma 3.19 Let F be a $\Pi\Sigma$ formula having the properties (P1)–(P4). Let

$$\Gamma = \{\gamma \in [n]_{k'+2} : F \textit{ accepts } \gamma\}.$$

Then there exists a set $\Delta \subseteq \Gamma$, $|\Delta| \leq n^{-\frac{2}{5}}(n)_{k'+2}$ such that

$$|\{\overline{b} \in [n]_{k'} : \exists x, y \text{ such that } \overline{b}xy \in \Gamma - \Delta\}| \leq n^{k'}e^{-\frac{k'}{2}}.$$

Proof: Let $\overline{b}xy \in \Gamma$. We wish to show that $\overline{b}$ is a derived π-sequence for some good partition π. To do this, we describe a procedure which reorders $\overline{b}$ to produce a π-sequence $\overline{a}$ for a good partition π. The procedure might fail for some elements of Γ; these we collect in the set Δ. In the end we show that $|\Delta| \leq n^{-\frac{2}{5}}(n)_{k'+2}$. The lemma will then follow from Lemma 3.17 because there are only $n^{k'}e^{-\frac{k'}{2}}$ derived π-sequences.

Let V^* be the variables in F that appear more than $\sqrt{n}$ times. Using property (P2), we have

$$|V^*| \leq \exp(\alpha(k))\sqrt{n}.$$

Let $\Delta_0 = \{\gamma \in \Gamma : \gamma \cap V^* \neq \emptyset\}$. That is, Δ_0 contains all the elements of Γ that include at least one variable in V^*. We have $k'+2$ possible positions for this variable, so

$$|\Delta_0| \leq (k'+2)\exp(\alpha(k))\sqrt{n}(n)_{k'+1}. \tag{10}$$

Let B be the formula obtained from F by fixing all the variables appearing in V^* at the value 0. Then, every sequence in $\Gamma - \Delta_0$ is accepted by B. If B is identically 0 then $\Gamma \subseteq \Delta_0$. Then the lemma follows easily because $|\Delta_0| \leq n^{-\frac{2}{5}}(n)_{k'+2}$. (Recall $k < (\log\log n)^2$, so that $(k'+2)\exp(\alpha(k)) \leq \log n$.) Hence we may assume that B is not identically 0. It follows that B has no empty S_j (that is, B has property (P3)). Since F is $(k'+1)$-immune, B is $(k'+1)$-immune. We conclude that B has properties (P1)–(P4) and no variable in B appears more than $\sqrt{n}$ times.

Let $\overline{b}xy \in \Gamma - \Delta_0$. We shall either show that $\overline{b}$ is a derived π-sequence or let $\overline{b}xy \in \Delta_1$. Finally, Δ will be $\Delta_0 \cup \Delta_1$. We construct a good partition $\pi = (m_1, m_2, \ldots, m_h, m_{h+1})$ and rearrange $\overline{b}$ into $\overline{a} = \overline{a}_1\overline{a}_2\ldots\overline{a}_{h+1}$, so that $\overline{a}$ is a π-sequence. Let the sequences r_i be defined by $\overline{r}_1 = \textit{empty}$, $\overline{r}_{i+1} = \overline{r}_i\overline{a}_i$. Recall (from Definition 3.12) that a pair of variables $\{x, y\}$ is *safe* for $B|_{\overline{r}Q_{\overline{r}}}$ if $\{x, y\} \cap \overline{r}Q_{\overline{r}} = \emptyset$ and $\{x, y\}$ does not include all the negated variables in any negative S_j of $B|_{\overline{r}Q_{\overline{r}}}$.

Initially set $i = 1$, $\overline{r} = \textit{empty}$ and $\hat{b} = b$. {Throughout we shall maintain that $\overline{r} = \overline{a}_1\overline{a}_2\ldots\overline{a}_{i-1}$. Also, $\hat{b}$ will be the sequence obtained by omitting $\overline{r}$ from $\overline{b}$.} Repeat the following four steps until some condition for stopping is met.

1. If $|\overline{r}| \geq \frac{k'}{2}$, then set $h = i-1$, $\overline{a}_{h+1} = \hat{b}$, $m_{h+1} = |\overline{a}_{h+1}|$ and **STOP**.

2. If there is no pair $\{\hat{x}, \hat{y}\}$ that is safe for $B|_{\overline{r}Q_{\overline{r}}}$ such that $\overline{b}\hat{x}\hat{y}$ is accepted by B, then let $\overline{b}xy \in \Delta_1$, and **STOP**. {In particular, we have that $\{x, y\}$ is not safe for $B|_{\overline{r}Q_{\overline{r}}}$. We will need this observation when we estimate $|\Delta_1|$.}

3. If such a pair exists, let $\{\hat{x}, \hat{y}\}$ be the lexicographically first such pair. Since B accepts $\overline{b}\hat{x}\hat{y}$, $B|_{\overline{r}}$ accepts $\hat{b}\hat{x}\hat{y}$. By our definition of exception sequence, $Q_{\overline{r}}$ does not contain all the negated variables of any negative S_j of $B|_{\overline{r}}$. Hence, $B|_{\overline{r}Q_{\overline{r}}}$ also accepts $\hat{b}\hat{x}\hat{y}$. Since $\{\hat{x}, \hat{y}\}$ is safe for $B|_{\overline{r}Q_{\overline{r}}}$, it follows from our definition of exception sequence that $\{\hat{x}, \hat{y}\}$ intersects at most $t^*(\overline{r})(1 - \frac{k}{2k'l(\overline{r})})$ of the positive S_j's of $B|_{\overline{r}Q_{\overline{r}}}$. Hence, $\hat{b}$ must intersect the remaining at least $t^*(\overline{r}) - t^*(\overline{r})(1 - \frac{k}{2k'l(\overline{r})}) = t^*(\overline{r})\frac{k}{2k'l(\overline{r})}$ positive S_j's. With this we invoke Lemma 3.18, setting $w(v) = w(\overline{r}, v)$, $\delta = \frac{t^*(\overline{r})}{l(\overline{r})}$, and with T as the set of variables appearing in $\hat{b}$. (By Lemma 3.13, $B|_{\overline{r}Q_{\overline{r}}}$ has at least one positive S_j, so $\delta > 0$. Since $\hat{b}$ intersects at least $t^*(\overline{r})\frac{k}{2k'l(\overline{r})}$ positive S_j's, $w(T) \geq t^*(\overline{r})\frac{k}{2k'l(\overline{r})} \geq \frac{\delta k}{2k'}$.) We conclude that there is a non-empty subsequence $\overline{a}_i$ of $\hat{b}$ such that

$$w(\overline{r}, v) \geq \frac{t^*(\overline{r})}{l(\overline{r})}e^3\binom{k'}{|\overline{a}_i|}^{\frac{1}{|\overline{a}_i|}}, \quad \forall v \in \overline{a}_i. \tag{11}$$

4. Set $m_i = |\overline{a}_i|$, $\overline{r} \leftarrow \overline{r}\,\overline{a}_i$, delete the elements of $\overline{a}_i$ from $\hat{b}$, and set $i \leftarrow i+1$.

Since none of the $\overline{a}_i$ is empty, the procedure does terminate, and $h \leq \frac{k'}{2}$. Suppose $\overline{b}xy$ was not put in Δ_1. From (11) it follows that $\overline{a} = \overline{a}_1\overline{a}_2 \ldots \overline{a}_{h+1}$ is a π-sequence for the good partition $\pi = (m_1, m_2, \ldots, m_{h+1})$. Our construction thus ensures that $\overline{b}$ is a derived π-sequence.

It remains only to bound the size of Δ_1. Suppose $\overline{b}xy \in \Delta_1$. Then at some stage i in the reordering process it was detected that $\{x, y\}$ is not safe for $B|_{\overline{r}_i Q_{\overline{r}_i}}$. Note that the value of i depends only on $\overline{b}$ and not on $\{x, y\}$. We have two possibilities (based on the definition of a safe pair): (1) $\{x, y\}$ intersects $\overline{r}_i Q_{\overline{r}_i}$; (2) $\{x, y\}$ includes all the negated variables of some negative S_j of $B|_{\overline{r}_i Q_{\overline{r}_i}}$.

1. $\{x, y\}$ intersects $\overline{r}_i Q_{\overline{r}_i}$. Since $\overline{r}_i$ does not intersect $\{x, y\}$, $\{x, y\}$ must intersect $Q_{\overline{r}_i}$. Let R_0 be the set of those sequences, $\overline{b}xy \in \Delta_1$, where $\{x, y\}$ intersects $Q_{\overline{r}_i}$. The number of possibilities for $\{x, y\}$ is at most $2|Q_{\overline{r}_i}|n$. Now by Lemma 3.13, $|Q_{\overline{r}_i}| \leq (\frac{k'}{2} + 1)$. Since there are only $(n)_{k'}$ possible values for $\overline{b}$, we get that

$$|R_0| \leq 2|Q_{\overline{r}_i}|n(n)_{k'} \leq 2(\frac{k'}{2} + 1)n(n)_{k'}. \tag{12}$$

2. Otherwise, $\{x, y\}$ includes all the variables that appear negated in some negative S_j of $B|_{\overline{r}_i Q_{\overline{r}_i}}$. In this case, $Q_{\overline{r}_i}xy$ included all the variables that appear negated in some negative S_j of $B|_{\overline{r}_i}$. We have two cases. In the first case, $\{x, y\}$ by itself includes all the variables negated in some negative S_j of $B|_{\overline{r}_i}$. Now $B|_{\overline{r}_i}$ accepts $\overline{b}xy$; hence, in this case, $\overline{b}xy$ must also include a non-negated variable in the S_j (so that $\bigvee_{q \in S_j} q$ evaluates to 1). For the second case, we have that $\{x, y\}$ and $Q_{\overline{r}_i}$ both contain a variable negated in some S_j of $B|_{\overline{r}_i}$. Let

$$R_1 = \{\overline{b}xy \in \Delta_1 : \overline{b}xy \textit{ includes all negations in some negative } S_j \textit{ of } B\}.$$

That is, R_1 includes all the sequences considered in the first case above. Let

$$P = \{(v, w) : \exists S_j \; v \textit{ is negated and } w \textit{ is non-negated in } S_j\}.$$

Arguing as in the proof of Lemma 3.7, we may assume that there are at most $k'+1$ negations in any S_j. Thus $|P| \leq \text{size}(B)(k'+1)$. Each $\overline{b}xy \in R_1$ contains at least one pair in P. As there are only $(k'+2)(k'+1)$ possible positions for the pair, we get, using property (P2), that

$$|R_1| \leq |P|(k'+2)(k'+1)(n)_{k'} \leq n^{-\frac{2}{3}}(n)_{k'+2}. \tag{13}$$

Let $R_2 = \Delta_1 - R_1$. That is, R_2 includes all the sequences considered in the second case above but not included in R_1. Consider any $\overline{b} \in [n]_{k'}$. We will estimate the number of extensions $\overline{b}xy$ that belong to R_2. Since we failed to rearrange $\overline{b}$, at stage i some variable in $Q_{\overline{r}_i}$ and some variable in $\{x, y\}$ both appear negated in a negative S_j of $B|_{\overline{r}_i}$. Let

$$X = \{v : \exists S_j \, \exists w \in Q_{\overline{r}_i} \; v \textit{ and } w \textit{ both appear negated in } S_j\}.$$

By Lemma 3.13, $|Q_{\overline{r}_i}| \leq (\frac{k'}{2} + 1)$. Since no variable occurs more than $\sqrt{n}$, times and since there are at most $(k'-1)$ negations in any S_j, we get that

$$|X| \leq |Q_{\overline{r}_i}|(k'-1)\sqrt{n} \leq (k'-1)(\frac{k'}{2} + 1)\sqrt{n}. \tag{14}$$

If $\overline{b}xy \in R_2$, then $\{x, y\} \cap X \neq \emptyset$. The number of such extension for any $\overline{b} \in [n]_{k'}$ is at most $2|X|n$. Since there are only $(n)_{k'}$ values for $\overline{b}$, we have that $|R_2| \leq 2|X|n(n)_{k'}$.

Now $k < (\log\log n)^2$, so that $k' \leq \log\log n$ and $\exp(\alpha(k)) \leq \log n$. Thus, using (10), (12), (13), and (14), we have

$$\begin{aligned}
|\Delta| &\leq |\Delta_0| + |\Delta_1| \\
&\leq |\Delta_0| + |R_0| + |R_1| + |R_2| \\
&\leq (k'+2)\exp(\alpha(k))\sqrt{n}(n)_{k'+1} + 2(\frac{k'}{2}+1)\sqrt{n}n(n)_{k'} + n^{-\frac{2}{3}}(n)_{k'+2} + 2|X|n(n)_{k'} \\
&\leq n^{-\frac{2}{5}}(n)_{k'+2}.
\end{aligned}$$

□

Lemma 3.20 Let $A = \bigwedge_{j=1}^{t} \bigvee_{q \in S_j} q$ be a $(k-1)$-immune $\Pi\Sigma$ formula. Let $\Gamma = \{\gamma \in [n]_k : A \textit{ accepts } \gamma\}$.

(a) Suppose $\text{size}(A) \leq \frac{n}{2}$. Let

$$\Psi = \{\overline{a} \in [n]_{k-2} : \exists\, x, y \text{ such that } \overline{a}xy \in \Gamma\}.$$

Then

$$|\Psi| \leq \left(\frac{\text{size}(A)}{n}\right)\exp(-\frac{k'}{3})n^{k-2}.$$

(b) Suppose $\text{size}(A) > \frac{n}{2}$. Then there exists a set $\Delta \subseteq \Gamma$, $|\Delta| \leq n^{-\frac{1}{3}}(n)_k$ such that if

$$\Psi = \{\overline{a} : \exists\, x, y \text{ such that } \overline{a}xy \in \Gamma - \Delta\},$$

then

$$|\Psi| \leq \left(\frac{\text{size}(A)}{n}\right)\exp(-\frac{k'}{3})n^{k-2}.$$

Proof: Suppose $\text{size}(A) \leq \frac{n}{2}$. Then the number of variables appearing explicitly in A is at most $\frac{n}{2}$. Let $\overline{a}xy \in \Gamma$. Then, since A is $(k-1)$-immune, $\overline{a}$ must contain only those variables that appear explicitly in A. Thus the number of choices for $\overline{a}$ is at most $(\text{size}(A))^{k-2} \leq \left(\frac{\text{size}(A)}{n}\right)2^{-(k-3)}n^{k-2}$, and the lemma follows easily.

Now suppose that $\text{size}(A) > \frac{n}{2}$. We may assume that $\text{size}(A) \leq n\exp(\frac{k'}{3})$ for otherwise the lemma provides a bound greater than n^{k-2} which is trivial.

We say that an S_j is *big* if it has more than $\frac{nk'}{k}$ non-negated variables. Thus A has at most $(\frac{\text{size}(A)k}{nk'})$ big S_j. Let $\overline{a} \in [n]_{k-2}$. Let $\overline{a} = \overline{a}_L\overline{a}_R$, where $|\overline{a}_L| = k-2-k'$ and $|\overline{a}_R| = k'$. Let

$$\begin{aligned}
\mathcal{B} &= \{\overline{a} \in [n]_{k-2} : A|_{\overline{a}_L} \textit{ has a big } S_j\}; \\
\mathcal{G} &= [n]_{k-2} - \mathcal{B}.
\end{aligned}$$

Here, read $\mathcal{G}$ as *good*, for there are no big ORs after $\overline{a}_L$, and read $\mathcal{B}$ as *bad*, for there are big ORs.

Now for $\overline{a}$ chosen randomly and a fixed big S_j

$$\Pr[\overline{a}_L \textit{ does not intersect } S_j] \leq (1-\frac{k'}{k})^{k-2-k'} \leq \exp(-\frac{k'}{k}(k-2-k')) \leq \exp(-\frac{k'}{2}).$$

Thus,

$$\Pr[\overline{a} \in \mathcal{B}] \leq (\frac{\text{size}(A)k}{nk'})\Pr[\overline{a}_L \textit{ does not intersect a fixed big } S_j] \leq (\frac{\text{size}(A)}{n})\exp(-\frac{k'}{2}).$$

It follows that

$$|\mathcal{B}| \leq (\frac{\text{size}(A)}{n}) \exp(-\frac{k'}{2}) n^{k-2}. \tag{15}$$

Let

$$\Gamma_0 = \{\overline{a}xy \in \Gamma : \overline{a} \in \mathcal{B}\}.$$

Let $\mathcal{G}_L = \{\overline{a}_L : \overline{a} \in \mathcal{G}\}$. Let $\overline{a} \in \mathcal{G}$. We claim that $A|_{\overline{a}_L}$ has property (P1), that is, $A|_{\overline{a}_L}$ is $(k'+1)$-immune. For suppose $A|_{\overline{a}_L}$ accepts $\overline{b}$ and $|\overline{b}| \leq (k'+1)$. We may assume that all variables appearing in $\overline{b}$ also appear in some S_j of $A|_{\overline{a}_L}$. But then A, a $(k-1)$-immune formula, accepts $\overline{a}_L\overline{b}$ which has length at most $k-1$. This contradiction establishes the claim. For $\overline{a}_L \in \mathcal{G}_L$, let

$$\begin{aligned}
\Gamma_{\overline{a}_L} &= \{\overline{a}_L\sigma xy : \overline{a}_L\sigma xy \in \Gamma\}; \\
\Gamma'_{\overline{a}_L} &= \{\sigma xy \in [n]_{k'+2} : A|_{\overline{a}_L} \textit{ accepts } \sigma xy\}.
\end{aligned}$$

Now,

$$\Gamma \subseteq \Gamma_0 \cup \bigcup_{\overline{a}_L \in \mathcal{G}_L} \Gamma_{\overline{a}_L}.$$

If $A|_{\overline{a}_L}$ has an empty S_j, then $\Gamma'_{\overline{a}_L} = \emptyset$. Suppose $A|_{\overline{a}_L}$ has no empty S_j. Then it is easy to see that $A|_{\overline{a}_L}$ has properties (P1)–(P4). By Lemma 3.19 we may find $\Delta'_{\overline{a}_L} \subseteq \Gamma'_{\overline{a}_L}$ of size at most $n^{-\frac{2}{5}}(n)_{k'+2}$ such that if $\Psi'_{\overline{a}_L} = \{\sigma : \exists\, xy\ \sigma xy \in \Gamma'_{\overline{a}_L} - \Delta'_{\overline{a}_L}\}$, then $|\Psi'_{\overline{a}_L}| \leq \exp(-\frac{k'}{2})n^{k'}$. Let

$$\begin{aligned}
\Psi_{\overline{a}_L} &= \{\overline{a}_L\sigma : \sigma \in \Psi'_{\overline{a}_L}\}; \\
\Delta_{\overline{a}_L} &= \{\overline{a}_L\sigma : \sigma \in \Delta'_{\overline{a}_L}\}; \\
\Delta &= \bigcup_{\overline{a}_L \in \mathcal{G}_L} \Delta_{\overline{a}_L}. \\
\Psi_1 &= \bigcup_{\overline{a}_L \in \mathcal{G}_L} \Psi_{\overline{a}_L}.
\end{aligned}$$

It follows from our definitions that (note that $k < (\log\log n)^2$)

$$\begin{aligned}
|\Delta| &\leq |\mathcal{G}_L| n^{-\frac{2}{5}}(n)_{k'+2} \\
&\leq (n)_{k-k'-2} n^{-\frac{2}{5}}(n)_{k'+2} \\
&\leq n^{-\frac{1}{3}}(n)_k \\
|\Psi_1| &\leq |\mathcal{G}_L| \exp(-\frac{k'}{2}) n^{k'} \\
&\leq \exp(-\frac{k'}{2}) n^{k-2}.
\end{aligned}$$

It is easy to verify that

$$\Psi = \{\overline{a} : \exists\, x, y\ \overline{a}xy \in \Gamma - \Delta\} \subseteq \mathcal{B} \cup \Psi_1.$$

Thus using (15) we get that

$$\begin{aligned}
|\Psi| \leq |\Psi'| &= |\mathcal{B}| + |\Psi_1| \\
&\leq (\frac{\text{size}(A)}{n} \exp(-\frac{k'}{2}) + \exp(-\frac{k'}{2})) n^{k-2} \\
&\leq 3\frac{\text{size}(A)}{n} \exp(-\frac{k'}{2}) n^{k-2} \\
&\leq \left(\frac{\text{size}(A)}{n}\right) \exp(-\frac{k'}{3}) n^{k-2}.
\end{aligned}$$

This completes the proof of the lemma. □

After this, the proof of our combinatorial lemma is straight forward.

Lemma 3.21 (Combinatorial Lemma) Let $A = \bigwedge_{j=1}^{t} \bigvee_{q \in S_j} q$ be a $(k-1)$-immune $\Pi\Sigma$ formula. Let $\Gamma = \{\gamma \in \binom{[n]}{k} : A \text{ accepts } \gamma\}$. Let $\alpha(k) = \frac{1}{6}\left\lfloor \sqrt{\frac{k}{e^4 \ln k}} \right\rfloor$.

(a) Suppose $\text{size}(A) \leq \frac{n}{2}$. Let

$$\Psi = \{\overline{a} \in \binom{[n]}{k-2} : \exists\, x, y \text{ such that } \overline{a} \cup \{x, y\} \in \Gamma\}.$$

Then $|\Psi| \leq \left(\frac{\text{size}(A)}{n}\right) e^{-\alpha(k)} \binom{n}{k-2}$.

(b) Suppose $\text{size}(A) > \frac{n}{2}$. Then there exists a set $\Delta \subseteq \Gamma$, $|\Delta| \leq n^{-\frac{1}{3}}\binom{n}{k}$, such that if

$$\Psi = \{\overline{a} \in \binom{[n]}{k-2} : \exists\, x, y \text{ such that } \overline{a} \cup \{x, y\} \in \Gamma - \Delta\},$$

then $|\Psi| \leq \left(\frac{\text{size}(A)}{n}\right) e^{-\alpha(k)} \binom{n}{k-2}$.

Proof: The only difference between this and the previous lemma is that here we consider sets instead of sequences. Note that that every k-set corresponds to precisely $(n)_k$ sequences. For us n is large and $k < (\log\log n)^2$, so $(n)_{k-2} \geq \frac{1}{2} n^{k-2}$. Our lemma is an immediate consequence of Lemma 3.20. □

4 The Upper Bound

In this section we show that there exist $\Sigma\Pi\Sigma$ formulas for computing T_k^n, when k is small, of size at most

$$e^{2\sqrt{k}\ln k} n \log n.$$

Assume that $k^{3/2}$ is an integer that divides n.

We construct the formulas in two stages. In the first stage we construct $\Pi\Sigma$ formulas. These formulas are $(k-1)$-immune and they accept a large proportion of all inputs that a formula computing T_k^n must accept. In the next stage, we take the disjunction of random copies of this formula and obtain a $\Sigma\Pi\Sigma$ formula computing T_k^n. Let $\binom{[n]}{k}$ denote the set of of all k sized subsets of $[n]$.

Lemma 4.1 There exists a $\Pi\Sigma$ formula computing $T_l^{l^2}$ of size at most

$$\binom{l^2}{l-1}(l^2 - l + 1).$$

Proof: Let

$$F = \bigwedge_{S \in \binom{[l^2]}{l^2 - l + 1}} \bigvee_{j \in S} x_j.$$

It is easy to verify that F computes $T_l^{l^2}$ correctly. Also,

$$\text{size}(F) = \binom{l^2}{l-1}(l^2 - l + 1).$$

□

Lemma 4.2 There exists a $(k-1)$-immune $\Pi\Sigma$ formula G such that $\text{size}(G) \leq \binom{k}{\sqrt{k}}n$ and G accepts at least $\exp(-\sqrt{k}(\ln\sqrt{k}+2))\binom{n}{k}$ sets of size k.

Proof: Let $l = \sqrt{k}$. Let $D_1, D_2, \ldots, D_l$ be a partition of $[n]$ into l equal parts. For each $i = 1, \ldots, l$, let $D_i^1, D_i^2, \ldots, D_i^{l^2}$ be a partition of D_i into l^2 equal parts. Thus $|D_i^j| = \frac{n}{l^3}$.

Let F_i be the formula obtained from the formula F in Lemma 4.1 by replacing the variable x_j by $\bigvee_{q \in D_i^j} x_q$. That is,

$$F_i = \bigwedge_{S \in \binom{[l^2]}{l^2-l+1}} \bigvee_{j \in S} \bigvee_{q \in D_i^j} x_q$$

Note that F_i is a $\Pi\Sigma$ formula and it is $(\sqrt{k}-1)$-immune. Let $G = \bigwedge_{i=1}^{l} F_i$. Note that G is a $\Pi\Sigma$ formula and it is $(k-1)$-immune. We have

$$\begin{aligned}
\text{size}(F_i) &= \binom{l^2}{l-1}(l^2-l+1)\frac{n}{l^3}; \\
\text{size}(G) &= \frac{l^2-l+1}{l^2}\binom{l^2}{l-1}n \\
&\leq \binom{l^2}{l}n.
\end{aligned}$$

The number of sets of size k accepted by G is given by

$$\begin{aligned}
\prod_{i=1}^{l}(\textit{ the number of sets of size } l \textit{ accepted by } F_i) &= \left[\binom{l^2}{l}(\frac{n}{l^3})^l\right]^l \\
&\geq \left[\frac{(l^2)_l}{l!}\frac{n^l}{l^{3l}}\right]^l \\
&\geq \left[\frac{l^{2l}(1-\frac{1}{l})^l}{l!}\frac{n^l}{l^{3l}}\right]^l \\
&\geq (\frac{1}{e^2 l})^l\binom{n}{l^2}.
\end{aligned}$$

Since $l^2 = k$, the proof is complete. □

Theorem 4.3 There exists a $\Sigma\Pi\Sigma$ formula of size at most $e^{3\sqrt{k}\log k}n\log n$ computing T_k^n.

Proof: Let r be a parameter to be chosen later. We take r independent copies of the formula G described in Lemma 4.2 by randomly permuting the variable set. Let these random copies be $G_1, G_2, \ldots, G_r$. For any fixed set T of size k,

$$\Pr[G_i \textit{ does not accept } T] \leq 1 - \exp(-\sqrt{k}(\ln\sqrt{k}+2)), \quad \textit{for } i = 1, 2, \ldots, r.$$

Since the G_i are independently chosen,

$$E[\textit{number of } k\textit{-sets accepted by none of } G_1, G_2, \ldots, G_r] \leq \binom{n}{k}(1-\exp(-\sqrt{k}(\ln\sqrt{k}+2)))^r.$$

For $r = k\exp(\sqrt{k}(\ln\sqrt{k}+2))\ln n$, this expected value is less than 1. Hence there must be some r copies $\hat{G}_1, \hat{G}_2, \ldots, \hat{G}_r$, such that every set of size k is accepted by at least one of them. Let our $\Sigma\Pi\Sigma$ formula for T_k^n be $\hat{F} = \bigvee_{i=1}^{r}\hat{G}_i$.

Clearly $\hat{F}$ is $(k-1)$-immune. It accepts every set of size k and by monotonicity every set of size at least k. Further,

$$\text{size}(\hat{F}) \leq \binom{k}{\sqrt{k}} (e^2\sqrt{k})^{\sqrt{k}} kn \log n \leq e^{2\sqrt{k}\ln k} n \log n.$$

□

Acknowledgment

I thank my advisor, Endre Szemerédi, for contributing so generously to this work. I am grateful to the referees for their comments and suggestions, especially to Referee 3 for her patient and sympathetic reading of the utterly unreadable first version of this paper. Shiva Chaudhuri checked the proofs carefully, pointed out many errors and suggested several improvements. Magnuús Halldòrsson's suggestions contributed greatly to the presentation in this paper. I thank them for their help. I thank Zoli Király, Ilan Newman, and Avi Wigderson for their comments. The final version of this paper was prepared while I was visiting the Japan Advanced Institute of Science and Technology, Hokuriku.

References

[1] R. B. Boppana. *Optimal Separations Between Concurrent-Write Parallel Machines.* Proceedings of the 23rd ACM STOC, 1989, pp. 320–326.

[2] R. B. Boppana. *Amplification of Probabilistic Boolean Formulas.* Advances in Computing Research, Vol. 5, 1989, pp. 27–45.

[3] R. B. Boppana and Michael Sipser. *The Complexity of Finite Functions.* Chapter 14, The Handbook of Theoretical Computer Science, (J. van Leeuven, ed.), Elsevier Science Publishers B. B., 1990, pp. 759–804.

[4] I. Csiszár and J. Körner. *Information Theory, Coding Theorems for Discrete Memoryless Systems.* Académia Kiadó, Budapest, 1981.

[5] M. Fredman and J. Komlós. *On the size of Separating Systems and Perfect Hash Functions.* Siam J. Alg. Disc. Meth., 1984, pp. 61–68.

[6] J. Hastad. *Computational Limitations for Small Depth Circuits.* MIT Press, 1986.

[7] G. Hansel. *Nombre minimal de contacts de fermature nécessaires pour réaliser une fonction booléenne symétrique de n variables.* C. R. Acad. Sci. Paris 258 (1964), pp. 6037–6040.

[8] L. S. Khasin. *Complexity Bounds for the Realization of Monotone Symmetrical Functions by Means of Formulas in the Basis* $\vee$, $\wedge$, $\neg$. Sov. Phys. Dokl. 14 (1970), pp. 1149–1151.

[9] V. M. Khrapchenko. *A method of obtaining lower bounds for the complexity* π*-schemes.* Math. Notes Acad. Sci. USSR 11 (1972), pp. 474–479.

[10] J. Körner. *Fredman–Komlós Bound and Information Theory.* Siam J. Alg. Disc. Meth., 1986, pp. 560–570.

[11] R. E. Krichevskii. *Complexity of contact circuits realizing a function of logical algebra.* Sov. Phys. Dokl. 8 (1964) pp. 770–772.

[12] I. Newman, P. Ragde, and A. Wigderson. *Perfect Hashing, Graph Entropy and Circuit Complexity* Proceedings of the 5th annual conference on Structure in Complexity Theory, 1990, pp. 91–99.

[13] M. S. Paterson, N. Pippenger, and U. Zwick. *Optimal carry save networks.* Boolean Function Complexity: selected papers for the LMS symposium, Durham 1990. Cambridge Univ. Press, 1992, pp. 174–201.

[14] J. Radhakrishnan. *Better Bounds for Threshold Formulas.* In the proceedings of the 32nd IEEE FOCS, 1991, pp. 314-323.

[15] J. Radhakrishnan. *Improved Bounds for Covering Complete Uniform Hypergraphs.* Information Processing Letters 41 (1992), pp. 203–207.

[16] M. Snir. *The Covering Problem of Complete Uniform Hypergraphs.* A note, Discrete Math. 27, 1979, pp. 103–105.

[17] L. G. Valiant. *Short monotone formulae for the majority function.* Journal of Algorithms 5, 1984, pp. 363–366.

[18] I. Wegener. *The Complexity of Boolean Functions.* Wiley-Teubner Series in Computer Science, 1987.