A ΣΠΣ formula has the form $\bigvee \bigwedge \bigvee L_{uvw}$, where each $L$ is either a variable or a negated variable. In this paper we study the computation of threshold functions by ΣΠΣ 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 ΣΠΣ formula computing the threshold function $T^n_k$ 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 ΣΠΣ formulas for computing $T^n_k$ with size at most $\exp(2\sqrt{k} \ln k) n \log n$.

1. Introduction

The $k$th threshold function, $T^n_k$, 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 $T^n_k$ using ΣΠΣ formulas. A ΣΠΣ formula has the form $\bigvee \bigwedge \bigvee q$, where each $S_{ij}$ is a subset of variables and their negations.

The complexity of computing the majority function, $T^n_{n/2}$, 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 $T^n_k$, $k \leq n/2$, 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 $T^n_k$.

The complexity of computing $T^n_k$ 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 $O(n^{4.57})$. Kharpchenko [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 $\left\lceil \frac{k}{2} \right\rceil n \log \left( \frac{n}{k-1} \right)$.

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 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 $\Omega \left( \frac{\exp(k)}{k^{1/k}} n \log n \right)$. 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 $\delta(k) = \frac{1}{30} \sqrt{\frac{k}{\ln k}}$.

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