Polynomial threshold functions and Boolean threshold circuits

Kristoffer Arnsfelt Hansen[1] Vladimir V. Podolskii[2]

[1]Aarhus University

[2]Steklov Mathematical Institute

MFCS 2013

Boolean Threshold Functions

Boolean function $f\colon \{a, b\}^n \to \{-1, 1\}$.

Polynomial threshold gate computing f is a polynomial $p \in \mathbb{R}[x_1, \ldots, x_n]$ such that for all $x \in \{a, b\}^n$ we have

$$f(x) = \operatorname{sign} p(x).$$

Complexity measures:

The degree of p is the degree of the polynomial.

The length of p is the number of its monomials.

Example

$\{a, b\} = \{1, 2\}.$

$p(x, y) = 16 - 15xy + 3x^2y^2.$

Example

$\{a, b\} = \{1, 2\}$.

$p(x, y) = 16 - 15xy + 3x^2y^2$.

$$
\begin{aligned}
x = y = 1 & \qquad 16 - 15 + 3 > 0 \\
x = 2, y = 1 & \qquad 16 - 30 + 12 < 0 \\
x = y = 2 & \qquad 16 - 60 + 48 > 0
\end{aligned}
$$

Example

$\{a, b\} = \{1, 2\}$.

$p(x, y) = 16 - 15xy + 3x^2y^2$.

$x = y = 1$	$16 - 15 + 3 > 0$
$x = 2, y = 1$	$16 - 30 + 12 < 0$
$x = y = 2$	$16 - 60 + 48 > 0$

$p(x, y)$ computes PARITY function:
$p(x, y) > 0$ iff $x + y$ is odd.

The domain

The most studied cases are $\{a, b\} = \{0, 1\}$ and $\{a, b\} = \{-1, 1\}$.

In these cases we can assume that $\deg p \leq n$.
Indeed, $x^2 = x$, if $x \in \{0, 1\}$ and
$x^2 = 1$, if $x \in \{-1, 1\}$.

The domain

The most studied cases are $\{a, b\} = \{0, 1\}$ and $\{a, b\} = \{-1, 1\}$.

In these cases we can assume that $\deg p \leq n$.
Indeed, $x^2 = x$, if $x \in \{0, 1\}$ and
$x^2 = 1$, if $x \in \{-1, 1\}$.

For general $\{a, b\}$ this is not the case, in principle degree greater than n can help to reduce the length.

Degree vs. length

Indeed, large degree can help.

Theorem (Basu et. al, 2004)

PARITY *over* $\{1, 2\}$ *requires length* 2^n *when the degree is bounded by* n*, but is computable by degree* n^2 *and length* $n + 1$ *threshold gate.*

Our example:

$$p(x, y) = 16 - 15xy + 3x^2y^2$$

$n = 2$, length is $n + 1 = 3$ degree is $n^2 = 4$.

The PTF complexity class

Given $l(n)$ and $d(n)$, we denote by

$$\mathrm{PTF}_{a,b}(l(n), d(n))$$

the class of Boolean functions over $\{a, b\}^n$ computable by polynomial threshold functions of length $l(n)$ and degree $d(n)$.

$\mathrm{PTF}_{a,b}(l(n), \infty)$ — no bound on the degree.

$\mathrm{PTF}_{a,b}(d(n)) = \mathrm{PTF}_{a,b}(\mathrm{poly}(n), d(n))$.

The PTF complexity class

Given $l(n)$ and $d(n)$, we denote by

$$\mathsf{PTF}_{a,b}(l(n), d(n))$$

the class of Boolean functions over $\{a, b\}^n$ computable by polynomial threshold functions of length $l(n)$ and degree $d(n)$.

$\mathsf{PTF}_{a,b}(l(n), \infty)$ — no bound on the degree.

$\mathsf{PTF}_{a,b}(d(n)) = \mathsf{PTF}_{a,b}(\mathrm{poly}(n), d(n))$.

Below we concentrate on $\{1, 2\}$-domain. Our results also hold for all $\{a, b\}$-domains, which are essentially different from $\{0, 1\}$ and $\{-1, 1\}$.

Circuit Classes Notation

We consider classes AND, OR, XOR, AC^0.

THR: $f(x) = \text{sign}(\sum_i w_i x_i + w_0)$.

MAJ: $f(x) = \text{sign}(\sum_i w_i x_i + w_0)$, where all w_i are integers bounded by polynomial in n.

Let $\mathcal{C}_1$ and $\mathcal{C}_2$ be two classes of Boolean circuits.
By $\mathcal{C}_1 \circ \mathcal{C}_2$ we denote the class of polynomial size circuits consisting of circuit from $\mathcal{C}_1$ with circuits from $\mathcal{C}_2$ as inputs.

Exponential form of PTFs

For a variable $y \in \{1, 2\}$ consider $x = \log_2 y \in \{0, 1\}$.
Then $y = 2^x$.

For monomials we have

$$y_1^{a_1} \dots y_n^{a_n} = 2^{a_1 x_1 + \dots + a_n x_n}$$

and for polynomials

$$P(y) = \sum_{j=1}^{l} c_j \prod_{i=1}^{n} y_i^{a_{ij}} = \sum_{j=1}^{l} c_j 2^{\sum_{i=1}^{n} a_{ij} x_i} = \sum_{j=1}^{l} 2^{\sum_{i=1}^{n} a_{ij} x_i + \log_2 c_j}$$

Initial results

Lemma
$\mathrm{PTF}_{1,2}(2, \infty) = \mathrm{THR}$ *and* $\mathrm{PTF}_{1,2}(2, \mathit{poly}(n)) = \mathrm{MAJ}$.

Initial results

Lemma

$\mathrm{PTF}_{1,2}(2, \infty) = \mathrm{THR}$ *and* $\mathrm{PTF}_{1,2}(2, poly(n)) = \mathrm{MAJ}$.

Proof.

Consider THR gate: $\sum_{i=1}^{n} w_i x_i - w_0 \geq 0$.
Raise each side to the power of 2.

In the other direction, consider

$$c_1 2^{\sum_{i=1}^{n} a_i x_i} + c_2 2^{\sum_{i=1}^{n} b_i x_i} \geq 0.$$

Interesting case: $\operatorname{sign} c_1 \neq \operatorname{sign} c_2$.
Move one summand to the other side and take a logarithm.

□

Bounded degree PTFs

Theorem

$$\mathsf{PTF}_{1,2}(\mathit{poly}(n)) = \mathsf{THR} \circ \mathsf{MAJ}$$

Bounded degree PTFs

Theorem

$$\mathsf{PTF}_{1,2}(\mathit{poly}(n)) = \mathsf{THR} \circ \mathsf{MAJ}$$

Note that

$$\mathsf{PTF}_{0,1}(\mathsf{poly}(n)) = \mathrm{THR} \circ \mathrm{AND}$$

and

$$\mathsf{PTF}_{-1,1}(\mathsf{poly}(n)) = \mathrm{THR} \circ \mathrm{XOR}.$$

Thus, threshold gates over $\{1, 2\}$ are strictly stronger.

Depth 2 Threshold Circuits

$\mathrm{THR} \circ \mathrm{THR}$

| ?

$\mathrm{THR} \circ \mathrm{MAJ}$

| Goldman et al., 92

$\mathrm{MAJ} \circ \mathrm{MAJ}$

Theorem (Goldman, Håstad, Razborov, 92)

$\mathrm{MAJ} \circ \mathrm{THR} = \mathrm{MAJ} \circ \mathrm{MAJ}.$

Bounded degree PTFs

Theorem (restated)

$$\mathsf{PTF}_{1,2}(poly(n)) = \mathsf{THR} \circ \mathsf{MAJ}$$

Main observation: linear form in each MAJ gate can obtain only polynomially many values.

We can precisely compute each MAJ gate by polynomial length $\{1, 2\}$-polynomial.

Byproduct

Lemma

Any polynomial size circuit in $\mathrm{THR} \circ \mathrm{MAJ}$ *is equivalent to a polynomial size circuit of the same form such that all majority gates on the bottom level are monotone.*

The same is true for $\mathrm{MAJ} \circ \mathrm{MAJ}$.

Lower bounds

Let $x, y \in \{0,1\}^n$.
Inner product function:

$$IP(x, y) = \bigoplus_i x_i \wedge y_i.$$

Lower bounds

Let $x, y \in \{0,1\}^n$.
Inner product function:

$$IP(x,y) = \bigoplus_i x_i \wedge y_i.$$

Theorem (restated)

$$\mathsf{PTF}_{1,2}(poly(n)) = \mathsf{THR} \circ \mathsf{MAJ}$$

Corollary

$\mathrm{IP} \notin \mathsf{PTF}_{1,2}(poly(n))$, $\mathrm{AND} \circ \mathrm{OR} \circ \mathrm{AND}_2 \notin \mathsf{PTF}_{1,2}(poly(n))$.

Lower bounds

Let $x, y \in \{0,1\}^n$.
Inner product function:

$$IP(x, y) = \bigoplus_i x_i \wedge y_i.$$

Theorem (restated)

$$\mathsf{PTF}_{1,2}(poly(n)) = \mathsf{THR} \circ \mathsf{MAJ}$$

Corollary

$\mathrm{IP} \notin \mathsf{PTF}_{1,2}(poly(n))$, $\mathrm{AND} \circ \mathrm{OR} \circ \mathrm{AND}_2 \notin \mathsf{PTF}_{1,2}(poly(n))$.

What about $\mathsf{PTF}_{1,2}(\infty)$?

Sign rank

Let $A = (a_{ij})$ be a real matrix with nonzero elements.
Sign rank of A is the minimal rank of the real matrix $B = (b_{ij})$ such that $\operatorname{sign} b_{ij} = \operatorname{sign} a_{ij}$ for all i, j.

For the Boolean function $f(x, y)$ consider the matrix $M_f = (f(x, y))_{x,y}$ of size $2^n \times 2^n$. The sign rank of $f(x, y)$ is the sign rank of M_f.

Theorem (Forster, 2002)

The sign rank of $\mathrm{IP}(x, y)$ *is* $2^{\Omega(n)}$.

Theorem (Razborov, Sherstov, 2010)

The sign rank of $\mathrm{AND} \circ \mathrm{OR} \circ \mathrm{AND}_2$ *is* $2^{\Omega(n^{1/3})}$.

From this: IP and $\mathrm{AND} \circ \mathrm{OR} \circ \mathrm{AND}_2$ require exponential size $\mathrm{THR} \circ \mathrm{MAJ}$ circuits.
Why: MAJ gates compute low rank matrices. Rank is subadditive.

Lower bounds for $\mathrm{PTF}_{1,2}(\infty)$

Lemma

Assume $f : \{0,1\}^n \times \{0,1\}^n \to \{-1,1\}$ is computed by a PTF of length s on the domain $\{1,2\}^n \times \{1,2\}^n$. Then the matrix M_f has sign rank at most s.

Proof.

Consider one monomial

$$\prod_i x_i^{a_i} y_i^{b_i} = \left(\prod_i x_i^{a_i}\right) \cdot \left(\prod_i y_i^{b_i}\right).$$

It defines rank 1 matrix. □

Lower bounds for $\mathrm{PTF}_{1,2}(\infty)$

Lemma

Assume $f : \{0,1\}^n \times \{0,1\}^n \to \{-1,1\}$ is computed by a PTF of length s on the domain $\{1,2\}^n \times \{1,2\}^n$. Then the matrix M_f has sign rank at most s.

Proof.

Consider one monomial

$$\prod_i x_i^{a_i} y_i^{b_i} = \left(\prod_i x_i^{a_i}\right) \cdot \left(\prod_i y_i^{b_i}\right).$$

It defines rank 1 matrix. □

Corollary

Any PTF on the domain $\{1,2\}^n \times \{1,2\}^n$ computing IP_2 requires length $2^{\Omega(n)}$. Any PTF on the domain $\{1,2\}^n \times \{1,2\}^n$ computing $\mathrm{AND} \circ \mathrm{OR} \circ \mathrm{AND}_2$ requires length $2^{\Omega(n^{1/3})}$.

Bounded Weight vs. Unbounded Weight

Is it true that $\mathrm{PTF}_{1,2}(\mathrm{poly}(n)) = \mathrm{PTF}_{1,2}(\infty)$?

Bounded Weight vs. Unbounded Weight

Is it true that $\mathrm{PTF}_{1,2}(\mathrm{poly}(n)) = \mathrm{PTF}_{1,2}(\infty)$?

Open problem!

Bounded Weight vs. Unbounded Weight

Is it true that $\mathrm{PTF}_{1,2}(\mathrm{poly}(n)) = \mathrm{PTF}_{1,2}(\infty)$?

Open problem!

Theorem

If $\mathrm{THR} \circ \mathrm{THR} \not\subseteq \mathrm{THR} \circ \mathrm{MAJ} \circ \mathrm{AND}_2$ *then* $\mathrm{PTF}_{1,2}(\infty) \not\subseteq \mathrm{PTF}_{1,2}(\mathit{poly}(n))$.

Bounded Weight vs. Unbounded Weight

Is it true that $\mathrm{PTF}_{1,2}(\mathrm{poly}(n)) = \mathrm{PTF}_{1,2}(\infty)$?

Open problem!

Theorem

If $\mathrm{THR} \circ \mathrm{THR} \not\subseteq \mathrm{THR} \circ \mathrm{MAJ} \circ \mathrm{AND}_2$ *then* $\mathrm{PTF}_{1,2}(\infty) \not\subseteq \mathrm{PTF}_{1,2}(\mathit{poly}(n))$.

To prove this we need the following lemma.

Lemma

$\mathrm{THR} \circ \mathrm{THR} \subseteq \mathrm{PTF}_{1,2}(\infty) \circ \mathrm{AND}_2$.

Proof of the lemma

Lemma (restated)

$\mathsf{THR} \circ \mathsf{THR} \subseteq \mathsf{PTF}_{1,2}(\infty) \circ \mathsf{AND}_2$.

Proof of the lemma

Lemma (restated)

$\mathsf{THR} \circ \mathsf{THR} \subseteq \mathsf{PTF}_{1,2}(\infty) \circ \mathsf{AND}_2$.

Proof of the lemma.

Definition. ETHR: $f(x) = 1$ iff $\sum_i w_i x_i + w_0 = 0$.

It is known that $\mathsf{THR} \circ \mathsf{THR} = \mathsf{THR} \circ \mathsf{ETHR}$ (Hansen, P., 2010).

Proof of the lemma

Lemma (restated)

$\mathrm{THR} \circ \mathrm{THR} \subseteq \mathrm{PTF}_{1,2}(\infty) \circ \mathrm{AND}_2$.

Proof of the lemma.

Definition. ETHR: $f(x) = 1$ iff $\sum_i w_i x_i + w_0 = 0$.

It is known that $\mathrm{THR} \circ \mathrm{THR} = \mathrm{THR} \circ \mathrm{ETHR}$ (Hansen, P., 2010).

Note that ETHR-gate defined by $L(x) = 0$ can be approximated by $2^{-c \cdot L(x)^2}$, where c is positive constant.

Thus we can rewrite $\mathrm{THR} \circ \mathrm{ETHR}$ in the form

$$\mathrm{sign}\left(\sum_i 2^{-c \cdot L_i(x)^2}\right),$$

where $L_i(x)$ are linear forms.
Opening the brackets in the exponent we get the circuit of the form $\mathrm{PTF}_{1,2}(\infty) \circ \mathrm{AND}_2$.

□

Bounded Weight vs. Unbounded Weight

Theorem (restated)

If $\mathrm{THR} \circ \mathrm{THR} \not\subseteq \mathrm{THR} \circ \mathrm{MAJ} \circ \mathrm{AND}_2$ *then* $\mathrm{PTF}_{1,2}(\infty) \not\subseteq \mathrm{PTF}_{1,2}(\mathit{poly}(n))$.

Proof.

Assume $\mathrm{PTF}_{1,2}(\mathrm{poly}(n)) = \mathrm{PTF}_{1,2}(\infty)$. Then

$$\mathrm{THR} \circ \mathrm{THR} \subseteq \mathrm{PTF}_{1,2}(\infty) \circ \mathrm{AND}_2 =$$

$$\mathrm{PTF}_{1,2}(\mathrm{poly}(n)) \circ \mathrm{AND}_2 = \mathrm{THR} \circ \mathrm{MAJ} \circ \mathrm{AND}_2,$$

□

Note that $\mathrm{THR} \circ \mathrm{THR} \subseteq \mathrm{THR} \circ \mathrm{MAJ} \circ \mathrm{AND}_2$ implies $\mathrm{THR} \circ \mathrm{THR} \circ \mathrm{AND} = \mathrm{THR} \circ \mathrm{MAJ} \circ \mathrm{AND}$.

Relations to Communication Complexity

$f\colon \{0,1\}^n \times \{0,1\}^n \to \{0,1\}$.
There are players Alice and Bob.
Alice gets x, Bob gets y.
They have to compute $f(x,y)$.
Communication complexity of f is the worst case bit size of their communication.

Unbounded error randomized communication complexity:
Each of Alice and Bob has an access to the source of random bits (separately). They have to output $f(x,y)$ correctly with probability $> 1/2$. For this version of complexity we use the notation $UCC(f)$.

Theorem (Paturi, Simon, 1986)

For any f $UCC(f)$ is equal to the logarithm of the sign rank of f up to an additive constant.

Three players, Number on the Forehead

Suppose now there are 3 players A, B and C and f depends on variables $x, y, z \in \{0,1\}^n$.

A has access to y, z, B has access to x, z, C has access to y, z. We can consider unbounded error case in this setting too.

We denote it by $UCC_3(f)$.

Tensor rank

Let $A = (a_{ijk})$ be an order 3 tensor, $i, j, k = 1, \ldots, n$.

A is a *cylinder tensor* if it does not depend on one of the coordinates.

A is a *cylinder product* if it can be written as a Hadamard product $A_1 \odot A_2 \odot A_3$ where A_1,A_2, and A_3 are cylinder tensors. That is, $a_{ijk} = a^{(1)}_{jk} a^{(2)}_{ik} a^{(3)}_{ij}$.

The *sign complexity* of an order 3 tensor $A = (a_{ijk})$ is the minimum r such that there exist cylinder product tensors $B_1, \ldots, B_r$, with $B_\ell = (b^{(\ell)}_{ijk})$, such that $\text{sign}(a_{ijk}) = \text{sign}\left(b^{(1)}_{ijk} + \cdots + b^{(r)}_{ijk}\right)$, for all i, j, k.

Note that we have a nonstandard notion of rank!

Tensor rank and Communication Complexity

Lemma

Consider $f : \{0,1\}^n \times \{0,1\}^n \times \{0,1\}^n \to \{-1,1\}$ *and let* s *be the uniform sign complexity of the associated communication tensor* T_f*. Then*

$$UCC_3(f) = \Theta(\log_2 s).$$

Lemma (restated)

Assume $f : \{0,1\}^n \times \{0,1\}^n \to \{-1,1\}$ *is computed by a PTF of length* s *on the domain* $\{1,2\}^n \times \{1,2\}^n$*. Then the matrix* M_f *has sign rank at most* s*.*

Thus, f above has communication complexity $\Omega(s)$.

Lemma (restated)

Assume $f : \{0,1\}^n \times \{0,1\}^n \to \{-1,1\}$ *is computed by a PTF of length* s *on the domain* $\{1,2\}^n \times \{1,2\}^n$*. Then the matrix* M_f *has sign rank at most* s*.*

Thus, f above has communication complexity $\Omega(s)$.

Lemma

Assume that $f : \{0,1\}^n \times \{0,1\}^n \times \{0,1\}^n \to \{-1,1\}$ *is computed by a* $\mathrm{PTF}_{1,2}(\infty) \circ \mathrm{AND}_2$*. Then the sign complexity of* T_f *is polynomial in* n*.*

The proof is analogous: $2^{p(x,y,z)} = 2^{p_1(x,y)} 2^{p_2(x,z)} 2^{p_3(y,z)}$.

Lemma (restated)

$\mathrm{THR} \circ \mathrm{THR} \subseteq \mathrm{PTF}_{1,2}(\infty) \circ \mathrm{AND}_2.$

Corollary

Assume that $f : \{0,1\}^n \times \{0,1\}^n \times \{0,1\}^n \to \{-1,1\}$ *has unbounded error 3-player communication complexity* c*. Then every* $\mathrm{THR} \circ \mathrm{THR}$ *computing* f *must contain* $2^c/poly(n)$ *gates.*

We do not know functions with large unbounded error 3-player communication complexity.

Relations between the domains

Our results works for all domains $\{a, b\}$ such that $a, b \neq 0$ and $|a| \neq |b|$.

But what is the relation of classes for different domains? Is it true that

$$\mathsf{PTF}_{1,2}(\infty) = \mathsf{PTF}_{1,3}(\infty)?$$

Relations between the domains

Our results works for all domains $\{a, b\}$ such that $a, b \neq 0$ and $|a| \neq |b|$.
But what is the relation of classes for different domains? Is it true that

$$\mathsf{PTF}_{1,2}(\infty) = \mathsf{PTF}_{1,3}(\infty)?$$

Open problem!

Relations between the domains

Our results works for all domains $\{a, b\}$ such that $a, b \neq 0$ and $|a| \neq |b|$.
But what is the relation of classes for different domains? Is it true that

$$\mathrm{PTF}_{1,2}(\infty) = \mathrm{PTF}_{1,3}(\infty)?$$

Open problem!

But we know that $\mathrm{PTF}_{1,2}(\infty) = \mathrm{PTF}_{1,-2}(\infty)$.

More generally,

Lemma

For all $a, b \in \mathbb{R}$ and for any natural number k we have $\mathrm{PTF}_{a,b}(\infty) = \mathrm{PTF}_{a^k,b^k}(\infty)$.

$$\mathrm{PTF}_{1,2}(\infty) = \mathrm{PTF}_{1,4}(\infty) = \mathrm{PTF}_{1,-2}(\infty).$$

Other results

We also consider max-plus version of PTFs:

- they are somewhere between $\mathrm{AND} \circ \mathrm{THR}$ and $\mathrm{AND} \circ \mathrm{OR} \circ \mathrm{THR}$;
- we know lower bounds for them (through usual PTFs);
- the class is still strong (can compute various “complex” functions).

Other partial results:

- Exponential degree implies doubly exponential weight and vice versa;
- Exponential degree upper bound for length 3 PTFs;
- Exponential degree lower bound for constant length PTFs.