Subspace-Invariant AC^0 Formulas

Benjamin Rossman*
University of Toronto

June 12, 2018

Abstract

We consider the action of a linear subspace U of $\{0,1\}^n$ on the set of AC^0 formulas with inputs labeled by literals in $\{X_1, \overline{X}_1, \ldots, X_n, \overline{X}_n\}$, where an element $u \in U$ acts by toggling negations on literals with indices i such that $u_i = 1$. A formula is *U-invariant* if it is fixed by this action. For example, the smallest known construction of depth $d+1$ formulas computing PARITY_n, of size $O(n2^{dn^{1/d}})$, are P-invariant where P is the subspace of even-weight elements of $\{0,1\}^n$. In this paper, we show a nearly matching $2^{d(n^{1/d}-1)}$ lower bound on the P-invariant depth $d+1$ formula size of PARITY_n. (Quantitatively this improves the best known $\Omega(2^{\frac{1}{84}d(n^{1/d}-1)})$ lower bound for unrestricted depth $d+1$ formulas [17].) More generally, for any linear subspaces $U \subset V$, we show that if a Boolean function is U-invariant and non-constant over V, then its U-invariant depth $d+1$ formula size is at least $2^{d(m^{1/d}-1)}$ where m is the minimum Hamming weight of a vector in $U^\perp \setminus V^\perp$.

1 Introduction

There are two natural group actions on the set of literals $\{X_1, \overline{X}_1, \ldots, X_n, \overline{X}_n\}$: the symmetric group S_n acts by permuting indices, while Z_2^n acts by toggling negations. These groups (which together generate $Z_2^n \rtimes S_n$ a.k.a. the wreath product $Z_2 \wr S_n$) have a similar action on the set of n-variable Boolean functions, as well as the set of n-variable Boolean circuits. Here we consider circuits with unbounded fan-in AND and OR gates and inputs labeled by literals, also known as AC^0 circuits, where groups S_n and Z_2^n act syntactically by reassigning the literals that label inputs. If G is subgroup of S_n or Z_2^n (or more generally of $Z_2^n \rtimes S_n$), we say that a Boolean function or circuit is *G-invariant* if it is fixed under the action of each element in G. Note that every G-invariant circuit computes a G-invariant function, and conversely every G-invariant function is computable by a G-invariant circuit. (If C is any circuit computing a G-invariant function f, then the AND (or the OR) of circuits $\{C^g : g \in G\}$ is a G-invariant circuit computing f.)

We define the *G-invariant circuit size* of a G-invariant function f as the minimum number of gates in a G-invariant circuit that computes f. This may be compared to the *unrestricted circuit size* of f, noting that f can be computed (possibly more efficiently) by circuits that are not necessarily G-invariant. Several questions arise. What gap, if any, exists between the G-invariant vs. unrestricted circuit size of G-invariant functions? Are lower bounds on G-invariant circuit size easier to obtain, and do they suggest new strategies for proving lower bounds for unrestricted

*Supported by NSERC. A preliminary version of this paper appeared in ICALP 2017.

circuits? Is there a nice characterization of functions computable by polynomial-size G-invariant circuits? The same questions may be asked with respect to G-invariant versions of other complexity measures, such as formula (leaf)size and bounded-depth versions of both circuit and formula size, noting that the action of G on circuits preserves both depth and fan-out.

The answer to these questions appears to be very different for subgroups of S_n and subgroups of Z_2^n. This is illustrated by considering the n-variable parity function, which maps an input $x \in \{0,1\}^n$ to the Hamming weight of x modulo 2. This function is both S_n-invariant (it is a so-called *symmetric function*) and P-invariant where $P \subset Z_2^n$ is the index-2 subgroup of even-weight elements in Z_2^n. The smallest known circuits and formulas for PARITY_n have size $O(n)$ and leafsize $O(n^2)$, respectively. These circuits and formulas turn out to be P-invariant, as do the smallest known bounded-depth circuits and formulas (which we describe in §2.3). In contrast, the S_n-invariant circuit size of PARITY_n is known to be exponential [2].

1.1 Invariance under subgroups of S_n

G-invariant circuit complexity for subgroups G of S_n has been previously studied from the standpoint of Descriptive Complexity, where the goal is to characterize complexity classes in terms of definability in different logics [11]. Here one considers Boolean functions that encode (isomorphism-invariant) properties of relational structures. Properties of m-vertex simple graphs, for instance, can be identified with G-invariant functions $\{0,1\}^n \to \{0,1\}$ of $n = \binom{m}{2}$ variables (each corresponding to a potential edge) and G is the group S_m acting on the set of potential edges. More generally, if σ is a finite relational signature σ, one considers the action of S_m on $n = \sum_{R \in \sigma} m^{\text{arity}(R)}$ variables (encoding the possible σ-structures with universe $[m]$).

Denenberg et al [8] showed that S_m-invariant circuits of polynomial size and constant depth (subject to a certain uniformity condition) capture precisely the first-order definable properties of finite σ-structures. Otto [13] introduced a certain limit object of finite circuits (imposing uniformity in a different way) and showed a correspondence between the logic $L^{\omega}_{\infty\omega}$ (infinitary logic with a bounded number of variables) and S_m-invariant circuits of polynomial size and arbitrary depth. Otto also gave characterizations of fixed-point logic and partial-fixed-point logic in terms of S_m-invariant Boolean networks. Recently, Anderson and Dawar [2] showed a correspondence between fixed-point logic and polynomial-size S_m-invariant circuits, as well between fixed-point logic with counting and polynomial-size S_m-invariant circuits in the basis that includes majority gates.

Choiceless Polynomial Time [3, 4, 6, 16] provides a different example of a G-invariant model of computation, where $G \subseteq S_n$ is the automorphism group of the input structure. Invariance under subgroups of S_n has been explored in other settings as well, see for instance [1, 15, 18].

1.2 Invariance under subgroups of Z_2^n

In this paper, we initiate a study of invariant complexity with respect to subgroups of Z_2^n. Since we will be using linear algebra, we henceforth identify Z_2^n with the $\mathbb{F}_2$-vector space $\{0,1\}^n$ under coordinate-wise addition modulo 2, denoted $\oplus$. We identify subgroups of Z_2^n with linear subspaces U of $\{0,1\}^n$. A function $f : \{0,1\}^n \to \{0,1\}$ is thus *U-invariant* if $f(x) = f(x \oplus u)$ for all $x \in \{0,1\}^n$ and $u \in U$. Note that U-invariant functions are in one-to-one correspondence with functions from the quotient space $\{0,1\}^n/U$ to $\{0,1\}$.

Our focus is on bounded-depth circuits and formulas. Returning to the example of the P-invariant function PARITY_n (where P is the even-weight subgroup of $\{0,1\}^n$), there is a well-known

recursive construction of depth $d+1$ circuits for PARITY_n, which we describe in §2.3. Roughly speaking, one combines a depth-2 circuit for $\text{PARITY}_{n^{1/d}}$ with depth-d circuits for $\text{PARITY}_{n^{(d-1)/d}}$ on disjoint blocks of variables. This results in a depth $d+1$ circuit of size $O(n \cdot 2^{n^{1/d}})$ or, when converted to a tree, a depth $d+1$ formula of leafsize $O(n \cdot 2^{dn^{1/d}})$. This circuit and formula are the smallest known computing PARITY_n. They are also easily seen to be P-invariant.

The main result of this paper gives a nearly matching lower bound of $2^{d(n^{1/d}-1)}$ on the P-invariant depth $d+1$ formula size of PARITY_n. This implies a $2^{n^{1/d}-1}$ lower bound on the P-invariant depth $d+1$ circuit size, via the basic fact that every (U-invariant) depth $d+1$ circuit of size s is equivalent to a (U-invariant) depth $d+1$ formula of size at most s^d. Quantitatively, the lower bounds are stronger than the best known $\Omega(2^{\frac{1}{10}n^{1/d}})$ and $\Omega(2^{\frac{1}{84}d(n^{1/d}-1)})$ lower bounds for unrestricted depth $d+1$ circuits [9] and formulas [17], respectively. Of course, P-invariance is a severe restriction for circuits and formulas, so it is no surprise that the lower bounds we obtain is stronger and significantly easier to prove. The linear-algebraic technique in this paper is entirely different from the "switching lemma" approach of [9, 17].

The general form of our lower bound is the following:

Theorem 1. *Let $U \subset V$ be linear subspaces of $\{0,1\}^n$, and suppose F is a U-invariant depth $d+1$ formula which is non-constant over V. Then F has size at least $2^{d(m^{1/d}-1)}$ where $m = \min\{|x| : x \in U^\perp \setminus V^\perp\}$, that is, the minimum Hamming weight of a vector x which is orthogonal to U and non-orthogonal to V.*

Here *size* refers to the number of depth-1 subformulas, as opposed to *leafsize*. Note that the bound in Theorem 1 does not depend on the dimension n of the ambient space. Also note that aforementioned $2^{d(n^{1/d}-1)}$ lower bound for PARITY_n follows from the case $U = P$ and $V = \{0,1\}^n$.

We remark that, since $\lim_{d\to\infty} d(m^{1/d}-1) = \ln(m)$, Theorem 1 implies an $m^{\ln(2)}$ lower bound on the size of *unbounded-depth* formulas which are U-invariant and non-constant over V. Theorem 1 also implies a $2^{m^{1/d}-1}$ lower bound for depth $d+1$ circuits; however, we get no nontrivial lower bound for unbounded-depth circuits, since $\lim_{d\to\infty} m^{1/d} - 1 = 0$.

2 Preliminaries

Let n and d range over positive integers. $[n]$ is the set $\{1,\ldots,n\}$. $\ln(n)$ is the natural logarithm and $\log(n)$ is the base-2 logarithm.

The Hamming weight of a vector $x \in \{0,1\}^n$, denoted $|x|$, is the cardinality of the set $\{i \in [n] : x_i = 1\}$. For vectors $x, y \in \{0,1\}^n$, let $x \oplus y$ denote the coordinate-wise sum modulo 2 and let $\langle x, y\rangle$ denote the inner product modulo 2.

Let $\mathcal{L}$ denote the lattice of linear subspaces of $\{0,1\}^n$. For $U, V \in \mathcal{L}$, let $U+V$ denote the subspace spanned by U and V. Let $V^\perp$ denote the dual subspace $V^\perp = \{x \in \{0,1\}^n : \langle x, v\rangle = 0 \text{ for all } v \in V\}$. Recall the following facts about duality of subspaces over finite fields:

$$\dim(V) + \dim(V^\perp) = n, \qquad U \subseteq V \iff V^\perp \subseteq U^\perp,$$
$$V = (V^\perp)^\perp, \qquad (U+V)^\perp = U^\perp \cap V^\perp, \qquad (U \cap V)^\perp = U^\perp + V^\perp.$$

2.1 AC^0 formulas

We write $\mathcal{F}$ for the set of n-variable AC^0 formulas (with unbounded fan-in AND and OR gates and leaves labeled by literals). Formally, let $\mathcal{F} = \bigcup_{d\in\mathbb{N}} \mathcal{F}_d$ where $\mathcal{F}_d$ is the set of *depth-d formulas*,

defined inductively:

- $\mathcal{F}_0$ is the set $\{X_1, \overline{X}_1, \ldots, X_n, \overline{X}_n\} \cup \{0, 1\}$,
- $\mathcal{F}_{d+1}$ is the set of ordered pairs $\{(gate, \mathcal{G}) : gate \in \{\text{AND}, \text{OR}\}$ and $\mathcal{G}$ is a nonempty subset of $\mathcal{F}_d\}$.

Every formula $F \in \mathcal{F}$ computes a Boolean function $\{0,1\}^n \to \{0,1\}$ in the usual way. For $x \in \{0,1\}^n$, we write $F(x)$ for the value of F on x. For a nonempty set $S \subseteq \{0,1\}^n$ and $b \in \{0,1\}$, notation $F(S) \equiv b$ denotes that $F(x) = b$ for all $x \in S$. We say that F is *non-constant* on S if $F(S) \not\equiv 0$ and $F(S) \not\equiv 1$ (i.e., there exist $x, y \in S$ such that $F(x) = 0$ and $F(y) = 1$).

The *depth* of F is the unique $d \in \mathbb{N}$ such that $F \in \mathcal{F}_d$. *Leafsize* is the number of depth-0 subformulas, and *size* is the number of depth-1 subformulas. Inductively,

$$\text{leafsize}(F) = \begin{cases} 1 & \text{if } F \in \mathcal{F}_0, \\ \sum_{G \in \mathcal{G}} \text{leafsize}(G) & \text{if } F = (gate, \mathcal{G}) \in \mathcal{F} \setminus \mathcal{F}_0, \end{cases}$$

$$\text{size}(F) = \begin{cases} 0 & \text{if } F \in \mathcal{F}_0, \\ 1 & \text{if } F \in \mathcal{F}_1, \\ \sum_{G \in \mathcal{G}} \text{size}(G) & \text{if } F = (gate, \mathcal{G}) \in \mathcal{F} \setminus (\mathcal{F}_0 \cup \mathcal{F}_1). \end{cases}$$

Clearly $\text{size}(F) \leq \text{leafsize}(F)$. Note that size is within a factor 2 of the number of gates in F, which is how one usually measures the size of circuits. Our lower bound naturally applies to size, while the upper bound that we present in §2.3 is naturally presented in terms of leafsize.

2.2 The action of $\{0,1\}^n$

We now formally define the action of $\{0,1\}^n$ (as the group Z_2^n) on the set $\mathcal{F}$. For $u \in \{0,1\}^n$ and $F \in \mathcal{F}$, let F^u be the formula obtained from F by exchanging literals X_i and $\overline{X}_i$ for every $i \in [n]$ with $u_i = 1$. Formally, this action is defined inductively by

$$F^u = \begin{cases} F & \text{if } F \in \{0,1\}, \\ X_i \text{ (resp. } \overline{X}_i) & \text{if } F = X_i \text{ (resp. } \overline{X}_i) \text{ and } u_i = 0, \\ \overline{X}_i \text{ (resp. } X_i) & \text{if } F = X_i \text{ (resp. } \overline{X}_i) \text{ and } u_i = 1, \\ (gate, \{G^u : G \in \mathcal{G}\}) & \text{if } F = (gate, \mathcal{G}). \end{cases}$$

Note that F^u has the same depth and size as F and computes the function $F^u(x) = F(x \oplus u)$ for all $x \in \{0,1\}^n$.

Let U be a linear subspace of $\{0,1\}^n$ (i.e., subgroup of Z_2^n). We say that an AC^0 formula F is:

- *U-invariant* if $F^u = F$ (i.e., these are syntactically identical formulas) for every $u \in U$,
- *semantically U-invariant* if F computes a U-invariant function (i.e., $F(x) = F(x \oplus u)$ for every $u \in U$ and $x \in \{0,1\}^n$).

Note that every U-invariant formula is semantically U-invariant, but not conversely.

2.3 Upper bound

We review the smallest known construction of bounded-depth formulas for PARITY_n and observe that these formulas are P-invariant.

Proposition 2. *For all $d, n \geq 1$, PARITY_n is computable by a P-invariant depth $d+1$ formula of leafsize at most $n \cdot 2^{dn^{1/d}}$ where P is the even-weight subspace of $\{0,1\}^n$. If $n^{1/d}$ is an integer, this bound improves to $n \cdot 2^{d(n^{1/d}-1)}$.*

Proof. For an optimal choice of $k, n_1, \ldots, n_k \geq 1$ with $n_1 + \cdots + n_k = n$, we construct a P_n-invariant depth $d+1$ formula for PARITY_n with output gate OR (resp. AND) by composing the brute-force DNF (resp. CNF) for PARITY_k (in which each variable occurs 2^{k-1} times) with P_{n_i}-invariant depth-d formulas with output gate AND (resp. OR) computing both PARITY_{n_i} and $1 - \mathrm{PARITY}_{n_i}$. The minimum leafsize $\beta(d+1, n)$ achieved by this construction is given by the following recurrence:

$$\beta(1,n) = \begin{cases} 1 & \text{if } n = 1, \\ \infty & \text{if } n > 1, \end{cases} \qquad \beta(d+1,n) = \min_{\substack{k, n_1, \ldots, n_k \geq 1 : \\ n_1 + \cdots + n_k = n}} 2^{k-1} \sum_{i=1}^{k} \beta(d, n_i).$$

If $n^{1/d}$ is an integer, we get the bound $\beta(d+1,n) \leq n \cdot 2^{d(n^{1/d}-1)}$ by setting $k = n^{1/d}$ and $n_1 = \cdots = n_k = n^{(d-1)/d}$. For arbitrary $d, n \geq 1$, we get the bound $\beta(d+1,n) \leq n \cdot 2^{dn^{1/d}}$ by setting $k = \lceil n/t \rceil$ and $n_1, \ldots, n_k \in \{t-1, t\}$ where $t = \lfloor n^{(d-1)/d} \rfloor$. □

Remark 3. PARITY_n is known computable by P-invariant formulas of depth $\lceil \log n \rceil + 1$ and leafsize $O(n^2)$. The $n \cdot 2^{dn^{1/d}}$ upper bound of Proposition 2 is therefore slack, as this equals n^3 when $d = \log n$, whereas $n \cdot 2^{d(n^{1/d}-1)} = n^2$. We suspect that the upper bound of Proposition 2 can be improved that $O(n \cdot 2^{d(n^{1/d}-1)})$ for all $d \leq \log n$, perhaps by a more careful analysis of the recurrence for $\beta(d+1, n)$.

3 Linear-algebraic lemmas

Recall that $\mathcal{L}$ denotes the lattice of linear subspaces of $\{0,1\}^n$. Let U, V, S, T range over elements of $\mathcal{L}$. If U is a subspace of V, recall that a *projection* from V to U is a linear map $\rho : V \to U$ such that $\rho(u) = u$ for every $u \in U$. We begin by showing that if U is a codimension-k subspace of V (i.e., $\dim(V) - \dim(U) = k$), then there there exists a projection $\rho : V \to U$ with "Hamming-weight stretch" $k+1$.

Lemma 4. *If U is a codimension-k subspace of V, then there exists a projection ρ from V to U such that $|\rho(v)| \leq (k+1)|v|$ for all $v \in V$.*

Proof. Greedily choose a basis $w_1, \ldots, w_k$ for V over U such that w_i has minimal Hamming weight among elements of $V \setminus \mathrm{Span}(U \cup \{w_1, \ldots, w_{i-1}\})$ for all $i \in [k]$. Each $v \in V$ has a unique representation $v = u \oplus a_1 w_1 \oplus \cdots \oplus a_k w_k$ where $u \in U$ and $a_1, \ldots, a_k \in \{0,1\}$. Let $\rho : V \to U$ be the map $v \mapsto u$ and observe that this is a projection.

To show that $|\rho(v)| \leq (k+1)|v|$, we first observe that $|a_i w_i| \leq |v|$ for all $i \in [k]$. If $a_i = 0$, this is obvious, as $|a_i w_i| = 0$. If $a_i = 1$, then $v \in V \setminus \mathrm{Span}(U \cup \{w_1, \ldots, w_{i-1}\})$, so by our choice of w_i

we have $|a_i w_i| = |w_i| \leq |v|$. Completing the proof, we have

$$\begin{aligned} |\rho(v)| &= |v \oplus a_1 w_1 \oplus \cdots \oplus a_k w_k| \\ &\leq |v| + |a_1 w_1| + \cdots + |a_k w_k| \\ &\leq (k+1)|v|. \end{aligned}$$

□

Definition 5. Define sets $\mathcal{L}_2$ and $\mathcal{L}_4$ as follows:

$$\mathcal{L}_2 = \big\{(U, V) \in \mathcal{L} \times \mathcal{L} : U \text{ is a codimension-1 subspace of } V\big\},$$
$$\mathcal{L}_4 = \big\{((S, T), (U, V)) \in \mathcal{L}_2 \times \mathcal{L}_2 : T \cap U = S \text{ and } T + U = V\big\}.$$

The next lemma shows that $\mathcal{L}_4$ is symmetric under duality.

Lemma 6. *For all $((S, T), (U, V)) \in \mathcal{L}_4$, we have $((V^\perp, U^\perp), (T^\perp, S^\perp)) \in \mathcal{L}_4$.*

Proof. We use the properties of dual subspaces states in §2. Consider any $((S, T), (U, V)) \in \mathcal{L}_4$. First note that $(V^\perp, U^\perp) \in \mathcal{L}_2$ by the fact that $U \subseteq V \implies V^\perp \subseteq U^\perp$ and $\dim(U^\perp) - \dim(V^\perp) = (n - \dim(U)) - (n - \dim(V)) = \dim(V) - \dim(U) = 1$. Similarly, we have $(T^\perp, S^\perp) \in \mathcal{L}_2$. We now have $((V^\perp, U^\perp), (T^\perp, S^\perp)) \in \mathcal{L}_4$ since $U^\perp \cap T^\perp = (T + U)^\perp = V^\perp$ and $U^\perp + T^\perp = (T \cap U)^\perp = S^\perp$. □

Finally, we state a dual pair of lemmas which play a key role in the proof of Theorem 1.

Lemma 7. *For all $(S, T) \in \mathcal{L}_2$ and $V \supseteq T$, there exists $U \supseteq S$ such that $((S, T), (U, V)) \in \mathcal{L}_4$ and*

$$\min_{x \in V \setminus U} |x| \geq \frac{1}{\dim(V) - \dim(T) + 1} \min_{y \in T \setminus S} |y|.$$

Proof. By Lemma 4, there exists a projection ρ from V onto T such that $|\rho(v)| \leq (\dim(V) - \dim(T) + 1)|v|$ for all $v \in V$. Let $U = \rho^{-1}(S)$ and note that U is a codimension-1 subspace of V. (This follows by applying the rank-nullity theorem to linear maps $\rho : V \to T$ and $\rho\restriction U : U \to S$ and noting that $\ker(\rho) = \ker(\rho\restriction U)$.) We have $S = T \cap U$ and $T + U = V$, hence $((S, T), (U, V)) \in \mathcal{L}_4$. Choosing x with minimum Hamming weight in $V \setminus U$, we observe that $\rho(x) \in T \setminus S$ and $|x| \geq |\rho(v)|/(\dim(V) - \dim(T) + 1)$, which proves the lemma. □

Lemma 8. *For all $(U, V) \in \mathcal{L}_2$ and $S \subseteq U$, there exists $T \subseteq V$ such that $((S, T), (U, V)) \in \mathcal{L}_4$ and*

$$\min_{x \in S^\perp \setminus T^\perp} |x| \geq \frac{1}{\dim(U) - \dim(S) + 1} \min_{y \in U^\perp \setminus V^\perp} |y|.$$

Proof. Follows directly from Lemmas 6 and 7. □

4 Proof of Theorem 1

We first prove the base case of Theorem 1 for depth-2 formulas, also known as DNFs and CNFs.

Lemma 9. *Suppose F is a depth-2 formula and $(U, V) \in \mathcal{L}_2$ such that $F(U) \equiv b$ and $F(V \setminus U) \equiv 1 - b$ for some $b \in \{0, 1\}$. Then $\mathrm{size}(F) \geq 2^{m-1}$ and $\mathrm{leafsize}(F) \geq m \cdot 2^{m-1}$ where $m = \min\{|x| : x \in U^\perp \setminus V^\perp\}$.*

Note that conditions $F(U) \equiv b$ and $F(V \setminus U) \equiv 1 - b$ imply that that F is semantically U-invariant (i.e., computes a U-invariant function). The stronger hypothesis that F is (syntactically) U-invariant is required to prove Theorem 1 for formulas of depth > 2.

Proof. Without loss of generality, assume that F is a DNF formula (i.e. an OR-of-ANDs formula) and $F(U) \equiv 0$ and $F(V \setminus U) \equiv 1$. (The argument is similar if we replace DNF with CNF, or if we assume that $F(U) \equiv 1$ and $F(V \setminus U) \equiv 0$.) We may further assume that F is minimal firstly with respect to the number of clauses and secondly with respect to the number of literals in each clause.

Consider a clause G of F. This clause G is the AND of some number ℓ of literals. Without loss of generality, suppose these literals involve the first ℓ coordinates. Let π be the projection $\{0,1\}^n \to \{0,1\}^\ell$ onto the first ℓ coordinates. There is a unique element $p \in \{0,1\}^\ell$ such that $G(x) = 1 \iff \pi(x) = p$ for all $x \in \{0,1\}^n$. Observe that $G(U) \equiv 0$ (since $F(U) \equiv 0$) and, therefore, $p \notin \pi(U)$.

We claim that $p \in \pi(V \setminus U)$. To see why, assume for contradiction that $p \notin \pi(V \setminus U)$. Then $G(V) \equiv 0$. But this means that the clause G can be removed from F and the resulting function F' would still satisfy $F'(U) \equiv 0$ and $F'(V \setminus U) \equiv 1$. This contradicts the minimality of F with respect to number of clauses.

For each $i \in [\ell]$, let $p^{(i)}$ be the neighbor of p in $\{0,1\}^\ell$ along the ith coordinate. We claim that $p^{(1)}, \ldots, p^{(\ell)} \in \pi(U)$. Without loss of generality, we give the argument showing $p^{(\ell)} \in \pi(U)$. Let G' be the AND of the first $\ell - 1$ literals in G, and let F' be the formula obtained from F by replacing G with G'. For all $x \in \{0,1\}^n$, we have $G(x) \leq G'(x)$ and hence $F(x) \leq F'(x)$. Therefore, $F'(V \setminus U) \equiv 1$. We now note that there exists $u \in U$ such that $F'(u) = 1$ (otherwise, we would have $F'(u) \equiv 0$, contradicting the minimality of F with respect to the width of each clause). Since $F(u) = 0$ and G' is the only clause of F' distinct from the clauses of F, it follows that $G'(u) = 1$. This means that $u_{\{1,\ldots,\ell-1\}} = p_{\{1,\ldots,\ell-1\}}$. We now have $\pi(u) = p^{(\ell)}$ (otherwise, we would have $\pi(u) = p$ and therefore $G(u) = 1$ and $F(u) = 1$, contradicting that fact that $F(U) \equiv 0$).

Since π is a linear function and $\pi(U) \neq \pi(V)$, it follows that $\pi(U)$ is a codimension-1 subspace of $\pi(V)$. The fact that $p \in \pi(V \setminus U)$ and $p^{(1)}, \ldots, p^{(\ell)} \in \pi(U)$ now forces $\pi(V) = \{0,1\}^\ell$ and $\pi(U) = \{q \in \{0,1\}^\ell : |q| \text{ is even}\}$. Therefore, $1^\ell \in \pi(U)^\perp \setminus \pi(V)^\perp$ (writing 1^ℓ for the all-1 vector in $\{0,1\}^\ell$). It follows that $1^\ell 0^{n-\ell} \in U^\perp \setminus V^\perp$ and, therefore, $\ell = |1^\ell 0^{n-\ell}| \geq m$ (by definition of m).

We now observe that

$$\mathop{\mathbb{P}}_{v \in V}[G(v) = 1] = \mathop{\mathbb{P}}_{v \in V}[\pi(v) = p] = \mathop{\mathbb{P}}_{q \in \pi(V)}[q = p] = \mathop{\mathbb{P}}_{q \in \{0,1\}^\ell}[q = p] = 2^{-\ell} \leq 2^{-m}.$$

That is, each clause in F has value 1 over at most 2^{-m} fraction of points in V. Since the set $V \setminus U$ has density $1/2$ in V, we see that 2^{m-1} clauses are required to cover $V \setminus U$.

Subject to the stated minimality assumptions on F (first with respect to the number of clauses and second to the width of each clause), we conclude that F contains $\geq 2^{m-1}$ clauses, each of width $\geq m$. Therefore, $\text{size}(F) \geq 2^{m-1}$ and $\text{leafsize}(F) \geq m \cdot 2^{m-1}$. □

The induction step of Theorem 1 makes use of the following inequality.

Lemma 10. *For all real $a, b, c \geq 1$, we have $a + c(b/a)^{1/c} \geq (c+1)b^{1/(c+1)}$. This holds with equality iff $a = b^{1/(c+1)}$.*

Proof. Taking the derivative of the lefthand side with respect to a, we get $\frac{\partial}{\partial a}\big(a + c(b/a)^{1/c}\big) = 1 - (b/a^{c+1})^{1/c}$. The function $a \mapsto a + c(b/a)^{1/c}$ is thus seen to have a unique minimum at $a = b^{1/(c+1)}$, where it takes values $(c+1)b^{1/(c+1)}$. □

Onto the main result:

Theorem 1 (restated). *Let $U \subset V$ be linear subspaces of $\{0,1\}^n$, and suppose F is a U-invariant depth $d+1$ formula which is non-constant over V. Then F has size at least $2^{d(m^{1/d}-1)}$ where $m = \min\{|x| : x \in U^\perp \setminus V^\perp\}$.*

Proof. We first observe that it suffices to prove the theorem in the case where $(U,V) \in \mathcal{L}_2$, that is, U has codimension-1 in V. To see why, note that for any $U \subset V$ such that F is U-invariant and non-constant over V, there must exist $U \subset W \subseteq V$ such that $(U,W) \in \mathcal{L}_2$ and F is non-constant over W. Assuming the theorem holds with respect to $U \subset W$, it also hold with respect to $U \subset V$, since $U^\perp \setminus V^\perp \subseteq U^\perp \setminus W^\perp$ and hence $\min\{|x| : x \in U^\perp \setminus V^\perp\} \geq \min\{|x| : x \in U^\perp \setminus W^\perp\}$.

Therefore, we assume $(U,V) \in \mathcal{L}_2$ and prove the theorem by induction on d. The base case $d = 1$ is established by Lemma 9. For the induction step, let $d \geq 2$ and assume $F \in \mathcal{F}_{d+1}$ is a U-invariant and non-constant over V. Without loss of generality, we consider the case where $F = (\mathrm{OR}, \mathcal{G})$ for some nonempty $\mathcal{G} \subseteq \mathcal{F}_d$. (The case where $F = (\mathrm{AND}, \mathcal{G})$ is symmetric, with the roles of 0 and 1 exchanged.)

Since F is U-invariant, we have $G^u \in \mathcal{G}$ for every $u \in U$ and $G \in \mathcal{G}$. We claim that it suffices to prove the theorem in the case where the action of U on $\mathcal{G}$ is transitive (i.e. $\mathcal{G} = \{G^u : u \in U\}$ for every $G \in \mathcal{G}$). To see why, consider the partition $\mathcal{G} = \mathcal{G}_1 \sqcup \cdots \sqcup \mathcal{G}_t$, $t \geq 1$, into orbits under U. For each $i \in [t]$, let F_i be the formula $(\mathrm{OR}, \mathcal{G}_i)$. Note that F_i is U-invariant and U acts transitively on $\mathcal{G}_i$. Clearly, we have $F(v) = \bigvee_{i \in [t]} F_i(v)$ for all $v \in V$. Since every U-invariant Boolean function is constant over sets U and $V \setminus U$ (using the fact that U has codimension-1 in V), this means that each F_i satisfies either $F_i(V) \equiv 0$ or $F(v) = F_i(v)$ for all $v \in V$. Because F is non-constant over V, it follows that there exists $i \in [t]$ such that $F(v) = F_i(v)$ for all $v \in V$. In particular, this F_i is non-constant over V. Since $\mathrm{size}(F) \geq \mathrm{size}(F_i)$, we have reduced proving the theorem for F to proving to theorem for F_i.

In light of the preceding paragraph, we proceed under the assumption that U acts transitively on $\mathcal{G}$. Fix an arbitrary choice of $G \in \mathcal{G}$. Let

$$\begin{aligned} S &= \mathrm{Stab}_U(G) \ (= \{u \in U : G^u = G\}), \\ a &= \dim(U) - \dim(S) + 1. \end{aligned}$$

By the orbit-stabilizer theorem,

$$|\mathcal{G}| = |\mathrm{Orbit}_U(G)| = [U : S] = |U|/|S| = 2^{a-1}.$$

Since $\mathrm{size}(G') = \mathrm{size}(G)$ for every $G' \in \mathcal{G}$, we have

$$\mathrm{size}(F) = \sum_{G' \in \mathcal{G}} \mathrm{size}(G') = |\mathcal{G}| \cdot \mathrm{size}(G) = 2^{a-1} \cdot \mathrm{size}(G). \tag{1}$$

We next observe that G^u is S-invariant for every $u \in U$ (in fact, $S = \mathrm{Stab}_U(G^u)$). This follows from the fact that $(G^u)^s = G^{u \oplus s} = (G^s)^u = G^u$ for every $s \in S$.

By Lemma 8, there exists T such that $((S,T),(U,V)) \in \mathcal{L}_4$ and

$$\min_{x \in S^\perp \setminus T^\perp} |x| \geq \frac{1}{\dim(U) - \dim(S) + 1} \min_{y \in U^\perp \setminus V^\perp} |y| = \frac{m}{a}.$$

We claim that there exists $u \in U$ such that G^u is non-constant on T. There are two cases to consider:

Case 1: *Suppose* $F(U) \equiv 0$ *and* $F(V \setminus U) \equiv 1$.

We have $G(U) \equiv 0$ and $G(V) \not\equiv 0$. Fix any $v \in V \setminus U$ such that $G(v) = 1$. In addition, fix any $w \in T \setminus U$ (noting that $T \setminus U$ is nonempty since $U + T = V$ and $U \subset V$). Let $u = v \oplus w$ and note that $u \in U$ (since U is a codimension-1 subspace of V and $v, w \in V \setminus U$). We have $G^u(U) \equiv 0$ and $G^u(w) = G(w \oplus u) = G(v) = 1$. By the S-invariance of G^u, it follows that $G^u(S) \equiv 0$ and $G^u(T \setminus S) \equiv 1$. In particular, G^u is non-constant on T.

Case 2: *Suppose* $F(U) \equiv 1$ *and* $F(V \setminus U) \equiv 0$.

We have $G(U) \not\equiv 0$ and $G(V \setminus U) \equiv 0$. Fix any $u \in U$ such that $G(u) = 1$. In addition, fix any $w \in T \setminus U$ and let $v = w \oplus u$. We have $G^u(v) = G(v \oplus u) = G(w) = 0$ (since $w \in V \setminus U$ and $G(V \setminus U) \equiv 0$). We also have $G^u(\vec{0}) = G(u) = 1$ where $\vec{0}$ is the origin in $\{0,1\}^n$. By S-invariance of G^u, it follows that $G^u(S) \equiv 1$ and $G^u(T \setminus S) \equiv 0$. In particular, G^u is non-constant on T.

Since G^u is S-invariant and non-constant on T and $\mathrm{depth}(G^u) = (d-1)+1$, we may apply the induction hypothesis to G^u. Thus, we have

$$\mathrm{size}(G) = \mathrm{size}(G^u) \geq 2^{(d-1)((m/a)^{1/(d-1)}-1)}. \tag{2}$$

Since $d \geq 2$, Lemma 10 tells us

$$a + (d-1)(m/a)^{1/(d-1)} \geq d(m/a)^{1/d}. \tag{3}$$

Putting together (1), (2), (3), we get the desired bound

$$\begin{aligned} \mathrm{size}(F) &\geq 2^{a-1} \cdot 2^{(d-1)((m/a)^{1/(d-1)}-1)} \\ &= 2^{a+(d-1)(m/a)^{1/(d-1)}-d} \\ &\geq 2^{d(m^{1/d}-1)}. \end{aligned}$$

This completes the proof of Theorem 1. □

5 Remarks and open questions

5.1 Another application of Theorem 1

Theorem 1 applies to interesting subspaces U of $\{0,1\}^n$ besides the even-weight subspace P. Here we describe one example. Let G be a simple graph with n edges, so that $\{0,1\}^n$ may be identified with the set of spanning subgraphs of G. The *cycle space* of G is the subspace $Z \subseteq \{0,1\}^n$ consisting of *even subgraphs* of G (i.e., spanning subgraphs in which every vertex has even degree). Consider the even-weight subspace $Z_0 = \{z \in Z : |z| \text{ is even}\}$. Provided that G is non-bipartite, Z_0 is a codimension-1 subspace of Z.

Let $m = \min\{|x| : x \in Z_0^\perp \setminus Z^\perp\}$ as in Theorem 1 with $U = Z_0$ and $V = Z$. This number m is seen to be equal to the minimum number of edges whose removal makes G bipartite. It follows that $m = n - c$ where c is the number edges in a maximum cut in G. Now suppose G is generated as a uniform random 3-regular graph with n edges (and $\frac{2}{3}n$ vertices). There is a constant $\varepsilon > 0$ such that $c \leq (1-\varepsilon)n$ (and hence $m \geq \varepsilon n$) holds asymptotically almost surely [5]. From these observations, we have

Corollary 11. *Every Z_0-invariant depth $d+1$ formula that computes* PARITY_n *over Z has size at least $2^{d((\varepsilon n)^{1/d}-1)}$ asymptotically almost surely.*

The AC^0 complexity of computing PARITY_n over the cycle space of a graph G is loosely related to the AC^0-Frege proof complexity of the Tseitin tautology on G, which has been explored recently in [10, 14]. In general, however, we do not have techniques to lower bound the (non-subspace-invariant) AC^0 complexity of PARITY_n over arbitrary subspaces of $\{0,1\}^n$.

5.2 The $V \setminus U$ search problem

For linear subspaces $U \subset V$ of $\{0,1\}^n$, consider the following "$V \setminus U$ search problem". There is a hidden vector $w \in V \setminus U$ and the goal is to learn a nonzero coordinate of w (any $i \in [n]$ such that $w_i = 1$) by asking queries (yes/no questions) in the form of linear functions $\{0,1\}^n \to \{0,1\}$. The *d-round query complexity* of this problem is the minimum number of queries required by a deterministic protocol which issues batches of queries over d consecutive rounds. (Thus, a 1-round protocol is a non-adaptive, while a fully adaptive protocol issues 1 query per round.) By a similar argument as the proof of Theorem 1, we get a $d(m^{1/d}-1)$ lower bound on the d-round query complexity of the $V \setminus U$-search problem where $m = \min\{|x| : x \in U^\perp \setminus V^\perp\}$. We remark that the "$V \setminus U$ search problem" may be viewed as an U-invariant version of the Karchmer-Wigderson game.

5.3 Open questions

We conclude by mentioning some open questions and challenges raised by this work:

1. Does the $2^{d(m^{1/d}-1)}$ lower bound of Theorem 1 (or even a weaker bound like $2^{\Omega(m^{1/d})}$ or $2^{m^{\Omega(1/d)}}$) apply to depth $d+1$ formulas which are *semantically U-invariant* and non-constant on V?

2. Counting leafsize instead of size, improve the lower bound of Theorem 1 from $2^{d(m^{1/d}-1)}$ to $m \cdot 2^{d(m^{1/d}-1)}$.

3. Improve the upper bound of Proposition 2 from $n \cdot 2^{dn^{1/d}}$ to $O(n \cdot 2^{d(n^{1/d}-1)})$ for all $d \leq \log n$.

4. What is the maximum gap, if any, between the U-invariant vs. unrestricted AC^0 complexity of a U-invariant Boolean function?

References

[1] Miklos Ajtai. Symmetric systems of linear equations modulo p. In *TR94-015 of the Electronic Colloquium on Computational Complexity*, 1994.

[2] Matthew Anderson and Anuj Dawar. On symmetric circuits and fixed-point logics. *Theory of Computing Systems*, pages 1–31, 2016.

[3] Andreas Blass, Yuri Gurevich, and Saharon Shelah. Choiceless polynomial time. *Ann. Pure & Applied Logic*, 100(1–3):141–187, 1999.

[4] Andreas Blass, Yuri Gurevich, and Saharon Shelah. On polynomial time computation over unordered structures. *Journal of Symbolic Logic*, 67(3):1093–1125, 2002.

[5] Béla Bollobás. The isoperimetric number of random regular graphs. *European Journal of combinatorics*, 9(3):241–244, 1988.

[6] Anuj Dawar. On symmetric and choiceless computation. In *International Conference on Topics in Theoretical Computer Science*, pages 23–29. Springer, 2015.

[7] Larry Denenberg, Yuri Gurevich, and Saharon Shelah. Definability by constant-depth polynomial-size circuits. *Information and Control*, 70(2-3):216–240, 1986.

[8] Johan Håstad. Almost optimal lower bounds for small depth circuits. In *STOC '86: Proceedings of the Eighteenth Annual ACM Symposium on Theory of Computing*, pages 6–20, 1986.

[9] Johan Håstad. On small-depth frege proofs for tseitin for grids. In *58th IEEE Annual Symposium on Foundations of Computer Science (FOCS), OCT 15-17, 2017, Berkeley, CA*, pages 97–108. IEEE, 2017.

[10] Neil Immerman. *Descriptive complexity*. Springer, 2012.

[11] V. M. Khrapchenko. Complexity of the realization of a linear function in the class of π-circuits. *Mathematical Notes of the Academy of Sciences of the USSR*, 9(1):21–23, 1971.

[12] Martin Otto. The logic of explicitly presentation-invariant circuits. In *International Workshop on Computer Science Logic*, pages 369–384. Springer, 1996.

[13] Toniann Pitassi, Benjamin Rossman, Rocco A Servedio, and Li-Yang Tan. Poly-logarithmic Frege depth lower bounds via an expander switching lemma. In *Proceedings of the 48th Annual ACM SIGACT Symposium on Theory of Computing*, pages 644–657. ACM, 2016.

[14] Søren Riis and Meera Sitharam. Generating hard tautologies using predicate logic and the symmetric group. *Logic Journal of IGPL*, 8(6):787–795, 2000.

[15] Benjamin Rossman. Choiceless computation and symmetry. In *Fields of logic and computation*, pages 565–580. Springer, 2010.

[16] Benjamin Rossman. The average sensitivity of bounded-depth formulas. In *Proc. 56th Annual IEEE Symposium on Foundations of Computer Science*, pages 424–430. IEEE, 2015.

[17] Steven Rudich and Leonard Berman. Optimal circuits and transitive automorphism groups. In *International Colloquium on Automata, Languages, and Programming*, pages 516–524. Springer, 1988.

[18] Jun Tarui. Smallest formulas for the parity of 2^k variables are essentially unique. *Theoretical Computer Science*, 411(26):2623–2627, 2010.