Subspace-Invariant AC^0 Formulas

Benjamin Rossman[*]

University of Toronto, Toronto, Canada
rossman@utoronto.ca

Abstract

The n-variable PARITY function is computable (by a well-known recursive construction) by AC^0 formulas of depth $d+1$ and leafsize $n{\cdot}2^{dn^{1/d}}$. These formulas are seen to possess a certain symmetry: they are syntactically invariant under the subspace P of even-weight elements in $\{0,1\}^n$, which acts (as a group) on formulas by toggling negations on input literals. In this paper, we prove a $2^{d(n^{1/d}-1)}$ lower bound on the size of syntactically P-invariant depth $d+1$ formulas for PARITY. Quantitatively, this beats the best $2^{\Omega(d(n^{1/d}-1))}$ lower bound in the non-invariant setting [16].

1998 ACM Subject Classification F.1.1 Models of Computation

Keywords and phrases lower bounds, size-depth tradeoff, parity, symmetry in computation

Digital Object Identifier 10.4230/LIPIcs.ICALP.2017.93

1 Introduction

Let U be a linear subspace of $\{0,1\}^n$. We say that a Boolean function $f : \{0,1\}^n \to \{0,1\}$ is *U-invariant* if $f(x) = f(x \oplus u)$ for all $u \in U$ and $x \in \{0,1\}^n$. (Note that U-invariant Boolean functions are in one-to-one correspondence with functions from the quotient space $\{0,1\}^n/U$ to $\{0,1\}$.) An obvious example is the PARITY function $x \mapsto \bigoplus_{i=1}^n x$, which is P-invariant where P is the linear subspace of even-weight elements in $\{0,1\}^n$.

We may also view U as a group that acts on the set of n-variable Boolean circuits (as well as the set of n-variable Boolean formulas). Here we consider circuits with unbounded fan-in AND and OR gates and inputs labeled by literals in the set $\{X_1, \overline{X}_1, \dots, X_n, \overline{X}_n\}$, also known as *AC^0 circuits* in the setting where depth is bounded. For a circuit C and an element $u \in U$, let C^u be the circuit obtained from C by negating the ith pair of literals (i.e. exchanging X_i and $\overline{X}_i$ as labels on inputs) for all coordinates $i \in [n]$ such that $u_i = 1$. This action of U on circuits is compatible with the action on Boolean functions: for all $u \in U$ and $x \in \{0,1\}^n$, we have $C^u(x) = C(x \oplus u)$.

There are two notions of U-invariance for circuits. We say that C is *syntactically U-invariant* if C is identical to C^u for every $u \in U$ (we define this notion precisely for formulas), while we say that C is *semantically U-invariant* if it computes a U-invariant function. Syntactic U-invariance clearly implies semantic U-invariance. However, the converse is false: a circuit may compute a U-invariant function without being syntactically U-invariant.

For a U-invariant Boolean function f, we define its *U-invariant circuit size* as the minimum number of gates in a syntactically U-invariant circuit that computes it. This quantity may be compared to the usual ("non-invariant") circuit size of f. There are several questions we may ask: What gap, if any, is there between the U-invariant circuit size and non-invariant circuit size of f? Are lower bounds for U-invariant circuit size easier to prove,

[*] Supported by NSERC.

44th International Colloquium on Automata, Languages, and Programming (ICALP 2017).
Editors: Ioannis Chatzigiannakis, Piotr Indyk, Fabian Kuhn, and Anca Muscholl;
Article No. 93; pp. 93:1–93:11
Leibniz International Proceedings in Informatics
LIPICS Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

and do they suggest new strategies for proving lower bound in the non-invariant setting? The same questions may be asked with respect to U-invariant versions of other complexity measures, such as formula size and bounded-depth versions of both circuit and formula size (noting that the action of U on circuits preserves fan-out and depth).[1]

In this paper, we focus on bounded-depth formula size. Our primary target is the P-invariant PARITY function where P is the linear subspace of even-weight elements in $\{0,1\}^n$. We start from the observation that the best known construction of bounded-depth circuits and formulas for PARITY are syntactically P-invariant. Here we refer to the well-known recursive construction, for all $d \geq 1$, of depth $d+1$ circuits and formulas for PARITY, of size at most $n{\cdot}2^{n^{1/d}}$ and $n{\cdot}2^{dn^{1/d}}$ respectively. The main result of this paper (Theorem 1) yields a nearly matching lower bound of $2^{d(n^{1/d}-1)}$ on the P-invariant depth $d+1$ formula size of PARITY. This implies a $2^{n^{1/d}-1}$ lower bound on P-invariant depth $d+1$ circuit size.[2] 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 non-invariant depth $d+1$ circuits [10] and formulas [16], respectively. Qualitatively, syntactic P-invariance appears to be a severe restriction and unnatural from the standpoint of computation.

The general form of our lower bound is the following theorem.

► **Theorem 1.** *Let $U \subset V$ be linear subspaces of $\{0,1\}^n$, and suppose F is a syntactically 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 $m = \min\{|x| : x \in U^\perp \setminus V^\perp\}$ (i.e. m is the minimum Hamming weight of a vector x which is orthogonal to U but non-orthogonal to V).*

Some observations: first, notice that the bound in Theorem 1 does not depend on the parameter n, i.e. the dimension of the ambient hypercube. The lower bound for PARITY described in the previous paragraph is the special case $U = P$ and $V = \{0,1\}^n$. Theorem 1 implies an $m^{1/\log_2(e)}$ lower bound for unbounded-depth formulas, since $\lim_{d\to\infty} d(m^{1/d}-1) = \ln(m)$. It also implies a $2^{m^{1/d}-1}$ lower bound for depth $d+1$ circuits. (However, we get no non-trivial lower bound for unbounded-depth circuits, since $\lim_{d\to\infty} m^{1/d} - 1 = 0$.)

The proof of Theorem 1 uses elementary linear algebra, in particular a small lemma on the existence of linear retractions with small Hamming-weight distortion (Lemma 5). Overall, this is much simpler than the random restriction and polynomial approximation methods typically used to prove AC^0 lower bounds.

1.1 Related Work

Syntactically invariant models of computation have been previously studied from the perspective of Descriptive Complexity, an area that characterizes complexity classes in terms of definability in different logics [11]. In this context, the notion of invariance pertains to the action of S_m on $n = \binom{m}{2}$ binary variables, encoding the edge relation of a simple graph on m vertices. More generally, for a finite relational signature σ, one may consider the action of S_m on $n = \sum_{R\in\sigma} m^{\mathrm{arity}(R)}$ binary variables (encoding the possible σ-structures with universe $\{1,\dots,m\}$). The action of S_m on the set of variables $\{X_1,\dots,X_n\}$ induces a

[1] These questions have been asked previously concerning, e.g., the action of the symmetric group S_n on n-invariable circuits. For S_n-invariant Boolean functions (a.k.a. symmetric functions) including PARITY and MAJORITY, there is known to be an exponential gap between U-invariant and non-invariant circuit and formula size. (See the Related Work section, below.)

[2] This follows from the observation that every [syntactically U-invariant] depth $d+1$ circuit of size s is equivalent to a [syntactically U-invariant] depth $d+1$ formula of size at most s^d.

syntactic action of S_m on the set of n-variable Boolean circuits (and many other concrete models of computation, such as branching programs, etc.)

An early result in this area, due to Denenberg et al [8], shows that syntactically 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. A decade later, Otto [13] introduced a certain limit object of finite circuits (also viewed as a form of uniformity) and showed a correspondence between infinitary logic with a bounded number of variables ($L^\omega_{\infty\omega}$) and syntactically S_m-invariant circuits of polynomial size and arbitrary depth. Otto also gives characterizations of fixed point and partial fixed point logic in terms of syntactically S_m-invariant networks. More recently, Anderson and Dawar [2] showed a correspondence (under a different uniformity condition) between fixed-point logic (FP) and syntactically S_m-invariant polynomial-size circuits, as well between fixed-point logic with counting (FPC) and syntactically S_m-invariant polynomial-size circuits in the basis that includes majority gates.

So far as I know, this paper is the first to study syntactic invariance under the action of linear subspaces of $\{0,1\}^n$ (i.e. subgroups on $\mathbb{Z}_2^n$) on n-variable Boolean circuits. A different notion of syntactic invariance — with respect to the automorphism group of the input structure — can be found in the literature on Choiceless Polynomial Time [3, 4, 6, 7, 9, 15]. On S_m-invariant tautologies in proof complexity, see [1, 14].

2 Preliminaries

Let $\mathbb{N} = \{0, 1, 2, \dots\}$. Let n and d be arbitrary positive integers. Let $[n] = \{1, \dots, n\}$.

Our lower bound makes use of the following inequality involving the function $n \mapsto dn^{1/d}$:

▶ **Lemma 2.** *For all real numbers $a, b, c > 0$, we have*

$$a + c(b/a)^{1/c} \geq (c+1)b^{1/(c+1)}$$

with equality iff $a = b^{1/(c+1)}$.

Proof. We have $\frac{\partial}{\partial a}\left(a + c(b/a)^{1/c}\right) = 1 - (b/a^{(c+1)})^{1/c}$. Thus, the function $a \mapsto a + c(b/a)^{1/c}$ is seen to have a unique minimum at $a = b^{1/(c+1)}$ where it takes value $(c+1)b^{1/(c+1)}$. ◀

2.1 Linear Algebra

For $x, y \in \{0,1\}^n$, we write $|x| := \sum_{i=1}^n x_i$ for the Hamming weight of x, we write $x \oplus y$ for the bitwise sum of x and y modulo 2 (i.e. the element $z \in \{0,1\}^n$ with $z_i := x_i \oplus y_i$), and we write $\langle x, y\rangle := \bigoplus_{i=1}^n x_i y_i$ for the inner product of x and y.

We write $\mathcal{L}$ for the lattice of linear subspaces of $\{0,1\}^n$. For $U, V \in \mathcal{L}$, we write $\dim(V)$ for the dimension of V, we write $V^\perp := \{x \in \{0,1\}^n : \langle x, v\rangle = 0 \text{ for all } v \in V\}$ for the orthogonal complement of V, and we write $U + V$ for the subspace spanned by U and V. We say that U is a *codimension-k subspace* of V if $U \subseteq V$ and $\dim(V) - \dim(U) = k$.

The orthogonal complement has the following properties:

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

2.2 $\mathbf{AC^0}$ 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 as follows:[3]

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

Every $F \in \mathcal{F}$ computes a Boolean function $\{0,1\}^n \to \{0,1\}$, defined 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 $F(S) \equiv b$ is the assertion 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 minimum d such that $F \in \mathcal{F}_d$. The *leafsize* of a formula is the number of depth-0 subformulas. Let *size* of a formula refer to the number of depth-1 subformulas. Inductively,

$$\mathrm{leafsize}(F) = \begin{cases} 0 & \text{if } F \in \mathcal{F}_0, \\ \sum_{G\in\mathcal{G}} \mathrm{size}(G) & \text{if } F = (\gamma,\mathcal{G}) \in \mathcal{F} \setminus \mathcal{F}_0, \end{cases}$$

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

Clearly $\mathrm{size}(F) \le \mathrm{leafsize}(F)$. (Note that size is within a factor 2 of the number of gates in F, which is how one usually measures size of circuits.) We define these two complexity measures since our lower bound naturally applies to size, while the upper bounds are naturally stated in terms of leafsize.

2.3 The Action of $\{0,1\}^n$

We define a group action of $\{0,1\}^n$ on $\mathcal{F}$ as follows. 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

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

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

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

- *syntactically U-invariant* if $F^u = F$ for every $u \in U$,
- *semantically U-invariant* if $F(x) = F(x \oplus u)$ for every $u \in U$ and $x \in \{0,1\}^n$.

As remarked in Section 1, syntactic U-invariance implies semantic U-invariance (but not conversely).

[3] As a minor convenience, we do not include constants 0 and 1 in $\mathcal{F}_0$, nor do we allow identical sibling subformulas (i.e. multisets $\mathcal{G}$) in the definition of $\mathcal{F}_{d+1}$. This is without loss of generality: the *depth-d formula size* of a Boolean function is unaffected by these restrictions.

2.4 Upper Bound

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

▶ **Proposition 3.** *For all $d, n \geq 1$, the n-variable* PARITY *function is computable by syntactically P-invariant depth $d+1$ formulas 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 $k, n_1, \dots, n_k \geq 1$ with $n_1 + \dots + n_k = n$, we construct a syntactically 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 syntactically P_{n_i}-invariant depth d formulas for PARITY$_{n_i}$ (or $1 -$ PARITY$_{n_i}$) with output gate AND (resp. OR). The minimum leafsize $\beta(d+1, n)$ achievable by this construction is given by the recurrence

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

We now observe:

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

An aside

I suspect that, by analyzing the above recurrence more carefully, the upper bound in Proposition 3 can be improved to $O(n{\cdot}2^{d(n^{1/d}-1)})$ for all $d \leq \lceil \log n \rceil$. This is suggested by the observation that PARITY is computable by syntactically P-invariant formulas of depth $\lceil \log n \rceil + 1$ and leafsize $O(n^2)$. Note that the upper bound of Proposition 3 is slack (except when $n^{1/d}$ is an integer), since setting $d = \log n$, we have $n{\cdot}2^{d(n^{1/d}-1)} = n^2$ and $n{\cdot}2^{dn^{1/d}} = n^3$. Also note that $O(n{\cdot}2^{d(n^{1/d}-1)})$ is *not* an upper bound for $d \gg \log n$, since $\Omega(n^2)$ is lower bound even for non-invariant formulas of unbounded depth [12].

3 Linear-Algebraic Lemmas

In this section, we prove a linear-algebraic lemma (Lemma 9) which plays a key role in our lower bound. Recall that S, T, U, V range over the set of linear subspaces of $\{0,1\}^n$, denoted by $\mathcal{L}$.

▶ **Definition 4.** For linear spaces $U \subseteq V$, a *linear retraction* from V to U is a linear function $\rho : V \to U$ such that $\rho(u) = u$ for every $u \in U$.

We next give a small lemma on the existence of linear retractions with small (one-sided) Hamming-weight distortion.

▶ **Lemma 5.** *If U is a codimension-k subspace of V, then there exists a linear retraction $\rho : V \to U$ such that $|\rho(v)|/|v| \leq k+1$ for all $v \in V$.*

Proof. Greedily choose a basis $w_1, \dots, w_k$ for V over U such that w_i has minimal Hamming weight among elements of $V \setminus \text{Span}(U \cup \{w_1, \dots, w_{i-1}\})$ 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 $a_1, \dots, a_k \in \{0,1\}$. Let $\rho : V \to U$ be the map $v \mapsto u$ and observe that this is a linear retraction.

To show that $|\rho(v)| \leq (k+1)|v|$, we first notice 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 $v \in V \setminus \mathrm{Span}(U \cup \{w_1, \dots, w_{i-1}\})$, so by our choice of w_i we have $|a_i w_i| = |w_i| \leq |v|$. Completing the proof, we have

$$\begin{aligned}|\rho(v)| &= |v \oplus a_1 v_1 \oplus \cdots \oplus a_k v_k|\\ &\leq |v| + |a_1 v_1| + \cdots + |a_k v_k|\\ &\leq (k+1)|v|.\end{aligned}$$ ◀

▶ **Definition 6.** 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 orthogonal complementation.

▶ **Lemma 7.** *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. This follows from the properties of the orthogonal complement listed in §2.1. 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 $U^\perp + T^\perp = (T \cap U)^\perp = S^\perp$. ◀

▶ **Lemma 8.** *For all $S \subset T \subseteq V$ such that $(S,T) \in \mathcal{L}_2$, 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 5, there exists a linear retraction $\rho : V \to T$ such that $|\rho(v)|/|v| \leq \dim(V) - \dim(T) + 1$ for all $v \in V$. Let $U = \rho^{-1}(S)$ and note that U is a codimension-1 subspace of V. (This follows from applying the Rank-Nullity Theorem to linear functions $\rho : V \to T$ 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 9.** *For all $S \subseteq U \subset V$ such that $(U,V) \in \mathcal{L}_2$, 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 7 and 8. ◀

4 Proof of Theorem 1

The following lemma gives the base case of Theorem 1 for depth-2 formulas (a.k.a. DNFs and CNFs). In this case, we merely require the hypothesis of semantic rather than syntactic U-invariance. The proof is similar to the standard argument showing that depth-2 formulas for PARITY require 2^{n-1} clauses of width n.

▶ **Lemma 10.** *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 $m = \min\{|x| : x \in U^\perp \setminus V^\perp\}$.*

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 further assume that F is minimal with respect to the number of clauses and the number of literals in any particular 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 map $\{0,1\}^n \to \{0,1\}^\ell$. Then there is a point $p \in \{0,1\}^\ell$ 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 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 $p^{(1)}, \dots, p^{(\ell)} \in \pi(U)$. 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 $u_{\{1,\dots,\ell-1\}} = p_{\{1,\dots,\ell-1\}}$. 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 $p^{(1)}, \dots, p^{(\ell)} \in \pi(U)$ 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 (with respect to the number of clauses and, secondarily, to the width of each clause), we conclude that F contains $\geq 2^{m-1}$ clauses, each of width $\geq m$. Therefore, $\mathrm{size}(F) \geq 2^{m-1}$ and $\mathrm{leafsize}(F) \geq m{\cdot}2^{m-1}$. ◀

On to our main result:

▶ **Theorem 1 (restated).** *Let $U \subset V$ be linear subspaces of $\{0,1\}^n$, and suppose F is a syntactically 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 $U \subset V$ where F is syntactically 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 10.[4] For the induction step, let $d \geq 2$ and assume $F \in \mathcal{F}_{d+1}$ is a syntactically 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 syntactically 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 syntactically 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

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

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 syntactically 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 9, 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.

[4] Note that Theorem 1 makes sense even when $d = 0$, if we interpret $0 \cdot (m^{1/0} - 1)$ as 0 if $m = 1$ and ∞ if $m > 1$.

- **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 syntactically 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 2 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 Further Remarks and Open Questions

5.1 Another Application of Theorem 1

Theorem 1 applies to interesting subspaces U besides the even-weight subspace P. Here we describe one example. Let G be a simple graph with n edges, so that $\{0,1\}^n$ is identified with the set of spanning subgraphs of G. The *cycle space* of G is the linear 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 G is non-bipartite, Z_0 is a codimension-1 subspace of Z.

Let $m = \min\{|z| : z \in Z_0^\perp \setminus Z^\perp\}$ as in Theorem 1 with $U = Z_0$ and $V = Z$. It is not hard to show that m is 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.) Moreover, if G is a uniform random 3-regular graph on $\frac{2}{3}n$ vertices, then $m = \Omega(n)$ asymptotically almost surely [5]. By these observations, we have:

▶ **Corollary 11.** *Let $Z \subseteq \{0,1\}^n$ be the cycle space of a random 3-regular graph with n edges, and let $Z_0 = \{z \in Z : |z|$ is even$\}$. Then a.a.s. every syntactically Z_0-invariant depth $d+1$ formula that computes* PARITY_n *over Z has size $2^{d(\Omega(n)^{1/d}-1)}$.*

5.2 The (U, V)-Search Problem

For linear subspaces $U \subset V$ of $\{0,1\}^n$, consider the following (U,V)*-search problem*: there is a hidden vector $x \in V \setminus U$ and the goal is to learn a nonzero coordinate of x (i.e. any $i \in [n]$ such that $w_i = 1$) by asking queries (i.e. 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 protocols that solve this problem on all $w \in V \setminus U$ by asking queries in d consecutive batches (thus, 1-round = non-adaptive). By a slightly simpler version of the argument in the proof of Theorem 1, we can show a $d(m^{1/d} - 1)$ lower bound on the d-round query complexity of the (U, V)-search problem for all $U \subset V$ where $m = \min\{|x| : x \in U^\perp \setminus V^\perp\}$.

This (U, V)-search problem is, in some sense, related to the Karchmer-Wigderson game where Alice gets $u \in U$ and Bob gets $v \in V \setminus U$ and their common goal is to learn a nonzero coordinate of $u \oplus v$. For an appropriate definition of "U-invariant protocols" (i.e. whatever comes from syntactically U-invariant formulas), we can translate the pair (u, v) to $(0, u \oplus v)$ without loss of generality and it becomes Alice's task to learn a nonzero coordinate of $u \oplus v$ by asking linear queries.

5.3 Open Questions

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

1. Does the lower bound of Theorem 1 (or something weaker like $2^{m^{\Omega(1/d)}}$) hold under the weaker assumption of semantic U-invariance, in place of syntactic U-invariance? What about Corollary 11?
2. Considering leafsize (rather than size, i.e. the number of depth-1 subformulas), 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 3 from $n{\cdot}2^{dn^{1/d}}$ to $O(n{\cdot}2^{d(n^{1/d}-1)})$ for all $d \leq \lceil \log n \rceil$.
4. What is the maximum gap, if any, between U-invariant [depth d] formula size and non-invariant [depth d] formula size?

Acknowledgements. I thank the anonymous referees for their helpful comments.

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 Anuj Dawar, David Richerby, and Benjamin Rossman. Choiceless polynomial time, counting and the Cai-Fürer-Immerman graphs. *Annals of Pure and Applied Logic*, 152:31–50, 2008.

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

9 Erich Grädel and Martin Grohe. Is polynomial time choiceless? In *Fields of Logic and Computation II*, pages 193–209. Springer, 2015.

10 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.

11 Neil Immerman. *Descriptive complexity*. Springer, 2012.

12 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.

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

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.