

On the Constant-Depth Complexity of k -Clique

Benjamin Rossman^{*}
Massachusetts Institute of Technology
32 Vassar St., Cambridge, MA 02139
brossman@theory.csail.mit.edu

ABSTRACT

We prove a lower bound of $\omega(n^{k/4})$ on the size of constant-depth circuits solving the k -clique problem on n -vertex graphs (for every constant k). This improves a lower bound of $\omega(n^{k/89d^2})$ due to Beame where d is the circuit depth. Our lower bound has the advantage that it does not depend on the constant d in the exponent of n , thus breaking the mold of the traditional size-depth tradeoff.

Our k -clique lower bound derives from a stronger result of independent interest. Suppose $f_n : \{0, 1\}^{\binom{n}{2}} \rightarrow \{0, 1\}$ is a sequence of functions computed by constant-depth circuits of size $O(n^t)$. Let G be an Erdős-Rényi random graph with vertex set $\{1, \dots, n\}$ and independent edge probabilities $n^{-\alpha}$ where $\alpha \leq \frac{1}{2t-1}$. Let A be a uniform random k -element subset of $\{1, \dots, n\}$ (where k is any constant independent of n) and let K_A denote the clique supported on A . We prove that $f_n(G) = f_n(G \cup K_A)$ asymptotically almost surely.

These results resolve a long-standing open question in finite model theory (going back at least to Immerman in 1982). The m -variable fragment of first-order logic, denoted by FO^m , consists of the first-order sentences which involve at most m variables. Our results imply that the *bounded variable hierarchy* $\text{FO}^1 \subset \text{FO}^2 \subset \dots \subset \text{FO}^m \subset \dots$ is strict in terms of expressive power on finite ordered graphs. It was previously unknown that FO^3 is less expressive than full first-order logic on finite ordered graphs.

Categories and Subject Descriptors

F.2.2 [Theory of Computation]: Nonnumerical Algorithms and Problems—*Computations on discrete structures*;
F.4.1 [Mathematical Logic and Formal Languages]:
Mathematical Logic—*Model theory*

^{*}Supported by a National Defense Science and Engineering Graduate Fellowship.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

Copyright 200X ACM X-XXXXX-XX-X/XX/XX ...\$5.00.

General Terms

Theory

Keywords

k -clique, constant-depth circuits, circuit complexity, AC^0 , first-order logic, bounded variable hierarchy

1. INTRODUCTION

Constant-depth circuits have been one of the most fruitful settings for complexity lower bounds. (Here we refer to logical circuits comprised of \neg , \wedge and \vee gates with unbounded fan-in.) Early results of Ajtai [1] and Furst, Saxe and Sipser [12] proved that the parity problem (given an n -bit string, are there an even number of 1's?) does not have polynomial-size constant-depth circuits. This result, often stated as $\text{PARITY} \notin \text{AC}^0$, was subsequently sharpened by Yao [28] and Håstad [15] among others, who eventually established that exponential-size constant-depth circuits are required.

In this paper we consider the k -clique problem for constant values of k . The k -clique problem asks, given a graph G on n vertices (specified by a bitstring of length $\binom{n}{2}$), does G contain a k -clique? Unlike the parity problem, the k -clique problem has polynomial-size constant-depth circuits (for every constant k). In particular, there is an obvious depth-2 circuit consisting of $\binom{n}{k}$ \wedge -gates at the bottom level (deciding each potential k -clique) feeding into a single \vee -gate on top. A natural question is whether there exist significantly smaller constant-depth circuits for the k -clique problem (say, of size $n^{o(k)}$).

Lower bounds on the complexity of problems within AC^0 traditionally involve *size-depth tradeoffs*. Lynch [23] in 1986 established the first size-depth tradeoff for the k -clique problem by proving a lower bound of $n^{\Omega(\sqrt{k}/d^{1.5})}$ on the size of depth- d circuits solving the k -clique problem on n -vertex graphs. This was improved to $\omega(n^{k/89d^2})$ by Beame [7] in 1990 (in the context of CRCW PRAM's). These results were based on developments in the *switching lemma* technology originally introduced by Håstad [15] for studying the parity problem. For his lower bound, Beame proved powerful switching lemmas in the setting of random graphs (see [8]).

Size-depth tradeoffs of this sort, however, are unsatisfactory inasmuch as they degrade in the exponent of n as d is taken to be a larger constant. (This seems to be an unavoidable consequence of using switching lemmas in the usual way.) For depth $d = \sqrt{k}$, the lower bounds of Lynch and Beame no longer beat the trivial lower bound of $\binom{n}{2}$ (the

number of input bits). These lower bounds thus do not rule out the possibility that the 100-clique problem has depth-10 circuits of size $O(n^2)$.

In this paper we create a more satisfactory state of affairs by proving a lower bound of $\omega(n^{k/4})$ on the size of depth- d circuits for the k -clique problem for all constants k and d . (In fact, our lower bound holds even for slightly increasing $d = d(n) = o(\sqrt{\log n})$ and perhaps moderately increasing $k = k(n)$ as well.) By contrast, our lower bound implies that 100-clique does not have size $O(n^{25})$ circuits of any constant depth. Thus it may now be said that, in some reasonable sense, no constant-depth circuits for the k -clique problem significantly beat the naive circuits of depth 2 and size $O(n^k)$.

Unlike traditional size-depth tradeoffs, our approach does not involve developing sharper switching lemmas for the problem at hand. In fact, we require nothing stronger than Håstad's original switching lemma. Rather, our technique breaks from past approaches in the innovative way that the switching lemma is used. A key technical notion we introduce is the s -bounded clique-sensitive core $\mathbb{T}_{(s)}^{f,G}(A)$ of a set A of vertices in a graph G with respect to a graph-function f (see §3.1). $\mathbb{T}_{(s)}^{f,G}(A)$ is a subset of A with certain nice properties. In particular, if we add to G the clique supported on $\mathbb{T}_{(s)}^{f,G}(A)$, then the value of f on the resulting graph is no longer sensitive to the further addition of any subclique of A up to size s . This new technical notion (which puts a twist on the familiar concept of *sensitive inputs*) allows for a novel inductive argument on circuits (Lemma 3.6). Ultimately, our approach leads to an effective new way of using the switching lemma in conjunction with a union bound over the nodes in a circuit.

We remark that the complexity of the k -clique problem has been studied in various models of computation besides constant-depth circuits. In the setting of *monotone circuits*, it is known that the k -clique problem (for specific increasing $k = k(n)$) requires exponential size [3, 13, 26]. A super-polynomial lower bound was even proved for circuits with a bounded number of negation gates [5]. Lower bounds have also been investigated in the context of branching programs and decision trees [27].

Of course, the greatest hope is eventually for a lower bound of $n^{\Omega(k)}$ on the size of circuits of *arbitrary depth* for the k -clique problem. Such a result would imply $\text{NP} \not\subseteq \text{P/poly}$ and hence $\text{P} \neq \text{NP}$.

Organization of the Paper.

In the rest of this section, we state our main results and discuss some interesting corollaries in logic. §2 fixes some notation and covers the preliminaries on graphs, random graphs and circuits. In §3 we present the technique behind our lower bound and prove weaker versions of our main theorems (modulo some technical lemmas). In §4 we prove those technical lemmas (concerning the sensitivity of randomly restricted AC^0 -computable functions). In §5 we prove our main theorems using the technique developed in §3. We state some conclusions and raise some open questions in §6.

1.1 Main Results

THEOREM 1.1. *Suppose $f_n : \{0, 1\}^{\binom{n}{2}} \rightarrow \{0, 1\}$ is a sequence of functions computed by constant-depth circuits of size $O(n^t)$ where $t > 1/2$. For any constants $k \in \mathbb{N}$ and*

$0 < \alpha \leq \frac{1}{2t-1}$, let $G = \text{ER}(n, n^{-\alpha})$ be an Erdős-Rényi random graph and let A be a uniform random set of k vertices of G . Then $f_n(G) = f_n(G \cup K_A)$ asymptotically almost surely.

This result directly implies our k -clique lower bound.

THEOREM 1.2. *For every constant k , the k -clique problem on n -vertex graphs requires constant-depth circuits of size $\omega(n^{k/4})$.*

In fact, Theorem 1.1 implies the even stronger assertion that for all $\ell > k$, no constant-depth circuits of size $O(n^{k/4})$ distinguish between the class of graphs which contain an ℓ -clique and the class of graphs which contain no k -clique.

1.2 Corollaries in Logic

Our k -clique lower bound has some nice corollaries in logic (finite model theory [18, 22]), which answer long-standing open questions, via the well-known descriptive complexity characterization of first-order logic in terms of the circuit class AC^0 [6, 11, 21].

The m -variable fragment of first-order logic, denoted by FO^m , consists of the first-order sentences which involve at most m variables. (Sentences of FO^m may contain more than m quantifiers, as variables may be reused.) For example, the following sentence (in the language of simple graphs with a symmetric binary relation \sim denoting adjacency) expresses “diameter is at most 4” using only 3 variables:

$$\forall x \forall y \exists z (\exists y (x \sim y) \wedge (y \sim z)) \wedge (\exists x (z \sim x) \wedge (x \sim y)).$$

More generally, 3 variables suffice to express “diameter is at most d ” for every constant d . Bounded variable fragments FO^m (and their infinitary counterparts $\mathcal{L}_{\infty, \omega}^m$) are important objects of study in both finite and classical model theory (see [14, 18]). The chain of m -variable fragments $\text{FO}^1 \subset \text{FO}^2 \subset \dots \subset \text{FO}^m \subset \dots$ is known as the *bounded (or finite) variable hierarchy*. On finite graphs without a linear order, this hierarchy is strict in terms of expressive power for the simplest of reasons: the sentence “there exist at least m vertices” is expressible in FO^m but not in FO^{m-1} .

Here, however, we are concerned with first-order logic on finite ordered graphs (where sentences may speak of order $<$ as well as adjacency \sim among elements). Exploiting order, we are now able to express “there exist at least m vertices” for every m using only 3 variables (similar to the diameter example above). Perhaps, then, every first-order sentence is equivalent on finite ordered graphs to a sentence with only 3 variables? It far from obvious how one might express “there exists a 4-clique” using only 3 variables. On the other hand, it is unclear how to prove that 3 variables do not suffice for the task. (Ehrenfeucht-Fraïssé games, the standard method for arguing inexpressibility in fragments of first-order logic, are greatly complicated by the presence of a linear order.) Indeed, it was (until now) an open problem whether, in terms of expressive power on finite ordered graphs, the bounded variable hierarchy is strict, or whether perhaps it collapses down to FO^3 . (It is long known that FO^2 is weaker than FO^3 .) This intriguing question was raised as early as 1982 by Immerman [16]. A survey article by Dawar [10] devoted to this very question was published in 2005. (Concerning some variations on this question: Poizat [25] showed that 3 variables suffice to express all first-order properties of finite colored linear orders, while Dawar [10] showed that the

hierarchy of m -variable existential sentences is expressively strict on finite ordered graphs.)

Our k -clique lower bound finally resolves the status of the bounded variable hierarchy on finite ordered graphs. The bridge between constant-depth circuits and bounded variable fragments of first-order logic is the following well-known result from descriptive complexity theory (discovered by various authors [6, 11] and recently refined [21]).

PROPOSITION 1.3. *For every m -variable first-order sentence Φ in the language of ordered graphs, there exist constant-depth circuits C_n , $n \in \mathbb{N}$, of size $O(n^m)$ on $\binom{n}{2}$ inputs such that C_n evaluates Φ on ordered graphs of size n .*

Theorem 1.2 and Proposition 1.3 directly imply:

COROLLARY 1.4. *No sentence in $\text{FO}^{\lfloor k/4 \rfloor}$ expresses the existence of a k -clique on finite ordered graphs.*

(In fact, Corollary 1.4 holds not just for the class of finite ordered graphs, but for classes of finite graphs with arbitrary numerical predicates; see the full paper.) Since one can express the existence of an m -clique in FO^m , namely by the sentence $\exists x_1 \dots \exists x_m \bigwedge_{i < j} (x_i \neq x_j) \wedge (x_i \sim x_j)$, we get:

COROLLARY 1.5. *The bounded variable hierarchy in non-collapsing in terms of expressive power on finite ordered graphs.*

Immerman [personal communication] has pointed out that Corollary 1.5 in fact implies strictness of the bounded variable hierarchy. Immerman's observation is that the expressive collapse of FO^{m+1} to FO^m on finite ordered graphs would imply the collapse of FO^{m+2} to FO^{m+1} and hence the full expressive collapse of FO to FO^m . The proof of this fact (included in the full paper) uses a clever Ehrenfeucht-Fraïssé game argument.

PROPOSITION 1.6 (IMMERMAN). *If FO^m and FO^{m+1} are equally expressive on finite ordered graphs, then so are FO^{m+1} and FO^{m+2} .*

Corollary 1.5 and Proposition 1.6 together yield:

COROLLARY 1.7. *The bounded variable hierarchy is strict in terms of expressive power on finite ordered graphs.*

2. PRELIMINARIES

Basic Notation.

For a positive integer n , let $[n]$ denote the set $\{1, \dots, n\}$. Letters A, B, C, D are reserved for subsets of $[n]$. For a set X and $k \in \mathbb{N}$, let $\binom{X}{k} = \{Y \subseteq X : |Y| = k\}$ and $\binom{X}{\leq k} = \bigcup_{j=0}^k \binom{X}{j}$. The abbreviation ‘‘a.a.s.’’ stands for ‘‘asymptotically almost surely (as $n \rightarrow \infty$)’’. Logarithms have base 2 unless otherwise indicated.

Graphs.

Graphs are by default finite simple graphs. Formally, a graph G is a pair (V_G, E_G) where V_G is a nonempty finite set and E_G is a subset of $\binom{V_G}{2}$. The quantity $\min_{H \subseteq G} |V_H|/|E_H|$ (where H ranges over induced subgraphs of G) is denoted by $\text{thres}(G)$ and called the *threshold exponent* of G (Lemma 2.1 explains why); we set $\text{thres}(G) = \infty$ in the event that

$|E_G| = \emptyset$. For example, if K_k is the complete graph on k vertices, then $\text{thres}(K_k) = \frac{2}{k-1}$.

The class of graphs with vertex set $[n]$ is denoted by \mathcal{G}_n . We identify \mathcal{G}_n with the set $\{0, 1\}^{\binom{n}{2}}$, particularly when viewing graphs as inputs to circuits. For a set $A \subseteq [n]$, we denote by K_A the clique supported on A (i.e., the complete graph with $V_{K_A} = A$ and $E_{K_A} = \binom{A}{2}$). For graphs $G \in \mathcal{G}_n$, we will often consider the union graph $G \cup K_A$ (i.e., G plus a clique on A).

Random Graphs.

A *random graph* is a probability distribution on \mathcal{G}_n (or a sequence of such distributions for each n). For $q \in [0, 1]$, we denote by $\text{ER}(n, q)$ the *Erdős-Rényi random graph* with vertex set $[n]$ in which each potential edge in $\binom{[n]}{2}$ is independently included with probability q . We write $G \in \text{ER}(n, q)$ to express that G is a random graph with distribution $\text{ER}(n, q)$. We are mainly interested in Erdős-Rényi random graphs with edge probability $q = n^{-\alpha}$ for constant values of $\alpha > 0$ (see [4] for background).

Given two graphs G and H , the number of induced subgraphs of G isomorphic to H is denoted by $\text{sub}(G, H)$. For a fixed graph H , taking $G = \text{ER}(n, n^{-\alpha})$, the following lemma (proved using Janson's inequality [19]) gives an asymptotic bound on the lower tail of the random variable $\text{sub}(G, H)$ as $n \rightarrow \infty$; we remark that stronger results are available. Note that $\text{sub}(G, H)$ has expectation approximately $n^{|V_H| - \alpha|E_H|}$, which is > 1 iff $\alpha < \text{thres}(H)$.

LEMMA 2.1 ([20]). *For every graph H and $\alpha > 0$, the following asymptotic bounds hold for random graphs $G \in \text{ER}(n, n^{-\alpha})$ as n grows to infinity.*

1. *If $\alpha > \text{thres}(H)$ then*

$$\Pr[\text{sub}(G, H) > 0] = \exp(-n^{\Omega(1)}).$$

2. *If $\alpha < \text{thres}(H)$ then for all $\varepsilon > 0$,*

$$\Pr[\text{sub}(G, H) < n^{|V_H| - \alpha|E_H| - \varepsilon}] = \exp(-n^{\Omega(1)}).$$

Circuits.

A *circuit* on n inputs (resp. $\binom{n}{2}$ inputs) is an acyclic directed graph C in which sources, called *input nodes*, are labeled by elements of $[n]$ (resp. $\binom{[n]}{2}$) and non-source nodes, called *gates*, are labeled by elements of $\{\neg, \wedge, \vee\}$. Gates labeled \neg are required to have fan-in (in-degree) 1, while \wedge and \vee gates have unrestricted fan-in. A subset of nodes in C (by default: the sinks in C) are designated as *output nodes*. The *size* and *depth* of C are respectively the number of gates it contains and the maximum number of gates on a path from an input node to an output node. The value of a circuit C on an input string $x \in \{0, 1\}^n$ (resp. an input graph $G \in \mathcal{G}_n$) is denoted by $C(x)$ (resp. $C(G)$). If C has m output nodes, then $C(x)$ is an element of $\{0, 1\}^m$. Thus, a circuit on n inputs (resp. $\binom{n}{2}$ inputs) and m outputs computes a function from $\{0, 1\}^n$ (resp. \mathcal{G}_n) to $\{0, 1\}^m$.

AC^0 .

(Non-uniform) AC^0 is the class of languages $\mathcal{L} \subseteq \{0, 1\}^*$ decided by a sequence of circuits (one for each input size n) of constant depth and size polynomial in n .

3. OUR TECHNIQUE

In this section we prove weaker versions of our main theorems (with slightly inferior parameters) modulo some technical lemmas on random restrictions (which we tackle later on in §4). This section illustrates our lower-bound technique in a somewhat simplified form. We eventually prove our main theorems in §5 using the technique introduced here.

3.1 s-Bounded Clique Sensitivity

We define a technical notion: the *clique-sensitive core* of a set of vertices in a given graph with respect to a given graph-function. This is closely related to the familiar notion of *sensitive vertices* of a graph-function (as we will see in §4.1). We also introduce the a more novel notion: the *s-bounded clique-sensitive core* (for a parameter $s \in \mathbb{N}$). This notion plays a central role in our technique.

DEFINITION 3.1. *Let f be a graph-function with domain \mathcal{G}_n (and arbitrary range) and let $G \in \mathcal{G}_n$. For all $A \subseteq [n]$ and $s \in \mathbb{N}$, sets $\mathbb{T}^{f,G}(A)$ and $\mathbb{T}_{(s)}^{f,G}(A)$ are defined by*

$$\mathbb{T}^{f,G}(A) = \left\{ a \in A : \begin{array}{l} \text{there exists } B \subseteq A \text{ such that} \\ f(G \cup K_B) \neq f(G \cup K_{B \setminus \{a\}}) \end{array} \right\},$$

$$\mathbb{T}_{(s)}^{f,G}(A) = \bigcup_{B \subseteq A : |B| \leq s} \mathbb{T}^{f,G}(B).$$

The set $\mathbb{T}^{f,G}(A)$ (resp. $\mathbb{T}_{(s)}^{f,G}(A)$) is called the (resp. s-bounded) *clique-sensitive core* of A under f in G . We say that A is *fully clique-sensitive* under f in G if $\mathbb{T}^{f,G}(A) = A$.

Intuitively, $\mathbb{T}^{f,G}(A)$ (resp. $\mathbb{T}_{(s)}^{f,G}(A)$) is the set of vertices $a \in A$ such that the graph-function f , when evaluated on G plus different subcliques of A (resp. of size $\leq s$), is sensitive to the inclusion of a in the added subclique. A simple example serves to illustrate this definition.

EXAMPLE 3.2. *Let $\text{clique}_k : \mathcal{G}_n \rightarrow \{0, 1\}$ be the function defined by $\text{clique}_k(G) = 1 \iff G$ contains a k -clique. Let G_{emp} denote the empty graph in \mathcal{G}_n (with no edges). For all $A \subseteq [n]$ and $s \in \mathbb{N}$, we have*

$$\mathbb{T}^{\text{clique}_k, G_{\text{emp}}}(A) = \begin{cases} A & \text{if } |A| \geq k, \\ \emptyset & \text{otherwise,} \end{cases}$$

$$\mathbb{T}_{(s)}^{\text{clique}_k, G_{\text{emp}}}(A) = \begin{cases} A & \text{if } |A| \geq k \text{ and } s \geq k, \\ \emptyset & \text{otherwise.} \end{cases}$$

Some elementary properties of operators $\mathbb{T}^{f,G}$ and $\mathbb{T}_{(s)}^{f,G}$ are stated below.

- (i) $\mathbb{T}_{(s)}^{f,G}(A) \subseteq \mathbb{T}^{f,G}(A) \subseteq A$ for all sets A .
- (ii) $\mathbb{T}^{f,G}$ and $\mathbb{T}_{(s)}^{f,G}$ are monotone. That is, $A \subseteq B$ implies $\mathbb{T}^{f,G}(A) \subseteq \mathbb{T}^{f,G}(B)$ and $\mathbb{T}_{(s)}^{f,G}(A) \subseteq \mathbb{T}_{(s)}^{f,G}(B)$.
- (iii) If $f : \mathcal{G}_n \rightarrow \{0, 1\}^m$ where $f_1, \dots, f_m : \mathcal{G}_n \rightarrow \{0, 1\}$ are the individual coordinate-functions of f , then $\mathbb{T}^{f,G}(A) = \bigcup_{i=1}^m \mathbb{T}^{f_i, G}(A)$ and $\mathbb{T}_{(s)}^{f,G}(A) = \bigcup_{i=1}^m \mathbb{T}_{(s)}^{f_i, G}(A)$.
- (iv) If sets A and B are both fully clique-sensitive under f in G , then so is their union $A \cup B$.

The next three lemmas give some additional properties of operators $\mathbb{T}^{f,G}$ and $\mathbb{T}_{(s)}^{f,G}$.

LEMMA 3.3. *Let $T = \mathbb{T}_{(s)}^{f,G}(A)$ and suppose B is a set such that $T \subseteq B \subseteq A$ and $|B| \leq s$. Then $f(G \cup K_T) = f(G \cup K_B)$.*

PROOF. Let b_1, \dots, b_m enumerate the set $B \setminus T$. For all $i \in \{1, \dots, m\}$, we have

$$f(G \cup K_{T \cup \{b_1, \dots, b_i\}}) = f(G \cup K_{T \cup \{b_1, \dots, b_{i-1}\}}).$$

(Otherwise, we would have $b_i \in \mathbb{T}^{f,G}(B) \subseteq T$.) It follows immediately that $f(G \cup K_T) = f(G \cup K_B)$. \square

It is of course possible that $|\mathbb{T}_{(s)}^{f,G}(A)| > s$, in which case the statement of Lemma 3.3 is vacuous.

LEMMA 3.4. *The set $\mathbb{T}_{(s)}^{f,G}(A)$ is the union of all fully clique-sensitive subsets of A of size at most s . That is,*

$$\mathbb{T}_{(s)}^{f,G}(A) = \bigcup \{ B \subseteq A : \mathbb{T}^{f,G}(B) = B \text{ and } |B| \leq s \}.$$

PROOF. The inclusion \supseteq is fairly obvious. To prove the opposite inclusion, we must find, for each $a \in \mathbb{T}_{(s)}^{f,G}(A)$, a fully-clique sensitive subset $B \subseteq A$ of size $\leq s$ which contains a . Because $a \in \mathbb{T}_{(s)}^{f,G}(A)$ we know there exists some subset $B \subseteq A$ such that $f(G \cup K_B) \neq f(G \cup K_{B \setminus \{a\}})$ and $|B| \leq s$. Take any minimal B satisfying these conditions. To show that B is fully clique-sensitive, we assume that there exists $b \in B \setminus \mathbb{T}^{f,G}(B)$ and derive a contradiction. First, we note that $f(G \cup K_{B \setminus \{b\}}) = f(G \cup K_{B \setminus \{a, b\}})$ by the minimality of our choice of B . However, the fact that $b \notin \mathbb{T}^{f,G}(B)$ implies

$$f(G \cup K_B) = f(G \cup K_{B \setminus \{b\}}),$$

$$f(G \cup K_{B \setminus \{a\}}) = f(G \cup K_{B \setminus \{a, b\}}).$$

But these three equalities imply $f(G \cup K_B) = f(G \cup K_{B \setminus \{a\}})$, contradicting our choice of B . Therefore, B is fully clique-sensitive. After noting that B contains a (since otherwise we again reach the contradiction $f(G \cup K_B) = f(G \cup K_{B \setminus \{a\}})$), we are done. \square

LEMMA 3.5.

1. $\mathbb{T}_{(s)}^{f,G}(A)$ is nonempty if, and only if, A has a fully clique-sensitive subset B of size $2 \leq |B| \leq s$.
2. $|\mathbb{T}_{(s)}^{f,G}(A)| > s$ if, and only if, A has fully clique-sensitive subsets B and C of size $\leq s$ such that $|B \cup C| \geq s + 1$.
3. If $|\mathbb{T}_{(s)}^{f,G}(A)| > s$, then A has a fully clique-sensitive subset D of size $s + 1 \leq |D| \leq 2s$.

PROOF. Implications (\Leftarrow) in both (1) and (2) are easy. We prove the implications (\Rightarrow) in statements (1)–(3). Let A_1, \dots, A_m , $m \in \mathbb{N}$, enumerate the nonempty fully clique-sensitive subsets of A of size $\leq s$. By Lemma 3.4, $\mathbb{T}_{(s)}^{f,G}(A) = A_1 \cup \dots \cup A_m$.

Suppose that $\mathbb{T}_{(s)}^{f,G}(A)$ is nonempty, that is, $m \geq 1$. Let $B = A_1$. We claim that $|B| \geq 2$. Toward a contradiction, assume $|B| < 2$. Then (since $B \neq \emptyset$) $B = \{b\}$ and hence $\mathbb{T}^{f,G}(\{b\}) = \{b\}$ for some $b \in A$. By definition of $\mathbb{T}^{f,G}$, there exists $C \subseteq \{b\}$ such that $f(G \cup K_C) \neq f(G \cup K_{C \setminus \{b\}})$. Clearly it must be that $C = \{b\}$. But then $f(G \cup K_{\{b\}}) \neq f(G \cup K_\emptyset)$, which is absurd since $K_{\{b\}}$ (the clique supported

on the single vertex b) has no edges and hence $G \cup K_{\{b\}} = G \cup K_\emptyset = G$. We conclude that $|B| \geq 2$, proving implication (\implies) of statement (1).

Now suppose that $|\mathbb{T}_{(s)}^{f,G}(A)| > s$. Let $\ell \in \{1, \dots, m\}$ be the least index such that $|A_1 \cup \dots \cup A_\ell| > s$, noting that ℓ is well-defined since $|A_1 \cup \dots \cup A_m| > s$. Let $B = A_1 \cup \dots \cup A_{\ell-1}$ and $C = A_\ell$. Clearly $|B|, |C| \leq s$ and $|B \cup C| \geq s + 1$. The set B is fully clique-sensitive (by observation (iv), as it is a union of fully clique-sensitive sets $A_1, \dots, A_{\ell-1}$), as is the set C . This proves implication (\implies) of statement (2). Statement (3) follows by taking $D = B \cup C$. \square

3.2 Circuit Lemma

Let C be a single-output circuit with $\binom{n}{2}$ inputs. For all nodes ν in C , we denote by C_ν the single-output subcircuit of C with output ν . We denote by C_ν^\bullet the circuit C_ν in which not only ν but also all of its children are designated as outputs. Thus C and C_ν compute functions $\mathcal{G}_n \rightarrow \{0, 1\}$, while C_ν^\bullet computes a function $\mathcal{G}_n \rightarrow \{0, 1\}^{\text{fan-in}(\nu)+1}$.

LEMMA 3.6 (CIRCUIT LEMMA). *Let $G \in \mathcal{G}_n$ and $A \subseteq [n]$ and $s \geq 1$. Suppose that $\mathbb{T}_{(s)}^{\mathsf{C},G}(A) = \emptyset$ and $|\mathbb{T}_{(s)}^{\mathsf{C}_\nu^\bullet,G}(A)| \leq s$ for all nodes ν in C . Then $\mathsf{C}(G) = \mathsf{C}(G \cup K_A)$.*

PROOF. To streamline notation, let $\mathcal{T}(\nu) = \mathbb{T}_{(s)}^{\mathsf{C}_\nu,G}(A)$ and $\mathcal{T}^\bullet(\nu) = \mathbb{T}_{(s)}^{\mathsf{C}_\nu^\bullet,G}(A)$ for all nodes ν in C . Note that

$$\mathcal{T}^\bullet(\nu) = \mathcal{T}(\nu) \cup \bigcup_{\text{children } \mu \text{ of } \nu} \mathcal{T}(\mu)$$

by observation (iii) in §3.1. We prove the following claim.

CLAIM 3.7. *For all nodes ν in C ,*

$$\mathsf{C}_\nu(G \cup K_{\mathcal{T}^\bullet(\nu)}) = \mathsf{C}_\nu(G \cup K_A).$$

PROOF OF CLAIM. The proof is by induction on the depth of ν . The claim is trivial when ν is an input node. Indeed, if ν is an input node corresponding to a potential edge $e \in \binom{[n]}{2}$, then (since $s \geq 1$) $\mathcal{T}^\bullet(\nu)$ is $\{e\}$ if $e \subseteq A$ and \emptyset if $e \not\subseteq A$. It follows that C_ν (the indicator function for edge e) has the same value on graphs $G \cup K_{\mathcal{T}^\bullet(\nu)}$ and $G \cup K_A$, as the edge e is either present in both graphs or absent from both graphs.

For the induction step, let ν be a non-input node and assume that $\mathsf{C}_\mu(G \cup K_{\mathcal{T}^\bullet(\mu)}) = \mathsf{C}_\mu(G \cup K_A)$ for all children μ of ν . Consider now a particular child μ of ν . Since $\mathcal{T}(\mu) \subseteq \mathcal{T}^\bullet(\nu) \subseteq A$ and $|\mathcal{T}^\bullet(\nu)| \leq s$, Lemma 3.3 implies

$$\mathsf{C}_\mu(G \cup K_{\mathcal{T}(\mu)}) = \mathsf{C}_\mu(G \cup K_{\mathcal{T}^\bullet(\nu)}).$$

Invoking Lemma 3.3 a second time, since $\mathcal{T}(\mu) \subseteq \mathcal{T}^\bullet(\mu) \subseteq A$ and $|\mathcal{T}^\bullet(\mu)| \leq s$,

$$\mathsf{C}_\mu(G \cup K_{\mathcal{T}(\mu)}) = \mathsf{C}_\mu(G \cup K_{\mathcal{T}^\bullet(\mu)}).$$

By the induction hypothesis,

$$\mathsf{C}_\mu(G \cup K_{\mathcal{T}^\bullet(\mu)}) = \mathsf{C}_\mu(G \cup K_A).$$

Therefore, $\mathsf{C}_\mu(G \cup K_{\mathcal{T}^\bullet(\nu)}) = \mathsf{C}_\mu(G \cup K_A)$. Because the value of C_ν on any input graph is determined by the values of C_μ on the various children μ of ν , we conclude that $\mathsf{C}_\nu(G \cup K_{\mathcal{T}^\bullet(\nu)}) = \mathsf{C}_\nu(G \cup K_A)$. \square_{CLAIM}

Let ν_{out} be the output node of C (so that $\mathsf{C}_{\nu_{\text{out}}} = \mathsf{C}$). To complete the proof of the lemma, we simply apply the

claim to ν_{out} . Thus $\mathsf{C}(G \cup K_{\mathcal{T}^\bullet(\nu_{\text{out}})}) = \mathsf{C}(G \cup K_A)$. But $\mathsf{C}(G \cup K_{\mathcal{T}(\nu_{\text{out}})}) = \mathsf{C}(G \cup K_{\mathcal{T}^\bullet(\nu_{\text{out}})})$ by Lemma 3.3 (since $\mathcal{T}(\nu_{\text{out}}) \subseteq \mathcal{T}^\bullet(\nu_{\text{out}}) \subseteq A$ and $|\mathcal{T}^\bullet(\nu_{\text{out}})| \leq s$) and $\mathcal{T}(\nu_{\text{out}}) = \mathbb{T}_{(s)}^{\mathsf{C},G}(A) = \emptyset$ by assumption. Hence $G \cup K_{\mathcal{T}(\nu_{\text{out}})} = G \cup K_\emptyset = G$. We conclude that $\mathsf{C}(G) = \mathsf{C}(G \cup K_A)$. \square_{LEMMA}

3.3 Proofs of (Almost) Main Theorems

In this section we prove versions of our main results, Theorems 1.1 and 1.2, with slightly inferior parameters. Theorems 1.1 and 1.2 (stated in the introduction) will be proved in §5 using an argument similar to the one present here. The advantage of first proving weaker results is a moderate gain in the simplicity of the argument (making things easier the second time around).

LEMMA 3.8 (MAIN TECHNICAL LEMMA). *Suppose $f = (f_n)_{n \in \mathbb{N}}$ is an AC^0 -computable sequence of functions $f_n : \{0, 1\}^{\binom{n}{2}} \rightarrow \{0, 1\}^{n^\beta}$ for some constant $\beta \geq 0$. Let $\alpha > 0$ and $k \in \mathbb{N}$. Then for a random graph $G \in \text{ER}(n, n^{-\alpha})$ and a uniform random set $A \in \binom{[n]}{k}$,*

$$\Pr[\mathbb{T}^{f,G}(A) = A] = n^{\alpha \binom{k}{2} + (\beta-1)k + o(1)}.$$

We prove this lemma at the end of §4.1 after developing the technical preliminaries concerning random restrictions.

COROLLARY 3.9. *Take $f, \beta, \alpha, k, G, A$ as in Lemma 3.8.*

1. $\Pr[\mathbb{T}_{(s)}^{f,G}(A) \neq \emptyset] = n^{\max(\alpha + 2(\beta-1), \alpha \binom{s}{2} + s(\beta-1)) + o(1)}$,
2. $\Pr[|\mathbb{T}_{(s)}^{f,G}(A)| > s] = n^{\max(\alpha \binom{s+1}{2} + (\beta-1)(s+1), \alpha \binom{2s}{2} + 2s(\beta-1)) + o(1)}$.

PROOF. We prove statement (2); the proof of (1) is similar.

$$\begin{aligned} & \Pr_{G \in \text{ER}(n, n^{-\alpha}), A \in \binom{[n]}{k}}[|\mathbb{T}_{(s)}^{f,G}(A)| > s] \\ & \leq \Pr_{G \in \text{ER}(n, n^{-\alpha}), A \in \binom{[n]}{k}} \left[\bigvee_{B \subseteq A : s < |B| \leq 2s} \mathbb{T}^{f,G}(B) = B \right] \\ & \quad \text{(by Lemma 3.5(3))} \\ & \leq \sum_{r=s+1}^{2s} \binom{k}{r} \Pr_{G \in \text{ER}(n, n^{-\alpha}), B \in \binom{[n]}{r}}[\mathbb{T}^{f,G}(B) = B] \\ & \quad \text{(union bound)} \\ & = \sum_{r=s+1}^{2s} \binom{k}{r} n^{\alpha \binom{r}{2} + (\beta-1)r + o(1)} \quad \text{(by Lemma 3.8)} \\ & = n^{\max(\alpha \binom{s+1}{2} + (\beta-1)(s+1), \alpha \binom{2s}{2} + 2s(\beta-1)) + o(1)}. \quad \square \end{aligned}$$

The beauty of this corollary is that k (the size of A) does not appear in these upper bounds outside of the $o(1)$ term. We now prove a weaker version of Theorem 1.1 (with $t > 2/3$ instead of $t > 1/2$ and $\alpha \leq \frac{4}{9t-3}$ instead of $\alpha \leq \frac{1}{2t-1}$).

THEOREM 3.10. *Suppose $f = (f_n)_{n \in \mathbb{N}}$ is a sequence of functions $f_n : \{0, 1\}^{\binom{n}{2}} \rightarrow \{0, 1\}$ computed by circuits of depth $O(1)$ and size $O(n^t)$ for constant $t > 1/2$. Let $G = \text{ER}(n, n^{-\alpha})$ be an Erdős-Rényi random graph and let A be a uniform random set in $\binom{[n]}{k}$ for any constants $0 < \alpha \leq \frac{4}{9t-3}$ and $k \in \mathbb{N}$. Then $f_n(G) = f_n(G \cup K_A)$ asymptotically almost surely as $n \rightarrow \infty$.*

PROOF. Let $s = \lfloor 3t/2 \rfloor$. By a routine calculation,

$$\begin{aligned} \max(\alpha - 2, \alpha \binom{s}{2} - s) &< 0, \\ \max(\alpha \binom{s+1}{2} - (s+1), \alpha \binom{2s}{2} - 2s) &< -t. \end{aligned}$$

So (by continuity) there exists $\beta > 0$ such that

$$\max(\alpha \binom{s+1}{2} + (\beta - 1)(s+1), \alpha \binom{2s}{2} + 2s(\beta - 1)) < -t.$$

Let $\mathbf{C} = (\mathbf{C}_n)_{n \in \mathbb{N}}$ be a sequence of circuits computing f in depth $O(1)$ and size $O(n^t)$. Without loss of generality, we can assume that all nodes of \mathbf{C} have fan-in n^β . This assumption only increases the depth of \mathbf{C} by a factor of $\lceil t/\beta \rceil$ without even doubling its size.

For all nodes ν in \mathbf{C} , let \mathbf{C}_ν^\bullet be (as defined at the beginning of §3.2) the subcircuit of \mathbf{C} in which ν and all its children are designated as outputs nodes. So \mathbf{C}_ν^\bullet computes a function $\mathcal{G}_\nu \rightarrow \{0, 1\}^{n^\beta+1}$ in depth $O(1)$ and size $O(n^t)$. Therefore, by Corollary 3.9(2),

$$\begin{aligned} \Pr[|\mathbb{T}_{(s)}^{\mathbf{C}_\nu^\bullet, G}(A)| > s] \\ &= n^{\max(\alpha \binom{s+1}{2} + (\beta-1)(s+1), \alpha \binom{2s}{2} + 2s(\beta-1)) + o(1)} \\ &= n^{-t - \Omega(1)}. \end{aligned}$$

Taking a union bound over all nodes ν in \mathbf{C} ,

$$\Pr[\bigvee_\nu |\mathbb{T}_{(s)}^{\mathbf{C}_\nu^\bullet, G}(A)| > s] = O(n^t) \cdot n^{-t - \Omega(1)} = n^{-\Omega(1)}.$$

In addition, by Corollary 3.9(1) we have

$$\Pr[\mathbb{T}_{(s)}^{\mathbf{C}, G}(A) \neq \emptyset] = n^{\max(\alpha+2, \alpha \binom{s}{2} + s) + o(1)} = n^{-\Omega(1)}.$$

Therefore, it holds asymptotically almost surely that $\mathbb{T}_{(s)}^{\mathbf{C}, G}(A) = \emptyset$ and $|\mathbb{T}_{(s)}^{\mathbf{C}_\nu^\bullet, G}(A)| \leq s$ for all nodes ν in \mathbf{C} . The Circuit Lemma (Lemma 3.6) now gives the desired result that a.a.s. $\mathbf{C}(G) = \mathbf{C}(G \cup K_A)$ (i.e., $f_n(G) = f_n(G \cup K_A)$). \square

The following theorem is a weaker version of Theorem 1.2 (with $\omega(n^{2k/9})$ instead of $\omega(n^{k/4})$).

THEOREM 3.11. *For every constant k , the k -clique problem on n -vertex graphs requires constant-depth circuits of size $\omega(n^{2k/9})$.*

PROOF. The lower bound of $\omega(n^{2k/9})$ is certainly trivial for $k \leq 2$, so we assume that $k \geq 3$. Let $\mathbf{C} = (\mathbf{C}_n)_{n \in \mathbb{N}}$ be a sequence of circuits on $\binom{n}{2}$ inputs of depth $O(1)$ and size $O(n^t)$ for some constant $t = \frac{2k}{9} (> \frac{2}{3})$. Let $\alpha = \frac{4}{9t-3}$ and note that $\alpha = \frac{2}{k-1.5} > \frac{2}{k-1} = \text{thres}(K_k)$. For random $G \in \text{ER}(n, n^{-\alpha})$ and $A \in \binom{[n]}{k}$, Theorem 3.10 tells us that asymptotically almost surely $\mathbf{C}(G) = \mathbf{C}(G \cup K_A)$. By Lemma 2.1(1), G almost surely contains no k -clique. On the other hand, $G \cup K_A$ contains a k -clique (with probability 1). It follows that \mathbf{C} does not define the property of containing a k -clique (for sufficiently large n). Therefore, the k -clique problem does not have constant-depth circuits of size $O(n^{2k/9})$. \square

4. RANDOM RESTRICTIONS AND AC^0

In this section we prove some technical results concerning the sensitivity of AC^0 functions when hit with random restrictions. The endgoal of this section is a proof of the main technical lemma from §3 (Lemma 3.8). We begin with some standard definitions.

DEFINITION 4.1 (SENSITIVE INPUTS). *Let I be a set (we mainly consider $I = [n]$ or $\binom{[n]}{2}$) and let f be a function with domain $\{0, 1\}^I$ (and arbitrary range). The set $\mathbb{S}(f)$ of sensitive inputs of f consists of all $i \in I$ for which there exist $x, y \in \{0, 1\}^I$ such that $f(x) \neq f(y)$ and $x_j = y_j$ for all $j \in I \setminus \{i\}$.*

DEFINITION 4.2 (RESTRICTION). *A restriction on I is a function $\varrho : I \rightarrow \{0, 1, \star\}$. Intuitively ϱ fixes a certain subset of input bits (assigning either 0 or 1) which leaving other input bits unrestricted (by assigning \star). A restriction ϱ thus may be applied to a function f with domain $\{0, 1\}^I$, resulting in a function $f|_\varrho$ with domain $\{0, 1\}^{\varrho^{-1}(\star)}$. The value of $f|_\varrho$ on an element $y \in \{0, 1\}^{\varrho^{-1}(\star)}$ is defined by $(f|_\varrho)(y) = f(x)$ where $x \in \{0, 1\}^I$ satisfies $x(i) = \varrho(i)$ if $i \in \varrho^{-1}(\{0, 1\})$ and $x(i) = y(i)$ if $i \in \varrho^{-1}(\star)$.*

We now consider a particular family of random restrictions (also studied in [7, 8]).

DEFINITION 4.3. *For a finite index set I and $p, q \in [0, 1]$, let $\mathcal{R}_I^{p,q}$ denote the random restriction $\varrho : I \rightarrow \{0, 1, \star\}$ where values $\varrho(i)$ are i.i.d. such that*

$$\begin{aligned} \Pr[\varrho(i) = \star] &= p, \\ \Pr[\varrho(i) = 1] &= (1-p)q, \\ \Pr[\varrho(i) = 0] &= (1-p)(1-q). \end{aligned}$$

As a matter of notation, we write $\mathcal{R}_n^{p,q}$ for $\mathcal{R}_{[n]}^{p,q}$ in the case $I = [n]$.

Our first lemma concerning random restrictions is proved by a standard application of the Håstad Switching Lemma. The proof is included in the full paper, but omitted here since results similar to Lemma 4.4 are found elsewhere in the literature (e.g., in proofs of $\text{PARITY} \notin \text{AC}^0$ [4, 15]).

LEMMA 4.4. *Suppose functions $f_n : \{0, 1\}^n \rightarrow \{0, 1\}$, $n \in \mathbb{N}$, are computed by circuits of constant depth d and size $O(n^t)$. Then for every (small) $\delta > 0$ and (large) $\ell > 0$, there is a constant $c = c(d, t, \delta, \ell)$ such that*

$$\Pr_{\varrho \in \mathcal{R}_n^{n^{-\delta}, 1/2}}[|\mathbb{S}(f|_\varrho)| > c] = O(1/n^\ell).$$

In plain language, Lemma 4.4 says that with high probability (i.e., failing with probability as small an inverse polynomial in n as one desires), an AC^0 -computable function will depend on only constant many bits after being hit with a random restriction which fixes (i.e., assigns either 0 or 1 to) all but a fractional power of n input bits. Besides the having $\text{PARITY} \notin \text{AC}^0$ as an immediate corollary, Lemma 4.4 also implies that the average sensitivity of AC^0 functions is $n^{o(1)}$ (though better is known [9]).

The next lemma extends Lemma 4.4 to biased random restrictions. We remark that the proof of Lemma 4.5, below, does not rely on biased versions of the Håstad Switching Lemma (e.g., Beame's Imbalanced Switching Lemma from [8]). Instead, we use a simple gadget (an \wedge -gate) to bootstrap from Lemma 4.4.

LEMMA 4.5. *Suppose functions $f_n : \{0, 1\}^n \rightarrow \{0, 1\}$, $n \in \mathbb{N}$, are computed by circuits of constant depth d and*

size $O(n^\ell)$. Then for all $\alpha, \delta > 0$ and $\ell > 0$, there is a constant $c = c(d, t, \alpha, \delta, \ell)$ such that

$$\Pr_{\varrho \in \mathcal{R}_n^{n^{-(\alpha+\delta)}, n^{-\alpha}}} [|\mathbb{S}(f[\varrho])| > c] = O(1/n^\ell).$$

PROOF SKETCH. Let $m = \lfloor \alpha \log_2 n \rfloor$ and consider the function $g : \{0, 1\}^{n \times m} \rightarrow \{0, 1\}^n$ which maps $y \in \{0, 1\}^{n \times m}$ to the element $x \in \{0, 1\}^n$ where $x(i) = \bigwedge_{j \in [m]} y(i, j)$.

Let $\varrho : [n] \times [m] \rightarrow \{0, 1, \star\}$ be a random restriction under $\mathcal{R}_{n \times m}^{(nm)^{-\delta}, 1/2}$. We lift ϱ to a restriction $\tilde{\varrho} : [n] \rightarrow \{0, 1, \star\}$ defined by

$$\tilde{\varrho}(i) = \begin{cases} 0 & \text{if } \exists j \in [m], \varrho(i, j) = 0, \\ 1 & \text{if } \forall j \in [m], \varrho(i, j) = 1, \\ \star & \text{otherwise.} \end{cases}$$

$\tilde{\varrho}$ is clearly a random restriction under $\mathcal{R}_n^{p, q}$ for some $p = p(n)$ and $q = q(n)$ in $[0, 1]$. By a straightforward calculation, $\log_n(p) \sim -(\alpha + \delta)$ and $\log_n(q) \sim -\alpha$.

Since f and g are both AC^0 -computable, so is their composition $f \circ g : \{0, 1\}^{n \times m} \rightarrow \{0, 1\}$. Therefore, by Lemma 4.4 there is a constant c such that $\Pr [|\mathbb{S}((f \circ g)[\varrho])| > c] = O(1/n^\ell)$. Note that

$$\mathbb{S}(f[\tilde{\varrho}]) = \{i : \exists j (i, j) \in \mathbb{S}((f \circ g)[\varrho])\}.$$

It follows that $|\mathbb{S}(f[\tilde{\varrho}])| \leq |\mathbb{S}((f \circ g)[\varrho])|$. As required, this proves $\Pr [|\mathbb{S}(f[\tilde{\varrho}])| > c] = O(1/n^\ell)$.

This argument, however, fails to prove the precise statement of the lemma, since p and q are not exactly equal to $n^{-(\alpha+\delta)}$ and $n^{-\alpha}$. This is a minor issue to get around. The resolution involves redoing the proof sketched here with a few additional parameters (see full paper). \square

4.1 Random Graph-Restrictions

We now shift our focus from random restrictions $[n] \rightarrow \{0, 1, \star\}$ to random *graph-restrictions* $\binom{[n]}{2} \rightarrow \{0, 1, \star\}$.

DEFINITION 4.6 (SENSITIVE VERTICES). *Suppose f is a graph-function with domain \mathcal{G}_n (and arbitrary range) and let $\varrho : \binom{[n]}{2} \rightarrow \{0, 1, \star\}$ be a graph-restriction. The set $\mathbb{S}(f[\varrho])$ of sensitive inputs of $f[\varrho]$ is a subset of $\binom{[n]}{2}$ (i.e., the potential edges of a graph in \mathcal{G}_n). We call elements of $\mathbb{S}(f[\varrho])$ sensitive edges of the restricted function $f[\varrho]$ (rather than sensitive inputs). We define the set $\mathbb{V}(f[\varrho]) \subseteq [n]$ by*

$$\mathbb{V}(f[\varrho]) = \{i \in [n] : \exists j \in [n] \{i, j\} \in \mathbb{S}(f[\varrho])\}.$$

We call elements of $\mathbb{V}(f[\varrho])$ sensitive vertices of $f[\varrho]$. That is, sensitive vertices are simply the endpoints of sensitive edges.

Notice that $|\mathbb{V}(f[\varrho])| \leq 2|\mathbb{S}(f[\varrho])|$ since each sensitive edge of $f[\varrho]$ contributes two sensitive vertices (with possible overlap).

DEFINITION 4.7. *For every graph H and $q \in [0, 1]$ and $n \geq |V_H|$, let $\mathcal{GR}_n^q(H)$ denote the random graph-restriction $\varrho : \binom{[n]}{2} \rightarrow \{0, 1, \star\}$ defined as follows. First, a one-to-one function $w : V_H \rightarrow [n]$ is chosen uniformly at random. For all edges $\{i, j\} \in E_H$, the element $\{w(i), w(j)\} \in \binom{[n]}{2}$ is mapped under ϱ to \star . All remaining $e \in \binom{[n]}{2} \setminus w(E_H)$ are then independently mapped under ϱ to 1 with probability q and to 0 with probability $1 - q$.*

We now prove the main result of this section. We then derive Lemma 3.8 (the main technical lemma of §3.3) as a corollary.

PROPOSITION 4.8. *Suppose $f : \mathcal{G}_n \rightarrow \{0, 1\}^{n^\beta}$ is AC^0 -computable where $\beta \geq 0$. Let H be any graph and let $0 < \alpha < \text{thres}(H)$. Then*

$$\Pr_{\varrho \in \mathcal{GR}_n^{n^{-\alpha}}(H)} [|\mathbb{V}(f[\varrho])| = |V_H|] = n^{\alpha|E_H| + (\beta-1)|V_H| + o(1)}.$$

We conveniently pretend that n^β is always an integer. To be correct, every instance of n^β in the following proof should be replaced with $\lceil n^\beta \rceil$.

PROOF. Fix $\varepsilon > 0$ and let

$$\delta = \min \left\{ \frac{\varepsilon}{2|E_H|}, \frac{\text{thres}(H) - \alpha}{2} \right\}. \quad (1)$$

We also pick a random restriction $\xi : \binom{[n]}{2} \rightarrow \{0, 1, \star\}$ from the distribution $\mathcal{R}_{\binom{[n]}{2}}^{n^{-(\alpha+\delta)}, n^{-\alpha}}$.

Let $f_1, \dots, f_{n^\beta} : \{0, 1\}^n \rightarrow \{0, 1\}$ be the individual coordinate-functions of f . By Lemma 4.5 (noting that $\delta > 0$ and taking $\ell = |V_H|$), there is a constant c such that $\Pr [|\mathbb{S}(f_j[\xi])| > c] = O(1/n^{|V_H|})$ for all $j \in [n^\beta]$.

Let $\check{\nu}_1$ stand for the event that $|\mathbb{V}(f[\xi])| \leq 2cn^\beta$. Our first order of business is to bound the probability of $\neg\check{\nu}_1$.

$$\begin{aligned} \Pr[\neg\check{\nu}_1] &\leq \Pr [|\mathbb{S}(f[\xi])| > cn^\beta] & (2) \\ &\quad (\text{since } |\mathbb{V}(f[\xi])| \leq 2|\mathbb{S}(f[\xi])|) \\ &\leq \Pr [\bigvee_{j \in [n^\beta]} |\mathbb{S}(f_j[\xi])| > c] \\ &\quad (\text{since } \mathbb{S}(f[\xi]) = \bigcup_{j \in [n^\beta]} \mathbb{S}(f_j[\xi])) \\ &\leq n^\beta \cdot O(1/n^{|V_H|}) \quad (\text{union bound}) \\ &= O(1/n^{(1-\beta)|V_H|}). \end{aligned}$$

Let G_ξ be the graph with vertex set $V_{G_\xi} = [n]$ and edge set $E_{G_\xi} = \xi^{-1}(\star)$. Note that G_ξ is a random graph with distribution $\text{ER}(n, n^{-(\alpha+\delta)})$.

Let $\check{\nu}_2$ denote the event that $\text{sub}(G_\xi, H) \geq n^{|V_H| - \alpha|E_H| - \varepsilon}$. We have $\alpha + \delta < \text{thres}(H)$ and $\delta|E_H| \leq \varepsilon/2$ by definition (1) of δ . We now bound the probability of $\neg\check{\nu}_2$.

$$\begin{aligned} \Pr[\neg\check{\nu}_2] &= \Pr [\text{sub}(G_\xi, H) < n^{|V_H| - \alpha|E_H| - \varepsilon}] & (3) \\ &= \Pr_{G \in \text{ER}(n, n^{-(\alpha+\delta)})} [\text{sub}(G, H) < n^{|V_H| - \alpha|E_H| - \varepsilon}] \\ &\leq \Pr_{G \in \text{ER}(n, n^{-(\alpha+\delta)})} [\text{sub}(G, H) < n^{|V_H| - (\alpha+\delta)|E_H| - \frac{\varepsilon}{2}}] \\ &= \exp(-n^{\Omega(1)}) \quad (\text{by Lemma 2.1(2)}). \end{aligned}$$

In the (very likely) event that $\check{\nu}_1$ and $\check{\nu}_2$ both hold, we proceed to pick a random graph H' and a random graph-restriction ϱ as follows. First we choose H' uniformly at random from among the induced subgraphs of G_ξ isomorphic to H (noting that $\check{\nu}_2 \implies \text{sub}(G_\xi, H) > 0$). Having chosen H' , we then pick a random-graph restriction $\varrho : \binom{[n]}{2} \rightarrow \{0, 1, \star\}$ subject to

- $\varrho^{-1}(\star) = E_{H'}$ (with probability 1),
- $\varrho(e) = \xi(e)$ (with probability 1) for all $e \in \xi^{-1}(\{0, 1\})$,
- $\Pr[\varrho(e) = 1] = n^{-\alpha}$ independently for all $e \in \xi^{-1}(\star) \setminus E_{H'}$.

At this point, some observations are in order.

- (i) Conditioned on $\sqrt{1}$ and $\sqrt{2}$ (so that ϱ is well-defined), ϱ has distribution $\mathcal{GR}_n^{\alpha^{-1}}(H)$.
- (ii) $\mathbb{S}(f[\varrho]) \subseteq \mathbb{S}(f[\xi])$ and $\mathbb{V}(f[\varrho]) \subseteq \mathbb{V}(f[\xi])$ by virtue of the fact that ϱ refines ξ (i.e., $\xi(e) \in \{0, 1\} \implies \varrho(e) = \xi(e)$).
- (iii) $\sqrt{1} \implies \mathbb{V}(f[\xi])$ contains $\leq \binom{2cn^\beta}{|V_H|}$ subsets of size $|V_H|$.
- (iv) $\sqrt{2} \implies$ no matter what ξ is, there are $\geq n^{|V_H| - \alpha|E_H| - \varepsilon}$ equally likely possibilities for the random graph H' .

Combining these observations, we have

$$\begin{aligned} \Pr[|\mathbb{V}(f[\varrho])| = |V_H| \mid \sqrt{1}, \sqrt{2}] &= \Pr[V_{H'} = \mathbb{V}(f[\varrho]) \mid \sqrt{1}, \sqrt{2}] \\ &\stackrel{\text{(by (ii))}}{\leq} \Pr[V_{H'} \subseteq \mathbb{V}(f[\xi]) \mid \sqrt{1}, \sqrt{2}] \\ &\stackrel{\text{(by (iii) and (iv))}}{\leq} \binom{2cn^\beta}{|V_H|} / n^{|V_H| - \alpha|E_H| - \varepsilon} \\ &= n^{\beta|V_H| + o(1)} / n^{|V_H| - \alpha|E_H| - \varepsilon} \\ &= n^{\alpha|E_H| + (\beta-1)|V_H| + \varepsilon + o(1)}. \end{aligned}$$

Putting this together with (2) and (3), we get

$$\begin{aligned} \Pr[|\mathbb{V}(f[\varrho])| = |V_H|] &\leq \Pr[|\mathbb{V}(f[\varrho])| = |V_H| \mid \sqrt{1}, \sqrt{2}] \\ &\quad + \Pr[\neg\sqrt{1}] + \Pr[\neg\sqrt{2}] \\ &= n^{\alpha|E_H| + (\beta-1)|V_H| + \varepsilon + o(1)} \\ &\quad + O(n^{(\beta-1)|V_H|}) + \exp(-n^{\Omega(1)}) \\ &= n^{\alpha|E_H| + (\beta-1)|V_H| + \varepsilon + o(1)}. \end{aligned}$$

The result follows, as this inequality holds for all $\varepsilon > 0$. \square

Finally, we conclude this section by proving our main technical lemma from §3.3.

PROOF OF LEMMA 3.8. We assume that $\alpha < \binom{2}{k-1}$, since otherwise the lemma is trivial. Suppose $f : \{0, 1\}^{\binom{n}{2}} \rightarrow \{0, 1\}^{n^\beta}$ is AC⁰-computable. Consider a random graph $G \in \text{ER}(n, n^{-\alpha})$ and a random set $A \in \binom{[n]}{k}$.

Let $\varrho_A^G : \binom{[n]}{2} \rightarrow \{0, 1, \star\}$ be the graph-restriction which maps $e \in \binom{[n]}{2}$ to \star if $e \in \binom{A}{2}$, to 1 if $e \in E_G \setminus \binom{A}{2}$, and to 0 otherwise. So in other words, ϱ_A^G is obtained from the characteristic function of the edge set $E_G \subseteq \binom{[n]}{2}$ by changing the value at every element of $\binom{A}{2}$ to \star . Note that ϱ_A^G is a random graph-restriction with distribution $\mathcal{GR}_n^{\alpha^{-1}}(K_k)$. It is easy to see that $\mathbb{T}^{f,G}(A) \subseteq \mathbb{V}(f[\varrho_A^G]) \subseteq A$. That is, every a in the clique-sensitive core of A under f in G is clearly a sensitive vertex of the restricted function $f[\varrho_A^G]$. Consequently, $\mathbb{T}^{f,G}(A) = A \implies |\mathbb{V}(f[\varrho_A^G])| = k$. Therefore, by Proposition 4.8 (noting that $\alpha < \binom{2}{k-1} = \text{thres}(K_k)$),

$$\begin{aligned} \Pr_{G,A}[\mathbb{T}^{f,G}(A) = A] &\leq \Pr_{G,A}[|\mathbb{V}(f[\varrho_A^G])| = k] \\ &= \Pr_{\varrho \in \mathcal{GR}_n^{\alpha^{-1}}(K_k)}[|\mathbb{V}(f[\varrho])| = k] \\ &= n^{\alpha|E_{K_k}| + (\beta-1)|V_{K_k}| + o(1)} \\ &= n^{\alpha \binom{k}{2} + (\beta-1)k + o(1)}. \quad \square \end{aligned}$$

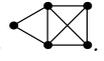
5. PROOFS OF THEOREMS 1.1 AND 1.2

In this section we prove our main results, Theorems 1.1 and 1.2. Recall that in §3.3, we proved versions of these theorems with weaker parameters (Theorems 3.10 and 3.11). The proofs we now present follow the same basic scheme as before. To pinpoint the difference this time around, recall statements (2) and (3), below, of Lemma 3.5.

$$\begin{aligned} |\mathbb{T}_{(s)}^{f,G}(A)| > s &\stackrel{(2)}{\iff} \exists \text{ fully clique-sensitive } B, C \subseteq A \text{ with} \\ &\quad |B|, |C| \leq s \text{ and } |B \cup C| \geq s+1 \\ &\stackrel{(3)}{\implies} \exists \text{ fully clique-sensitive } D \subseteq A \text{ with} \\ &\quad s+1 \leq |D| \leq 2s. \end{aligned}$$

In §3.3 we merely exploited statement (3) of Lemma 3.5, specifically in the proof of Corollary 3.9(2). This time we use the stronger statement (2) of Lemma 3.5 to derive superior bounds (without too much added complexity).

DEFINITION 5.1. For $a, b, c \in \mathbb{N}$ such that $a \leq \min(b, c)$ and $b + c - a \geq 1$, let $H_{a,b,c}$ denote the graph formed by gluing together a b -clique and a c -clique over a subclique of

size a . For example, $H_{2,3,4}$ is the graph 

Clearly $|V_{H_{a,b,c}}| = b + c - a$ and $|E_{H_{a,b,c}}| = \binom{b}{2} + \binom{c}{2} - \binom{a}{2}$. Also note that every induced subgraph of $H_{a,b,c}$ is a graph $H_{a',b',c'}$ for some smaller a', b', c' .

LEMMA 5.2. $\text{thres}(H_{a,b,c}) > (1 + \frac{1}{\sqrt{2}}) / \max(b, c)$.

PROOF. Let $s = \max(b, c)$. We assume $s \geq 2$, since otherwise $\text{thres}(H_{a,b,c}) = \infty$ (as $H_{a,b,c}$ has no edges). We have

$$\begin{aligned} \text{thres}(H_{a,b,c}) &= \min_{\substack{\text{induced subgraphs} \\ H_{a',b',c'} \subseteq H_{a,b,c}}} \frac{|V_{H_{a',b',c'}}|}{|E_{H_{a',b',c'}}|} \\ &\geq \min_{\substack{a', b', c' \in \mathbb{N}: \\ a' \leq b', c' \leq s \\ b' + c' - a' \geq 1}} \frac{b' + c' - a'}{\binom{b'}{2} + \binom{c'}{2} - \binom{a'}{2}}. \end{aligned}$$

This last quantity is clearly minimal for $b' = c' = s$. Notice that a' can be at most $s - 1$. Letting $\lambda = s/a'$, we have

$$\begin{aligned} \text{thres}(H_{a,b,c}) &\geq \min_{a' \in \{0, \dots, s-1\}} \frac{2s - a'}{2 \binom{s}{2} - \binom{a'}{2}} \\ &> \frac{1}{s} \min_{0 \leq \lambda \leq 1} \frac{2(2 - \lambda)}{2 - \lambda^2} = \left(1 + \frac{1}{\sqrt{2}}\right) \frac{1}{s}. \quad \square \end{aligned}$$

DEFINITION 5.3. For $s \in \mathbb{N}$, let W_s denote the set of triples $(a, b, c) \in \mathbb{N}^3$ such that $a \leq \min(b, c)$ and $\max(b, c) \leq s$ and $b + c - a \geq s + 1$.

The next lemma will play the same role in proving Theorem 1.1 as our main technical lemma (Lemma 3.8) played in proving the weaker Theorem 3.10.

LEMMA 5.4. Suppose $f = (f_n)_{n \in \mathbb{N}}$ is an AC⁰-computable sequence of functions $f_n : \{0, 1\}^{\binom{n}{2}} \rightarrow \{0, 1\}^{n^\beta}$ for some constant $\beta \geq 0$. Let $s \in \mathbb{N}$ and $(a, b, c) \in W_s$ and $0 < \alpha \leq 1/s$. Then for a random graph $G \in \text{ER}(n, n^{-\alpha})$ and uniform random sets $B \in \binom{[n]}{b}$ and $C \in \binom{[n]}{c}$ subject to $|B \cap C| = a$,

$$\begin{aligned} \Pr[\mathbb{T}^{f,G}(B) = B \text{ and } \mathbb{T}^{f,G}(C) = C] \\ = n^{\alpha \left(\binom{b}{2} + \binom{c}{2} - \binom{a}{2} \right) + (\beta-1)(b+c-a) + o(1)}. \end{aligned}$$

PROOF. Consider the graph-restriction $\varrho_{B,C}^G : \binom{[n]}{2} \rightarrow \{0, 1, \star\}$ which maps $e \in \binom{[n]}{2}$ to \star if $e \in \binom{B}{2} \cup \binom{C}{2}$, to 1 if $e \in E_G \setminus (\binom{B}{2} \cup \binom{C}{2})$, and to 0 otherwise. That is, $\varrho_{B,C}^G$ is obtained from the characteristic function of the edge set $E_G \subseteq \binom{[n]}{2}$ by changing the value at every element of $\binom{B}{2} \cup \binom{C}{2}$ to \star . Note that $\varrho_{B,C}^G$ is a random graph-restriction with distribution $\mathcal{GR}_n^{n-\alpha}(H_{a,b,c})$.

Every element of $\mathbb{T}^{f,G}(B) \cup \mathbb{T}^{f,G}(C)$ must clearly be a sensitive vertex with respect to the restricted graph-function $f \upharpoonright_{\varrho_{B,C}^G}$. In particular,

$$\mathbb{T}^{f,G}(B) \cup \mathbb{T}^{f,G}(C) \subseteq \mathbb{V}(f \upharpoonright_{\varrho_{B,C}^G}) \subseteq B \cup C.$$

By Proposition 4.8 (noting that $\alpha \leq 1/s < \text{thres}(H_{a,b,c})$ by Lemma 5.2), we have

$$\begin{aligned} & \Pr_{G,B,C} [\mathbb{T}^{f,G}(B) = B \text{ and } \mathbb{T}^{f,G}(C) = C] \\ & \leq \Pr_{G,B,C} [|\mathbb{V}(f \upharpoonright_{\varrho_{B,C}^G})| = |B \cup C|] \\ & \leq \Pr_{e \in \mathcal{GR}_n^{n-\alpha}(H_{a,b,c})} [|\mathbb{V}(f \upharpoonright_e)| = |V_{H_{a,b,c}}|] \\ & = n \alpha |E_{H_{a,b,c}}| + (\beta - 1) |V_{H_{a,b,c}}| + o(1) \\ & = n \alpha \binom{b}{2} + \binom{c}{2} - \binom{a}{2} + (\beta - 1)(b + c - a) + o(1). \quad \square \end{aligned}$$

The following corollary of Lemma 5.4 directly strengthens Corollary 3.9(2).

COROLLARY 5.5. *Take f, β, s, α as in Lemma 5.4. Let k be a constant (independent of n). For a random graph $G \in \text{ER}(n, n^{-\alpha})$ and a uniform random set $A \in \binom{[n]}{k}$,*

$$\Pr [|\mathbb{T}_{(s)}^{f,G}(A)| > s] = n^{\gamma+o(1)}$$

where $\gamma = \max_{(a,b,c) \in W_s} \alpha \binom{b}{2} + \binom{c}{2} - \binom{a}{2} + (\beta - 1)(b + c - a)$.

PROOF. Similar to the proof of Corollary 3.9, we have

$$\begin{aligned} & \Pr_{G \in \text{ER}(n, n^{-\alpha}), A \in \binom{[n]}{k}} [|\mathbb{T}_{(s)}^{f,G}(A)| > s] \\ & = \Pr_{\substack{G \in \text{ER}(n, n^{-\alpha}) \\ A \in \binom{[n]}{k}}} \left[\bigvee_{\substack{B, C \in \binom{[n]}{\leq s} \\ |B \cup C| > s}} \mathbb{T}^{f,G}(B) = B \text{ and } \mathbb{T}^{f,G}(C) = C \right] \end{aligned}$$

(by Lemma 3.5(2))

$$\leq \sum_{(a,b,c) \in W_s} \binom{k}{b+c-a} \frac{(b+c-a)!}{a!(b-a)!(c-a)!} \times \Pr_{\substack{G \in \text{ER}(n, n^{-\alpha}) \\ B \in \binom{[n]}{b}, C \in \binom{[n]}{c} \\ |B \cap C| = a}} [\mathbb{T}^{f,G}(B) = B \text{ and } \mathbb{T}^{f,G}(C) = C]$$

(union bound)

$$= \sum_{(a,b,c) \in W_s} \overbrace{\binom{k}{b+c-a} \frac{(b+c-a)!}{a!(b-a)!(c-a)!}}^{o(1)} \times n \alpha \binom{b}{2} + \binom{c}{2} - \binom{a}{2} + (\beta - 1)(b + c - a) + o(1)$$

(by Lemma 5.4)

$$= n^{\gamma+o(1)}. \quad \square$$

We are finally ready to prove our main theorems.

PROOF OF THEOREM 1.1. Let $s = \lceil 2t - 1 \rceil$ and note that $\alpha = \frac{1}{2t-1} \leq 1/s$. By a straightforward calculation (included in the full paper), we have

$$\begin{aligned} & \max (\alpha - 2, \alpha \binom{s}{2} - s) < 0, \\ & \max_{(a,b,c) \in W_s} \alpha \binom{b}{2} + \binom{c}{2} - \binom{a}{2} - (b + c - a) < -t. \end{aligned}$$

(The maximizing $(a, b, c) \in W_s$ in the second inequality is the triple $(s - 1, s, s)$.)

From this point forward, the argument continues exactly as in the proof of Theorem 3.10 (but using Corollary 5.5 instead of Corollary 3.9). That is, we choose a small positive $\beta > 0$ such that $\gamma < -t$ where

$$\gamma = \max_{(a,b,c) \in W_s} \alpha \binom{b}{2} + \binom{c}{2} - \binom{a}{2} + (\beta - 1)(b + c - a).$$

We next choose circuits $C = (C_n)_{n \in \mathbb{N}}$ with fan-in n^β which compute functions f_n in depth $O(1)$ and size $O(n^t)$ with fan-in n^β . For each node $\nu \in C$, Corollary 5.5 gives us

$$\Pr [|\mathbb{T}_{(s)}^{C_\nu, G}(A)| > s] = n^{\gamma+o(1)} = n^{-t-\Omega(1)}.$$

So by a union bound

$$\Pr [\bigvee_\nu |\mathbb{T}_{(s)}^{C_\nu, G}(A)| > s] = O(n^t) \cdot n^{-t-\Omega(1)} = n^{-\Omega(1)}.$$

Also, by Corollary 3.9(1) we have

$$\Pr [\mathbb{T}_{(s)}^{C, G}(A) \neq \emptyset] = n^{\max(\alpha+2, \alpha \binom{s}{2} + s) + o(1)} = n^{-\Omega(1)}.$$

Therefore, it holds asymptotically almost surely that $\mathbb{T}_{(s)}^{C, G}(A) = \emptyset$ and $|\mathbb{T}_{(s)}^{C_\nu, G}(A)| \leq s$ for all $\nu \in C$. The Circuit Lemma (Lemma 3.6) now gives the desired result that a.a.s. $C(G) = C(G \cup K_A)$ (i.e., $f_n(G) = f_n(G \cup K_A)$). \square

PROOF OF THEOREM 1.2. The lower bound of $\omega(n^{k/4})$ is certainly trivial for $k \leq 3$, so we assume that $k \geq 4$. Let $C = (C_n)_{n \in \mathbb{N}}$ be a sequence of circuits on $\binom{[n]}{2}$ inputs of depth $O(1)$ and size $O(n^t)$ for some constant $t = \frac{k}{4} (> \frac{1}{2})$. Let $\alpha = \frac{1}{2t-1}$ and note that $\alpha = \frac{2}{k-2} > \frac{2}{k-1} = \text{thres}(K_k)$. For random $G \in \text{ER}(n, n^{-\alpha})$ and $A \in \binom{[n]}{k}$, Theorem 1.1 tells us that $C(G) = C(G \cup K_A)$ asymptotically almost surely. By Lemma 2.1(1), G a.a.s. contains no k -clique as $\alpha > \text{thres}(K_k) = \frac{2}{k-1}$. On the other hand, $G \cup K_A$ contains a k -clique (with probability 1). It follows that C does not define the property of containing a k -clique (for sufficiently large n). Therefore, the k -clique problem does not have constant-depth circuits of size $O(n^{k/4})$. \square

6. CONCLUSION

Superconstant Depth.

The constant-depth assumption is really only used in one place in this paper, namely Lemma 4.4. In fact, Lemma 4.4 (and hence all other results in this paper) actually holds for slightly increasing circuit depth $d = d(n) = o(\sqrt{\log n})$. To accommodate this change, the statement of Lemma 4.4 changes from “there exists a constant $c = c(d, t, \delta, \ell)$ ” to “there exists a function $c_{d,t,\delta,\ell}(n) = n^{o(1)}$ ”. The proof of Proposition 4.8 (the only place where Lemma 4.4 is used) goes through unchanged when $c = n^{o(1)}$.

Though we have not yet looked into the question, we believe that our lower bound of $\omega(n^{k/4})$ should also hold for superconstant clique-size $k = k(n)$, perhaps up to $k \leq \log n$ (as is the case with the lower bounds of Lynch [23] and Beame [7] mentioned in the introduction).

Open Questions.

What is the maximal $\alpha(t)$ such that for all $k \in \mathbb{N}$ and every sequence $(f_n)_{n \in \mathbb{N}}$ of functions $f_n : \{0, 1\}^{\binom{[n]}{k}} \rightarrow \{0, 1\}$ computed by constant-depth circuits of size $O(n^t)$, it holds that $f_n(G) = f_n(G \cup K_A)$ asymptotically almost surely where $G \in \text{ER}(n, n^{\alpha(t) - \Omega(1)})$ and A is a random k -element subset of $[n]$? We showed that $\alpha(t) \geq \frac{1}{2t-1}$, but it is conceivable that $\alpha(t) = \frac{2}{t-1}$.

What is the minimal $t(k)$ such that the k -clique problem has constant-depth circuits of size $n^{t(k)+o(1)}$? We showed that $t(k) \geq k/4$, but it is conceivable that $t(k) = k$.

What is the fewest number of variables needed to define the class of finite ordered graphs which contain a k -clique in first-order logic? We showed that $\lfloor k/4 \rfloor$ variables are necessary, but it is conceivable that k variables are necessary.

Finally, can our technique be used to obtain lower bound for problems other than k -clique?

Acknowledgements.

I am grateful to Neil Immerman for pointing out that our non-collapse result (Corollary 1.5) implies strictness of the bounded variable hierarchy on finite ordered graphs (Corollary 1.7) and for allowing me to include his proof of Proposition 1.6 in the full version of this paper. I am also grateful to Madhu Sudan for many helpful discussions and to Philipp Weis for his comments on an earlier draft of this paper. I also thank the anonymous referees for their detailed comments.

7. REFERENCES

- [1] M. Ajtai. Σ_1^1 formulae on finite structures. *Annals of Pure and Applied Logic*, 24:1–48, 1983.
- [2] M. Ajtai. First-order definability on finite structures. *Ann. Pure Appl. Logic*, 45(3):211–225, 1989.
- [3] N. Alon and R. B. Boppana. The monotone circuit complexity of boolean functions. *Combinatorica*, 7(1):1–22, 1987.
- [4] N. Alon and J. Spencer. *The Probabilistic Method*. John Wiley, 1992.
- [5] K. Amano and A. Maruoka. A superpolynomial lower bound for a circuit computing the clique function with at most $(1/6) \log \log n$ negation gates. *SIAM J. Comput.*, 35(1):201–215, 2005.
- [6] D. M. Barrington, N. Immerman, and H. Straubing. On uniformity within NC^1 . *Journal of Computer and System Sciences*, 41(3):274–306, 1990.
- [7] P. Beame. Lower bounds for recognizing small cliques on CRCW PRAM's. *Discrete Appl. Math.*, 29(1):3–20, 1990.
- [8] P. Beame. A switching lemma primer. Technical Report UW-CSE-95-07-01, Department of Computer Science and Engineering, University of Washington, November 1994.
- [9] R. B. Boppana. The average sensitivity of bounded-depth circuits. *Inf. Process. Lett.*, 63(5):257–261, 1997.
- [10] A. Dawar. How many first-order variables are needed on finite ordered structures? In *We Will Show Them: Essays in Honour of Dov Gabbay*, pages 489–520, 2005.
- [11] L. Denenberg, Y. Gurevich, and S. Shelah. Definability by constant-depth polynomial-size circuits. *Information and Control*, 70(2/3):216–240, 1986.
- [12] M. L. Furst, J. B. Saxe, and M. Sipser. Parity, circuits, and the polynomial-time hierarchy. *Mathematical Systems Theory*, 17:13–27, 1984.
- [13] M. Goldmann and J. Håstad. A simple lower bound for the depth of monotone circuits computing clique using a communication game. *Information Processing Letters*, 41(4):221–226, 1992.
- [14] M. Grohe. Finite variable logics in descriptive complexity theory. *Bulletin of Symbolic Logic*, 4(4):345–398, 1998.
- [15] J. 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.
- [16] N. Immerman. Upper and lower bounds for first order expressibility. *J. Comput. Syst. Sci.*, 25(1):76–98, 1982.
- [17] N. Immerman. Expressibility and parallel complexity. *SIAM J. Comput.*, 18(3):625–638, 1989.
- [18] N. Immerman. *Descriptive Complexity*. Graduate Texts in Computer Science. Springer-Verlag, New York, 1999.
- [19] S. Janson. Poisson approximation for large deviations. *Random Struct. Algorithms*, 1(2):221–230, 1990.
- [20] S. Janson, T. Luczak, and A. Rucinski. An exponential bound for the probability of nonexistence of a specified subgraph in a random graph. In *Random Graphs '87*, pages 73–87, 1990.
- [21] M. Koucky, C. Lautemann, S. Poloczek, and D. Therien. Circuit lower bounds via Ehrenfeucht-Fraïssé games. In *CCC '06: Proceedings of the 21st Annual IEEE Conference on Computational Complexity*, pages 190–201, 2006.
- [22] L. Libkin. *Elements of Finite Model Theory*. Springer-Verlag, 2004.
- [23] J. F. Lynch. A depth-size tradeoff for boolean circuits with unbounded fan-in. In *Structure in Complexity Theory Conference*, pages 234–248, 1986.
- [24] M. Otto. *Bounded variable logics and counting – A study in finite models*, volume 9. Springer-Verlag, 1997.
- [25] B. Poizat. Deux ou trois choses que je sais de L_N . *J. Symb. Log.*, 47(3):641–658, 1982.
- [26] A. A. Razborov. Lower bounds on the monotone complexity of some boolean functions. *Doklady Akademii Nauk SSSR*, 281:798–801, 1985. English translation in Soviet Math. Doklady 31 (1985), 354–357.
- [27] I. Wegener. On the complexity of branching programs and decision trees for clique functions. *J. ACM*, 35(2):461–471, 1988.
- [28] A. C.-C. Yao. Separating the polynomial-time hierarchy by oracles. In *Proc. 26th Annual Symposium on Foundations of Computer Science*, pages 1–10, 1985.