Average-Case Complexity of Detecting Cliques
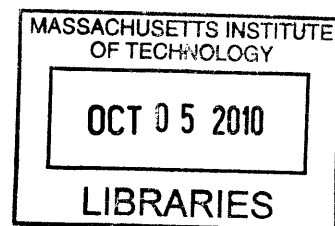
by

Benjamin Rossman

Submitted to the Department of Electrical Engineering and Computer Science
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

September 2010

Author ....
Department of Electrical Engineering and Computer Science
September 3, 2010

Certified by...
Madhu Sudan
Professor of Electrical Engineering and Computer Science
Thesis Supervisor

Accepted by...
Terry P. Orlando
Chairman, Department of Electrical Engineering and Computer Science

Average-Case Complexity of Detecting Cliques

by

Benjamin Rossman

Abstract

The computational problem of testing whether a graph contains a complete subgraph of size k is among the most fundamental problems studied in theoretical computer science. This thesis is concerned with proving *lower bounds* for k-CLIQUE, as this problem is known. Our results show that, in certain models of computation, solving k-CLIQUE in the average case requires $\Omega(n^{k/4})$ resources (moreover, $k/4$ is tight). Here the models of computation are bounded-depth Boolean circuits and unbounded-depth monotone circuits, the complexity measure is the number of gates, and the input distributions are random graphs with an appropriate density of edges. Such random graphs (the well-studied Erdős-Rényi random graphs) are widely believed to be a source of computationally hard instances for clique problems (as Karp suggested in 1976). Our results are the first unconditional lower bounds supporting this hypothesis.

For bounded-depth Boolean circuits, our average-case hardness result significantly improves the previous worst-case lower bounds of $\Omega(n^{k/\text{poly}(d)})$ for depth-d circuits. In particular, our lower bound of $\Omega(n^{k/4})$ has no noticeable dependence on d for circuits of depth $d \leqslant k^{-2} \log n / \log \log n$, thus bypassing the previous "size-depth tradeoffs". As a consequence, we obtain a novel Size Hierarchy Theorem for uniform AC^0. A related application answers a longstanding open question in finite model theory (raised by Immerman in 1982): we show that the hierarchy of bounded-variable fragments of first-order logic is strict on finite ordered graphs. Additional results of this thesis characterize the average-case descriptive complexity of k-CLIQUE through the lens of first-order logic.

Thesis Supervisor: Madhu Sudan
Title: Professor of Electrical Engineering and Computer Science

Acknowledgments

I dedicate this thesis to my parents, David and Lynne Rossman.

Bibliographic note

The main results in this thesis appear in two conference papers:

- "On the constant-depth complexity of k-CLIQUE" [64], containing preliminary versions of the results in Chapter 3, and

- "The monotone complexity of k-CLIQUE on random graphs" [66] (forthcoming), which constitutes Chapter 4.

I thank Christoph Berkholz for pointing out a minor mistake in [64] (fixed in this thesis) and Tomoyuki Hayasaka and Koutaro Nakagawa for helpful feedback on [66]. Chapter 6 of this thesis (on descriptive complexity) includes an unpublished result due to Neil Immerman (Theorem 6.11), as well as the fruits of a collaboration with Joel Spencer (Theorem 6.23).

Contents

Chapter 1

Introduction

The computational problem of testing whether a graph contains a complete subgraph of size k is among the most fundamental problems studied in theoretical computer science. This thesis is concerned with proving *lower bounds* for k-CLIQUE, as this problem is known. That is, our aim is to show that k-CLIQUE cannot be solved with certain limited computational resources.

Understanding the complexity of k-CLIQUE is key to unlocking the relationship between P and NP. In the traditional view where k is part of the input, the problem of detecting k-cliques is famously NP-complete (one Karp's 21 NP-complete problems [45]). In this thesis, however, we consider the problem for fixed but arbitrary values of k. Here the brute-force $O(n^k)$ algorithm places k-CLIQUE in P. A crucial observation is that to separate P from NP, it suffices to show that k-CLIQUE requires time $\Omega(n^{c_k})$ where $c_1, c_2, \ldots$ is any sequence of real numbers which grows to infinity.[1] Conceptually, it might be easier to prove such a sequence of increasing polynomial lower bounds, than a single super-polynomial lower bound (for all k simultaneously), even though the former formally entails the latter. This hope motivates studying the complexity of k-CLIQUE for fixed values of k.

The results of this thesis are unconditional lower bounds of $\Omega(n^{k/4})$ on the complexity of k-CLIQUE in two restricted models of computations (bounded-depth circuits and monotone circuits). Moreover, these lower bounds are achieved in the average-case setting for random graphs with an appropriate density of edges. The next few sections contain additional background on lower bounds in computational complexity (§1.1), clique problems on random graphs (§1.2), and a question arising in descriptive complexity (§1.3). Our results are formally stated in §1.4, followed by a discussion of applications (§1.5) and techniques (§1.6).

1.1 Lower bounds in computational complexity

Lower bounds for the most natural complexity measures (time, space, ...) in the most general models of computation (Turing machines, Boolean circuits, ...) are the holy grail of complexity theory. The famous "barriers" in complexity theory—relativization, natural proofs and algebrization—demonstrate that existing techniques are inadequate for even some modest-sounding goals in complexity theory. Indeed, no one has yet proved a super-linear lower bound on the size of logarithmic-depth circuits computing an explicit problem

---

[1]This hypothesis, known by the equation FPT $\neq$ W[1], is well-studied in parameterized complexity theory [24].

in NP. Since *unconditional* lower bounds in the most *general* settings appear beyond reach, the two main directions in complexity theory are:

- <u>conditional</u> lower bounds (showing that problem X is at least as hard as problem Y via a reduction), and

- unconditional lower bounds in <u>restricted</u> models of computation.

The *conditional hardness* of detecting k-cliques is firmly established: in the view where k is the part of the input, the problem is NP-complete [45] as well as NP-hard to approximate [34, 75]; in the view of parameterized complexity, k-CLIQUE is W[1]-complete [24] and it is known that an upper bound of $n^{o(k)}$ implies the failure of the exponential time hypothesis [20]. In contrast to these conditional hardness results, in this thesis we are after *unconditional* lower bounds for k-CLIQUE in restricted models of computation. Here much less is known. (We discuss some previous results in §1.1.1 and §1.1.2, below. Other results include [13, 29, 72].)

The restricted models of computation that we consider are special classes of Boolean circuits. Recall that Boolean circuits (formally defined in §2.2) are comprised of AND and OR gates (with unbounded fan-in) in addition to NOT gates, on top of $\binom{n}{2}$ variables representing the possible edges in an n-vertex graph. *Size* refers to the number of gates and *depth* refers to the length of the longest path from a variable to the output gate. A circuit is *monotone* if it contains no NOT gates. (To illustrate these definitions, note that k-CLIQUE is computed by a depth-2 monotone circuit (an OR of AND's) of size $\binom{n}{k} + 1$.) Boolean circuits, like Turing machines, are a fully general model of computation. Since Boolean circuits are more combinatorial than Turing machines, they have a natural appeal from the perspective of unconditional lower bounds.

The setting for our results are two restricted classes of Boolean circuits: **bounded-depth circuits** and **monotone circuits**. These two classes of circuits have a storied history in complexity theory, which we summarize in the next two subsections. In contrast to Turing machines and unrestricted Boolean circuits, there has been significant progress over the years toward understanding the limitations of bounded-depth and monotone circuits (see [17] for an excellent survey from 1990, still largely up-to-date).

1.1.1 Bounded-depth circuits

The first major breakthrough on bounded-depth circuits occurred in the early 1980's when Furst, Saxe and Sipser [28] and independently Ajtai [1] proved that PARITY (the problem of determining whether a string of 0's and 1's contains an even number of 1's) is not in the complexity class AC^0 of problems solvable by circuits of polynomial size and constant depth. This result was sharpened by Yao [74] and then Håstad [33], who eventually showed that PARITY requires depth-d circuits of size $2^{\Omega(n^{1/(d-1)})}$ (which is essentially tight). Building on the method of random restrictions, Håstad introduced a tool known as the *switching lemma*, of which many versions presently exist. Switching lemmas (which convert DNF's to much smaller CNF's, and vice-versa, via a random restriction) have played a key role in numerous results on bounded-depth circuits (for example, the LMN Theorem [51]).[2]

---

[2]Another essential technique on bounded-depth circuits, which must be mentioned alongside the switching lemma, is the method of approximation by low-degree polynomials [61, 69].

Lynch [52] in 1986 gave the first lower bound for k-CLIQUE on bounded-depth circuits, showing that it requires depth-d circuits of size $n^{\Omega(k^{1/2}d^{-3/2})}$. This was followed by a lower bound of $n^{\Omega(kd^{-2})}$ proved by Beame [11] in 1990 using a special-purpose switching lemma. Both of these lower bounds exhibit a "size-depth tradeoff" whereby the strength of the bound decreases exponentially as depth increases. At first glance, this appears to be an inescapable limitation of arguments based on a switching lemma.

1.1.2 Monotone circuits

A Boolean function $f : \{0,1\}^m \rightarrow \{0,1\}$ is *monotone* if $f(x) \leqslant f(y)$ for all $x \leqslant y$ (under the standard partial order on $\{0,1\}^m$). It is easy to show that every monotone circuit computes a monotone functions, and conversely, every monotone function is computed by some monotone circuit. It is thus natural to consider the *monotone complexity* of a monotone function f, defined as the size of the smallest monotone circuit computing f.

Razborov [60] in 1985 achieved a major breakthrough by showing that k-CLIQUE has monotone complexity $\Omega((n/\log^2 n)^k)$ (for fixed values of k) and $n^{\Omega(\log n)}$ (when k is a part of the input). Alon and Boppana [2] improved these bounds to $\Omega((n/\log n)^k)$ (for fixed k) and $\Omega((n/\log n)^{1/3})$ (for variable k). Amano and Maruoka [7] extended these bounds to non-monotone circuits with a limited number (up to $(1/6)\log\log n$) of NOT gates. These results, while nearly optimal for the worst-case complexity of k-CLIQUE, leave open the question of the monotone complexity of k-CLIQUE in the average case.

1.2 Average-case lower bounds

It is a well-noted phenomenon that computational problems which may be hard in special cases, turn out to be easy usually or in practice. In complexity theory, this is modeled by the notion of *average-case analysis* (introduced by Levin [49]) whereby the measure of an algorithm's complexity is its expected running time when the input is randomly drawn from a probability distribution representing "typical" instances. Upper bounds for the worst-case performance of an algorithm are stronger than corresponding average-case upper bounds (for any input distribution); conversely, average-case lower bounds (for any input distribution) are stronger than corresponding worst-case lower bounds.

There is another, conceptual argument for pursuing lower bounds in the average-case setting. With rare exceptions, proving lower bounds for a problem X requires understanding something about the nature of X. It is a huge advantage if one can identify inputs distributions where solving X seems to be hard. For any monotone problem (such as k-CLIQUE or k-SAT), there is a canonical guess one can make. For every non-constant monotone function $f : \{0,1\}^m \rightarrow \{0,1\}$, there is a unique $p_f \in [0,1]$ such that for a random input $x \in \{0,1\}^m$ in which each coordinate equals 1 with probability p_f, we have $\Pr[f(x) = 1] = \frac{1}{2}$. This value p_f is known as the *threshold* of f. For k-SAT, it turns out that random inputs at the threshold indeed seem to be a source of hard instances. (This belief is bolstered by work in statistical physics on the complex phase transition of k-SAT.) The results of this thesis support a similar intuition that *random graphs at the threshold are a source of hard instances for k-CLIQUE.*

1.2.1 Karp's question

The idea that random graphs may be a source of hard instances for clique problems goes back to a question raised by Karp [46] in 1976. The random graphs considered here, denoted $G(n, p)$, are graphs on n vertices in which each edge is independent present with probability $p \in [0, 1]$. Such random graph, known as *Erdős-Rényi random graphs*, are well-studied in combinatorics and theoretical computer science (see [4, 15, 41]). Karp considered the problem of *efficiently finding a large clique*, the larger the better, in the balanced random graph $G(n, \frac{1}{2})$. One obvious approach is the following *greedy algorithm*: starting with any vertex v_1, choose any neighbor v_2 of v_1, then choose any common neighbor v_3 of v_1 and v_2, and continue in this manner until $v_1, \ldots, v_t$ are found which have no common neighbor. The output of this algorithm, the set $\{v_1, \ldots, v_t\}$, is a clique in the input graph which is *maximal*, though not necessarily *maximum*. Karp observed that with high probability (w.h.p.) the greedy algorithm, executed in a randomized fashion on $G(n, \frac{1}{2})$, outputs a maximal clique of size $\sim \log n$; this is roughly half the expected maximum clique size $\sim 2 \log n$ (these logarithms have base 2).[3] Karp posed the following question:

Question 1.1. *Is there a polynomial-time algorithm which w.h.p. finds a clique of size $(1 + \varepsilon) \log n$ in the balanced random graph $G(n, \frac{1}{2})$ for any constant $\varepsilon > 0$?*

Karp's question remains wide open today despite receiving considerable attention over the years. Much work has focused on a variant of Question 1.2.2 in which a very large clique is planted in $G(n, \frac{1}{2})$. Kučera [48] showed that a planted clique of size $\Omega(\sqrt{n \log n})$ is likely to show up among the vertices of highest degree and is therefore easy to find. Using spectral techniques, Alon, Krivelevich and Sudakov [3] gave an efficient algorithm that w.h.p. finds a planted clique of size $\Omega(\sqrt{n})$. On the flipside, Jerrum [42] showed that the Metropolis algorithm, a fixed-temperature variant of simulated annealing (basically a random walk in a particular Markov chain on the cliques), fails to find a clique in $G(n, \frac{1}{2})$ of size $(1 + \varepsilon) \log n$ in polynomial time, even if there is a planted clique of size $n^{(1/2)-\varepsilon}$. Different randomized algorithms were similarly shown to fail by Peinado in [56, 57]. In addition, various cryptographic schemes have been proposed (e.g. [43, 48]) based on the presumed hardness of finding planted cliques in $G(n, \frac{1}{2})$.

1.2.2 Scaling Karp's question to $G(n, n^{-\alpha})$

Although most of the work around Karp's question has focused on the balanced random graph $G(n, \frac{1}{2})$, we wish to point out that this question scales nicely to the Erdős-Rényi random graph $G(n, n^{-\alpha})$ where $\alpha > 0$ is a fixed constant.[4] For convenience, we restrict attention to the random graph $G(n, p)$ where $p(n) = n^{-2/(k-1)}$ is a threshold for the existence of k-cliques (that is, $\Pr[G(n, p)$ contains a k-clique] is bounded away from 0 and 1). It is

---

[3]Here we offer an informal explanation. For any distinct vertices $v_1, \ldots, v_t$, the probability that $v_1, \ldots, v_t$ have no common neighbor is $(1 - 2^{-t})^{n-t}$. If $t \leqslant (1 - \varepsilon) \log n$, this probability is $o(1/t)$; so we expect to find common neighbors for all t up to $(1 - \varepsilon) \log n$. If $t \geqslant (1 + \varepsilon) \log n$, this probability is close to 1; so we no longer expect to find a common neighbor. Indeed, the probability $(1 - 2^{-t})^{n-t}$ is balanced for $t \sim \log n$, which explains why the greedy algorithm is likely to find a clique of size $\sim \log n$.

As for the maximum clique having size $\sim 2 \log n$, note that the expected number $\binom{n}{s} 2^{-\binom{s}{2}}$ of s-cliques in $G(n, \frac{1}{2})$ equals 1 for $s \sim 2 \log n$. This gives roughly the right estimate if we pretend that the different s-element sets of vertices are *independently* cliques in $G(n, \frac{1}{2})$. While not truly independent, these events are independent enough for this intuition to be reasonable.

[4]We are not aware if similar observations have been made elsewhere.

not hard to show that w.h.p. the random greedy algorithm on $G(n,p)$ outputs a maximal clique of size approximately $\frac{k}{2}$ (that is, half the size of the maximum clique). Note that the greedy algorithm has linear running time.

What about the complexity of finding a clique of size $(\frac{1}{2}+\varepsilon)k$ for some $\varepsilon \in (0,\frac{1}{2})$? This can be accomplished in time $O(n^{(\frac{1}{2}+\varepsilon)k})$ by brute-force search. The question is, can we do better? In fact, it turns out that we can: by running the greedy algorithm $n^{\varepsilon^2 k + O(1)}$ times, w.h.p. we find a clique of size $(\frac{1}{2}+\varepsilon)k$. This suggests the following question, which can be seen as a version of Question for the random graph $G(n,p)$:

Question 1.2. *For constants $\varepsilon \in (0,\frac{1}{2})$ and $\delta > 0$, is there an $O(n^{(1-\delta)\varepsilon^2 k})$ time algorithm which w.h.p. find a clique of size at least $(\frac{1}{2}+\varepsilon)k$ in $G(n,p)$?*

Similarly, by running the greedy algorithm $n^{k/4 + O(1)}$ times, we are able to determine w.h.p. whether $G(n,p)$ contains a k-clique.[5] Corresponding to the limit case $\varepsilon = \frac{1}{2}$ in Question 1.2, we ask:

Question 1.3. *Does any $O(n^{(1-\delta)k/4})$ time algorithm solve k-CLIQUE w.h.p. on $G(n,p)$?*

A negative answer to Question 1.3 for all fixed k would imply $\mathsf{P} \neq \mathsf{NP}$ (as we noted earlier). Let us not fail to mention that the best known method for solving k-CLIQUE in the *worst case* (due to Nešetřil and Poljak [54]) has running time $n^{(\omega/3)k + O(1)}$ where $n^{\omega + o(1)}$ is the complexity of multiplying two $n \times n$ matrices. With the current best known bound of $\omega < 2.376$ [21], this gives an upper bound of $n^{.792k + O(1)}$ on the worst-case complexity of k-CLIQUE. We point out that, even assuming $\omega = 2$ (which is best possible), the method of solving k-CLIQUE using fast matrix multiplication can only produce an upper bound of $n^{(2/3)k + O(1)}$. This is safely above the hypothetical lower bound of $n^{(1/4)k - o(k)}$ suggested by Question 1.3.

1.3 Bounded-variable logics

This thesis also investigates the descriptive complexity of detecting k-cliques from the perspective of first-order logic. (Recall that *formulas* of first-order logic are built up from expressions of the form $x = y$ and $R(x_1, \ldots, x_r)$, where R comes from a fixed set of relation symbols, via connectives $\land$, $\lor$ and $\neg$ and quantifiers $\exists x$ and $\forall x$. Variables $x, y, \ldots$ range over elements of a structure in which a formula is evaluated.) Descriptive complexity views first-order logic as a model of computation (analogous to Turing machines or Boolean circuits). Formulas of first-order logic are seen as algorithms, and various parameters of formulas (such as length, the nesting depth of quantifiers, the number of variables, ...) are seen as complexity measures. It turns out that the complexity measures for first-order formulas are closely related to the complexity measures (size, depth, ...) for Boolean circuits [32, 23, 36, 37]. (Descriptive complexity studies other systems of logic as well. For additional background, the definitive reference is [38]. Concerning the larger subject of finite model theory, see [25, 30, 50].)

---

[5]There are well-known deterministic algorithms (e.g. the Bron-Kerbosch algorithm [18]) which enumerate all maximal cliques of an arbitrary graph G in time $n^{O(1)}|\{\text{maximal cliques in } G\}|$. This can be highly inefficient for arbitrary graphs, which can have up to $3^{n/3}$ maximal cliques by the Moon-Moser Theorem [53]. However, w.h.p. $G(n,p)$ has only $n^{k/4 + O(1)}$ maximal cliques. Thus, we have another method of solving k-CLIQUE w.h.p. on $G(n,p)$ in time $n^{k/4 + O(1)}$. In Chapter 5, using an idea of Amano [6], we show how to implement such an algorithm on monotone circuits of depth $O(k)$.

The complexity measure that interests us here is the *number of variables* in a first-order formula. It is important that a single variable may be quantified multiple times in a formula. To illustrate this concept, consider the following formula (without requantification) in the first-order language of linearly ordered graphs (with adjacency relation $\sim$ and linear order $<$):

$$\exists x_1 \exists x_2 \exists x_3 \exists x_4 \ (x_1 \sim x_2) \wedge (x_1 < x_2) \wedge (x_2 \sim x_3) \wedge (x_2 < x_3) \wedge (x_3 \sim x_4) \wedge (x_3 < x_4).$$

This formula contains 4 distinct variables and expresses "there exists an increasing path of length at least 4". Note that we can embed quantifiers $\exists x_3$ and $\exists x_4$ to get an equivalent formula:

$$\exists x_1 \exists x_2 \ (x_1 \sim x_2) \wedge (x_1 < x_2) \wedge \exists x_3 \ (x_2 \sim x_3) \wedge (x_2 < x_3) \wedge \exists x_4 \ (x_3 \sim x_4) \wedge (x_3 < x_4).$$

Now consider the following formula with only 2 variables, x and y:

$$\exists x \exists y \ (x \sim y) \wedge (x < y) \wedge \exists x \ (y \sim x) \wedge (y < x) \wedge \exists y \ (x \sim y) \wedge (x < y).$$

Notice that we have substituted x for both x_1 and x_3, and y for both x_2 and x_4. Although this formula would be considered bad style in expository mathematics, the semantics is perfectly unambiguous: it too expresses that there exists an increasing path of length 4. Indeed, for any positive integer ℓ, there is a similar 2-variable formula expressing "there exists an increasing path of length at least ℓ". This simple example demonstrates the expressive power of formulas with only 2 variables on finite ordered graphs.[6]

The collection of formulas with at most m variables (including both free and bound variables) is known as the *m-variable fragment* of first-order logic and denoted L^m. The chain of m-variable fragments $L^1 \subseteq L^2 \subseteq \cdots$ is known as the *variable hierarchy*. Bounded-variable logics are a major subject in finite model theory (see e.g. [31, 55]). One natural question is the following:

Question 1.4. *For which classes of structures is the variable hierarchy strict in terms of expressive power (i.e., for every $m \geqslant 1$ there is a property defined by a formula of L^m but not by any formula of L^{m-1})?*

Question 1.4 has been studied in the context of many different classes of structures. In some cases, the hierarchy is known to be strict; in other cases, it is known to collapse. For instance, the hierarchy is strict on the class of [finite] graphs (where $\sim$ is the only relation); this is seen by the fact that "there exist $\geqslant m$ vertices" can be expressed with m but not $m - 1$ variables. However, on the class of finite linear orders (where $<$ is the only relation), the hierarchy collapses to L^2 (i.e., every first-order property can be expressed using only 2 variables); this is because every first-order property is a finite Boolean combination of properties "there exist $\geqslant m$ vertices", each expressed by a formula with only 2 variables similar to the example above. By more sophisticated results of Poizat [58], the hierarchy collapses to L^3 on the class of [finite] linear orders with any number of unary relations (in addition to $<$). The next natural question along these lines is:

Question 1.5. *Is the variable hierarchy strict on finite ordered graphs?*

---

[6]The linear order is essential in this example. On graphs without a linear order, "there exists a path of length at least ℓ" cannot be expressed with fewer than ℓ variables.

Question 1.5 goes back at least to Immerman [35] in 1982 (who in particular asked whether the property "there exists a k-clique" can be expressed with $< k$ variables in the presence of a linear order). Due to the importance of both ordered structures and bounded-variable logics in finite model theory and descriptive complexity, this question was widely investigated. Dawar [22] in 2005 summarized progress at the time in an article entitled "*How many first-order variables are needed on finite ordered structures?*" Despite some related results (including strictness of a variable hierarchy for *existential* formulas), it was then still unknown whether just 3 variables are sufficient to express every first-order property of finite ordered graphs (for instance, "there exists a clique of size 1000").

1.4   Our results

The main results of this thesis are lower bounds of $\Omega(n^{k/4})$ for the average-case complexity of k-CLIQUE on both bounded-depth Boolean circuits and unbounded-depth monotone circuits. In the following, let k be an arbitrary but fixed integer $\geqslant 5$ and let $p(n) = \Theta(n^{-2/(k-1)})$ (i.e., $p(n)$ is any threshold function for the existence of k-cliques in $G(n,p)$).

- **Lower Bound for k-Clique on Bounded-Depth Circuits** (Theorem 3.1)

  Boolean circuits of size $O(n^{k/4})$ and depth at most $k^{-2} \log n / \log \log n$ cannot solve k-CLIQUE w.h.p. on $G(n,p)$.

- **Lower Bound for k-Clique on Monotone Circuits** (Theorem 4.1)

  Monotone circuits of size $O(n^{k/4})$ cannot solve k-CLIQUE w.h.p. on both $G(n,p)$ and $G(n, p + p^{1+k^{-2}})$.

These lower bounds in fact follow from stronger results, which we describe in §1.6.

Furthermore, we show that the exponent $k/4$ is tight (up to an additive constant) in both of these lower bounds simultaneously.

- **Matching Upper Bound**[7] (Theorem 5.1)

  There exist monotone circuits of size $n^{k/4+O(1)}$ and depth $3k$ which solve k-CLIQUE w.h.p. on $G(n,p)$ for all functions $p : \mathbb{N} \to [0,1]$.

Our results are the first unconditional *average-case* lower bounds for k-CLIQUE on random graphs with an appropriate density of edges.[8] In the setting of monotone circuits, our lower bound of $\Omega(n^{k/4})$ does not improve the best known *worst-case* lower bound of $\Omega((n/\log n)^k)$ [2]. However, the fact that $n^{k/4+O(1)}$ is tight for the average case (by our upper bound) reveals an interesting and perhaps surprising gap between worst-case and average-case complexities of this important problem.

In contrast to the situation for monotone circuits, our lower bound of $\Omega(n^{k/4})$ for bounded-depth circuits *significantly improves* the previous worst-case lower bounds of $n^{\Omega(k/d^2)}$ [11] and $n^{\Omega(\sqrt{k/d^3})}$ [52] for depth-d circuits. In particular, our lower bound has

---

[7]A similar upper bound for *non-monotone* constant-depth circuits was shown by Amano [6].

[8]The notion of "average case" is slightly different in the two results (see the discussion in §1.4). Both results, however, clearly convey the message that k-CLIQUE is hard on random graphs at the threshold.

no noticeable "size-depth tradeoff" for $d \leqslant k^{-2} \log n / \log \log n$. (Instead, we obtain a trade-off of the form $\Omega(n^{-(1-\delta)k/4})$ at depth $\delta k^{-1} \log n / \log \log n$, as we explain in §7.1.1.) This fact has some interesting consequences in both complexity and logic, which we describe in the next section.

Additional results of this thesis characterize the number of variables needed to express the property "there exists a k-clique" in first-order logic in the average case setting (i.e., with high probability on $G(n, p)$), both with a linear order and without a linear order.

- **Number of Variables to Express k-Clique in the Average Case** (Theorems 6.16 and 6.23)

  With a linear order, $k/4$ variables are necessary. Without a linear order, $k/2$ variables are necessary and $k/2 + O(1)$ variables are sufficient.

The $k/4$ variable lower bound bound in the presence of a linear order follows from our bounded-depth circuit lower bound. In the absence of a linear order, the stronger $k/2$ variable lower bound and $k/2 + O(1)$ variable upper bound rely on subtle results of Shelah and Spencer [67] on the "almost-sure theory" of $G(n, p)$.

1.5 Applications in complexity and logic

One immediate corollary of our bounded-depth circuit lower bound is the following:

- **Size Hierarchy Theorem**

  The hierarchy of complexity classes uniform $\mathsf{AC}^0(\text{size } O(n^s))$, *parameterized by $s \geqslant 1$, is infinite.*

The proof of this result is that k-CLIQUE is computed by uniform depth-2 circuits of size $O(n^k)$, but not by constant-depth circuits of size $O(n^{k/4})$ (by our lower bound).[9] The only previous result in this direction, due to Chaudhuri and Radhakrishnan [19], showed that not every function in uniform AC^0 has linear-size constant-depth circuits. However, it had been open whether every function in uniform AC^0 has constant-depth circuits of size $O(n^{1+\varepsilon})$ for some constant $\varepsilon > 0$.[10] (For context, we add that an analogous Depth Hierarchy Theorem is a classic result of Sipser [68].)

Another (closely related) corollary is:

- **Variable Hierarchy Theorem** (Corollary 6.12)

  The hierarchy $L^1 \subseteq L^2 \subseteq \cdots$ of bounded-variable fragments of first-order logic is strict in terms of expressive power on finite order graphs.

The context and significance of this result was already discussed in §1.3. Let us only mention that this result follows from a combination of our bounded-depth circuit lower bound, which implies $L^m \neq L^{4m}$ (i.e., this hierarchy is infinite), and an argument due to Neil Immerman showing $L^{m-1} = L^m \implies L^m = L^{m+1}$.

---

[9] Amano [6] subsequently separated all levels of this hierarchy, by applying our technique to obtain lower bounds for the AC^0 complexity of the k-clique problem on ℓ-uniform hypergraphs (see §7.1.3).

[10] Uniformity is essential here, as a counting argument shows that for all $\beta > \alpha > 0$, there exist functions computed by depth-2 circuits of size n^β that cannot be computed by circuits of size n^α (of any depth). This observation is attributed in [19] to Eric Allender.

1.6 Techniques

The lower bounds stated in §1.4 in fact derive from stronger results, which better illustrate our techniques. Let $f : \{n\text{-vertex graphs}\} \to \{0,1\}$ be a Boolean function on n-vertex graphs (really, a sequence of such functions for $n = 1, 2, \dots$).

- **Stronger Lower Bound for Bounded-Depth Circuits** (Theorem 3.2)

  Suppose f is computed by Boolean circuits of size $O(n^{k/4})$ and depth at most $k^{-2} \log n / \log \log n$. Let $G \sim G(n,p)$ and $K_k \sim \text{Plant}(n, K_k)$ (that is, a k-clique planted uniformly at random among n vertices). Then w.h.p. $f(G) = f(G \cup K_k)$.

- **Stronger Lower Bound for Monotone Circuits** (Theorem 4.3)

  Suppose f is computed by monotone circuits of size $O(n^{k/4})$. Let $G^- \sim G(n, p^{1+k^{-2}})$ and $K_k \sim \text{Plant}(n, K_k)$. If $\text{E}[f(K_k)] = 1 - o(1)$, then $\text{E}[f(G^-)] = 1 - \exp(-\Omega(n^{k^{-2}}))$.

It is not very hard to show that these results imply the lower bounds stated in §1.4.

This result on bounded-depth circuits is proved by a novel method. There are two components:

(1) a new notion of "sensitivity" (the *f-sensitive subgraph*, Definition 3.10) together with a result (Proposition 3.13) that broadly generalizes the well-known fact that polynomial-size bounded-depth circuits have low average-sensitivity, and

(2) a novel inductive argument on circuits (Lemmas 3.17 and 3.25) which shows that, for a circuit to have "large" sensitivity, some node in the circuit must have "medium" sensitivity.

Like other results on bounded-depth circuits, part (1) relies on Håstad's switching lemma [33]. However, due to part (2), the structure of our proof is notably different from previous k-CLIQUE lower bounds [11, 52] where the use of a switching lemma is the main thrust in the proof.

Our result on monotone circuits follows the general framework of Razborov's approximation method [60] (similar to previous lower bounds for k-CLIQUE on monotone circuits [2, 7, 60]). An essential new ingredient in the proof—and technical contribution of independent interest—is a novel variant of sunflowers, called *quasi-sunflowers* (Definition 4.9), in which petals may overlap, but appear "disjoint" on average. We prove a "quasi-sunflower lemma" (Theorem 4.11) along the lines of the Erdős-Rado sunflower lemma [26]. Since quasi-sunflowers naturally generalize a useful property of sunflowers, we believe there is broad potential for applications.

Finally, among new techniques in the thesis, we mention two very different methods (described in §3.4 and §4.7) for converting number-of-wires lower bounds into number-of-gates lower bounds with only a very small loss.

Chapter 2

Definitions and Preliminaries

This chapter presents the basic notation and definitions for this thesis. This is mostly standard, with a few notable exceptions (relating to terminology "graphs" and "patterns") which we summarize below:

- (§2.3) As a matter of terminology, we set up a useful distinction between *graphs* (which have vertex set $\{1, \ldots, n\}$ by default) and *patterns* (defined as constant-size graphs, independent of n, with no isolated vertices).

- (§2.3.1) The *threshold exponent* of a pattern P, denoted $\theta(P)$, is defined as the minimum of $|V_{P_0}|/|E_{P_0}|$ over subpatterns P_0 of P.

- (§2.4.1) $\mathrm{Plant}(n, P)$ denotes the random planted copy of P on n vertices, viewed as a random graph with vertex set $\{1, \ldots, n\}$.

- (§2.5) We define a useful classification of graphs and patterns into *small*, *medium* and *large*. (For example, "medium graph" has the following technical meaning: a graph with $\geqslant \frac{k}{2}$ non-isolated vertices which is the union of two graphs with $< \frac{k}{2}$ non-isolated vertices.)

2.1   Basic notation

Throughout this thesis, k is an arbitrary but fixed integer $\geqslant 5$. $\mathbb{N}$ is the set of nonnegative integers, n is an arbitrary nonnegative integer, and $[n]$ is the set $\{1, \ldots, n\}$. Nearly all asymptotic statements in this thesis refer to growing n. Expressions *with high probability* (*w.h.p.*) and *almost surely* mean with probability tending to 1 as $n \to \infty$. We have attempted to ensure that the hidden constants in this asymptotic notation ($O(\cdot)$, $o(\cdot)$, $\Omega(\cdot)$, etc.) are universal (and in particular independent of k, though this might not always be the case).

We have the following notation for sets: $|X|$ denotes the cardinality X. $\wp(X)$ denotes the power set of X. For $t \in \mathbb{N}$, $\binom{X}{t}$ denotes the set of t-element subsets of X. $X \bigtriangleup Y$ denotes the symmetric difference of X and Y, that is, the union of set-theoretic differences $X \setminus Y$ and $Y \setminus X$.

$\log(\cdot)$ denotes the base-2 logarithm and $\ln(\cdot)$ denotes the natural logarithm.

2.2 Circuits

In this thesis *circuits* are Boolean circuits with unbounded fan-in. We make the simplifying assumption that all NOT gates are at input level. (This assumption at most doubles the size of circuits without increasing depth. Any loss in our bounds is therefore eaten by the O's and Ω's.) Formally, a *circuit* C on variables $x_1, \ldots, x_N$ is a finite acyclic directed graph in which:

- each source (node of in-degree 0, also called an *input*) is labelled by either a variable x_i, its negation $\neg x_i$, the constant 0 or 1,

- each non-input node (also called a *gate*) is labelled by either AND and OR, and

- a subset of nodes are designated as *outputs*.

For nodes ν and μ in C, we say that μ is a *child* of ν if there is a directed edge (also called a *wire*) from μ to ν. Children(ν) denotes the set of all children of ν. The *height* of a node ν is the length of the longest path from an input to ν.

Each node ν in a circuit C computes a Boolean function $\{0,1\}^N \to \{0,1\}$ (defined in the natural way, that is, inductively according to the label of ν and the functions computed by its children). As a matter of notation, we identify ν with the function it computes by for instance writing $\nu(x)$ for the value computed by ν on $x \in \{0,1\}^N$. The circuit C itself computes a function $\{0,1\}^N \to \{0,1\}^{\{\text{outputs of C}\}}$. When $N = \binom{n}{2}$, we view C as computing a function of n-vertex graphs (via the natural bijection between $\{n$-vertex graphs$\}$ and $\{0,1\}^{\binom{n}{2}}$).

A circuit is *monotone* if no input node is labeled by a negated variable. Every monotone circuit clearly computes a monotone function $\{0,1\}^N \to \{0,1\}$ (and conversely, every monotone function is computed by some monotone circuit).

2.2.1 Circuit parameters

The following circuit parameters are studied in this thesis:

- size $=$ the number of gates,

- wires $=$ the number of wires $(= \sum_{\nu \in C} |\text{Children}(\nu)|)$,

- depth $=$ the length of the longest path from an input to an output,

- fan-in $=$ the maximum number of children of a node.

Observation 2.1. For every circuit C, we have

$$\text{size}(C) \leqslant \text{wires}(C) \leqslant \text{size}(C) \cdot \text{fanin}(C) \leqslant \text{size}(C) \cdot (\text{size}(C) + 2(\# \text{ of variables}) - 1).$$

Thus, for instance, a lower bound of $\Omega(n^{20})$ on wires implies a lower bound of $\Omega(n^{10})$ on size.

2.2.2 Complexity

For a Boolean function $f : \{0,1\}^n \to \{0,1\}$, the *Boolean complexity* of f is the size of the smallest circuit which computes f. If f is monotone, the *monotone complexity* of f is the size of the smallest monotone circuit which computes f.[1]

AC^0 denotes the complexity class of (sequences of) Boolean functions computed by (sequences of) polynomial-size constant-depth circuits. By default, we consider the *non-uniform* version of AC^0.

2.3 Graphs and patterns

Graphs in this thesis are finite simple graphs. Formally, a *graph* is a pair $G = (V_G, E_G)$ where V_G is a finite set and $E_G \subseteq \binom{V_G}{2}$. $\mathscr{G}^n$ denotes the set of graphs with vertex set $[n]$. By default a **graph** is an <u>element of $\mathscr{G}^n$</u>. By distinction, a **pattern** is a constant-size graph with <u>no isolated vertices</u>.

The distinction between graphs and patterns will be convenient throughout this thesis. One important difference is that we care about graphs *up to equality*, whereas we care about patterns *up to isomorphism*. We will often make asymptotic statements involving a fixed pattern P and a random graph $G \in \mathscr{G}^n$ where n is growing.

On both graphs and patterns, $\cup$ is the union operation and $\subseteq$ is the subgraph/subpattern relation. For a pattern P, a graph H is a *P-subgraph* of G if $H \subseteq G$ and the induced pattern on the non-isolated vertices of H is isomorphic to P. The number of P-subgraphs of G is denoted $\mathsf{sub}(P, G)$.

For $\ell \in \mathbb{N}$, K_ℓ denotes the complete pattern with vertex set $\{1, \dots, \ell\}$ and edge set $\binom{\{1,\dots,\ell\}}{2}$. $\varnothing$ denotes both the empty pattern (with no vertices) and the empty graph (with no edges).

2.3.1 Threshold exponent

The *threshold exponent* of a nonempty pattern P, denoted $\theta(P)$, is defined by

$$\theta(P) = \min_{P_0 \subseteq P} \frac{|V_{P_0}|}{|E_{P_0}|}.$$

As an important example, note that $\theta(K_k) = 2/(k-1)$. For the empty pattern $\varnothing$, we set $\theta(\varnothing) = \infty$.[2] (The threshold exponent $\theta(P)$ is closely related to the threshold probability $p(n)$ for the event that the random graph $\mathrm{G}(n,p)$ contains a P-subgraph. This connection is made precise by Lemma 2.3.)

2.3.2 Graph functions, monotonicity and minterms

In this thesis, "graph function" refers to any function with domain $\mathscr{G}^n$.[3] A graph function with range $\{0,1\}$ is said to be *Boolean*. A *monotone graph function* is a Boolean graph

---

[1]The gap between the Boolean and monotone complexities of a monotone function is known to be exponential in some instances [71].

[2]In the random graph literature (for instance [4, 15, 41]), it is more common to find results stated in terms of *maximum average degree* defined by $m(P) = 2\max_{P_0 \subseteq P} |E_{P_0}|/|V_{P_0}|$. Note that $\theta(P) = 2/m(P)$.

[3]According to our usage, graph functions need not be isomorphism-invariant (contrary to a common definition in the literature).

function f such that $f(G) \leqslant f(H)$ for all graphs $G \subseteq H$.

For a graph function f and a graph G, we denote by f^G the graph function defined by $f^G(H) = f(G \cup H)$. (Since this notation is not standard, we will remind the reader before using it.)

A graph H is a *minterm* of monotone graph function f if $f(H) = 1$ and $f(H') = 0$ for every proper subgraph $H' \subset H$. For a pattern P, a minterm H is a *P-minterm* if the induced pattern on the non-isolated vertices of H is isomorphic to P. The set of minterms (resp. P-minterms) of f is denoted $\mathcal{M}(f)$ (resp. $\mathcal{M}(f, P)$).

2.4   Probability

We consistently represent random objects by boldface symbols ($\boldsymbol{G}$, $\boldsymbol{W}$, etc.). For a set X and $p \in [0,1]$, notation $\boldsymbol{W} \subseteq_p X$ expresses that $\boldsymbol{W}$ is a random subset of X where each $x \in X$ belongs to $\boldsymbol{W}$ independently with probability p.

2.4.1   Random graphs

A *random graph* is a random variable $\boldsymbol{G}$ taking values in $\mathscr{G}^n$. We also refer to distribution of $\boldsymbol{G}$ as a "random graph". An important class are the *Erdős-Rényi random graphs*. For $p : \mathbb{N} \to [0,1]$, we denote by $\boldsymbol{G} \sim \mathrm{G}(n,p)$ the random graph in which each element of $\binom{[n]}{2}$ is an edge independently with probability $p(n)$ (that is, $E_{\boldsymbol{G}} \subseteq_p \binom{[n]}{2}$).

Another important class of random graphs are the *random planted patterns*. For a pattern P, we denote $\boldsymbol{H} \sim \mathrm{Plant}(n, P)$ by the random graph with edge set $\{\{\boldsymbol{\pi}(v), \boldsymbol{\pi}(w)\} : \{v, w\} \in E_P\}$ where $\boldsymbol{\pi}$ is a uniform random one-to-one function from V_P to $[n]$. In other words, $\boldsymbol{H}$ is uniformly distributed among graphs H such that the induced pattern on the non-isolated vertices of H is isomorphic to P. As an important special case, $\boldsymbol{K}_k \sim \mathrm{Plant}(n, K_k)$ denotes the random planted k-clique.

We will often consider the union $\boldsymbol{G} \cup \boldsymbol{H}$ of an Erdős-Rényi random graph $\boldsymbol{G} \sim \mathrm{G}(n,p)$ and a random planted pattern $\boldsymbol{H} \sim \mathrm{Plant}(n, P)$.

2.4.2   Background lemmas

We state two background lemmas, whose proofs can be found in any of the books [4, 15, 41]. First, a simple calculation of the expected number of P-subgraphs of $\mathrm{G}(n, p)$.

Lemma 2.2. *For every pattern P,*

$$\mathsf{E}[\mathsf{sub}(P, \mathrm{G}(n,p))] = \frac{|V_P|!}{|\{automorphisms\ of\ P\}|} \binom{n}{|V_P|} p^{|E_P|}.$$

In particular, if $p = \Theta(n^{-\alpha})$ for constant $\alpha > 0$ then

$$\mathsf{E}[\mathsf{sub}(P, \mathrm{G}(n,p))] = \Theta(n^{|V_P| - \alpha |E_P|}).$$

The next lemma justifies calling $\theta(P)$ the "threshold exponent" of pattern P.

Lemma 2.3. *For every pattern P, the function $n^{-\theta(P)}$ is a threshold function for event that $G(n,p)$ contains a P-subgraph. That is,*

$$\lim_{n\to\infty} \Pr\left[G(n,p) \text{ contains a } P\text{-subgraph}\right] = \begin{cases} 0 & \text{if } p(n) = o(n^{-\theta(P)}), \\ 1 & \text{if } p(n) = \omega(n^{-\theta(P)}). \end{cases}$$

In particular,

$$p(n) = o(n^{-2/(k-1)}) \implies \text{w.h.p. } G(n,p) \text{ is } k\text{-clique-free,}$$
$$p(n) = \omega(n^{-2/(k-1)}) \implies \text{w.h.p. } G(n,p) \text{ has a } k\text{-clique.}$$

To state the next background lemma, let $\mathrm{Pois}(\lambda)$ denote the Poisson distribution with mean λ (recall that $\Pr[\mathrm{Pois}(\lambda) = t] = \lambda^t e^{-\lambda}/t!$ for $t \in \mathbb{N}$) and let $d_{\mathrm{TV}}(\cdot,\cdot)$ denote *total variation distance* ($= 1/2$ the ℓ_1-distance between two distributions/random variables). In particular, for random graphs $\boldsymbol{G}_1, \boldsymbol{G}_2 \in \mathscr{G}^n$,

$$d_{\mathrm{TV}}(\boldsymbol{G}_1, \boldsymbol{G}_2) = \frac{1}{2} \sum_{H \in \mathscr{G}^n} \left| \Pr[\boldsymbol{G}_1 = H] - \Pr[\boldsymbol{G}_2 = H] \right|.$$

For random variables $\boldsymbol{X}_1, \boldsymbol{X}_2$ supported on $\mathbb{N}$ (such as $\mathrm{Pois}(\lambda)$),

$$d_{\mathrm{TV}}(\boldsymbol{X}_1, \boldsymbol{X}_2) = \frac{1}{2} \sum_{i \in \mathbb{N}} \left| \Pr[\boldsymbol{X}_1 = i] - \Pr[\boldsymbol{X}_2 = i] \right|.$$

The following lemma concerns random graphs $G(n,p)$ where $p(n) = \Theta(n^{-2/(k-1)})$ is a threshold function for the existence of k-cliques. It includes the fact that the number of k-cliques in $G(n,p)$ is asymptotically Poisson.

Lemma 2.4. *Denote by $\kappa(G)$ the number of k-cliques in a graph G. Fix $c > 0$ and let $\boldsymbol{G} \sim G(n, cn^{-2/(k-1)})$ and $\boldsymbol{K}_k \sim \mathrm{Plant}(n, K_k)$ and $\boldsymbol{X} \sim \mathrm{Po}(c^{\binom{k}{2}}/k!)$. For $t \in \mathbb{N}$, let $\boldsymbol{G}_t \sim G(n, cn^{-2/(k-1)})$ conditioned on $\kappa(\boldsymbol{G}_t) = t$. Then*

$$d_{\mathrm{TV}}(\kappa(\boldsymbol{G}), \boldsymbol{X}) = o(1),$$
$$d_{\mathrm{TV}}(\boldsymbol{G}_{t+1}, \boldsymbol{G}_t \cup \boldsymbol{K}_k) = o(1),$$
$$\lim_{n\to\infty} d_{\mathrm{TV}}(\kappa(\boldsymbol{G}), \kappa(\boldsymbol{G} \cup \boldsymbol{K}_k)) = d_{\mathrm{TV}}(\boldsymbol{X}, \boldsymbol{X} + 1) < 1.$$

2.4.3 Janson's inequality

The following probabilistic inequality is due to [39] (see also Ch. 2 of [41] and Ch. 8 of [4]).

Lemma 2.5 (Janson's Inequality). *Let $\mathcal{F}$ be a nonempty family of subsets of X. Let $\boldsymbol{W}$ be a random subset of X such that events $x \in \boldsymbol{W}$ are mutually independent for $x \in X$ (for example, $\boldsymbol{W} \subseteq_p X$). Define μ and Δ by*

$$\mu = \sum_{U \in \mathcal{F}} \Pr\left[U \subseteq \boldsymbol{W}\right],$$

$$\Delta = \sum_{\substack{U, V \in \mathcal{F}: \\ U \neq V, \, U \cap V \neq \emptyset}} \Pr\left[U \cup V \subseteq \boldsymbol{W}\right].$$

23

$$\text{Then} \qquad \Pr\left[\bigwedge_{U \in \mathcal{F}} U \nsubseteq \boldsymbol{W}\right] \leqslant \exp\left(-\min\left\{\frac{\mu}{2}, \frac{\mu^2}{2\Delta}\right\}\right).$$

The following lemma states that in the highly supercritical regime where $p(n) = n^{\Omega(1)-\theta(P)}$, the random graph $\mathrm{G}(n,p)$ is extremely likely to have at least half its expected number of P-subgraphs. This concentration-of-measure result follows from a "lower tail" version of Janson's inequality (see Theorem 2.14 of [41]).

Lemma 2.6. *Let P be a pattern and let $\boldsymbol{G} \sim \mathrm{G}(n,p)$ where $p(n) = n^{\Omega(1)-\theta(P)}$. Then*

$$\Pr\left[\mathsf{sub}(P,\boldsymbol{G}) \leqslant \tfrac{1}{2} \, \mathrm{E}[\mathsf{sub}(P,\boldsymbol{G})]\right] = \exp(-n^{\Omega(1)}).$$

2.5 Small, medium, large

Let $\boldsymbol{G} \sim \mathrm{G}(n, \Theta(n^{-2/(k-1)}))$ be a random graph at a threshold function for containing k-cliques. It is instructive to calculate the expected number of ℓ-cliques in $\boldsymbol{G}$ for $\ell \in \{0, \dots, k\}$:

$$\mathrm{E}[\# \text{ of } \ell\text{-cliques in } \boldsymbol{G}] = \Theta\left(n^{\ell - \frac{2}{k-1}\binom{\ell}{2}}\right).$$

Letting $\lambda = \ell/k$, we have

$$\ell - \frac{2}{k-1}\binom{\ell}{2} = \lambda(1-\lambda)k + O(1).$$

Note that $\lambda(1-\lambda)k$ is maximal with value $k/4$ for $\lambda = 1/2$. (Indeed, $\ell - \frac{2}{k-1}\binom{\ell}{2}$ is maximal for $\ell \in \{\lfloor k/2 \rfloor, \lceil k/2 \rceil\}$.)

The fact that $\boldsymbol{G}$ has many cliques of "intermediate" size $\sim k/2$ and few cliques of size $\leqslant \varepsilon k$ or $\geqslant (1-\varepsilon)k$ for small $\varepsilon > 0$ motivates the following definition. (The large number of "intermediate" subgraphs plays an important part in our lower bounds.)

Definition 2.7. *A pattern P is:*

- *small*   *if $|V_P| < k/2$,*

- *medium*   *if $|V_P| \geqslant k/2$ and there exist small patterns P_1, P_2 such that $P = P_1 \cup P_2$,*

- *large*   *otherwise.*

A graph is small, medium *or* large *according to the induced pattern on its non-isolated vertices. That is, a graph G is:*

- *small*   *if it has $< k/2$ non-isolated vertices,*

- *medium*   *if it has $\geqslant k/2$ non-isolated vertices and is the union of two small graphs,*

- *large*   *otherwise.*

A key fact to keep in mind is that the union of two small patterns/graphs is small or medium (but never large). Note that the complete pattern K_ℓ is small if $\ell < k/2$ and large otherwise (but never medium). An important example of medium pattern is

$$P = K_{\lceil k/2 \rceil} - \{\text{a single edge}\}.$$

Note that P is the union of two overlapping copies of the small pattern $K_{\lceil k/2 \rceil - 1}$. In fact, this pattern P gives the optimal bound in the following lemma.

Lemma 2.8. *For every medium pattern P,*

$$|V_P| - \frac{2}{k-1}|E_P| \geqslant \frac{k+1}{4} + \frac{2}{k-1}.$$

Proof. Let P be a medium pattern which minimizes $|V_P| - \frac{2}{k-1}|E_P|$. By definition of medium, P is the union of two small patterns P_1 and P_2. We can assume that P_1 and P_2 are complete, since we only decrease $|V_{P_1 \cup P_2}| - \frac{2}{k-1}|E_{P_1 \cup P_2}|$ by replacing P_1 and P_2 with the (also small) complete patterns with the same vertices. Let $a = |V_P|$, $b = |V_{P_1}|$, $c = |V_{P_2}|$ and note that

$$|V_P| - \frac{2}{k-1}|E_P| = a - \frac{2}{k-1}\left(\binom{b}{2} + \binom{c}{2} - \binom{b+c-a}{2}\right).$$

First, suppose $k = 2t + 1$ is odd. Integers a, b, c satisfy $1 \leqslant b, c \leqslant t$ and $t + 1 \leqslant a \leqslant b + c$. Relaxing integrality, let α, β, γ be reals minimizing $\alpha - \frac{1}{t}\left(\binom{\beta}{2} + \binom{\gamma}{2} - \binom{\beta+\gamma-\alpha}{2}\right)$ subject to $1 \leqslant \beta, \gamma \leqslant t$ and $t + 1 \leqslant \alpha \leqslant \beta + \gamma$. Note that $\beta = \gamma$ since, if not, by replacing β and γ with their mean $(\beta + \gamma)/2$ we reduce the objective function while still satisfying the constraints. Thus, our task becomes minimizing the function $f(\alpha, \beta)$ defined by

$$f(\alpha, \beta) = \alpha + \frac{1}{t}\binom{2\beta - \alpha}{2} - \frac{2}{t}\binom{\beta}{2}$$

subject to $1 \leqslant \beta \leqslant t$ and $t + 1 \leqslant \alpha \leqslant 2\beta$. Since $\frac{d}{d\alpha}f(\alpha, \beta) > 0$ and $\frac{d}{d\beta}f(\alpha, \beta) < 0$ for all α, β satisfying these constraints, it follows that $\alpha = t + 1$ and $\beta = t$. Therefore,

$$|V_P| - \frac{2}{k-1}|E_P| \geqslant f(t+1, t) = \frac{t+1}{2} + \frac{1}{t} = \frac{k+1}{4} + \frac{2}{k-1}.$$

In the case where k is even, we get

$$|V_P| - \frac{2}{k-1}|E_P| \geqslant \frac{k+1}{4} + \frac{9}{4(k-1)} > \frac{k+1}{4} + \frac{2}{k-1}$$

by a similar calculation. $\qquad\square$

25

Chapter 3

Lower Bound for Bounded-Depth Circuits

In this chapter we prove lower bounds on the average-case complexity of k-CLIQUE for bounded-depth circuits. For convenience, we state our results in terms of the random graph $G(n, n^{-2/(k-1)})$. However, these results hold just the same for the random graph $G(n, p)$ where $p(n)$ is any function which is $\Theta(n^{-2/(k-1)})$. Our main result is the following lower bound:

Theorem 3.1. *Boolean circuits of size $O(n^{k/4})$ and depth at most $k^{-2} \log n / \log\log n$ cannot solve k-CLIQUE w.h.p. on $G(n, n^{-2/(k-1)})$.*

Theorem 3.1 follows from the even stronger result that such circuits almost surely fail to distinguish between a random graph $G \sim G(n, n^{-2/(k-1)})$ and the graph G with a random planted k-clique.

Theorem 3.2. *Suppose $f : \mathscr{G}^n \to \{0,1\}$ is computed by Boolean circuits of size $O(n^{k/4})$ and depth at most $k^{-2} \log n / \log\log n$. Let $G \sim G(n, n^{-2/(k-1)})$ and $K_k \sim \mathrm{Plant}(n, K_k)$. Then w.h.p. $f(G) = f(G \cup K_k)$.*

Before giving an overview of the proof of Theorem 3.2, we show how it implies Theorem 3.1.

Proof of Theorem 3.2 $\implies$ Theorem 3.1. Suppose that $f : \mathscr{G}^n \to \{0,1\}$ agrees with k-CLIQUE w.h.p. on $G \sim G(n, n^{-2/(k-1)})$ (that is, w.h.p. $f(G) = 1$ if and only if G contains a k-clique). Then the following hold:

- $\Pr[f(G) = 1]$ is within $o(1)$ of $\Pr[G$ contains a k-clique], which is bounded away from 1 (since $n^{-2/(k-1)}$ is a threshold function for k-CLIQUE, see Lemma 2.3). Therefore, $\Pr[f(G) = 1]$ is also bounded away from 1.

- Lemma 2.4 implies that $\Pr[f(G \cup K_k) = 1]$ is within $o(1)$ of $\Pr[f(G) = 1 \mid G$ contains a k-clique]. Since f agree with k-CLIQUE w.h.p. on G and $\Pr[G$ contains a k-clique] is bounded away from 0, it follows that $\Pr[f(G) = 1 \mid G$ contains a k-clique$] = 1 - o(1)$. Therefore, $\Pr[f(G \cup K_k) = 1] = 1 - o(1)$.

Assuming Theorem 3.2, it follows that f is not computed by Boolean circuits of size $O(n^{k/4})$ and depth at most $k^{-2} \log n / \log\log n$. Therefore, Theorem 3.1 holds. $\qquad\square$

The proof of Theorem 3.2 uses a novel argument (which differs significantly from standard arguments on bounded-depth circuits where a switching lemma is applied repeatedly). An informal sketch of the proof follows. Let C be a circuit of size $O(n^{k/4})$ and depth at most $k^{-2} \log n / \log \log n$. Let G and K_k be as in Theorem 3.2. For each node ν in the circuit C, we consider a particular subgraph of K_k called the ν^G-*sensitive subgraph* (defined formally in §3.2) whose edges represent the variables on which the Boolean function {subgraphs of K_k} $\to \{0,1\}$ defined by $\nu^G(H) = \nu(G \cup H)$ depends. Using a technical lemma on random restrictions (Lemma 3.7 in §3.1), we bound the probability that the ν^G-sensitive subgraph has a given pattern.[1] One consequence of this bound is that the ν^G-sensitive subgraph is medium with probability $o(n^{-k/4})$.[2] By a union bound, it follows that w.h.p. the ν^G-sensitive subgraph is not medium for any node ν in the circuit C. By a novel inductive argument on circuits (Lemma 3.17 in §3.3 and Lemma 3.25 in §3.4), we conclude that w.h.p. C has the same value on inputs G and $G \cup K_k$.

In fact, we present two version of this argument. A preliminary version of the argument (given in §3.3) is easier to state, but only produces a lower bound on the *number of wires*. The actual proof of Theorem 3.2 (given in §3.4) is more complicated, but produces the stated lower bound on *size* (i.e., the number of gates).

3.1 A lemma on random restrictions

In this section, we prove a technical lemma on the decision-tree depth of Boolean functions computed by small bounded-depth circuits when subject to random restrictions. The result, while new, follows the well-establish technique of repeatedly applying the famous switching lemma of Håstad [33].

We recall some standard definitions. A *decision tree* is rooted binary tree in which leaves are labelled either 0 or 1, interior nodes are labelled by Boolean-valued variables, and the edges between an interior node and its two children are labelled 0 and 1 respectively. A decision tree computes a Boolean function in the natural way: given an assignment of variables to $\{0,1\}$, we follow a branch of the decision tree (starting from the root) according the value of the variables we encounter (i.e., value $i \in \{0,1\}$ mean we follow the edge labelled i) and we output the value of the leaf where we end up. For a Boolean function $f : \{0,1\}^N \to \{0,1\}$, the *decision-tree depth* of f, denoted $\mathsf{DT}_{\mathsf{depth}}(f)$, is the height of the shortest tree-decision that computes f.

Observation 3.3. Note that a Boolean function with decision-tree depth D depends on at most 2^D variables.

A *restriction* on N variables is a function $\rho : \{1, \ldots, N\} \to \{0, 1, *\}$ where

- $\rho(i) = 0$ (resp. $\rho(i) = 1$) means that variable x_i is set to 0 (resp. 1),

- $\rho(i) = *$ means that variable x_i is left unassigned.

For a function $f : \{0,1\}^N \to X$ (where X is any set), we can apply ρ to f to get a function $f\lceil_\rho : \{0,1\}^{\rho^{-1}(*)} \to X$ (defined in the natural way).

---

[1] Here *pattern* means the isomorphism type of the non-isolated part of a graph (see §2.3).

[2] Recall that a graph is *medium* if it has $\geqslant k/2$ non-isolated vertices, but is the union of two graphs with $< k/2$ non-isolated vertices (see §2.5).

Definition 3.4. *For $q, p \in [0, 1]$, we denote by $\mathcal{R}(q, p)$ the random restriction ρ where $\rho(i)$ are independent such that*

$$\Pr[\rho(i) = *] = q, \qquad \Pr[\rho(i) = 1] = (1 - q)p, \qquad \Pr[\rho(i) = 0] = (1 - q)(1 - p).$$

The following lemma is a restatement of Håstad's original switching lemma [33] in terms of decision trees (as opposed to DNFs and CNFs).

Lemma 3.5 (Switching Lemma). *Suppose Boolean function f is an AND or OR of (arbitrary many) depth-r decision trees. Then for all $q \in [0, 1]$ and $r \in \mathbb{N}$,*

$$\Pr_{\rho \sim \mathcal{R}(q, 1/2)} \left[\mathsf{DT}_{\mathsf{depth}}(f\lceil_\rho) > r \right] \leqslant (5qr)^r.$$

Remark 3.6. The original result of [33] is slightly stronger. Rather than decision-tree depth, it speaks about r-DNFs and s-CNFs, that is, propositional formulas in disjunctive normal form (resp. conjunctive normal form) with clauses of width r (resp. width s). The precise statement is: if f is equivalent to an r-DNF, then for $\rho \sim \mathcal{R}(q, 1/2)$, $f\lceil_\rho$ is not equivalent to an s-DNF with probability $\leqslant (5pr)^s$. Lemma 3.5 is a special case of this statement (with $r = s$), since a function with decision-tree depth r is equivalent to both an r-DNF and an r-CNF.

We now present the technical lemma that we will need (a fairly straightforward corollary of Lemma 3.5).

Lemma 3.7. *Let $p \in [0, \frac{1}{2}]$ and $q \in [0, 1]$ and $c \geqslant 5$ and $t \geqslant 1$. Suppose Boolean function f is computed by a circuit of size S and depth at most $(-\log q)/(tc + \log \log S)$. Then*

$$\Pr_{\rho \sim \mathcal{R}(pq, p)} \left[\mathsf{DT}_{\mathsf{depth}}(f\lceil_\rho) > \frac{\log S}{c} \right] = S^{1-t}.$$

Proof. Let $d = \mathsf{depth}(\mathsf{C})$ and generate a sequence $\rho_0, \ldots, \rho_d$ of random restrictions as follows:

- let $\rho_0 \sim \mathcal{R}(p, \lambda)$ applied to variables of C where $\lambda \in [0, 1]$ is a value to be determined,

- for $i = 1, \ldots, d$, let $\rho_i \sim \mathcal{R}(q^{1/d}, 1/2)$ applied to the variables left unrestricted by $\rho_0, \ldots, \rho_{i-1}$.

For $i \in \{0, \ldots, d\}$, let ρ^i denote the composition of restrictions $\rho_0, \ldots, \rho_i$ (defined in the natural way).

We have
$$\Pr\left[\rho^d = *\right] = \Pr\left[\rho_0 = \cdots = \rho_d = *\right] = p(q^{1/d})^d = pq$$

and

$$\Pr\left[\boldsymbol{\rho}^d = 1 \mid \boldsymbol{\rho}^d \neq *\right] = \Pr\left[\boldsymbol{\rho}^d = 1\right] / \Pr\left[\boldsymbol{\rho}^d \neq *\right]$$

$$= \frac{1}{1-pq}\left(\Pr\left[\boldsymbol{\rho}_0 = 1\right] + \sum_{i=1}^{d}\Pr\left[\boldsymbol{\rho}_d = 1 \text{ and } \boldsymbol{\rho}_0 = \cdots = \boldsymbol{\rho}_{d-1} = *\right]\right)$$

$$= \frac{1}{1-pq}\left((1-p)\lambda + \sum_{i=1}^{d}\frac{pq^{(i-1)/d}(1 - q^{1/d})}{2}\right)$$

$$= \frac{1}{1-pq}\left((1-p)\lambda + \frac{p(1-q)}{2}\right).$$

We now set $\lambda = (1-p)^{-1}\left(p(1-pq) - \frac{1}{2}p(1-q)\right)$ (and check that indeed $\lambda \in [0,1]$), so that $\boldsymbol{\rho}^d \sim \mathcal{R}(pq, p)$.

For each node ν of height h in C, let $\boldsymbol{X}_\nu$ be the event that $\mathsf{DT}_{\mathsf{depth}}(\nu\lceil_{\boldsymbol{\rho}^h}) \leqslant (\log S)/c$. If ν is a variable (i.e., if ν has height 0), then $\boldsymbol{X}_\nu$ holds with probability 1. If ν is a gate at height $h \geqslant 1$, then

$$\Pr\left[\neg\boldsymbol{X}_\nu \;\middle|\; \bigwedge_{\text{children } \mu \text{ of } \nu} \boldsymbol{X}_\mu\right]$$

$$= \Pr\left[\mathsf{DT}_{\mathsf{depth}}(\nu\lceil_{\boldsymbol{\rho}^h}) > \frac{\log S}{c} \;\middle|\; \bigwedge_{\text{children } \mu \text{ of } \nu} \mathsf{DT}_{\mathsf{depth}}(\mu\lceil_{\boldsymbol{\rho}^{h-1}}) \leqslant \frac{\log S}{c}\right]$$

$$= \Pr\left[\mathsf{DT}_{\mathsf{depth}}((\nu\lceil_{\boldsymbol{\rho}^{h-1}})\lceil_{\boldsymbol{\rho}_h}) > \frac{\log S}{c} \;\middle|\; \bigwedge_{\text{children } \mu \text{ of } \nu} \mathsf{DT}_{\mathsf{depth}}(\mu\lceil_{\boldsymbol{\rho}^{h-1}}) \leqslant \frac{\log S}{c}\right]$$

$$\leqslant \left(\frac{5q^{1/d}\log S}{c}\right)^{(\log S)/c} \qquad \text{(by Lemma 3.5)}$$

$$= \left(q^{(\log S)/(-\log q)}\right)^{(tc+\log\log S)/c}\left(\log S\right)^{(\log S)/c}\left(\frac{5}{c}\right)^{(\log S)/c}$$

$$= S^{-(tc+\log\log S)/c}S^{(\log\log S)/c}\left(\frac{5}{c}\right)^{(\log S)/c}$$

$$\leqslant S^{-t}.$$

It follows that

$$\Pr\left[\mathsf{DT}_{\mathsf{depth}}(f\!\upharpoonright_\rho) > \frac{\log S}{c}\right] = \Pr\left[\neg \boldsymbol{X}_{\mathsf{output\ gate}}\right]$$

$$\leqslant \Pr\left[\bigvee_{\mathsf{gates}\ \nu} \neg \boldsymbol{X}_\nu\right]$$

$$\leqslant \sum_{\mathsf{gates}\ \nu} \Pr\left[\neg \boldsymbol{X}_\nu \ \middle|\ \bigwedge_{\mathsf{children}\ \mu\ \mathsf{of}\ \nu} \boldsymbol{X}_\mu\right]$$

$$\leqslant S^{1-t}. \qquad \qquad \square$$

Remark 3.8. Lemma 3.7 is the only place in our proof of Theorem 3.2 where the depth of the circuit matters. That is, the remainder of the proof is "depth-independent". Interestingly, Lemma 3.7 itself is "size-independent" in a certain sense: even though our main results concern circuits of size $O(n^{k/4})$, when invoking Lemma 3.7, we only need to assume that the circuits in question have size $O(n^c)$ for some constant $c > 0$ independent of k. This separation into depth-independent and size-independent portions appears to be a new feature of our proof. This separation also explains how Theorem 3.2 breaks out of the traditional "size-depth tradeoff".

3.2 The f-sensitive subgraph

In this section, we introduce a key concept in the proof of Theorem 3.2: the f-*sensitive subgraph* of graph H. After giving the definition, we list some basic properties of the f-sensitive subgraph in Lemma 3.12. We then prove a key result (Proposition 3.13) about f^G-sensitive subgraphs where f is computed by polynomial-sze bounded-depth circuits and G is a random graph (this result uses Lemma 3.7, our technical lemma concerning random restrictions).

Lemma 3.9. *For every graph function f and graph H, there is a unique minimal graph T such that $f(H') = f(H' \cap T)$ for every $H' \subseteq H$.*

Proof. Let $\mathcal{T}$ be the class of all T such that $f(H') = f(H' \cap T)$ for every $H' \subseteq H$. To show that $\mathcal{T}$ has a unique minimal element, it suffices to show that it is nonempty and closed under intersection. It is nonempty since $H \in \mathcal{T}$. It is closed under intersection since for all $T_1, T_2 \in \mathcal{T}$ and $H' \subseteq H$, we have $f(H') = f(H' \cap T_1 \cap T_2)$ as $f(H') = f(H' \cap T_1)$ (since $T_1 \in \mathcal{T}$) and $f(H' \cap T_1) = f(H' \cap T_1 \cap T_2)$ (since $T_2 \in \mathcal{T}$). $\square$

Definition 3.10 (f-sensitive subgraphs and f-cores). *For a graph function f and graph H, we denote by $\mathbb{T}(f, H)$ the unique minimal graph T such that $f(H') = f(H' \cap T)$ for every $H' \subseteq H$. We call $\mathbb{T}(f, H)$ the f-sensitive subgraph of H. If $\mathbb{T}(f, H) = H$, then we say that H is an f-core.*[3]

The following example illustrates these definitions.

---

[3]It would be more descriptive, but also excessively wordy, to say that H is an "f-sensitive core" or is "fully f-sensitive".

Example 3.11. Let f be the k-CLIQUE function. That is, f is the Boolean graph function defined by $f(G) = 1$ if and only if G contains a k-clique. Then $\mathbb{T}(f, G)$ is the union of all k-cliques in G, and G is an f-core if and only if every every edge in G belongs to a k-clique.

The next lemma lists some elementary properties of $\mathbb{T}(f, H)$. A proof is omitted, since these properties all follow easily from definitions.

Lemma 3.12 (Properties of $\mathbb{T}(f, H)$).

1. The edges of $\mathbb{T}(f, H)$ are precisely $e \in E_H$ such that there exist $H', H'' \subseteq H$ satisfying $E_{H'} \triangle E_{H''} = \{e\}$ and $f(H') \neq f(H'')$.

2. $\mathbb{T}(f, H)$ is an f-core (that is, $\mathbb{T}(f, H) = \mathbb{T}(f, \mathbb{T}(f, H))$, or equivalently $\mathbb{T}(f, \cdot)$ is an idempotent operator on graphs).

3. $f(H) = f(\mathbb{T}(f, H))$.

4. $\mathbb{T}(f, H') \subseteq \mathbb{T}(f, H)$ for all $H' \subseteq H$ (that is, $\mathbb{T}(f, \cdot)$ is a monotone operator on graphs).

5. The union of f-cores is an f-core.

6. If $f_1, \ldots, f_t$ are graph functions such that f is completely determined by the values of $f_1, \ldots, f_t$ (e.g., $f_1, \ldots, f_t$ are Boolean and $f = \text{AND}_{i=1}^{t} f_i$), then $\mathbb{T}(f, H) \subseteq \mathbb{T}(f_1, H) \cup \cdots \cup \mathbb{T}(f_t, H)$.

Recall a piece of notation introduced in §2.3.2: for any graph function f and graph G, we denote by f^G the graph function defined by $f^G(H) = f(G \cup H)$. The next proposition is our main technical result on f-sensitivity (actually, $f^{\boldsymbol{G}}$-sensitivity for a random graph $\boldsymbol{G}$). The proof uses Lemma 3.7 on random restrictions (from the previous section).

Proposition 3.13. Fix a pattern P and functions $p, q : \mathbb{N} \to [0, \frac{1}{2}]$ such that $p(n)q(n) = n^{\Omega(1) - \theta(P)}$. Suppose $f : \mathscr{G}^n \to \{0, 1\}^{n^{o(1)}}$ is computed by circuits of size $n^{O(1)}$ and depth $(-\log q)/(\omega(1) + \log \log n)$ (with $n^{o(1)}$ output gates). Then

$$\Pr_{\substack{\boldsymbol{G} \sim \mathrm{G}(n,p) \\ \boldsymbol{H} \sim \mathrm{Plant}(n,P)}} \left[\boldsymbol{H} \text{ is an } f^{\boldsymbol{G}}\text{-core} \right] \leqslant \frac{n^{o(1)}}{\mathbb{E}\left[\mathsf{sub}(P, \mathrm{G}(n, pq))\right]}.$$

Remark 3.14. Proposition 3.13 can be viewed a generalization of the fact—essentially given by the special case where P is a single edge—that polynomial-size bounded-depth circuits have low average sensitivity.[4]

Proof. Let $\boldsymbol{Q} \sim \mathrm{G}(n, pq)$ be random graph (independent of $\boldsymbol{G}$ and $\boldsymbol{H}$) and let

$$\boldsymbol{\mathcal{E}} = \left\{ e \in E_{\boldsymbol{Q}} : \exists \text{subgraphs } Q', Q'' \subseteq \boldsymbol{Q} \text{ with } E_{Q'} \triangle E_{Q''} = \{e\} \text{ and } f(\boldsymbol{G} \cup Q') \neq f(\boldsymbol{G} \cup Q'') \right\}.$$

That is, $\boldsymbol{\mathcal{E}}$ is a particular random subset of $\binom{[n]}{2}$ which depends on $\boldsymbol{G}$ and $\boldsymbol{Q}$ (but is independent of $\boldsymbol{H}$).

---

[4]The *average sensitivity* of a Boolean function $f : \{0, 1\}^N \to \{0, 1\}$ is the expectation, over uniform random $\boldsymbol{x} \in \{0, 1\}^N$, of the number of coordinates $i \in [N]$ such that $f(\boldsymbol{x}) \neq f(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_{i-1}, 1-\boldsymbol{x}_i, \boldsymbol{x}_{i+1}, \ldots, \boldsymbol{x}_N)$. Boolean functions computed by polynomial-size depth-d circuits are known to have average sensitivity $O(\log^{d-1} N)$ [16].

We first outline the plan of the proof. The first step is defining a coupling $\widetilde{H}$ of H which is independent of G (though it depends on Q). The idea is to select $\widetilde{H}$ uniformly from the P-subgraphs of Q. (We can assume that $\mathsf{sub}(P,Q)$ is at least half its expected value, since this happens with probability $1 - \exp(-n^{\Omega(1)})$ by Lemma 2.6 as $pq = n^{\Omega(1)-\theta(P)}$.) As desired, $\widetilde{H}$ has distribution $\mathrm{Plant}(n,P)$ (forgetting about Q) and is moreover independent of G. Our goal now is to bound the probability that $\widetilde{H}$ (rather than H) is an f^G-core. We next observe that $\widetilde{H}$ can only be an f^G-core in the event that $E_{\widetilde{H}} \subseteq \mathcal{E}$. (This follows directly from the definitions of $\mathbb{T}(f^G,\cdot)$ and the set $\mathcal{E}$.) Thus, it suffices to show that $|\mathcal{E}| = n^{o(1)}$ (i.e., for every constant $\varepsilon > 0$, w.h.p. $|\mathcal{E}| \leqslant n^{\varepsilon}$), since it follows that at most $\binom{|\mathcal{E}|}{|E_P|} = n^{o(1)}$ P-subgraphs of Q are f^G-cores. Finally, we show that $|\mathcal{E}| = n^{o(1)}$ by first noting that $\mathcal{E}$ is precisely the set of variables on which the function $f\lceil_\rho$ depends, where ρ is a particular random restriction (determined by G and Q) with distribution $\mathcal{R}(pq,p)$. We then apply Lemma 3.7 to bound the decision-tree depth $\mathsf{DT}_{\mathsf{depth}}(f_i\lceil_\rho)$ of the coordinate functions $f_1,\ldots,f_m$. This leads to a bound on $|\mathcal{E}|$, since each element of $\mathcal{E}$ is a variable on which some f_i depends.

We will now make this argument precise. Let $\varepsilon > 0$ be a fixed, but arbitrary, constant. Let $\checkmark_1$ be the event that $\mathsf{sub}(P,Q) \geqslant \frac{1}{2}\,\mathrm{E}[\mathsf{sub}(P,Q)]$ and let $\checkmark_2$ be the event that $|\mathcal{E}| \leqslant n^{\varepsilon/|E_P|}$. Generate a random graph $\widetilde{H} \sim \mathrm{Plant}(n,H)$ (which will be independent of G and H, but dependent on Q) as follows: if $\checkmark_1$ holds, let $\widetilde{H}$ be a uniform random P-subgraph of Q; otherwise, let $\widetilde{H} \sim \mathrm{Plant}(n,P)$ (independent of everything else). Observe that (G,H) and $(G,\widetilde{H})$ have exactly the same joint distribution. It follows that

$$\Pr\left[H \text{ is an } f^G\text{-core}\right] = \Pr\left[\widetilde{H} \text{ is an } f^G\text{-core}\right]$$
$$\leqslant \Pr\left[\widetilde{H} \text{ is an } f^G\text{-core} \mid \checkmark_1, \checkmark_2\right] + \Pr[\neg\checkmark_1] + \Pr[\neg\checkmark_2].$$

The remainder of the proof consists of two claims:

(i) $\Pr\left[\widetilde{H} \text{ is an } f^G\text{-core} \mid \checkmark_1, \checkmark_2\right] \leqslant \dfrac{n^{\varepsilon}}{\frac{1}{2}\,\mathrm{E}[\mathsf{sub}(P,Q)]}$,

(ii) $\checkmark_1$ and $\checkmark_2$ fail with negligible probability (i.e., $\Pr[\neg\checkmark_1 \vee \neg\checkmark_2] = n^{-\omega(1)}$).

Since ε can be chosen arbitrarily small and $\mathrm{E}[\mathsf{sub}(P,\mathrm{G}(n,pq))] = n^{\Omega(1)}$, the result follows from (i) and (ii), which we now prove.

Proof of (i): Conditioned on $\checkmark_1$ and $\checkmark_2$, note that $\widetilde{H}$ is uniformly distributed among the P-subgraphs of Q. If $\checkmark_1$ holds and $\widetilde{H}$ is an f^G-core, then $E_{\widetilde{H}} \subseteq \mathcal{E}$. If $\checkmark_2$ holds, then there are at most $\binom{\mathcal{E}}{|E_P|} \leqslant \binom{n^{\varepsilon/|E_P|}}{|E_P|} \leqslant n^{\varepsilon}$ P-subgraphs H of Q such that $E_H \subseteq \mathcal{E}$. It follows that

$$\Pr\left[\widetilde{H} \text{ is an } f^G\text{-core} \mid \checkmark_1, \checkmark_2\right] \leqslant \frac{|\{P\text{-subgraphs } H \text{ of } Q \text{ with } E_H \subseteq \mathcal{E}\}|}{\mathsf{sub}(P,Q)} \text{ given } \checkmark_1, \checkmark_2$$
$$\leqslant \frac{n^{\varepsilon}}{\frac{1}{2}\,\mathrm{E}[\mathsf{sub}(P,Q)]}.$$

Proof of (ii): By Lemma 2.6, not only is $\Pr[\neg\checkmark_1]$ negligible (i.e., $n^{-\omega(1)}$), but in fact $\Pr[\neg\checkmark_1] = \exp(-n^{\Omega(1)})$. Thus, we have only to show that $\Pr[\neg\checkmark_2]$ negligible.

Let $f_1, \ldots, f_m : \mathscr{G}^n \to \{0,1\}$ be the coordinate functions of $f : \mathscr{G}^n \to \{0,1\}^m$ where $m = n^{o(1)}$. Let S and d be the size and depth of the circuit computing f. We assume that $S = n^{\Theta(1)}$ (since $S = n^{O(1)}$ by assumption and we can pad the circuit by adding extraneous gates in case $S = n^{o(1)}$).

Define $\boldsymbol{\rho} : \binom{[n]}{2} \to \{0,1,*\}$ by

$$\rho(x_e) = \begin{cases} * & \text{if } e \in E_{\boldsymbol{Q}}, \\ 1 & \text{if } e \in E_{\boldsymbol{G}} \setminus E_{\boldsymbol{Q}}, \\ 0 & \text{otherwise.} \end{cases}$$

Note that $\boldsymbol{\rho}$ is random restriction with distribution $\mathcal{R}(pq, p)$. Let

$$c = \frac{\log S}{(\varepsilon/|E_P|)\log n - \log m} + 5.$$

Note that $5 \leqslant c \leqslant O(1)$ and

$$d = \frac{\log(1/q)}{\omega(1) + \log\log n} \leqslant \frac{\log(1/q)}{\omega(1)\cdot c + \log\log S}.$$

We may therefore apply Lemma 3.7: for every $i \in \{1, \ldots, m\}$, we have

$$\Pr\left[\mathsf{DT}_{\mathsf{depth}}(f_i{\restriction}_{\boldsymbol{\rho}}) > \frac{\varepsilon\log n}{|E_P|} - \log m\right] \leqslant \Pr\left[\mathsf{DT}_{\mathsf{depth}}(f_i{\restriction}_{\boldsymbol{\rho}}) > \frac{\log S}{c}\right] = S^{-\omega(1)}.$$

Moreover, this $\omega(1)$ is the same for every $i \in \{1, \ldots, m\}$. To complete the proof, we show that $\Pr[\neg\checkmark_2]$ is negligible as follows:

$$\begin{aligned} \Pr[\neg\checkmark_2] &= \Pr\left[f{\restriction}_{\boldsymbol{\rho}} \text{ depends on} > n^{\varepsilon/|E_P|} \text{ variables from } \boldsymbol{\rho}^{-1}(*)\right] \\ &\leqslant \Pr\left[\bigvee_{i \in \{1,\ldots,m\}} f_i{\restriction}_{\boldsymbol{\rho}} \text{ depends on} > \frac{n^{\varepsilon/|E_P|}}{m} \text{ variables}\right] \\ &\leqslant \sum_{i \in \{1,\ldots,m\}} \Pr\left[\mathsf{DT}_{\mathsf{depth}}(f_i) > \frac{\varepsilon\log n}{|E_P|} - \log m\right] \quad \text{(using Obs. 3.3)} \\ &= mS^{-\omega(1)} \\ &= n^{-\omega(1)} \quad \text{(since } S = n^{\Theta(1)} \text{ and } m = n^{o(1)}\text{).} \qquad \square \end{aligned}$$

The next lemma applies Proposition 3.13 to the case where P is a small or medium pattern and $p(n)$ is a threshold function for k-CLIQUE.

Lemma 3.15. *Let P be a fixed small or medium pattern and let $p(n) = \Theta(n^{-2/(k-1)})$. Suppose that $f : \mathscr{G}^n \to \{0,1\}^{n^{o(1)}}$ is computed by circuits of size $n^{O(1)}$ and depth at most $k^{-2}\log n/\log\log n + O(1)$. Then*

$$\Pr_{\substack{\boldsymbol{G}\sim\mathrm{G}(n,p) \\ \boldsymbol{H}\sim\mathrm{Plant}(n,P)}}\left[\boldsymbol{H} \text{ is an } f^{\boldsymbol{G}}\text{-core}\right] = \begin{cases} O(n^{-1}) & \text{if } P \text{ is nonempty and small,} \\ O(n^{-\frac{k}{4}-\frac{1}{k}}) & \text{if } P \text{ is medium.} \end{cases}$$

34

Proof. Let $q(n) = n^{-(k^{-2}+k^{-3})}$. Note that $p(n)q(n) = \Theta(n^{\frac{2}{k-1}-(k^{-2}+k^{-3})}) = n^{\Omega(1)-\theta(P)}$ since

$$\theta(P) \geqslant \theta(K_{k-1}) = \frac{2}{k-2} > \frac{2}{k-1} + k^{-2} + k^{-3}.$$

using the fact that K_{k-1} contains every small and medium pattern up to isomorphism.[5] Also note that

$$\frac{k^{-2}\log n}{\log\log n} + O(1) \leqslant \frac{-\log q}{\omega(1) + \log\log n}.$$

The circuit computing f, together with $p(n)$ and $q(n)$, thus satisfy the hypotheses of Proposition 3.13 with respect to the pattern P. We have

$$\Pr\left[\boldsymbol{H} \text{ is an } f^{\boldsymbol{G}}\text{-core}\right] \leqslant \frac{n^{o(1)}}{\mathrm{E}\left[\mathsf{sub}(P, \mathrm{G}(n, pq))\right]} \qquad \text{(by Proposition 3.13)}$$

$$= \frac{n^{o(1)}}{n^{|V_P|}(pq)^{|E_P|}} \qquad \text{(by Lemma 2.2)}$$

$$= n^{-|V_P|+(\frac{2}{k-1}+k^{-2}+k^{-3})|E_P|+o(1)}$$

$$\leqslant n^{-|V_P|+\frac{2}{k-1}|E_P|+\frac{1}{4}+\frac{1}{4k}+o(1)}. \qquad \text{(since } |E_P| \leqslant k^2/4\text{).}$$

(The fact that $|E_P| \leqslant k^2/4$ follows from the observation that among small and medium patterns, the union of two disjoint $\lfloor\frac{k-1}{2}\rfloor$-cliques has the most edges.) To prove the lemma, it therefore suffices to show

$$|V_P| - \frac{2}{k-1}|E_P| - \frac{1}{4} - \frac{1}{4k} > \begin{cases} 1 & \text{if } P \text{ is nonempty and small,} \\ \frac{k}{4} + \frac{1}{k} & \text{if } P \text{ is medium.} \end{cases}$$

Among nonempty small patterns P, $|V_P| - \frac{2}{k-1}|E_P| - \frac{1}{4} - \frac{1}{4k}$ is maximal with value $\frac{7}{4} - \frac{2}{k-1} - \frac{1}{4k} > 1$ (when P is a single edge). For medium P, we have $|V_P| - \frac{2}{k-1}|E_P| \geqslant \frac{k+1}{4} + \frac{2}{k-1}$ by Lemma 2.8, from which $|V_P| - \frac{2}{k-1}|E_P| - \frac{1}{4} - \frac{1}{4k} > \frac{k}{4} + \frac{1}{k}$ follows. $\square$

3.3 Preliminary result: lower bound on wires

In this section, we prove a weaker preliminary version of Theorem 3.2 (using a simpler argument that will serve a warm-up to the proof of Theorem 3.2 in §3.4). The difference between Proposition 3.16, below, and Theorem 3.2 is that "number of wires" replaces "size" (i.e., "number of gates").

Proposition 3.16. *Suppose* $f : \mathscr{G}^n \to \{0,1\}$ *is computed by circuits with* $O(n^{k/4})$ *wires and depth at most* $k^{-2}\log n/\log\log n$. *Let* $\boldsymbol{G} \sim \mathrm{G}(n, n^{-2/(k-1)})$ *and* $\boldsymbol{K}_k \sim \mathrm{Plant}(n, K_k)$. *Then w.h.p.* $f(\boldsymbol{G}) = f(\boldsymbol{G} \cup \boldsymbol{K}_k)$.

The proof of Proposition 3.13 uses a simple but seemingly new inductive argument on circuits.

---

[5] In fact, $\theta(P) > (1 + \frac{1}{\sqrt{2}})\frac{2}{k-1}$ for every small or medium P (see Lemma 5.2 of [64]).

Lemma 3.17 (First circuit induction). *Let* C *be a circuit with maximum fan-in 2 computing a Boolean graph function* f. *Suppose* H *is a graph such that for every gate* ν *in* C, $\mathbb{T}(\nu, H)$ *is not medium. Then* $\mathbb{T}(f, H)$ *is small.*

(In the next section we give a second, more sophisticated version (Lemma 3.25) of this inductive argument, which does not require the fan-in 2 restriction.)

Proof. We argue by induction on ν that $\mathbb{T}(\nu, H)$ is small for every node ν in C (in particular, for the output node of C whereby $\mathbb{T}(f, H)$ is small). Consider first the base case where ν is an input node (labelled by a constant 0 or 1 or a variable x_e or its negation $\neg x_e$ for some edge $e \in \binom{[n]}{2}$). Note that $\mathbb{T}(\nu, H)$ has at most one edge, so it is small, as required.

For the induction step, suppose ν is an AND or OR gate with children μ_1 and μ_2 and assume that $\mathbb{T}(\mu_1, H)$ and $\mathbb{T}(\mu_2, H)$ are small. Since the value of ν is completely determined by the values of μ_1 and μ_2, we have $\mathbb{T}(\nu, H) \subseteq \mathbb{T}(\mu_1, H) \cup \mathbb{T}(\mu_2, H)$ (by Lemma 3.12(6)). Since $\mathbb{T}(\nu, H)$ is the union of two small graphs, it is either small or medium (by Definition 2.7 of "small" and "medium"). By assumption, $\mathbb{T}(\nu, H)$ is not medium. Therefore, $\mathbb{T}(\nu, H)$ is small. $\qquad\square$

We are ready to prove Proposition 3.16. The proof will exploit that fact a circuit C with W wires is equivalent—via a transformation which does not increase the number of *alternations* between AND and OR gates—to a circuit C′ with size $\leqslant W$ and fan-in 2. This fact lets us apply Lemma 3.17.

Proof of Proposition 3.16. Consider the circuit computing f and replace each AND / OR gate with a binary tree of AND / OR gates (i.e., for each original gate with fan-in m, we get $m - 1$ new gates of the same type with fan-in 2). In this way, we obtain an equivalent circuit C which also computes f and has size $O(n^{k/4})$. Although the depth of C is no longer bounded, note that the function computed at each gate of C is computed by a circuit of size $O(n^{k/4})$ and depth $k^{-2} \log n / \log \log n$ (we simply collapses C back down to depth $k^{-2} \log n / \log \log n$ by combining adjacent gates of the same kind). We may therefore apply Lemma 3.15 to the gates of the circuit C.

Consider an arbitrary gate ν in C. We have

$$\Pr\left[\mathbb{T}(\nu^G, K_k) \text{ is medium}\right] \leqslant \sum_{\substack{\text{medium patterns } P \\ \text{(up to isomorphism)}}} \Pr\left[\mathbb{T}(\nu^G, K_k) \text{ is a } P\text{-subgraph of } K_k\right]$$

$$\leqslant \sum_{\text{medium } P} \mathrm{E}\left[\# \text{ of } P\text{-subgraphs of } K_k \text{ which are } \nu^G\text{-cores}\right]$$

$$= \sum_{\text{medium } P} \mathsf{sub}(P, K_k) \Pr_{H \sim \mathrm{Plant}(n, P)}\left[H \text{ is a } \nu^G\text{-core} \mid H \subseteq K_k\right]$$

$$= \sum_{\text{medium } P} \mathsf{sub}(P, K_k) \Pr_{H \sim \mathrm{Plant}(n, P)}\left[H \text{ is a } \nu^G\text{-core}\right].$$

(The last equality is due to the obvious independence of the event that H is a $\widetilde{\nu}^G$-core and the event that $H \subseteq K_k$.) Note that

$$\sum_{\text{medium } P} \mathsf{sub}(P, K_k) = |\{\text{medium subpatterns of } K_k\}| \leqslant 2^{k^2}.$$

36

By Lemma 3.15, we have

$$\Pr_{H \sim \mathrm{Plant}(n,P)} \left[H \text{ is a } \nu^G\text{-core} \right] = O(n^{-\frac{k}{4}-\frac{1}{k}}).$$

Putting these inequalities together (along with the fact that $2^{k^2} = o(n^{1/k})$), we have

$$(*) \qquad\qquad \Pr \left[\mathbb{T}(\nu^G, K_k) \text{ is medium} \right] = o(n^{-k/4}).$$

A similar calculation shows

$$(**) \qquad\qquad \Pr \left[\mathbb{T}(\nu^G, K_k) \text{ is nonempty and small} \right] = o(1).$$

Apply a union bound over the gates of C to $(*)$, it follows that w.h.p. $\mathbb{T}(\nu^G, K_k)$ is not medium for any gate ν in C. Lemma 3.17 therefore implies that w.h.p. $\mathbb{T}(f^G, K_k)$ is small. It follows from $(**)$ (applied to the output gate of C) that w.h.p. $\mathbb{T}(f^G, K_k)$ is not both nonempty and small. Therefore, w.h.p. $\mathbb{T}(f^G, K_k)$ is the empty graph $\varnothing$. This means that w.h.p. $f^G(\mathbb{T}(f^G, K_k)) = f^G(K_k)$, or equivalently, $f(G \cup K_k) = f(G)$. $\qquad\square$

Remark 3.18. This $\omega(n^{k/4})$ lower bound on the number of wires implies a corresponding bound of $\omega(n^{k/8})$ on size (quadratically worse than Theorem 3.2), since a circuit with $\omega(n^{k/4})$ wires has $\omega(n^{k/8})$ gates (see Obs. 2.1).

3.4 Main result: lower bound on size

In the previous section, we proved Proposition 3.16, which serves as a warm-up for the proof of our main result, Theorem 3.2, in this section. The argument here is somewhat more complicated. First, we introduce a variant of the f-sensitive subgraph:

Definition 3.19. *For a graph function f and a graph H, we denote by $\mathbb{S}(f, H)$ the unique minimal graph S such that $f(H') = f(H' \cap S)$ for every small or medium $H' \subseteq H$.*[6]

Remark 3.20. Analogous to Lemma 3.9, to see that $\mathbb{S}(f, H)$ is well-defined, consider the set $\mathcal{S}$ of graphs S such that $f(H') = f(H' \cap S)$ for every small or medium $H' \subseteq H$. We have $H \in \mathcal{S}$, so it suffices to show that $\mathcal{S}$ is closed under intersection. For $S_1, S_2 \in \mathcal{S}$ and small or medium $H \subseteq H$, we have $f(H') = f(H' \cap S_1)$ (since $S_1 \in \mathcal{S}$) and $f(H' \cap S_1 \cap S_2)$ (since $S_2 \in \mathcal{S}$ and $H' \cap S_1$ is also small-or-medium).

Observation 3.21. $\mathbb{S}(f, H)$ is the union of f-sensitive subgraphs $\mathbb{T}(f, H')$ over small-or-medium subgraphs H' of H.

The following example shows that $\mathbb{S}(f, H)$ can be a proper subgraph of $\mathbb{T}(f, H)$.

Example 3.22. Let f be the k-CLIQUE function, i.e., $f(G) = 1$ iff $\mathsf{sub}(K_k, G) \geqslant 1$. Suppose H is k-clique. Then $\mathbb{S}(f, H)$ is the empty graph, since $f(H') = 0$ for every small or medium $H' \subseteq H$. At the same time, we have $\mathbb{T}(f, H) = H$ (see Example 3.11).

The next lemma gives a crucial property of $\mathbb{S}(f, H)$.

---

[6]To avoid defining an important concept without a name, we propose calling $\mathbb{S}(f, H)$ the *f-s.m.-sensitive subgraph* of H where "s.m." stands for "small or medium". We will not, however, use this name (notation $\mathbb{S}(f, H)$ being sufficient).

Lemma 3.23. *If $\mathbb{S}(f, H)$ is not small, then some medium subgraph of H is an f-core.*

Proof. Suppose $\mathbb{S}(f, H)$ is not small. Let $H_1, \ldots, H_t$ enumerate the small and medium subgraphs of H. For $i \in \{0, \ldots, t\}$, let $T_i = \mathbb{T}(f, H_1) \cup \cdots \cup \mathbb{T}(f, H_i)$. Note that $\mathbb{S}(f, H) = T_t$ (by Obs. 3.21). Let $j \geqslant 2$ be the least index such that T_j is not small. Since $T_j = T_{j-1} \cup \mathbb{T}(f, H_j)$ and both T_{j-1} and $\mathbb{T}(f, H_j)$ are small, T_j is medium. Moreover, T_j is an f-core (since the union of f-cores is an f-core by Lemma 3.12(5)). $\square$

A second key concept we must introduce for the proof of Theorem 3.2 the *value/witness function* associated with a node in a circuit.

Definition 3.24 (Value/witness function of a node). *Let C be a circuit whose nodes are given in some linear order. That is, for each node ν, we can arrange the children of ν from "left" to "right". For each node ν, we define a function $\widetilde{\nu}$ with values in $\{0, 1\} \cup \mathsf{Children}(\nu)$ as follows.*

- *$\widetilde{\nu}(x) = \nu(x)$ in the following cases: ν is an input node, or ν is an AND gate and $\nu(x) = 1$, or ν is an OR gate and $\nu(x) = 0$.*

- *If ν is an AND gate (resp. OR gate) and $\nu(x) = 0$ (resp. $\nu(x) = 1$), then $\widetilde{\nu}(x)$ is the minimal (i.e., "leftmost") child $\mu \in \mathsf{Children}(\nu)$ such that $\mu(x) = 0$ (resp. $\mu(x) = 1$).*

We call $\widetilde{\nu}$ the *value/witness function* of ν, since it not only encodes the value $\nu(x)$, but in the cases of an AND gate with value 0 or an OR gate with value 1, the function $\widetilde{\nu}$ identifies a particular witness among its children. The following lemma plays a similar role to Lemma 3.17 in the previous section.

Lemma 3.25 (Second circuit induction). *Let C be a circuit computing a Boolean graph function f. Suppose H is a graph such that $\mathbb{S}(\widetilde{\nu}, H)$ is small for every gate ν in C. Then $f(H) = f(\mathbb{S}(f, H))$.*

Remark 3.26. Suppose H is a large graph. Note that the hypothesis of Lemma 3.25 (that $\mathbb{S}(\widetilde{\nu}, H)$ is small for every gate ν) depends only on the value of gates in C on small and medium subgraphs of H. In particular, this hypothesis never explicitly mentions the value of any gate on H itself. Yet, remarkably, the conclusion of Lemma 3.25 says something nontrivial about the value of C on H.

We mention that the inductive argument in Lemma 3.25 was originally conceived in the context of Ehrenfeucht-Frïssé games (see [65] for an explanation).

Proof of Lemma 3.25. To simplify notation in the proof: for nodes ν in C, let $S_\nu = \mathbb{S}(\widetilde{\nu}, H)$.

<u>Claim 1.</u> $\widetilde{\nu}(S_\nu) = \widetilde{\nu}(S_\nu \cup H')$ and $\nu(S_\nu) = \nu(S_\nu \cup H')$ for every node ν and small $H' \subseteq H$.

Since graphs S_ν and H' are both small, the union $S_\nu \cup H'$ is either small or medium. It follows that $\widetilde{\nu}(S_\nu) = \widetilde{\nu}(S_\nu \cup H')$ (by Def. 3.19). Further, $\nu(S_\nu) = \nu(S_\nu \cup H')$ (since $\widetilde{\nu}$ completely determines the value of ν).

<u>Claim 2.</u> $\nu(S_\nu) = \nu(H)$ for every node ν.

We argue by induction on ν. In the base case where ν is an input node (labelled by a constant 0 or 1 or a variable x_e or its negation $\neg x_e$ for some edge $e \in \binom{[n]}{2}$), $\mathbb{T}(\nu, H)$ is small

38

(since it has at most one edge). It follows that $\mathbb{T}(\nu, H) = \mathbb{S}(\nu, H) = S_\nu$ (since $\tilde{\nu} = \nu$) and hence $\nu(S_\nu) = \nu(\mathbb{T}(\nu, H)) = \nu(H)$.

For the induction step, let ν be a gate and assume that $\mu(H) = \mu(S_\mu)$ for all children μ of ν. With loss of generality, assume that ν is an AND gate (the argument for OR gates is the same, but with the roles of 0 and 1 exchanged). We consider two cases, according to the value of $\nu(S_\nu)$:

- Assume $\nu(S_\nu) = 0$. Let $\mu = \tilde{\nu}(S_\nu)$ (i.e., μ is the leftmost child of ν with value 0 on S_ν). By Claim 1, $\tilde{\nu}(S_\nu \cup S_\mu) = \tilde{\nu}(S_\nu) = \mu$ (i.e., μ is the leftmost child of ν with value 0 on $S_\nu \cup S_\mu$). We have

$$\begin{aligned}
\mu(H) &= \mu(S_\mu) && \text{(induction hypothesis)} \\
&= \mu(S_\nu \cup S_\mu) && \text{(by Claim 1)} \\
&= 0.
\end{aligned}$$

Since ν is an AND gate, it follows that $\nu(H) = 0$.

- Assume $\nu(S_\nu) = 1$. Let μ be any child of ν. By Claim 1, $\nu(S_\nu \cup S_\mu) = \nu(S_\nu) = 1$. We have

$$\begin{aligned}
\mu(H) &= \mu(S_\mu) && \text{(induction hypothesis)} \\
&= \mu(S_\nu \cup S_\mu) && \text{(by Claim 1)} \\
&= 1 && \text{(since } \nu \text{ is an AND gate and } \nu(S_\nu \cup S_\mu) = 1\text{)}.
\end{aligned}$$

Since μ is an arbitrary child of ν, it follows that $\mu(H) = 1$ for all children of ν. Therefore, $\nu(H) = 1$.

This completes the proof of Claim 2.

To finish the argument, let ν be the output node which computes the function f and note that

$$\begin{aligned}
f(H) &= f(S_\nu) && \text{(by Claim 2)} \\
&= f(S_\nu \cap \mathbb{S}(f, H)) && \text{(by definition of } \mathbb{S}(f, H), \text{ since } S_\nu \text{ is a small subgraph of } H\text{)} \\
&= f(\mathbb{S}(f, H)) && \text{(since } \mathbb{S}(f, H) \subseteq S_\nu, \text{ as } \tilde{\nu} \text{ determines the value of } f\text{)}. \qquad \square
\end{aligned}$$

The following lemma applies Lemma 3.25 to the function f^G.

Lemma 3.27. *Let* C *be a circuit computing a Boolean graph function f. Suppose G, H are graphs such that $\mathbb{S}(\tilde{\nu}^G, H)$ is small for every gate ν in C. Then $f(G \cup H) = f(G \cup \mathbb{S}(f^G, H))$.*

(*Nota bene:* Consistent with our usual notation, $\tilde{\nu}^G$ denotes the graph function $\tilde{\nu}^G(H) = \tilde{\nu}(G \cup H)$.)

Proof. For each edge in G, replace the corresponding variable in C with the constant 1 and its negations with 0. In this way we obtain a new circuit C_G which computes the function f^G. For a node ν in C, let ν_G denote the corresponding node in C_G. Note that $\widetilde{\nu_G} = \tilde{\nu}^G$ (where $\widetilde{\nu_G}$ denotes the value/witness function of the node ν_G in the circuit C_G). The result follows by applying Lemma 3.25 to the circuit C_G. $\square$

Equipped with Lemma 3.27, we are ready to prove Theorem 3.2. The proof closely parallels the proof of Proposition 3.13 in §3.3.

39

Proof of Theorem 3.2. Recall the statement to be proved. Suppose $f : \mathscr{G}^n \to \{0,1\}$ is computed by Boolean circuits of size $O(n^{k/4})$ and depth at most $k^{-2} \log n / \log \log n$. Let $G \sim \mathrm{G}(n, n^{-2/(k-1)})$ and $K_k \sim \mathrm{Plant}(n, K_k)$. We must show that w.h.p. $f(G) = f(G \cup K_k)$.

Let C be the circuit computing f. For each node ν in C, with children $\mu_1, \ldots, \mu_m$, we reinterpret the value/witness function $\tilde{\nu} : \mathscr{G}^n \to \{\mu_1, \ldots, \mu_m\} \cup \{0,1\}$ as a function $\mathscr{G}^n \to \{0,1\}^{\lceil \log(m+1) \rceil}$ as follows:

- if $\tilde{\nu}$ is an AND gate with value 0 (resp. an OR gate with value 1), then the value of $\tilde{\nu}$ in $\{0,1\}^{\lceil \log(m+1) \rceil}$ is the base-2 representation of the least index $j \in \{1, \ldots, m\}$ such that the μ_j has value 0 (resp. 1),

- otherwise, the value of $\tilde{\nu}$ is the all-zero string in $\{0,1\}^{\lceil \log(m+1) \rceil}$.

This binary encoding of $\tilde{\nu}$ contains exactly the same information as the original value/witness function. The advantage of this encoding is that we can now view $\tilde{\nu}$ as being computed by a circuit. It is easy to see that $\tilde{\nu}$ is computed by a circuit of size at most $\mathsf{size}(\mathsf{C}) + O(m) = O(n^{k/4})$ and depth at most $\mathsf{depth}(\mathsf{C}) + 3 \leqslant k^{-2} \log n / \log \log n + O(1)$ with $\lceil \log(m+1) \rceil = n^{o(1)}$ outputs. This circuit computing $\tilde{\nu}$ thus satisfies the hypotheses of Lemma 3.15.

Consider an arbitrary gate ν in C.

<u>Claim 1</u>: $\Pr \left[\mathbb{S}(\tilde{\nu}^G, K_k) \text{ is not small} \right] = o(n^{-k/4})$

To prove Claim 1, first note that by Lemma 3.23,

$$\Pr \left[\mathbb{S}(\tilde{\nu}^G, K_k) \text{ is not small} \right] = \Pr \left[\text{some medium subgraph of } \mathbb{S}(\tilde{\nu}^G, K_k) \text{ is a } \tilde{\nu}^G\text{-core} \right]$$

$$\leqslant \sum_{\substack{\text{medium patterns } P \\ \text{(up to isomorphism)}}} \Pr \left[\text{some } P\text{-subgraph of } \mathbb{S}(\tilde{\nu}^G, K_k) \text{ is a } \tilde{\nu}^G\text{-core} \right].$$

For each medium pattern P, we have

$$\Pr \left[\text{some } P\text{-subgraph of } \mathbb{S}(\tilde{\nu}^G, K_k) \text{ is a } \tilde{\nu}^G\text{-core} \right]$$

$$\leqslant \mathrm{E} \left[\text{number of } P\text{-subgraphs of } \mathbb{S}(\tilde{\nu}^G, K_k) \text{ which are } \tilde{\nu}^G\text{-cores} \right]$$

$$= \mathsf{sub}(P, K_k) \Pr_{H \sim \mathrm{Plant}(n,P)} \left[H \text{ is a } \tilde{\nu}^G\text{-core} \mid H \subseteq K_k \right]$$

$$= \mathsf{sub}(P, K_k) \Pr_{H \sim \mathrm{Plant}(n,P)} \left[H \text{ is a } \tilde{\nu}^G\text{-core} \right].$$

(The last equality is due to the obvious independence of the event that H is a $\tilde{\nu}^G$-core and the event that $H \subseteq K_k$.) By Lemma 3.15 (applied to the circuit computing $\tilde{\nu}$), we have

$$\Pr_{H \sim \mathrm{Plant}(n,P)} \left[H \text{ is a } \tilde{\nu}^G\text{-core} \right] = O(n^{-\frac{k}{4} - \frac{1}{k}}).$$

Putting these inequalities together, we have

$$\Pr \left[\mathbb{S}(\tilde{\nu}^G, K_k) \text{ is not small} \right] \leqslant \sum_{\substack{\text{medium patterns } P \\ \text{(up to isomorphism)}}} \mathsf{sub}(P, K_k) \cdot O(n^{-\frac{k}{4} - \frac{1}{k}}).$$

40

Claim 1 now follows from the fact that there are $\leqslant 2^{k^2} = o(n^{1/k})$ medium subpatterns of K_k.

Taking a union bound over gates in C, Claim 1 implies that w.h.p. $\mathbb{S}(\widetilde{\nu}^G, K_k)$ is small for every gate ν in C. By Lemma 3.27, it follows that w.h.p. $f(G \cup K_k) = f(G \cup \mathbb{S}(f^G, K_k))$. Therefore, to establish that w.h.p. $f(G) = f(G \cup K_k)$, it suffices to prove:

<u>Claim 2</u>: W.h.p. $\mathbb{S}(f^G, K_k) = \varnothing$.

To prove Claim 2, let ν be the output node of C. By Claim 1, w.h.p. $\mathbb{S}(\widetilde{\nu}, K_k)$ is small. Note that $\mathbb{S}(f, K_k) \subseteq \mathbb{S}(\widetilde{\nu}, K_k)$ (since $\widetilde{\nu}$ determines the value of f). Thus, w.h.p. $\mathbb{S}(f, K_k)$ is also small. Therefore, it suffices to show that w.h.p. $\mathbb{S}(f, K_k)$ is not both nonempty and small. This is shown by a similar argument to the above. First, we show:

$$\Pr\left[\mathbb{S}(f, K_k) \text{ is nonempty and small}\right] \leqslant \sum_{\substack{\text{nonempty small} \\ P \text{ (up to isom.)}}} \mathsf{sub}(P, K_k) \Pr_{H \sim \mathrm{Plant}(n,P)}\left[H \text{ is a } \widetilde{\nu}^G\text{-core}\right].$$

By Lemma 3.27, the quantity on the right is $\leqslant 2^{k^2} \cdot O(n^{-1}) = o(1)$. This proves Claim 2 and concludes the proof of Theorem 3.2. $\qquad\square$

We end this chapter by mentioning that a few extensions of Theorems 3.1 and 3.2 are discuss in §7.1. In particular, we show that there is a "size-depth tradeoff" of the form $n^{(1-\lambda)k/4}$ for circuits of depth $\lambda k^{-1} \log n / \log\log n$ (see §7.1.1).

Chapter 4

Lower Bound for Monotone Circuits

In the previous chapter we proved lower bounds on the average-case complexity of k-CLIQUE for bounded-depth circuits. In this chapter we prove similar lower bounds for monotone circuits. The notion of "average case" here is slightly weaker here, as we will explain. (In short, our results apply to monotone circuits which solve k-CLIQUE w.h.p. on both $\mathrm{G}(n, p_1)$ and $\mathrm{G}(n, p_2)$ where $p_1(n), p_2(n)$ are two sufficiently separated threshold functions, such as $n^{-2/(k-1)}$ and $2n^{-2/(k-1)}$.)

Before formally stating the results of this chapter in §1.4, we first fix some conventions and notation. In this chapter, circuits are assumed to be monotone, to have a single output node, and to have gates of <u>fan-in 2</u>. In the final section of this chapter (§4.7), we show how to achieve the same lower bounds without the fan-in 2 restriction. (This is similar to situation in the previous chapter where we first proved a lower bound on wires, followed by a strong lower bound on size.)

We recall the relevant definitions and notation from Chapter 2 concerning monotone circuits and minterms. Formally, a *monotone circuit* is an acyclic directed graph C with $\binom{n}{2}$ sources and one sink in which each non-sources node is labeled $\wedge$ or $\vee$ and has in-degree 2.[1] As usual, *size* refers to the number of gates (which is roughly the number of wires for fan-in 2 circuits). C computes a monotone graph function $\mathscr{G}^n \to \{0, 1\}$ in the natural way, as does each node ν in C. $\mathsf{C}(G)$ denotes the value of C on graph G (likewise $\nu(G)$). $\mathcal{M}(\mathsf{C})$ (resp. $\mathcal{M}(\mathsf{C}, P)$) denotes the set of minterms (resp. P-minterms) of the function computed by C (likewise $\mathcal{M}(\nu)$ and $\mathcal{M}(\nu, P)$).

4.1 Results of this chapter

Throughout this chapter, let $p = n^{-2/(k-1)}$ (to fix a particular threshold function[2]) and $\delta = k^{-2}$ (it is fine to regard δ as a sufficiently small constant). Let G^-, G, G^+ be independent Erdős-Rényi random graphs

$$G^- \sim \mathrm{G}(n, p^{1+\delta}), \qquad G \sim \mathrm{G}(n, p), \qquad G^+ \sim \mathrm{G}(n, p^{1-\delta}).$$

---

[1] In this chapter, labels $\wedge$ and $\vee$ substitute AND and OR.

[2] Our results hold for any $p(n) = \Theta(n^{-2/(k-1)})$.

(Note that $+$ and $-$ are not backwards here: $\boldsymbol{G}^-$ is sparser than $\boldsymbol{G}$, which is sparser than $\boldsymbol{G}^+$. Correspondingly $p^{1+\delta} < p < p^{1-\delta}$.) By Lemma 2.3, w.h.p. $\boldsymbol{G}^-$ is k-clique-free and $\boldsymbol{G}^+$ contains a k-clique. That is, with respect to the property of containing a k-clique, $\boldsymbol{G}^-$ and $\boldsymbol{G}^+$ and *subcritical* and *supercritical*. As usual, $\boldsymbol{K}_k \sim \mathrm{Plant}(n, K_k)$ is the random planted k-clique.

Our main theorem is a lower bound for monotone circuits which solve k-Clique w.h.p. on both $\boldsymbol{G}$ and $\boldsymbol{G} \cup \boldsymbol{G}^-$.

Theorem 4.1. *No monotone circuit of size $O(n^{k/4})$ solves k-Clique w.h.p. on both $\boldsymbol{G}$ and $\boldsymbol{G} \cup \boldsymbol{G}^-$.*

Note that $\boldsymbol{G} \cup \boldsymbol{G}^-$ is an Erdős-Rényi random graph $\mathrm{G}(n, \tilde{p})$ where $\tilde{p} = p + (1-p)p^{1+\delta}$, which is also a threshold function for k-Clique. Moreover, since $\tilde{p} = p + o(p)$, the numbers of k-cliques in $\boldsymbol{G}$ and $\boldsymbol{G} \cup \boldsymbol{G}^-$ are asymptotically equivalent Poisson random variables (by Lemma 2.4).[3]

Remark 4.2. Theorem 4.1 implies that no monotone circuit of size $O(n^{k/4})$ solves k-Clique w.h.p. on both $\mathrm{G}(n, p)$ and $\mathrm{G}(n, 2p)$. This follows from the observation that if monotone graph functions f and g agrees w.h.p. on both $\mathrm{G}(n, p_1)$ and $\mathrm{G}(n, p_2)$ for $p_1, p_2 : \mathbb{N} \to [0, 1]$ such that $p_1(n) \leqslant p_2(n)$, then f and g also agree w.h.p. on $\mathrm{G}(n, q)$ for every $q : \mathbb{N} \to [0, 1]$ such that $p_1(n) \leqslant q(n) \leqslant p_2(n)$. By the same observation, Theorem 4.1 may be stated as an average-case hardness result on a single distribution $\mathrm{G}(n, \boldsymbol{q})$ where $\boldsymbol{q}$ equals p with probability $1/2$ and $\tilde{p}$ with probability (or, alternatively, where $\boldsymbol{q}$ is uniformly distributed in $[p, \tilde{p}]$).

It would be nice to reduce the "gap" of $\tilde{p} - p \sim p^{1+\delta}$ between threshold functions p and $\tilde{p}$ in Theorem 4.1. We even conjecture that the gap can be eliminated entirely (Conjecture 7.6 in §7.2). However, there is reason to believe that this gap may be hard to close, since a single-threshold version of Theorem 4.1 seems to require techniques that go beyond the approximation method.[4]

Preliminary to Theorem 4.1, we prove the following lower bound:

Theorem 4.3. *If C is a monotone circuit of size $O(n^{k/4})$ such that $\mathrm{E}[\mathsf{C}(\boldsymbol{K}_k)] = 1 - o(1)$, then $\mathrm{E}[\mathsf{C}(\boldsymbol{G}^-)] = 1 - \exp(-\Omega(n^\delta))$.*

Theorem 4.3 should be compared with the following fact (a consequence of Janson's inequality (Lemma 2.5)), in which subcritical $\boldsymbol{G}^-$ is replaced by supercritical $\boldsymbol{G}^+$.

Fact 4.4. *If f is a monotone graph function such that $\mathrm{E}[f(\boldsymbol{K}_k)] = 1 - o(1)$, then $\mathrm{E}[f(\boldsymbol{G}^+)] = 1 - \exp(-\Omega(n^\delta))$ (irrespective of the monotone circuit complexity of f).*

In the final section of this chapter (§4.7), we strengthen Theorems 4.1 and 4.3 by removing the fan-in 2 restriction.

Theorem 4.5. *Theorems 4.1 and 4.3 hold for monotone circuits with $\bigwedge$ and $\bigvee$ gates of unbounded fan-in.*

---

[3]Notwithstanding, the total variation distance between random graphs $\boldsymbol{G}$ and $\boldsymbol{G} \cup \boldsymbol{G}^-$ is $1 - o(1)$.

[4]What about the random graph with exactly $\lceil \binom{n}{2} p \rceil$ edges? Note that the monotone complexity of k-Clique on this distribution is polynomially equivalent to the non-monotone complexity, since we are deal with a slice function.

44

We mention that the result of the next chapter show that the exponent $k/4$ is tight up to an additive constant in Theorem 4.1 (and simultaneously also in Theorem 3.1 on bounded-depth circuits). In particular, we construct monotone circuits of size $n^{k/4+O(1)}$ and depth $3k$ that solve k-CLIQUE w.h.p. on $G(n,p)$ for all functions $p : \mathbb{N} \to [0,1]$. In view of Theorem 4.6, this demonstrates a gap between the worst-case and average-case monotone complexity of k-CLIQUE.

4.2   Razborov's approximation method

In a seminal paper [60], Razborov proved the first lower bounds on the monotone complexity of k-CLIQUE.

Theorem 4.6. k-CLIQUE *has monotone circuit complexity* $\Omega(n^k/\log^{2k} n)$.[5]

Razborov in fact shows something stronger. Let H be the uniform random complete $(k-1)$-partite graph with vertex set $[n]$ (that is, $E_H = \{\{i,j\} \in \binom{[n]}{2} : \pi(i) \neq \pi(j)\}$ for uniform random function $\pi : [n] \to \{1,\ldots,k\}$). The following result is also from [60] (note the similarity to Theorem 4.3):

Theorem 4.7. *If* C *is a monotone circuit of size* $O(n^k/\log^{2k} n)$ *such that* $E[C(K_k)] = 1 - o(1)$, *then* $E[C(H)] = 1 - o(1)$.[6]

The technique introduced in [60] to prove Theorem 4.7 is known as the *approximation method*. (Note: The following summary is for background only. Our lower bounds do not explicitly follow this framework.) The idea of the approximation method is to replace the lattice $(\mathfrak{M}, \wedge, \vee)$ of monotone functions $\{0,1\}^m \to \{0,1\}$ with a smaller lattice $(\overline{\mathfrak{M}}, \overline{\wedge}, \overline{\vee})$ where $\overline{\mathfrak{M}} \subset \mathfrak{M}$ such that

- $\overline{\mathfrak{M}}$ contains the function $x \mapsto x_i$ for every $i \in \{1,\ldots,m\}$, and

- $\overline{\wedge}$ and $\overline{\vee}$ are the g.l.b. and l.u.b. operations in $\overline{\mathfrak{M}}$ with respect to the natural partial order on functions (i.e., $f \leqslant g$ iff $f(x) \leqslant g(x)$ for all $x \in \{0,1\}^m$).

For every monotone circuit C on m variables, there is a corresponding $\{\overline{\wedge}, \overline{\vee}\}$-circuit $\overline{C}$ in which the $\wedge$ and $\vee$ gates are replaced by $\overline{\wedge}$ and $\overline{\vee}$ gates. Note that $\overline{C}$ computes a function in $\overline{\mathfrak{M}}$.

Let Δ_0 and Δ_1 be two distributions on $\{0,1\}^m$ (e.g., the random graphs H and K_k from Theorem 4.7). Suppose our goal is to prove that no monotone circuit C of size S separates Δ_0 and Δ_1 in the sense that $E[C(\Delta_0)] = o(1)$ and $E[C(\Delta_1)] = 1 - o(1)$. Then it suffices to show that:

1. no function $f \in \overline{\mathfrak{M}}$ satisfies $E[f(\Delta_0)] = o(1)$ and $E[f(\Delta_1)] = 1 - o(1)$,

2. for all $f, g \in \overline{\mathfrak{M}}$,

$$E[(f \overline{\vee} g)(\Delta_0)] - E[(f \vee g)(\Delta_0)] = o(1/S),$$
$$E[(f \wedge g)(\Delta_1)] - E[(f \overline{\wedge} g)(\Delta_1)] = o(1/S).$$

---

[5]This bound is for constant k. [60] also gives lower bounds for k which depends on n.

[6]Moreover, if C is a monotone circuit of size $n^{k-\Omega(1)}$ such that $E[C(K_k)] = 1 - o(1)$, then $E[C(H)] = 1 - \exp(-n^{\Omega(1)})$.

By bounding "local errors" in this way, (2) shows that for any C of size S,

$$\mathrm{E}[\mathsf{C}(\Delta_0)] \leqslant \mathrm{E}[\overline{\mathsf{C}}(\Delta_0)] + o(1),$$
$$\mathrm{E}[\mathsf{C}(\Delta_1)] \geqslant \mathrm{E}[\overline{\mathsf{C}}(\Delta_1)] - o(1).$$

It follows that C cannot satisfy both $\mathrm{E}[\mathsf{C}(\Delta_0)] = o(1)$ and $\mathrm{E}[\mathsf{C}(\Delta_1)] = 1 - o(1)$.

Of course, being able to show (1) and (2) for given Δ_0 and Δ_1 depends on a clever choice of the lattice $\overline{\mathfrak{M}}$. To prove Theorem 4.7, Razborov defines a lattice $\overline{\mathfrak{M}}$ where the l.u.b. operation $\overline{\vee}$ involves "plucking" large sunflowers among the minterms of the function $f \vee g$. (For a full description of $\overline{\mathfrak{M}}$, see [60] or [2].)

Our proof of Theorem 4.3 does not precisely follow this framework. Rather, we work with a "one-sided version" of the approximation method (via a closure operator $\mathsf{cl} : \mathfrak{M} \to \mathfrak{M}$ defined in §4.4). The difference is merely a matter of exposition: our proof could easily be formulated in terms of an approximating lattice $\overline{\mathfrak{M}}$.

4.3 Quasi-sunflowers

In this section we introduce a new relaxation of sunflowers called *quasi-sunflowers* (parameterized by $p \in [0,1]$ and $\gamma \geqslant 0$). Like sunflowers, quasi-sunflowers are special hypergraphs. Some definitions: A *hypergraph* is a family $\mathcal{F}$ of subsets of a set X (i.e., $\mathcal{F} \subseteq \wp(X)$). Elements of $\mathcal{F}$ are called *hyperedges*. For an integer $s \geqslant 1$, $\mathcal{F}$ is *s-uniform* if every hyperedge has size s (i.e., $\mathcal{F} \subseteq \binom{X}{s}$).

A *sunflower* is a hypergraph $\mathcal{F}$ such that the intersection of any two distinct hyperedges coincides with the intersection $\bigcap \mathcal{F}$ $(= \bigcap_{U \in \mathcal{F}} U)$ of all hyperedges. The set $\bigcap \mathcal{F}$ is called the *core* and sets $U \setminus \bigcap \mathcal{F}$ where $U \in \mathcal{F}$ are called *petals* (note that petals are mutually disjoint). An essential fact about sunflowers is:

Fact 4.8 (Erdős-Rado Sunflower Lemma [26]). *Every s-uniform hypergraph $\mathcal{F}$ of size $>$ $s!(N-1)^s$ contains a sunflower of size N.*

Quasi-sunflowers are a relaxation of sunflowers in which petals may overlap slightly on average. While other variants of sunflowers are studied in extremal combinatorics (see Ch. 7 of [44]), the following definition appears to be new.

Definition 4.9. *Let $\mathcal{F}$ be a hypergraph on a set X and let $Y \subseteq \bigcap \mathcal{F}$. For $p \in [0,1]$ and $\gamma \geqslant 0$, we say that $\mathcal{F}$ is (p,γ)-quasi-sunflower over Y if for the random set $\boldsymbol{W} \subseteq_p X$,*

$$\Pr\left[\boldsymbol{W} \cup Y \text{ contains a hyperedge of } \mathcal{F}\right] \geqslant 1 - e^{-\gamma}.$$

Observation 4.10. Let $\mathcal{F} \subseteq \binom{X}{s}$ be an s-uniform sunflower of size n. Then $\mathcal{F}$ is a (p, np^s)-quasi-sunflower for every $p \in [0,1]$. To see this, let $Y = \bigcap \mathcal{F}$ and note that for $\boldsymbol{W} \subseteq_p X$, the probability that $\boldsymbol{W} \cup Y$ contains a hyperedge of $\mathcal{F}$ is

$$1 - (1 - p^{s-|Y|})^n \geqslant 1 - \exp(np^{s-|Y|}) \geqslant 1 - \exp(np^s).$$

For small p, this $\gamma = np^s$ is nearly tight if $Y = \emptyset$, but (as we will see) is far from tight if $Y \neq \emptyset$.

We suspect that wherever s-uniform sunflowers are used in monotone circuit lower bounds (e.g., [2, 7, 60]), one could just as well work with $(1/2, N/2^s)$-quasi-sunflowers

46

instead. That is, Definition 4.9 captures the essential property of sunflowers for these applications. Perhaps one even gets stronger bounds (as we do in this paper) by virtue of the following result.

Theorem 4.11 ("Quasi-sunflower lemma"). *For all $p \in [0,1]$ and $\gamma \geqslant 1$ and $s \geqslant 1$, every s-uniform hypergraph of size $\geqslant s!(2.47\gamma/p)^s$ contains a (p,γ)-quasi-sunflower.*

Remark 4.12. It follows from Fact 4.8 and Obs. 4.10 that every s-uniform hypergraph of size $\geqslant s!(\gamma/p^s)^s$ contains a (p,γ)-quasi-sunflower (namely, a sunflower of size γ/p^s). Theorem 4.11 is a significant quantitative improvement of this observation.

The proof of Theorem 4.11 uses Janson's inequality (Lemma 2.5) within an inductive argument resembling proofs of the Erdős-Rado Sunflower Lemma.

Proof of "Quasi-sunflower Lemma" (Theorem 4.11). Consider the sequence $\ell_1, \ell_2, \ldots$ defined by $\ell_1 = 1$ and $\ell_s = 2\sum_{t=1}^{s-1} \binom{s}{t}\ell_t$ for $s \geqslant 2$. We have $\ell_s \leqslant s! \ln^{-s}(3/2)$ ($< s! 2.47^s$) by induction: for $s \geqslant 2$, assuming $\ell_t \leqslant t! \ln^{-t}(3/2)$ for every $t \in \{1, \ldots, s-1\}$, we have

$$
\begin{aligned}
\ell_s &\leqslant 2\sum_{t=1}^{s-1} \binom{s}{t} t! \ln^{-t}(3/2) \\
&= 2\left(\sum_{t=1}^{s-1} \frac{\ln^{s-t}(3/2)}{(s-t)!}\right) s! \ln^{-s}(3/2) \\
&\leqslant 2\left(-1 + \sum_{j=0}^{\infty} \frac{\ln^j(3/2)}{j!}\right) s! \ln^{-s}(3/2) \\
&= s! \ln^{-s}(3/2).
\end{aligned}
$$

Suppose $\mathcal{F}$ is an s-uniform hypergraph of size $\geqslant \ell_s(\gamma/p)^s$. Arguing by induction on s, we claim that $\mathcal{F}$ contains an (p,γ)-quasi-sunflower (proving the theorem). In the base case where $s = 1$, let $\boldsymbol{W} \subseteq_p X$ and note that events $U \subseteq \boldsymbol{W}$ for $U \in \mathcal{F}$ are mutually independent. Therefore,

$$
\Pr\left[\bigwedge_{U \in \mathcal{F}} U \nsubseteq \boldsymbol{W}\right] = (1-p)^{|\mathcal{F}|} \leqslant (1-p)^{\gamma/p} \leqslant e^{-\gamma},
$$

so $\mathcal{F}$ itself is a (p,γ)-quasi-sunflower over the empty set.

For the induction step, let $s \geqslant 2$ and assume the claim holds for $t \in \{1, \ldots, s-1\}$. For every $A \subseteq X$ with $1 \leqslant |A| \leqslant s-1$, let

$$
\mathcal{F}_A = \{U \setminus A : U \in \mathcal{F} \text{ such that } A \subseteq U\}.
$$

Note that $\mathcal{F}_A$ is an $(s - |A|)$-uniform hypergraph. We now consider two cases.

First Case Suppose there exist $t \in \{1, \ldots, s-1\}$ and $A \in \binom{X}{t}$ such that $|\mathcal{F}_A| \geqslant \ell_{s-t}(\gamma/p)^{s-t}$. By the induction hypothesis, $\mathcal{F}_A$ contains a (p,γ)-quasi-sunflower $\mathcal{F}'$ over some $Y' \subseteq \bigcap \mathcal{F}'$. Note that $\{U \cup A : U \in \mathcal{F}'\} \subseteq \mathcal{F}$ is a (p,γ)-quasi-sunflower over $Y' \cup A$.

Second Case Suppose $|\mathcal{F}_A| \leqslant \ell_{s-t}(\gamma/p)^{s-t}$ for all $t \in \{1, \ldots, s-1\}$ and $A \in \binom{X}{t}$. We will show that $\mathcal{F}$ itself is a (p, γ)-quasi-sunflower over the empty set. Let $\boldsymbol{W} \subseteq_p X$ and define μ and Δ exactly as in the statement of Janson's inequality (Lemma 2.5), which says:

$$\Pr\left[\bigwedge_{U \in \mathcal{F}} U \not\subseteq \boldsymbol{W}\right] \leqslant \exp\left(-\min\left\{\frac{\mu}{2}, \frac{\mu^2}{2\Delta}\right\}\right).$$

Thus, it suffices to show that $\min\{\mu/2, \mu^2/2\Delta\} \geqslant \gamma$.

Clearly $\mu = |\mathcal{F}|p^s$ since $\Pr[U \subseteq \boldsymbol{W}] = p^s$ for every $U \in \mathcal{F}$. Since $|\mathcal{F}| \geqslant \ell_s(\gamma/p)^s$ and $\ell_s \geqslant 2$ (as $s \geqslant 1$) and $\gamma^s \geqslant \gamma$ (as $\gamma \geqslant 1$), we have $\mu/2 \geqslant \gamma$.

It remains to show that $\mu^2/2\Delta \geqslant \gamma$. For every $t \in \{1, \ldots, s-1\}$, we have $\sum_{A \in \binom{X}{t}} |\mathcal{F}_A| = \binom{s}{t}|\mathcal{F}|$ since each hyperedge in $\mathcal{F}$ is counted $\binom{s}{t}$ times in this summation. Therefore,

$$\sum_{A \in \binom{X}{t}} |\mathcal{F}_A|^2 \leqslant |\mathcal{F}| \sum_{A \in \binom{X}{t}} |\mathcal{F}_A|$$

$$\leqslant \mu \binom{s}{t} \ell_{s-t} \gamma^{s-t} p^{t-2s}$$

(using $|\mathcal{F}| = \mu p^{-s}$ and $|\mathcal{F}_A| \leqslant \ell_{s-t}(\gamma/p)^{s-t}$). Noting that $\Pr\left[U \cup V \subseteq \boldsymbol{W}\right] = p^{2s-|U \cap V|}$ for all $U, V \in \mathcal{F}$, we bound Δ as follows:

$$\Delta = \sum_{\substack{A \subseteq X: \\ 1 \leqslant |A| \leqslant s-1}} \sum_{\substack{U,V \in \mathcal{F}: \\ U \cap V = A}} \Pr\left[U \cup V \subseteq \boldsymbol{W}\right]$$

$$\leqslant \sum_{t=1}^{s-1} \left(\sum_{A \in \binom{X}{t}} |\mathcal{F}_A|^2\right) p^{2s-t}$$

$$\leqslant \mu \sum_{t=1}^{s-1} \binom{s}{t} \ell_{s-t} \gamma^{s-t}$$

$$\leqslant \mu \gamma^{s-1} \sum_{t=1}^{s-1} \binom{s}{t} \ell_t \qquad\qquad \text{(using } \gamma^t \leqslant \gamma^{s-1}\text{)}$$

$$= \frac{\mu \gamma^{s-1} \ell_s}{2} \qquad\qquad \text{(by definition of } \ell_s\text{)}.$$

Completing the proof, we have

$$\frac{\mu^2}{2\Delta} \geqslant \frac{\mu}{\gamma^{s-1}\ell_s} = \frac{|\mathcal{F}|p^s}{\gamma^{s-1}\ell_s} \geqslant \gamma. \qquad\qquad \square$$

4.4   The approximation via a closure operator

In this section we define a closure operator in the lattice of monotone graph functions. Closed functions will be combinatorially "nice" in the sense of having few P-minterms for small and medium patterns P (Lemma 4.22).

Remark 4.13. This is essentially one half of Razborov's approximation method. Typically, one also defines a "truncation" operator which cuts out large minterms. Although we find

it more natural to work with a one-sided version of the approximation method, our proof can be translated into Razborov's original framework (as described in §4.2).

Recall that we have fixed $p = n^{-2/(k-1)}$ (a threshold function for the existence of k-cliques) and $\delta = k^{-2}$ (just think of δ as "sufficiently small"). Also recall that $G \sim \mathrm{G}(n,p)$ (at the k-clique threshold) and $G^- \sim \mathrm{G}(n, p^{1+\delta})$ (below the k-clique threshold, i.e., G^- is almost surely k-clique-free).

Definition 4.14. *A monotone graph function $f : \mathscr{G}^n \to \{0,1\}$ is* closed *if for every small-or-medium graph H,*

$$\mathrm{E}[f(G^- \cup H)] \geqslant 1 - e^{-n^\delta} \implies f(H) = 1.$$

Observation 4.15. If f and g are both closed, then so is $f \wedge g$.

Definition 4.16. *For a monotone graph function f, we denote by $\mathsf{cl}(f)$ the unique minimal closed function such that $f \leqslant \mathsf{cl}(f)$, called the* closure *of f.*

Note that $\mathsf{cl}(f)$ is well-defined in view of Obs. 4.15 and the fact that the constant function 1 is closed.

Remark 4.17. Viewed as an operation on the set of monotone graph functions, $\mathsf{cl}(\cdot)$ is a closure operator in the usual sense. That is, it satisfies:

- (increasing) $f \leqslant \mathsf{cl}(f)$,

- (monotone) $f \leqslant g \implies \mathsf{cl}(f) \leqslant \mathsf{cl}(g)$,

- (idempotent) $\mathsf{cl}(\mathsf{cl}(f)) = \mathsf{cl}(f)$.

Definition 4.18. *We denote by $\overline{\vee}$ the operation on monotone graph functions defined by $f \overline{\vee} g = \mathsf{cl}(f \vee g)$. For a monotone circuit C, we denote by $\overline{\mathsf{C}}$ denote the corresponding circuit with basis $\{\wedge, \overline{\vee}\}$ in which the $\vee$-gates in C are replaced by $\overline{\vee}$-gates. For nodes ν in C, we denote by $\overline{\nu}$ the corresponding node in $\overline{\mathsf{C}}$.*

Note that $\mathsf{cl}(\mathsf{C})$ (i.e., $\mathsf{cl}(f)$ where f is the function computed by C) is not necessarily the same function as $\overline{\mathsf{C}}$, although $\overline{\mathsf{C}}$ is indeed a closed function satisfying $\mathsf{C} \leqslant \overline{\mathsf{C}}$ (i.e., $\mathsf{C}(G) \leqslant \overline{\mathsf{C}}(G)$ for all graphs G).

Lemma 4.19. *For every monotone graph function f,*

$$\Pr\left[f(G^-) \neq (\mathsf{cl}(f))(G^-) \right] \leqslant 2^{k^2} n^k e^{-n^\delta}.$$

Proof. We claim that there exist $t \in \mathbb{N}$ and small-or-medium graphs $H_1, \dots, H_t$ and monotone functions $f_0, \dots, f_t : \mathscr{G}^n \to \{0,1\}$ such that

- $f_0 = f$,

- $\mathrm{E}[f_{i-1}(G^- \cup H_i)] \in [1 - e^{-n^\delta}, 1)$,

- $f_i = f_{i-1} \vee \mathsf{Ind}_{H_i}$ where $\mathsf{Ind}_{H_i} : \mathscr{G}^n \to \{0,1\}$ is the function $\mathsf{Ind}_{H_i}(G) = 1$ iff $H_i \subseteq G$,

- f_t is closed.

49

To see this, note that we can generate such a sequence (a priori indefinitely) simply by choosing any suitable H_{i+1} so long as f_i is not closed. This process eventually terminates, since each small or medium graph H appears at most once in the sequence $H_1, H_2, \ldots$. In particular,

$$t \leqslant |\{\text{small and medium graphs in } \mathscr{G}^n\}| \leqslant 2^{k^2} n^k.$$

An inductive argument shows that $f_i \leqslant \mathsf{cl}(f)$ for $i = 1, \ldots, t$. In particular $f_t \leqslant \mathsf{cl}(f)$. Since f_t is closed, this means that $f_t = \mathsf{cl}(f)$. We now have

$$\Pr\left[f(\boldsymbol{G}^-) \neq (\mathsf{cl}(f))(\boldsymbol{G}^-)\right] \leqslant \sum_{i=1}^{t} \Pr\left[f_{i-1}(\boldsymbol{G}^-) \neq f_i(\boldsymbol{G}^-)\right]$$

$$= \sum_{i=1}^{t} \Pr\left[f_{i-1}(\boldsymbol{G}^-) = 0 \text{ and } H_i \subseteq \boldsymbol{G}^-\right]$$

$$\leqslant \sum_{i=1}^{t} \Pr\left[f_{i-1}(\boldsymbol{G}^- \cup H_i) = 0\right]$$

$$\leqslant 2^{k^2} n^k e^{-n^\delta}. \qquad \square$$

The next two lemmas follow immediately from Lemma 4.19.

Lemma 4.20. *For every monotone graph function f, $\mathcal{M}(\mathsf{cl}(f)) \setminus \mathcal{M}(f)$ contains only small and medium graphs.*

Proof. The proof of Lemma 4.19 shows that there exist small-or-medium graphs $H_1, \ldots, H_t$ such that $\mathsf{cl}(f) = f \vee \bigvee_{i=1}^{t} \mathsf{Ind}_{H_i}$. Thus, $\mathcal{M}(\mathsf{cl}(f)) \subseteq \mathcal{M}(f) \cup \{H_1, \ldots, H_t\}$. $\square$

Lemma 4.21. *For every monotone circuit C of size $\exp(o(n^\delta))$, $\mathrm{E}[\overline{\mathsf{C}}(\boldsymbol{G}^-)] - \mathrm{E}[\mathsf{C}(\boldsymbol{G}^-)] = \exp(-\Omega(n^\delta))$.*

Proof. For any graph H, note that if $\mathsf{C}(H) \neq \overline{\mathsf{C}}(H)$ then there exists an $\vee$-gate ν with children μ_1 and μ_2 in C such that $\overline{\nu}(H) \neq (\overline{\mu_1} \vee \overline{\mu_2})(H)$ (equivalently: $f(H) \neq (\mathsf{cl}(f))(H)$ where f is the function $\overline{\mu_1} \vee \overline{\mu_2}$). It follows that

$$\mathrm{E}[\overline{\mathsf{C}}(\boldsymbol{G}^-)] - \mathrm{E}[\mathsf{C}(\boldsymbol{G}^-)] = \Pr\left[\mathsf{C}(\boldsymbol{G}^-) \neq \overline{\mathsf{C}}(\boldsymbol{G}^-)\right]$$

$$\leqslant \sum_{\substack{\vee\text{-gates } \nu \text{ in } \mathsf{C} \text{ with} \\ \text{children } \mu_1 \text{ and } \mu_2}} \Pr\left[\overline{\nu}(\boldsymbol{G}^-) \neq (\overline{\mu_1} \vee \overline{\mu_2})(\boldsymbol{G}^-)\right]$$

$$\leqslant \mathsf{size}(\mathsf{C}) 2^{k^2} n^k e^{-n^\delta} \quad \text{(by Lemma 4.19)}$$

$$= \exp(-\Omega(n^\delta)). \qquad \square$$

The last lemma of this section gives an essential property of closed functions (using Theorem 4.11 on quasi-sunflowers).

Lemma 4.22. *A closed monotone graph function has at most $k^{k^2}(n^\delta/p^{1+\delta})^{|E_P|}$ P-minterms for every small or medium pattern P.*

Proof. Let f be a closed monotone graph function and let P be a small or medium pattern. Toward a contradiction, assume that $|\mathcal{M}(f, P)| \geqslant k^{k^2}(n^\delta/p^{1+\delta})^{|E_P|}$. Let $X = \binom{[n]}{2}$ and consider the $|E_P|$-uniform hypergraph $\mathcal{F} \subseteq \binom{X}{|E_P|}$ defined by $\mathcal{F} = \{E_F : F \in \mathcal{M}(f, P)\}$. Since

50

$|E_P| \leqslant k^2/4$ (i.e., no medium pattern has more than $k^2/4$ edges), we have $|E_P|!2.47^{|E_P|} \leqslant k^{k^2}$ and hence

$$|\mathcal{F}| = |\mathcal{M}(f, P)| \geqslant |E_P|!2.47^{|E_P|}(n^\delta/p^{1+\delta})^{|E_P|}.$$

By Theorem 4.11, there exists a $(p^{1+\delta}, n^\delta)$-quasi-sunflower $\mathcal{F}_0 \subseteq \mathcal{F}$ over some $Y \subseteq \bigcap \mathcal{F}$. Let H be the graph with edge set $E_H = Y$. Let $\boldsymbol{W} \subseteq_{p^{1+\delta}} X$ and note that $\boldsymbol{W}$ has the same distribution as $E_{\boldsymbol{G}^-}$. We have

$$\begin{aligned}
\mathrm{E}[f(\boldsymbol{G}^- \cup H)] &\geqslant \Pr\left[\boldsymbol{G}^- \cup H \text{ contains a } P\text{-minterm of } f\right] \\
&\geqslant \Pr\left[\boldsymbol{W} \cup Y \text{ contains a hyperedge of } \mathcal{F}_0\right] \\
&\geqslant 1 - e^{-n^\delta}.
\end{aligned}$$

Since f is closed and H is small or medium, it follows that $f(H) = 1$. Note that H has fewer than $|E_P|$ edges, so in particular H is a proper subgraph of some $F \in \mathcal{M}(f, P)$ such that $E_F \in \mathcal{F}_0$. However, this contradicts the fact that F is a minterm of f. □

4.5   K vs. $\mathbf{G}^-$

In the previous section, we defined a closure operator $\mathrm{cl}(\cdot)$ on monotone graph functions and an operation $\mathsf{C} \mapsto \overline{\mathsf{C}}$ transforming a monotone circuit C into a $\{\wedge, \bigtriangledown\}$-circuit $\overline{\mathsf{C}}$. In this section, we prove Theorem 4.3. We begin by noting a basic fact about minterms.

Observation 4.23. For all monotone graph functions f and g,

$$\begin{aligned}
\mathcal{M}(f \vee g) &\subseteq \mathcal{M}(f) \cup \mathcal{M}(g), \\
\mathcal{M}(f \wedge g) &\subseteq \{F \cup G : F \in \mathcal{M}(f),\ G \in \mathcal{M}(g)\}.
\end{aligned}$$

That is, every minterm of $f \vee g$ is a minterm of f or a minterm of g and every minterm of $f \wedge g$ is the union of a minterm of f and a minterm of g.

Lemma 4.24. *Let* C *be a monotone circuit. For every* $H \in \mathcal{M}(\overline{\mathsf{C}}, K_k)$, *there exist a gate* ν *in* C *and a medium subgraph* H' *of* H *such that* $H' \in \mathcal{M}(\overline{\nu})$.

Proof. Suppose $H \in \mathcal{M}(\overline{\mathsf{C}}, K_k)$ and for notational convenience let

$$\mathcal{H} = \{\text{subgraphs of } H\}, \quad \mathcal{A} = \{\text{small graphs}\}, \quad \mathcal{B} = \{\text{medium graphs}\}.$$

Toward a contradiction, assume that $\mathcal{M}(\overline{\nu}) \cap \mathcal{H} \cap \mathcal{B} = \emptyset$ for every gate ν in C. We will show, by induction on ν, that $\mathcal{M}(\overline{\nu}) \cap \mathcal{H} \subseteq \mathcal{A}$ for every node ν in C. This yields a contradiction, since $H \in (\mathcal{M}(\overline{\nu_{\mathrm{out}}}) \cap \mathcal{H}) \setminus \mathcal{A}$ where ν_{out} is the output gate of C.

Consider first the base case where ν is an input node labelled by either 0 or 1 or the indicator function for some edge $e \in \binom{[n]}{2}$. Note that $\mathcal{M}(\nu)$ is respectively either the empty set or $\{\text{the empty graph}\}$ or $\{\text{the graph with only edge } e\}$. In any case, ν has only small minterms. Since $\overline{\nu} = \nu$, we have $\mathcal{M}(\overline{\nu}) \cap \mathcal{H} \subseteq \mathcal{A}$ as required.

For the induction step, suppose ν is a gate in C with children μ_1 and μ_2 and assume

that $\mathcal{M}(\overline{\mu_i}) \cap \mathcal{H} \subseteq \mathcal{A}$ for $i \in \{1,2\}$. If ν is an $\wedge$-gate, then

$$
\begin{aligned}
\mathcal{M}(\overline{\nu}) \cap \mathcal{H} &= \mathcal{M}(\overline{\mu_1} \wedge \overline{\mu_2}) \cap \mathcal{H} \\
&= \{F_1 \cup F_2 : F_1 \in \mathcal{M}(\overline{\mu_1}),\ F_2 \in \mathcal{M}(\overline{\mu_2})\} \cap \mathcal{H} \quad \text{(Obs. 4.23)} \\
&= \{F_1 \cup F_2 : F_1 \in \mathcal{M}(\overline{\mu_1}) \cap \mathcal{H},\ F_2 \in \mathcal{M}(\overline{\mu_2}) \cap \mathcal{H}\} \\
&\subseteq \{F_1 \cup F_2 : F_1, F_2 \in \mathcal{A}\} \quad \text{(since } \mathcal{M}(\overline{\mu_i}) \cap \mathcal{H} \subseteq \mathcal{A}) \\
&\subseteq \mathcal{A} \cup \mathcal{B} \quad \text{(the union of two small graphs cannot be large)} \\
&\subseteq \mathcal{A} \quad \text{(by assumption } \mathcal{M}(\overline{\nu}) \cap \mathcal{H} \cap \mathcal{B} = \emptyset).
\end{aligned}
$$

Finally, if ν is a $\vee$-gate, then

$$
\begin{aligned}
\mathcal{M}(\overline{\nu}) \cap \mathcal{H} &= \mathcal{M}(\overline{\mu_1} \,\overline{\vee}\, \overline{\mu_2}) \cap \mathcal{H} \\
&= \mathcal{M}(\mathsf{cl}(\overline{\mu_1} \vee \overline{\mu_2})) \cap \mathcal{H} \quad \text{(definition of } \overline{\vee}) \\
&\subseteq \big(\mathcal{M}(\overline{\mu_1} \vee \overline{\mu_2}) \cup \mathcal{A} \cup \mathcal{B}\big) \cap \mathcal{H} \quad \text{(Lemma 4.20)} \\
&\subseteq \big(\mathcal{M}(\overline{\mu_1}) \cup \mathcal{M}(\overline{\mu_2}) \cup \mathcal{A} \cup \mathcal{B}\big) \cap \mathcal{H} \quad \text{(Obs. 4.23)} \\
&\subseteq \mathcal{A} \cup \mathcal{B} \quad \text{(since } \mathcal{M}(\overline{\mu_i}) \cap \mathcal{H} \subseteq \mathcal{A} \text{ for } i \in \{1,2\}) \\
&\subseteq \mathcal{A} \quad \text{(by assumption } \mathcal{M}(\overline{\nu}) \cap \mathcal{H} \cap \mathcal{B} = \emptyset). \qquad \square
\end{aligned}
$$

Lemma 4.25. *For every monotone circuit* C, *there exists a medium pattern* P *such that*

$$
\mathsf{size}(\mathsf{C}) \geqslant \frac{|\mathcal{M}(\overline{\mathsf{C}}, K_k)|}{(2k)^{k^2} n^{k-|V_P|} (n^\delta/p^{1+\delta})^{|E_P|}}.
$$

Proof. By Lemma 4.24, for each $H \in \mathcal{M}(\overline{\mathsf{C}}, K_k)$, there exists a gate μ_H in C and a medium subgraph H' of H such that $H' \in \mathcal{M}(\overline{\mu_H})$. Fix choices of μ_H and H' for all $H \in \mathcal{M}(\overline{\mathsf{C}}, K_k)$. For every gate ν in C and medium pattern P, let

$$
t(\nu, P) = |\{H \in \mathcal{M}(\overline{\mathsf{C}}, K_k) : \mu_H = \nu \text{ and } H' \in \mathcal{M}(\overline{\nu}, P)\}|.
$$

By a simple counting argument, there exist ν and P such that

$$
\frac{|\mathcal{M}(\overline{\mathsf{C}}, K_k)|}{\mathsf{size}(\mathsf{C}) \cdot |\{\text{medium patterns up to isomorphism}\}|} \leqslant t(\nu, P).
$$

For each $H' \in \mathcal{M}(\overline{\nu}, P)$, there are at most $n^{k-|V_P|}$ different $H \in \mathcal{M}(\overline{\mathsf{C}}, K_k)$ of which H' is a subgraph. It follows that

$$
t(\nu, P) \leqslant n^{k-|V_P|} |\mathcal{M}(\overline{\nu}, P)|.
$$

Since $\overline{\nu}$ is closed and P is medium, Lemma 4.22 implies

$$
|\mathcal{M}(\overline{\nu}, P)| \leqslant k^{k^2} (n^\delta/p^{1+\delta})^{|E_P|}.
$$

The result follows by combining these three inequalities, together with the bound 2^{k^2} on the number of medium patterns up to isomorphism. $\qquad \square$

Onto the main result:

Proof of Theorem 4.3. Suppose $f : \mathscr{G}^n \to \{0,1\}$ is computed by monotone circuits of size

52

$O(n^{k/4})$ and satisfies $\mathrm{E}[f(\boldsymbol{K}_k)] = 1 - o(1)$. We must show that $\mathrm{E}[f(\boldsymbol{G}^-)] = 1 - \exp(-\Omega(n^\delta))$. Let C be the circuit computing f. By Lemma 4.21,

$$\mathrm{E}[\overline{\mathsf{C}}(\boldsymbol{G}^-)] - \mathrm{E}[f(\boldsymbol{G}^-)] = \Pr[f(\boldsymbol{G}^-) \neq \overline{\mathsf{C}}(\boldsymbol{G}^-)] = \exp(-\Omega(n^\delta)).$$

Therefore, it suffices to show that $\mathrm{E}[\overline{\mathsf{C}}(\boldsymbol{G}^-)] = 1$. We will assume that $\mathrm{E}[\overline{\mathsf{C}}(\boldsymbol{G}^-)] \neq 1$ and derive a contradiction.

We claim that $|\mathcal{M}(\overline{\mathsf{C}}, K_k)| = (1 - o(1))\binom{n}{k}$. To show this, we consider the pattern $Q = K_k - \{\text{single edge}\}$ and let $\boldsymbol{H} \sim \mathrm{Plant}(n, Q)$. Since $\mathrm{E}[\overline{\mathsf{C}}(\boldsymbol{K}_k)] \geqslant \mathrm{E}[f(\boldsymbol{K}_k)] = 1 - o(1)$, it is enough to show that $\mathrm{E}[\overline{\mathsf{C}}(\boldsymbol{H})] = o(1)$ (i.e., these two inequalities imply that almost every planted k-clique is a minterm of $\overline{\mathsf{C}}$). The argument goes as follows: if we assume that $\mathrm{E}[\overline{\mathsf{C}}(\boldsymbol{H})] = \Omega(1)$, then $\Pr[\overline{\mathsf{C}}(\boldsymbol{G}^-)] = 1 - \exp(\Omega(n^{1/k})) \geqslant 1 - \exp(n^\delta)$ for sufficiently large n (recall that $\delta = k^{-2}$) by an straightforward application of Janson's inequality (Lemma 2.5; but since $\overline{\mathsf{C}}$ is closed, it follows that $\overline{\mathsf{C}}$(the empty graph) $= 1$ (contradicting $\mathrm{E}[\overline{\mathsf{C}}(\boldsymbol{G}^-)] \neq 1$).

We now invoke Lemma 4.25, which gives us a medium pattern P such that

$$\mathrm{size}(\mathsf{C}) \geqslant \frac{|\mathcal{M}(\overline{\mathsf{C}}, K_k)|}{(2k)^{k^2} n^{k - |V_P|} (n^\delta/p^{1+\delta})^{|E_P|}} = (1 - o(1)) \binom{n}{k} \frac{n^{|V_P|} (p^{1+\delta}/n^\delta)^{|E_P|}}{n^k (2k)^{k^2}}$$

$$= \Omega\left(\frac{n^{|V_P| - (\frac{2}{k-1}(1+\delta)+\delta)|E_P|}}{k^k (2k)^{k^2}}\right)$$

(using $p = n^{-2/(k-1)}$). Recall that $\delta = 1/k^2$ and note that $|E_P| < k^2/4$ (obs: among medium patterns, the disjoint union of two $\lfloor \frac{k-1}{2} \rfloor$-cliques has the most edges). By Lemma 2.8, $|V_P| - \frac{2}{k-1}|E_P| \geqslant \frac{k}{4} + \frac{1}{4} + \frac{2}{k-1}$. Thus,

$$|V_P| - \left(\frac{2}{k-1}(1+\delta) + \delta\right)|E_P| > |V_P| - \frac{2}{k-1}|E_P| - \frac{1}{4}\left(1 + \frac{2}{k-1}\right) > \frac{k}{4} + \frac{1}{k}.$$

Therefore,

$$\mathrm{size}(\mathsf{C}) = \Omega\left(\frac{n^{(k/4)+(1/k)}}{k^k (2k)^{k^2}}\right).$$

But since k is a constant, this contradicts the hypothesis that C has size $O(n^{k/4})$. $\qquad\square$

4.6 $\mathbf{G} \cup \mathbf{K}$ vs. $\mathbf{G} \cup \mathbf{G}^-$

In this section, we prove Theorem 4.1 using Theorem 4.3 together with the following lemma.

Lemma 4.26. *Let f be a graph function (not necessarily monotone) and let $\boldsymbol{G}_0 \sim \mathrm{G}(n, p)$ conditioned on $\boldsymbol{G}_0$ being k-clique-free.*

1. *If f solves k-Clique w.h.p. on $\boldsymbol{G}$, then $\mathrm{E}[f(\boldsymbol{G}_0 \cup \boldsymbol{K}_k)] = 1 - o(1)$.*

2. *If f solves k-Clique w.h.p. on $\boldsymbol{G} \cup \boldsymbol{G}^-$, then $\mathrm{E}[f(\boldsymbol{G}_0 \cup \boldsymbol{G}^-)] = o(1)$.*

Proof. Denote by $\kappa(G)$ the number of k-cliques in a graph G.

For (1): Suppose f solves k-Clique w.h.p. on $\boldsymbol{G}$. This means, in particular, that $\mathrm{E}[f(\boldsymbol{G}) \mid \kappa(\boldsymbol{G}) = 1] = 1 - o(1)$. Let $\boldsymbol{G}_1 \sim \mathrm{G}(n, p)$ conditioned on $\kappa(\boldsymbol{G}_1) = 1$. Note that $\mathrm{E}[f(\boldsymbol{G}_1)] = 1 - o(1)$ (using the fact that $\Pr[\kappa(\boldsymbol{G}) = 1] = \Omega(1)$). By Lemma 2.4,

random graphs $G_0 \cup K_k$ and G_1 have total variation distance $o(1)$. Therefore, w.h.p. $\mathrm{E}[f(G_0 \cup K_k)] = 1 - o(1)$.

For (2): Suppose f solves k-CLIQUE w.h.p. on $G \cup G^-$. In particular,

$$(*) \qquad \mathrm{E}[f(G \cup G^-) \mid \kappa(G \cup G^-) = 0] = o(1).$$

Since $G \sim \mathrm{G}(n,p)$ and $G \cup G^- \sim \mathrm{G}(n, p + o(p))$, random variables $\kappa(G)$ and $\kappa(G \cup G^-)$ converge in distribution to the same Poisson distribution by Lemma 2.4. In particular, we have

$$(**) \qquad \Pr[\kappa(G) = 0] = (1 + o(1)) \Pr[\kappa(G \cup G^-) = 0].$$

Thus, we have

$$
\begin{aligned}
\mathrm{E}[f(G_0 \cup G^-)] &= \Pr[f(G \cup G^-) = 1 \mid \kappa(G) = 0] \\
&= \frac{\Pr[f(G \cup G^-) = 1 \ \& \ \kappa(G) = 0]}{\Pr[\kappa(G) = 0]} \\
&\geqslant \frac{\Pr[f(G \cup G^-) = 1 \ \& \ \kappa(G \cup G^-) = 0]}{\Pr[\kappa(G) = 0]} \\
&\overset{(**)}{=} \frac{\Pr[f(G \cup G^-) = 1 \ \& \ \kappa(G \cup G^-) = 0]}{(1 + o(1)) \Pr[\kappa(G \cup G^-) = 0]} \\
&= (1 - o(1)) \Pr[f(G \cup G^-) = 1 \mid \kappa(G \cup G^-) = 0] \\
&\overset{(*)}{=} 1 - o(1).
\end{aligned}
$$

(Under the assumption that f is monotone, (2) can also be proved using the Holley inequality.) $\qquad\square$

Proof of Theorem 4.1. Let C be a monotone circuit of size $O(n^{k/4})$. Toward a contradiction, assume that C solves k-CLIQUE w.h.p. on both G and $G \cup G^-$. For a graph G, let C^G be the circuit obtained from C by substituting 1 for each input corresponding to an edge in G. Note that C^G computes the function $\mathsf{C}^G(H) = \mathsf{C}(G \cup H)$.

Let $G_0 \sim \mathrm{G}(n,p)$ conditioned on G_0 being k-clique-free. Lemma 4.26 implies that for every constant $\varepsilon > 0$,

$$
\Pr_{G_0}\left[\mathop{\mathrm{E}}_{G^-}[\mathsf{C}^{G_0}(K_k)] \geqslant 1 - \varepsilon \right] = o(1),
$$

$$
\Pr_{G_0}\left[\mathop{\mathrm{E}}_{K_k}[\mathsf{C}^{G_0}(G^-)] \leqslant \varepsilon \right] = 1 - o(1).
$$

It follows that there is a sequence of monotone circuits of size $O(n^{k/4})$ (namely, C^{G_0} for almost every G_0) with expected value $1 - o(1)$ on K_k and $o(1)$ on G^-. However, Theorem 4.3 says this is impossible, giving the desired contradiction. $\qquad\square$

4.7 Removing the fan-in restriction

We conclude this chapter by showing how to remove the fan-in 2 restriction in our results. In what follows, let C be a fixed monotone circuit of size $O(n^{k/4})$ with $\bigwedge$-gates and $\bigvee$-gates of <u>unbounded fan-in</u>. We will show that the size lower bounds of Theorems 4.1 and 4.3 still

hold in this setting.

We first note that there is an obvious generalization of the binary operation $\overline{\vee}$ (defined by $f \mathbin{\overline{\vee}} g = \mathsf{cl}(f \vee g)$) to a multi-ary operation $\overline{\bigvee}$ on functions $f_1, \ldots, f_m$, which we define by $\overline{\bigvee}_{i=1}^{m} f_i = \mathsf{cl}\big(\bigvee_{i=1}^{m} f_i\big)$. Denote by $\overline{\mathsf{C}}$ the $\{\wedge, \overline{\bigvee}\}$-circuit obtained by replacing $\bigvee$-gates in C with $\overline{\bigvee}$-gates.

There is only one place in the proof of Theorem 4.3 where the fan-in 2 assumption comes into play: namely in Lemma 4.24. To be precise, this lemma relies on the fact that if f_1 and f_2 are monotone graph functions with only small minterms, then $f_1 \wedge f_2$ has no large minterms. This is a consequence of facts:

- the union of two small graphs is either small or medium, and

- $\mathcal{M}(f_1 \wedge f_2) \subseteq \{F_1 \cup F_2 : F_1 \in \mathcal{M}(f_1), F_1 \in \mathcal{M}(f_2)\}$ (see Obs. 4.23).

The trouble is that the union of three or more small graphs can be large. So Lemma 4.24 is invalid for the circuit C.

We get around this problem as follows. Denote by $[k \log n]$ the set $\{1, \ldots, \lceil k \log n \rceil\}$ and by $[n^{1/2k}]$ the set $\{1, \ldots, \lceil n^{1/2k} \rceil\}$. For every $\bigwedge$-gate ν in C and $i \in [k \log n]$ and $j \in [n^{1/2k}]$, generate a random set $\boldsymbol{S}_{\nu,i,j} \subseteq_{2^{-i}} \mathsf{Children}(\nu)$ (that is, $\boldsymbol{S}_{\nu,i,j}$ independently contains each child of ν with probability 2^{-i}). We replace Lemma 4.24 with the following:

Lemma 4.27. *With probability $1 - \exp(-\Omega(n^{1/3k}))$, the following holds: for every $H \in \mathcal{M}(\overline{\mathsf{C}}, K_k)$, there exist a gate ν in C and a medium subgraph H' of H such that either*

- *ν is an $\bigvee$-gate and $H' \in \mathcal{M}(\overline{\nu})$, or*

- *ν is an $\bigwedge$-gate and $H' \in \mathcal{M}(\bigwedge_{\mu \in \boldsymbol{S}_{\nu,i,j}} \overline{\mu})$ for some $i \in [k \log n]$ and $j \in [n^{1/2k}]$.*

Proof. Suppose $H \in \mathcal{M}(\overline{\mathsf{C}}, K_k)$ and for notational convenience let

$$\mathcal{H} = \{\text{subgraphs of } H\}, \quad \mathcal{A} = \{\text{small graphs}\}, \quad \mathcal{B} = \{\text{medium graphs}\}.$$

An easy argument (along the lines of the proof of Lemma 4.24) shows that there exists a gate ν in C, with children $\mu_1, \ldots, \mu_m$, such that

1. $\mathcal{M}(\overline{\mu_\ell}) \cap \mathcal{H} \subseteq \mathcal{A}$ for all $\ell \in [m]$, and

2. $\big(\mathcal{M}(\overline{\nu}) \cap \mathcal{H}\big) \setminus \mathcal{A}$ is nonempty.

Fix any $H' \in \big(\mathcal{M}(\overline{\nu}) \cap \mathcal{H}\big) \setminus \mathcal{A}$.

In the case where ν is $\bigvee$-gate, we have

$$
\begin{aligned}
\mathcal{M}(\overline{\nu}) &= \mathcal{M}\big(\overline{\bigvee}_{\ell \in [m]} \overline{\mu_\ell}\big) && \\
&= \mathcal{M}\big(\mathsf{cl}(\bigvee_{\ell \in [m]} \overline{\mu_\ell})\big) && \text{(definition of } \overline{\bigvee}) \\
&\subseteq \mathcal{M}\big(\bigvee_{\ell \in [m]} \overline{\mu_\ell}\big) \cup \mathcal{A} \cup \mathcal{B} && \text{(Lemma 4.20)} \\
&\subseteq \bigcup_{\ell \in [m]} \mathcal{M}(\overline{\mu_\ell}) \cup \mathcal{A} \cup \mathcal{B} && \text{(Obs. 4.23)}.
\end{aligned}
$$

Since $H' \notin \mathcal{A}$ and $\mathcal{M}(\overline{\mu_\ell}) \cap \mathcal{H} \subseteq \mathcal{A}$ for all $\ell \in [m]$, it follows that $H' \in \mathcal{B}$ (i.e., H' is medium, so we are done).

Now suppose ν is an $\bigwedge$-gate. We have

$$\mathcal{M}(\overline{\nu}) = \mathcal{M}(\bigwedge_{\ell \in [m]} \overline{\mu_\ell})$$
$$\subseteq \{F_1 \cup \cdots \cup F_m : F_1 \in \mathcal{M}(\overline{\mu_1}), \ldots, F_m \in \mathcal{M}(\overline{\mu_m})\} \quad \text{(Obs. 4.23)}.$$

Hence there exist $F_1 \in \mathcal{M}(\overline{\mu_1}), \ldots, F_m \in \mathcal{M}(\overline{\mu_m})$ such that $H' = F_1 \cup \cdots \cup F_m$. Fix any such $F_1, \ldots, F_m$.

We next fix an enumeration $H_1, \ldots, H_t$ of the set $\{F_1, \ldots, F_m\}$ subject to

$$|\{\ell \in [m] : H_{t'} = F_\ell\}| \geqslant |\{\ell \in [m] : H_{t'+1} = F_\ell\}|$$

for all $t' \in \{1, \ldots, t-1\}$. (That is, $H_{t'}$ are ranked in decreasing order according to their frequency among $F_1, \ldots, F_m$.) Note that $t \leqslant 2^{k^2}$ (even though m may be as large as $n^{k/4}$), since there are $\leqslant 2^{k^2}$ distinct subgraphs of H.

Let s be the least index in $\{2, \ldots, t\}$ such that $H_1 \cup \cdots \cup H_s \notin \mathcal{A}$. (Such s is well-defined since $H_1 \cup \cdots \cup H_t = F_1 \cup \cdots \cup F_m = H' \notin \mathcal{A}$.) Note that $H_1 \cup \cdots \cup H_s \in \mathcal{B}$, since $H_1 \cup \cdots \cup H_{s-1} \in \mathcal{A}$ and $H_s \in \mathcal{A}$ (using the fact that the union of two small graphs is either small or medium). Let i be the unique integer in $[k \log n]$ such that

$$2^{i-1} \leqslant |\{\ell \in [m] : H_s = F_\ell\}| < 2^i.$$

(Such i exists since $m \leqslant \mathsf{fanin}(\mathsf{C}) \leqslant \mathsf{size}(\mathsf{C}) = O(n^{k/4})$.)

We now show that, with extremely high probability, there exists $j \in [n^{1/2k}]$ such that $H_1 \cup \cdots \cup H_s \in \mathcal{M}(\bigwedge_{\mu \in \boldsymbol{S}_{\nu,i,j}} \overline{\mu})$. Consider a random set $\boldsymbol{S} \subseteq_{2^{-i}} [m]$. For $t' \in \{1, \ldots, t\}$, denote by $\boldsymbol{X}_{t'}$ the event that there exists $\ell \in \boldsymbol{S}$ such that $H_{t'} \in \mathcal{M}(\overline{\mu_\ell})$. Note the following:

- for $t' \in \{1, \ldots, s\}$,

$$\Pr[\boldsymbol{X}_{t'}] = 1 - (1 - 2^{-i})^{|\{\ell \in [m]: H_{t'} = F_\ell\}|} \geqslant 1 - (1 - 2^{-i})^{2^{i-1}} > 1 - \frac{1}{\sqrt{e}} > \frac{1}{4},$$

- for $t' \in \{s+1, \ldots, t\}$,

$$\Pr[\neg \boldsymbol{X}_{t'}] \geqslant (1 - 2^{-i})^{|\{\ell \in [m]: H_{t'} = F_\ell)\}|} > (1 - 2^{-i})^{2^i} \geqslant \frac{1}{4}.$$

Note that $\boldsymbol{X}_1, \ldots, \boldsymbol{X}_t$ are independent. It follows that

$$\Pr\left[H_1 \cup \cdots \cup H_s \in \mathcal{M}(\bigwedge_{\ell \in \boldsymbol{S}} \overline{\mu_\ell})\right] \geqslant \Pr\left[(\boldsymbol{X}_1 \wedge \cdots \wedge \boldsymbol{X}_s) \wedge (\neg \boldsymbol{X}_{s+1} \wedge \cdots \wedge \neg \boldsymbol{X}_t)\right]$$
$$= \Big(\prod_{t' \in \{1, \ldots, s\}} \Pr[\boldsymbol{X}_{t'}]\Big)\Big(\prod_{t' \in \{s+1, \ldots, t\}} \Pr[\neg \boldsymbol{X}_{t'}]\Big)$$
$$\geqslant 4^{-t}$$
$$\geqslant 4^{-2^{k^2}}.$$

By independence (for different $j \in [n^{1/2k}]$) of sets $\boldsymbol{S}_{\nu,i,j}$, we have

$$\Pr\left[\forall j \in [n^{1/2k}],\ H_1 \cup \cdots \cup H_s \notin \mathcal{M}(\bigwedge_{\mu \in \boldsymbol{S}_{\nu,i,j}} \overline{\mu})\right] \leqslant \prod_{j \in [n^{1/2k}]} \Pr\left[H_1 \cup \cdots \cup H_s \notin \mathcal{M}(\overline{\sigma_{\nu,i,j}})\right]$$

$$\leqslant \left(1 - 4^{-2^{k^2}}\right)^{n^{1/2k}}$$

$$\leqslant \exp(-4^{-2^{k^2}} n^{1/2k}).$$

Taking a union bound over all $\leqslant \binom{n}{k}$ graphs $H \in \mathcal{M}(\overline{C}, K_k)$, we upper bound the total failure probability by

$$\binom{n}{k} \exp(-4^{-2^{k^2}} n^{1/2k}) = \exp(-\Omega(n^{1/3k})),$$

which proves the lemma. $\qquad\square$

We now get the following modified version of Lemma 4.25.

Lemma 4.28. *With probability* $1 - \exp(-\Omega(n^{1/3k}))$, *there exists a medium pattern P such that*

$$\mathsf{size}(\mathsf{C}) \geqslant \frac{1}{(k \log n)n^{1/2k}} \cdot \frac{|\mathcal{M}(\overline{\mathsf{C}}, K_k)|}{(2k)^{k^2} n^{k-|V_P|}(n^\delta/p^{1+\delta})^{|E_P|}}.$$

Proof. The proof is a simple counting argument (which uses Lemma 4.27), completely analogous to the proof of Lemma 4.25 (which uses Lemma 4.24). (Note: we use the fact that $\bigwedge_{\mu \in \boldsymbol{S}_{\nu,i,j}} \overline{\mu}$ are closed functions.) $\qquad\square$

Compared with Lemma 4.25, the bound of Lemma 4.28 has a loss of $1/(k \log n)n^{1/2k}$. This is tolerable if we desire a lower bound of $\omega(n^{k/4})$, since this loss is eaten up by the slack factor of $n^{1/k}/k^k(2k)^{k^2}$ in the actual lower bound

$$\mathsf{size}(\mathsf{C}) = \Omega\left(\frac{n^{(k/4)+(1/k)}}{k^k(2k)^{k^2}}\right).$$

given by the proof of Theorem 4.3. The fact that the bottleneck of Lemma 4.28 fails with probability $\exp(-\Omega(n^{1/3k}))$ is also tolerable, since the error $\mathrm{E}[\overline{\mathsf{C}}(G^-)] - \mathrm{E}[\mathsf{C}(G^-)]$ allowed by the proof of Theorem 4.3 is $\exp(-\Omega(n^\delta))$ where $\delta = 1/k^2 = o(1/3k)$. By this argument, it is seen that the lower bounds proved in this chapter remain valid for monotone circuits with unbounded fan-in.

Chapter 5

Matching Upper Bound

In this chapter we prove the following upper bound:

Theorem 5.1. *There exist monotone circuits of size $n^{k/4+O(1)}$ and depth $3k$ which solve k-CLIQUE w.h.p. on $G(n,p)$ for all functions $p : \mathbb{N} \to [0,1]$.*

Theorem 5.1 demonstrates that the exponent $k/4$ is tight up to an additive constant *simultaneously* in both of our lower bounds (Theorems 3.2 and 4.1). The circuits in Theorem 5.1 are based on a construction of non-monotone constant-depth circuits due to Amano [6]. An extra trick (involving perfect families of hash functions) is employed to make these circuits monotone.

5.1 Auxiliary subcircuits

We begin by defining some auxiliary subcircuits. First, a useful definition:

Definition 5.2. *For a graph $G \in \mathscr{G}^n$ and sets $U_1,\dots,U_\ell \subseteq [n]$ where $\ell \geqslant 1$, define $\Gamma_G(U_1,\dots,U_\ell) \subseteq U_\ell$ inductively as follows:*

- $\Gamma_G(U_1) = U_1$,

- *for $\ell \geqslant 2$, $\Gamma_G(U_1,\dots,U_\ell)$ is the set of $u_\ell \in U_\ell$ such that for every $i \in \{1,\dots,\ell-1\}$, there exists $u_i \in \Gamma_G(U_1,\dots,U_i)$ such that $\{u_i,u_\ell\}$ is an edge in G.*

Lemma 5.3. *If $\Gamma_G(U_1,\dots,U_i)$ is a singleton $\{u_i\}$ for each $i \in \{1,\dots,\ell\}$, then $u_1,\dots,u_\ell$ form an ℓ-clique in G.*

Proof. Immediate from the definition of $\Gamma_G(\cdot)$. $\qquad\square$

The next lemma describes our first family of auxiliary circuits.

Lemma 5.4. *There exist monotone circuits $\mathsf{A}_{U_1,\dots,U_\ell}$ of size $\leqslant \ell^2 n^2$ and depth $3(\ell-1)$ with n output nodes, denoted $\mathsf{A}^v_{U_1,\dots,U_\ell}$ for $v \in [n]$, such that for every graph $G \in \mathscr{G}^n$ and vertex $v \in [n]$,*

$$\mathsf{A}^v_{U_1,\dots,U_\ell}(G) = 1 \iff v \in \Gamma_G(U_1,\dots,U_\ell).$$

Proof. The proof is by induction on ℓ. In the base case where $\ell = 1$, the circuit A_{U_1} consists of n isolated output nodes where $\mathsf{A}^v_{U_1}$ is labeled by the constant 1 if $v \in U_1$ and by the constant 0 otherwise.

59

For the induction step, assume $\ell \geqslant 2$. Starting with the circuit $\mathsf{A}_{U_1,\ldots,U_{\ell-1}}$, for each $v \in [n]$ create new gates ν_v and $(\mu_{v,i})_{i\in\{1,\ldots,\ell-1\}}$ and $(\xi_{v,i,w})_{i\in\{1,\ldots,\ell-1\},w\in[n]\setminus\{v\}}$. The labels and wires are as follows:

- $\nu_v = \underset{i\in\{1,\ldots,\ell-1\}}{\mathrm{AND}} \mu_{v,i}$,

- $\mu_{v,i} = \underset{w\in[n]\setminus\{v\}}{\mathrm{OR}} \xi_{v,i,w}$,

- $\xi_{v,i,w} = \mathrm{AND}(x_{\{v,w\}}, \mathsf{A}^w_{U_1,\ldots,U_i})$ where $x_{\{v,w\}}$ is the indicator variable for the edge $\{v,w\}$.

The output node $\mathsf{A}^v_{U_1,\ldots,U_{\ell-1}}$ is of course ν_v.

It is easy to see that $\mathsf{A}_{U_1,\ldots,U_\ell}$ correctly computes the set $\Gamma_G(U_1,\ldots,U_\ell)$ (assuming that $\mathsf{A}_{U_1,\ldots,U_i}$ correctly computes $\Gamma_G(U_1,\ldots,U_i)$ for every $i \in \{1,\ldots,\ell-1\}$). Note that we have added 3 to the depth and created $(\ell-1)n^2 + n \leqslant \ell n^2$ new gates. Thus, as required we have

- $\mathsf{depth}(\mathsf{A}_{U_1,\ldots,U_\ell}) = \mathsf{depth}(\mathsf{A}_{U_1,\ldots,U_{\ell-1}}) + 3 = 3(\ell-1)$,

- $\mathsf{size}(\mathsf{A}_{U_1,\ldots,U_\ell}) \leqslant \mathsf{size}(\mathsf{A}_{U_1,\ldots,U_{\ell-1}}) + \ell n^2 \leqslant (\ell-1)^2 n^2 + \ell n^2 \leqslant \ell^2 n^2$. $\qquad\square$

Next comes another useful definition, followed by our second family of auxiliary circuits:

Definition 5.5. *A sequence $(U_1,\ldots,U_\ell)$ is c-bounded in graph G if for all $i \in \{1,\ldots,\ell\}$ and distinct $u_1 \in U_1$, $\ldots$, $u_{i-1} \in U_{i-1}$, there are at most c different $u_i \in U_i$ such that $\{u_j, u_i\}$ is an edge in G for all $j \in \{1,\ldots,i-1\}$.*

Lemma 5.6. *For all $c, \ell \in \mathbb{N}$, there exist single-output monotone circuits $\mathsf{B}_{c;U_1,\ldots,U_\ell}$ of size $O(n^2 \log n)$ and depth $3\ell - 1$ such that for every graph G, if $(U_1,\ldots,U_\ell)$ is c-bounded in G then*

$$\mathsf{B}_{c;U_1,\ldots,U_\ell}(G) = 1 \iff U_1 \times \cdots \times U_\ell \text{ contains an } \ell\text{-clique in } G.$$

(The hidden constant in this $O(\cdot)$ term actually looks like $\ell c^\ell 4^{\ell c^\ell}$.)

Proof. Let $\mathcal{H}$ be a ℓc^ℓ-perfect family of $O(4^{\ell c^\ell} \log n)$ hash functions from $[n]$ to $[\ell c^\ell]$. That is, $\mathcal{H}$ is a set of functions $[n] \to [\ell c^\ell]$ such that for every $X \subseteq [n]$ such that $|X| \leqslant \ell c^\ell$, there exists $h \in \mathcal{H}$ such that $|h(X)| = |X|$. The existence of an ℓc-perfect family $|\mathcal{H}|$ of size $O(4^{\ell c^\ell} \log n)$ is established by a probabilistic argument: simply choose $4^{\ell c^\ell} \log n$ functions at random for sufficiently large n (see [5]).

Define $\mathsf{B}_{c;U_1,\ldots,U_\ell}$ by

$$\mathsf{B}_{c;U_1,\ldots,U_\ell} = \underset{h\in\mathcal{H}}{\mathrm{AND}} \ \underset{z_1,\ldots,z_\ell\in[\ell c^\ell],\, v\in[n]}{\mathrm{OR}} \ \mathsf{A}^v_{U_1\cap h^{-1}(z_1),\ldots,U_\ell\cap h^{-1}(z_\ell)}.$$

Finally, note that $\mathsf{B}_{c;U_1,\ldots,U_\ell}$ has size $O(n^2 \log n)$ and depth $3\ell - 1$.

To see that $\mathsf{B}_{c;U_1,\ldots,U_\ell}$ computes a suitable function, let G be any graph such that $(U_1,\ldots,U_\ell)$ is c-bounded in G. We will now show that $\mathsf{B}_{c;U_1,\ldots,U_\ell}(G) = 1 \iff U_1 \times \cdots \times U_\ell$ contains an ℓ-clique in G.

($\impliedby$) Assume that $u_1,\ldots,u_\ell$ form an ℓ-clique in G where $u_i \in U_i$ for every $i \in \{1,\ldots,\ell\}$. To show that $\mathsf{B}_{c;U_1,\ldots,U_\ell}(G) = 1$, let h be any function in $\mathcal{H}$ (chosen adversarially). Select any $z_i = h^{-1}(u_i)$ for each $i \in \{1,\ldots,\ell\}$ and note that $\mathsf{A}^{u_\ell}_{U_1\cap h^{-1}(z_1),\ldots,U_\ell\cap h^{-1}(z_\ell)}(G) = 1$.

60

($\Longrightarrow$) Assume that $\mathsf{B}_{c;U_1,\ldots,U_\ell}(G) = 1$. Let $X = \bigcup_{i\in\{1,\ldots,\ell\}} \Gamma_{U_1,\ldots,U_i}(G)$. Since $(U_1,\ldots,U_\ell)$ is c-bounded in G, we have $|X| \leqslant c + c^2 + \cdots + c^\ell \leqslant \ell c^\ell$. Since $\mathcal{H}$ is an ℓc^ℓ-perfect family of hash functions, there exists $h \in \mathcal{H}$ such that $|h(X)| = |X|$. By definition of $\mathsf{B}_{c;U_1,\ldots,U_\ell}$, there exist $z_1,\ldots,z_\ell \in [\ell c^\ell]$ such that

$$\operatorname*{OR}_{v\in[n]} \mathsf{A}^v_{U_1\cap h^{-1}(z_1),\ldots,U_\ell\cap h^{-1}(z_\ell)}(G) = 1.$$

An inductive argument shows that $\Gamma_G(U_1\cap h^{-1}(z_1),\ldots,U_i\cap h^{-1}(z_i))$ is a singleton for every $i \in \{1,\ldots,\ell\}$. By Lemma 5.3, it follows that G contains an ℓ-clique in $U_1 \times \cdots \times U_\ell$. $\qquad\square$

5.2 Some random sets

Using the circuits $\mathsf{B}_{c;U_1,\ldots,U_\ell}$ defined in the last section, we now present a monotone constant-depth circuit of size $n^{k/4+O(1)}$ which solves k-CLIQUE on $\mathrm{G}(n,p)$ for all functions $p : \mathbb{N} \to [0,1]$.

Definition 5.7. *For $i \in \{1,\ldots,k\}$, let $p_i = \min\{n^{(i-2)\frac{2}{k}-1}, 1\}$ and let $U_i \subseteq_{p_i} [n]$ (that is, $\Pr[v \in U_i] = p_i$ independently for all $v \in [n]$).*

This choice of random sets U_i is motivated by the following lemma:

Lemma 5.8. *For all $\alpha \geqslant \frac{2}{k}$ and $i \in \{1,\ldots,k\}$ and distinct $v_1,\ldots,v_{i-1} \in [n]$,*

$$\Pr_{G\sim\mathrm{G}(n,n^{-\alpha})} \left[U_i \text{ contains} > c \text{ common neighbors of } v_1,\ldots,v_{i-1} \text{ in } G\right] = O(n^{-\frac{2}{k}(c+1)}).$$

Proof. For each $w \in [n] \setminus \{v_1,\ldots,v_{i-1}\}$, the probability that w belongs to U_i and is a common neighbor of $v_1,\ldots,v_{i-1}$ in G is precisely $p_i(n^{-\alpha})^{i-1}$. Since $p_i \leqslant n^{(i-2)\frac{2}{k}-1}$ and $n^{-\alpha} \leqslant n^{-\frac{2}{k}}$, we have $p_i(n^{-\alpha})^{i-1} \leqslant n^{-1+\frac{2}{k}}$. By a union bound and independence,

$$\Pr\left[U_i \text{ contains} > c \text{ common neighbors of } v_1,\ldots,v_{i-1} \text{ in } G\right]$$

$$\leqslant \sum_{\text{distinct } w_1,\ldots,w_{c+1}\in[n]\setminus\{v_1,\ldots,v_{i-1}\}} \Pr\left[\begin{array}{l} w_1,\ldots,w_{c+1} \in U_i \text{ are common} \\ \text{neighbors of } v_1,\ldots,v_{i-1} \text{ in } G \end{array}\right]$$

$$= \binom{n-i+1}{c+1} \left(n^{-1+\frac{2}{k}}\right)^{c+1}$$

$$= O(n^{\frac{2}{k}(c+1)}). \qquad\square$$

We now proceed with three lemmas and a corollary giving further properties of $U_1,\ldots,U_k$.

Lemma 5.9. $\displaystyle\prod_{i=1}^{k} p_i \geqslant n^{-\frac{k}{4}-3}.$

Proof. We have $\prod_{i=1}^{k} p_i = n^{-\beta}$ where

$$\beta = \sum_{i=1}^{k} \max\left\{0, 1 + \tfrac{2}{k} - (i-1)\tfrac{2}{k}\right\} = \sum_{i=1}^{\lceil \frac{k+1}{2} \rceil} \left(1 + \tfrac{2}{k} - (i-1)\tfrac{2}{k}\right)$$

$$= \left\lceil \tfrac{k+1}{2} \right\rceil \left(1 + \tfrac{2}{k}\right) - \tfrac{2}{k}\binom{\lceil \frac{k+1}{2} \rceil + 1}{2}$$

$$< \tfrac{k}{4} + 3. \qquad \square$$

Lemma 5.10. *For all $\alpha \geqslant \frac{2}{k}$ and $G \sim \mathrm{G}(n, n^{-\alpha})$,*

$$\Pr\left[(U_1, \ldots, U_k) \text{ is not } k^2\text{-bounded in } G\right] = O(n^{-k}).$$

Proof.

$$\Pr\left[(U_1, \ldots, U_k) \text{ is not } k^2\text{-bounded in } G\right]$$

$$\leqslant \sum_{\text{distinct } v_1,\ldots,v_k \in [n]} \sum_{i \in \{1,\ldots,k\}} \Pr\left[\begin{matrix} U_i \text{ contains} > k^2 \text{ common} \\ \text{neighbors of } v_1, \ldots, v_{i-1} \text{ in } G \end{matrix}\right]$$

$$\leqslant \binom{n}{k} k \cdot O(n^{-\frac{2}{k}(k^2+1)}) \qquad \text{(by Lemma 5.8)}$$

$$= O(n^{-k}). \qquad \square$$

Definition 5.11. *Let $S = \lceil n^{\frac{k}{4}+4} \rceil$ and let $U_1^{(t)}, \ldots, U_k^{(t)}$ be independent copies of $U_1, \ldots, U_k$ for $t \in \{1, \ldots, S\}$.*

We choose S large enough so that the following lemma holds.

Lemma 5.12. *W.h.p.,* $\bigcup_{t \in \{1,\ldots,S\}} U_1^{(t)} \times \cdots \times U_k^{(t)} = [n]^k$.

Proof.

$$\Pr\left[\bigcup_{t \in \{1,\ldots,S\}} U_1^{(t)} \times \cdots \times U_k^{(t)} \neq [n]^k\right] = \Pr\left[\bigvee_{v_1,\ldots,v_k \in [n]} \bigwedge_{t \in \{1,\ldots,S\}} \bigvee_{i \in \{1,\ldots,k\}} v_i \notin U_i^{(t)}\right]$$

$$\leqslant \sum_{v_1,\ldots,v_k} \Pr\left[\bigwedge_t \bigvee_i v_i \notin U_i^{(t)}\right] \qquad \text{(union bound)}$$

$$= \sum_{v_1,\ldots,v_k} \left(\Pr\left[\bigvee_i v_i \notin U_i\right]\right)^S \qquad \text{(independence)}$$

$$= \sum_{v_1,\ldots,v_k} \left(1 - \Pr\left[\bigwedge_i v_i \in U_i\right]\right)^S$$

$$= \sum_{v_1,\ldots,v_k} \left(1 - \prod_i p_i\right)^S$$

$$\leqslant n^k \left(1 - n^{-(k/4)-3}\right)^{n^{(k/4)+4}} \qquad \text{(Lemma 5.9)}$$

$$= o(1). \qquad \square$$

Lemma 5.13. *There exist sets $U_1^{(t)}, \ldots, U_k^{(t)}$, for $t = 1, \ldots, S$, such that*

1. $$\bigcup_{t \in \{1, \ldots, S\}} U_1^{(t)} \times \cdots \times U_k^{(t)} = [n]^k, \text{ and}$$

2. *for all $\alpha \geq \frac{2}{k}$, w.h.p. for $G \sim G(n, n^{-\alpha})$,*

$$(U_1^{(t)}, \ldots, U_k^{(t)}) \text{ is } k^2\text{-bounded in } G \text{ for all } t \in \{1, \ldots, S\}.$$

Proof. Follows from Lemma 5.10 (taking a union bound over t) and Lemma 5.12. $\qquad\square$

5.3 The full construction

In this section we prove Theorem 5.1 by constructing the circuit C. Fix sets $U_1^{(t)}, \ldots, U_k^{(t)}$, for $t = 1, \ldots, S$, as in Lemma 5.13. Let $\mathcal{E}$ be an arbitrary subset $\binom{[n]}{2}$ of size $\lceil n^{2/(k-(1/2))} \rceil$. Define the circuit C by

$$\mathsf{C} = \left(\underset{t \in \{1, \ldots, S\}}{\mathrm{OR}} \mathsf{B}_{k^2; U_1^{(t)}, \ldots, U_k^{(t)}} \right) \mathrm{OR} \left(\underset{\{v, w\} \in \mathcal{E}}{\mathrm{OR}} x_{\{v, w\}} \right).$$

(Here the subcircuit $\mathrm{OR}_{\{v,w\} \in \mathcal{E}} \, x_{\{v,w\}}$ has value 1 on a graph G if and only if some element of $\mathcal{E}$ is an edge in G.) We first check that the circuit C has the correct size and depth. We see that C has size $O(Sn^3 \log n) = n^{k/4 + O(1)}$ (as each B subcircuit has size only $O(n^3 \log n)$). Combining the three OR gates at the top of C, we see that C has depth $3k$ (since B has depth $3k - 1$).

It remains to show that C solves k-CLIQUE w.h.p. on $G \sim G(n, p)$ for all functions $p : \mathbb{N} \to [0, 1]$. First, consider the case that $p \geq n^{-2/k}$. In this case, w.h.p. G contains both a k-clique (by Lemma 2.3, since $n^{-2/k} = \omega(n^{-2/(k-1)})$). So it suffices to show that w.h.p. $\mathsf{C}(G) = 1$. Indeed, this is true since w.h.p. G contains an edge in $\mathcal{E}$.

Next, consider p in the range $n^{-2/(k-(1/4))} < p < n^{-2/k}$. Also in this case, w.h.p. G contains a k-clique. Let $\checkmark$ $(= \checkmark(G))$ denote the event that $(U_1^{(t)}, \ldots, U_k^{(t)})$ is k^2-bounded in G for all $t \in \{1, \ldots, S\}$. Since $p \leq n^{-2/k}$, note that $\checkmark$ holds w.h.p. by Lemma 5.13(2). *Now condition on $\checkmark$ holding and G containing a k-clique.* It suffices to that $\mathsf{C}(G) = 1$ (i.e., with conditional probability 1). Let $\{v_1, \ldots, v_k\}$ be a k-clique in G. By Lemma 5.13(1), there exists $t \in \{1, \ldots, S\}$ such that $(v_1, \ldots, v_k) \in U_1^{(t)} \times \cdots \times U_k^{(t)}$. Since $(U_1^{(t)}, \ldots, U_k^{(t)})$ is k^2-bounded (by $\checkmark$), Lemma 5.6 implies $\mathsf{B}_{k^2; U_1^{(t)}, \ldots, U_k^{(t)}}(G) = 1$. It follows that $\mathsf{C}(G) = 1$ as required.

Finally, consider the case that $p \leq n^{-2/(k-(1/4))}$. Again $\checkmark$ holds w.h.p. Since $p = o(n^{-2/(k-(1/2))})$, w.h.p. G contains no edge in $\mathcal{E}$. *Now condition on $\checkmark$ holding and G not containing any edge in $\mathcal{E}$.* It suffices to show that $\mathsf{C}(G) = 1 \iff G$ contains a k-clique (i.e., with conditional probability 1). Note that $\mathsf{C}(G) = \mathrm{OR}_{t \in \{1, \ldots, S\}} \mathsf{B}_{k^2; U_1^{(t)}, \ldots, U_k^{(t)}}$. If G contains a k-clique $\{v_1, \ldots, v_k\}$, then $\mathsf{C}(G) = 1$ by same reason as before (that is, there exists $t \in \{1, \ldots, S\}$ such that $(v_1, \ldots, v_k) \in U_1^{(t)} \times \cdots \times U_k^{(t)}$, etc.) Conversely, if $\mathsf{C}(G) = 1$, then there exists $t \in \{1, \ldots, S\}$ such that $(v_1, \ldots, v_k) \in U_1^{(t)} \times \cdots \times U_k^{(t)}$, which means that $\{v_1, \ldots, v_k\}$ is a k-clique in G by Lemma 5.6. $\qquad\square$

Chapter 6

Descriptive Complexity

In this chapter we study the descriptive complexity of the k-CLIQUE problem. In particular, we are interested in the question of how many variables are required to express the property "there exists a k-clique" in the average-case setting. As usual, we fix a threshold function $p(n) = \Theta(n^{-2/(k-1)})$. Let φ be a first-order formula with no free variables in either the language of graphs (with adjacency relation $\sim$ only) or the language of ordered graphs (with linear order $<$ in addition to $\sim$). We say that φ defines k-CLIQUE w.h.p. on $\mathrm{G}(n,p)$ if $\lim_{n \to \infty} \Pr_{G \sim \mathrm{G}(n,p)}[G \models \varphi] = 1$.

We discuss some background on bounded-variable logics in §1.3. Recall that L^m denotes the m-variable fragment of first-order logic, defined as the collection of first-order formulas with at most m variables (including both free and bound variables). $L^1 \subseteq L^2 \subseteq \ldots$ is known as the variable hierarchy. The question of whether the variable hierarchy is strict in terms of expressive power on a given class of structures (Question 1.4) is widely studied. In particular, for the class of finite ordered graphs, this question (Question 1.5) was a major open problem in finite model theory.

This chapter contains the following results:

- (§6.1) *No formula with $\leqslant k/4$ variables defines k-CLIQUE w.h.p. on $\mathrm{G}(n,p)$ in the language of ordered graphs.*

- (§6.2) *The variable hierarchy is strict on finite ordered graphs.*

  A key step of showing *infinite* $\implies$ *strict* for the variable hierarchy on finite ordered graph is due to Neil Immerman.

- (§6.3.1) *No formula with $\leqslant k/2$ variables defines k-CLIQUE w.h.p. on $\mathrm{G}(n,p)$ in the language of graphs (without a linear order).*

- (§6.3.3) *There is a formula with $\frac{k}{2} + O(1)$ which defines k-CLIQUE w.h.p. on $\mathrm{G}(n,p)$ in the language of graphs.*

  A preliminary $\frac{k}{2} + \log k + O(1)$ variable upper bound is first presented in §6.3.2. The improvement to $\frac{k}{2} + O(1)$ was achieved jointly with Joel Spencer.

I am grateful to Neil Immerman and Joel Spencer for these contributions.

6.1 $k/4$ variable lower bound

There is a well-known correspondence in descriptive complexity between first-order logic and the complexity class AC^0 of polynomial-size constant-depth circuits [36, 37]. Under this correspondence, the m-variable L^m corresponds to AC^0 circuits of size $O(n^m)$ (see [38] for a detailed explanation). In the context of m-variable formulas in the language of ordered graphs, one direction of this correspondence is as follows:

Lemma 6.1. *For every m-variable formula $\varphi(x_1,\ldots,x_\ell)$ (with $\ell \leqslant m$ free variables $x_1,\ldots,x_\ell$) in the language of ordered graphs, there exists a circuit C_φ of size $O(n^m)$ and depth $O(1)$ with n^ℓ output nodes, denoted $\mathsf{C}_\varphi^{\vec{v}}$ for $\vec{v} \in [n]^\ell$, such that for every graph $G \in \mathscr{G}^n$ and ℓ-tuple of vertices $\vec{v} \in [n]^\ell$, we have $\mathsf{C}_\varphi^{\vec{v}}(G) = 1 \iff G \models \varphi(\vec{v})$.*

The proof is a simple inductive argument on formulas. As an immediate corollary:

Corollary 6.2. *No formula with $\leqslant k/4$ variables solves k-Clique w.h.p. on $\mathrm{G}(n,p)$ in the language of ordered graphs.*

Proof. Assume such a formula exists. Then by Lemma 6.1 there exists a circuit of size $O(n^{k/4})$ and depth $O(1)$ which solves k-Clique w.h.p. on $\mathrm{G}(n,p)$, contradicting Theorem 3.1. $\square$

Corollary 6.3. *The variable hierarchy is infinite on finite ordered graphs.*

Proof. The k-variable formula $\exists x_1 \ldots \exists x_k \bigwedge_{1 \leqslant i < j \leqslant k}(x_i \sim x_j)$ expresses "there exists a k-clique" on the class of finite ordered graphs (without even mentioning the linear order). By Corollary 6.2, no formula with $\leqslant k/4$ variables can express property on finite ordered graphs (even making use of the linear order). It follows that $L^{\lfloor k/4 \rfloor}$ is less expressive than L^k for every $k \in \mathbb{N}$. Therefore, the variable hierarchy is infinite. $\square$

Remark 6.4. Corollaries 6.2 and 6.3 are valid not only for the class of finite ordered graphs, but for classes of finite graphs with *arbitrary background relations* on the set $\{1,\ldots,n\}$ for every $n \in \mathbb{N}$ (for instance, arithmetic operations $+$ and $\times$). This is due to the fact that our bounded-depth circuit lower bound apply to *non-uniform* families of circuits. We state our results in terms of finite ordered graphs since even in this relatively simple setting, questions about the variable hierarchy had been wide open and widely studied.

6.2 Strictness of the variable hierarchy

Upon learning of our result that the variable hierarchy is infinite, Neil Immerman (personal communication) devised a clever proof that the hierarchy must therefore be strict on finite ordered graphs. In this section we present Immerman's proof that *infinite $\Longrightarrow$ strict*, which he was kind enough to let us include here (as this result is unpublished elsewhere).

In order to present the proof, we need some basic definitions and lemma from model theory. In particular, we present a version of Ehrenfeucht-Fraïssé game for the m-variable logic L^m. (For additional background, we again refer the reader to [25, 38, 50].) In what follows, let $\mathcal{A}$ and $\mathcal{B}$ be finite ordered graphs with vertex sets A and B of possibly different sizes (we assume $A \cap B = \emptyset$ for convenience). Recall that the *quantifier rank* of a formula φ is defined as the maximum nesting depth of quantifiers in φ. Also recall that formulas with no free variables are called *sentences*.

The *r-round m-pebble game* (independently defined by Barwise [10], Immerman [35] and Poizat [58]) gives a convenient characterization of when $\mathcal{A}$ and $\mathcal{B}$ satisfy the same first-order m-variable sentences of quantifier rank $\leqslant r$. The game, which is a modified version of the more familiar Ehrenfeucht-Fraïssé game, is played as follows. There are two players, Spoiler and Duplicator. The "game board" is the set $A \cup B$ and the "game pieces" are m matching pairs of pebbles $(\alpha_1, \beta_1), \ldots, (\alpha_m, \beta_m)$. Initially, all pebbles are off the board (that is, not placed anywhere in A or B). In the course of the game, pebbles α_i will be placed on elements of A and eventually moved from one element of A to another; pebbles β_i are similarly placed in B. In each round of the game:

- first, Spoiler selects any pebble (α_i or β_i for any $i \in \{1, \ldots, m\}$) and places it—or moves it, if has already been placed once—onto any element in the appropriate set (A or B, resp.);

- second, Duplicator places—or moves—the opposite pebble (β_i or α_i, resp.) in the other structure (B or A, resp.).

The game lasts for r rounds. Note that, at any point in time, the pebbles sitting in A and B describe a set of (up to m) pairs in $A \times B$. (Before round 1, this is the empty set.) Duplicator *wins* the game if and only if, at every point in time, the subset of $A \times B$ describe by the placement of pebbles constitutes a *partial isomorphism* from $\mathcal{A}$ to $\mathcal{B}$.[1] Duplicator has a *winning strategy* if—in the obvious sense—there exists a strategy for Duplicator which guarantees a win no matter how Spoiler plays.

We introduce some useful notation. Let $\mathcal{A} \equiv_r^m \mathcal{B}$ stand for the fact that Duplicator has a winning strategy in the r-round m-pebble game on $\mathcal{A}$ and $\mathcal{B}$. For $\ell \in \{1, \ldots, m\}$ and ℓ-tuples $\vec{a} \in A^\ell$ and $\vec{b} \in B^\ell$, let $(\mathcal{A}, \vec{a}) \equiv_r^m (\mathcal{B}, \vec{b})$ stand for the fact that Duplicator has a winning strategy in the r-round m-pebble game on $\mathcal{A}$ and $\mathcal{B}$ from the *starting configuration* with pebbles $\alpha_1, \ldots, \alpha_\ell, \beta_1, \ldots, \beta_\ell$ on $a_1, \ldots, a_\ell, b_1, \ldots, b_\ell$ (and all other pebbles off the board). The following fundamental fact about the r-round m-pebble game is proved in [10, 35, 58].

Lemma 6.5. $(\mathcal{A}, \vec{a}) \equiv_r^m (\mathcal{B}, \vec{b})$ *if, and only if,* $\mathcal{A} \models \varphi(\vec{a}) \iff \mathcal{B} \models \varphi(\vec{b})$ *for every m-variable formula $\varphi(\vec{x})$ of quantifier rank $\leqslant r$. (As a special case: $\mathcal{A} \equiv_r^m \mathcal{B}$ if, and only if, $\mathcal{A}$ and $\mathcal{B}$ satisfy exactly the same m-variable sentences of quantifier rank $\leqslant r$.)*

The next lemma is an easy corollary of Lemma 6.5.

Lemma 6.6. *Let $\mathscr{P}$ be a subclass of $\{$finite ordered graphs$\}$ (that is, a property of finite ordered graphs). $\mathscr{P}$ is <u>not</u> defined by any m-variable sentence if, and only if, for every r there exist finite ordered graphs $\mathcal{A}$ and $\mathcal{B}$ such that $\mathcal{A} \in \mathscr{P}$ and $\mathcal{B} \notin \mathscr{P}$ and $\mathcal{A} \equiv_r^m \mathcal{B}$.*

Note that $\equiv_r^m$ constitutes an equivalence relation on $\{$finite ordered graphs$\}$. A crucial fact (familiar to anyone acquainted with finite model theory) is:

Lemma 6.7. *For every m and r, there are only finitely many $\equiv_r^m$-equivalence classes of finite ordered graphs. Moreover, for every $\equiv_r^m$-equivalence class $\mathscr{E}$, there is a single m-variable sentence φ of quantifier rank $\leqslant r$ such that $\mathcal{A} \in \mathscr{E} \iff \mathcal{A} \models \varphi$.*

---

[1]Since $\mathcal{A}$ and $\mathcal{B}$ are ordered graphs, this means that if pebbles $\alpha_i, \alpha_j, \beta_i, \beta_j$ are siting on elements a_i, a_j, b_i, b_j, then $a_i = a_j \iff b_i = b_j$ and $a_i <^A a_j \iff b_i <^B b_j$ and $a_i \sim^A a_j \iff b_i \sim^B b_j$.

Lemma 6.7 has nothing to do with finite ordered graphs. Rather, it is a consequence of the fact that language of finite ordered graphs has *finite many* relation symbols. We remark that the number $f(m, r)$ of $\equiv_r^m$-equivalence relations is a non-elementary function of r for every sufficiently large fixed m.

Let "$L^{m-1} = L^m$" stand for the assertion that L^{m-1} and L^m express exactly the same properties of finite ordered graphs. We will eventually show that $L^{m-1} = L^m$ is <u>false</u> for every m. First, we note the following consequence of $L^{m-1} = L^m$.

Lemma 6.8. *Assume* $L^{m-1} = L^m$. *Then there exists a function* $\gamma : \mathbb{N} \to \mathbb{N}$ *such that* $\mathcal{A} \equiv_{\gamma(r)}^{m-1} \mathcal{B} \implies \mathcal{A} \equiv_r^m \mathcal{B}$ *for every* $r \in \mathbb{N}$.

Proof. Let $r \in \mathbb{N}$. By Lemma 6.7 there are only finitely many $\equiv_r^m$-equivalence classes $\mathscr{E}_1, \ldots, \mathscr{E}_t$, as well as m-variable sentences φ_i characterizing each $\mathscr{E}_i$. By the assumption that $L^{m-1} = L^m$, each φ_i is equivalent on finite ordered graphs to a sentence ψ_i with $\leqslant m - 1$ variables. Let s be the maximum quantifier rank among $\psi_1, \ldots, \psi_t$. We claim that this s gives a suitable value for the function γ on r. Indeed, suppose $\mathcal{A} \equiv_s^m \mathcal{B}$. There is a unique $i \in \{1, \ldots, t\}$ such that $\mathcal{A} \in \mathscr{E}_i$. Thus, $\mathcal{A} \models \varphi_i$ and so $\mathcal{A} \models \psi_i$. Since ψ_i has $\leqslant m - 1$ variable and quantifier rank $\leqslant s$ and $\mathcal{A} \equiv_s^{m-1} \mathcal{B}$, it follows that $\mathcal{B} \models \psi_i$. Thus, $\mathcal{B} \models \varphi_i$ and so $\mathcal{B} \in \mathscr{E}_i$. It follows that $\mathcal{A} \equiv_r^m \mathcal{B}$, since both belong to $\mathscr{E}_i$. $\qquad\square$

We introduce a small useful item of notation.

Definition 6.9. *For a finite ordered graph* $\mathcal{A}$ *and a vertex* $a \in A$, *let* $\mathcal{A}^a$ *be the finite ordered graph obtained from* $\mathcal{A}$ *by adding a new vertex — call it* a^* *— which is maximal in the linear order and adjacent only to* a. *For multiple vertices* $a_1, \ldots, a_\ell \in A$, *let* $\mathcal{A}^{a_1, \ldots, a_\ell}$ *be the finite ordered graph* $(\ldots (\mathcal{A}^{a_1})^{a_2} \ldots)^{a_\ell}$ *with vertex set* $A \cup \{a_1^*, \ldots, a_\ell^*\}$.

Lemma 6.10. *For all finite ordered graphs* $\mathcal{A}$ *and* $\mathcal{B}$ *and vertices* $a \in A$ *and* $b \in B$,

1. $(\mathcal{A}, a) \equiv_r^m (\mathcal{B}, b) \implies \mathcal{A}^a \equiv_r^{m-1} \mathcal{B}^b$, *and*

2. $\mathcal{A}^a \equiv_r^m \mathcal{B}^b \implies \mathcal{A} \equiv_{r-1}^m \mathcal{B}$.

Proof. (1) Duplicator plays the game for $\mathcal{A}^a \equiv_r^{m-1} \mathcal{B}^b$ according his winning strategy in the game for $(\mathcal{A}, a) \equiv_r^m (\mathcal{B}, b)$. (As a special case, whenever Spoiler plays the extra vertex a^* or b^*, Duplicator replies by playing the extra vertex in the opposite structure.)

(2) Consider Duplicator's winning strategy for $\mathcal{A}^a \equiv_r^m \mathcal{B}^b$. Note that in the first $r - 1$ rounds, Duplicator never plays the extra vertex a^* or b^* unless Spoiler does. Duplicator thus has a winning strategy for $\mathcal{A} \equiv_{r-1}^m \mathcal{B}$. $\qquad\square$

Finally, we present the key part of Immerman's argument.

Theorem 6.11. *Suppose that* $L^{m-1} = L^m$ *where* $m \geqslant 3$. *Then* $L^m = L^{m+1}$.

Proof. Toward a contradiction, suppose $L^m \neq L^{m+1}$. Then there exists a sentence φ in L^{m+1} that is not equivalence to any sentence in L^m. Let r be the quantifier-rank of φ. By Lemma 6.8, there is a function $\gamma : \mathbb{N} \to \mathbb{N}$ such that $\mathcal{A} \equiv_{\gamma(t)}^{m-1} \mathcal{B} \implies \mathcal{A} \equiv_t^m \mathcal{B}$ for all $t \in \mathbb{N}$ and finite ordered graphs $\mathcal{A}$ and $\mathcal{B}$. Fix any such γ and define sequence $t(0), \ldots, t(r)$ inductively by $t(0) = 2r$ and $t(i) = \gamma(t(i - 1)) + 1$ for $i = 1, \ldots, r$.

By Lemma 6.6, there exist finite ordered graphs $\mathcal{A}$ and $\mathcal{B}$ such that $\mathcal{A} \equiv_{t(r)}^m \mathcal{B}$ and $\mathcal{A} \models \varphi$ and $\mathcal{B} \models \neg\varphi$. We will show that $\mathcal{A} \equiv_r^{m+1} \mathcal{B}$ (thereby obtaining a contradiction) by

describing a winning strategy for Duplicator in the r-round $(m + 1)$-pebble game on $\mathcal{A}$ and $\mathcal{B}$.

Round 1: In round 1 of the game, suppose that Spoiler plays an element $a_1 \in A$ (without loss of generality). Duplicator consults his winning strategy for the game showing $\mathcal{A} \equiv^m_{t(r)}$ $\mathcal{B}$ and replies with an element $b_1 \in B$ such that $(\mathcal{A}, a_1) \equiv^m_{\gamma(t(r-1))} (\mathcal{B}, b_1)$ (here using $\gamma(t(r - 1)) = t(r) - 1$). By Lemma 6.10(1), we have $\mathcal{A}^{a_1} \equiv^{m-1}_{\gamma(t(r-1))} \mathcal{B}^{b_1}$. By definition of γ, it follows that $\mathcal{A}^{a_1} \equiv^m_{t(r-1)} \mathcal{B}^{b_1}$.

Round 2: In round 2, suppose that Spoiler plays $a_2 \in A$ (again without loss of generality, as the symmetric argument works if Spoiler instead plays an element of B). Duplication consults his winning strategy for $\mathcal{A}^{a_1} \equiv^m_{t(r-1)} \mathcal{B}^{b_1}$ and finds $b_2 \in B \cup \{b_1^*\}$ such that $(\mathcal{A}^{a_1}, a_2) \equiv^m_{\gamma(t(r-2))} (\mathcal{B}^{b_1}, b_2)$. Note that we are guaranteed that $b_2 \neq b_1^*$, since $a_2 \neq a_1^*$ and it takes only one quantifier to express that a given element is maximal in a finite linear order. Duplicator replies by playing b_2. Again by Lemma 6.10(1), we have $\mathcal{A}^{a_1,a_2} \equiv^{m-1}_{\gamma(t(r-2))} \mathcal{B}^{b_1,b_2}$. Again by definition of γ, we have $\mathcal{A}^{a_1,a_2} \equiv^m_{t(r-2)} \mathcal{B}^{b_1,b_2}$.

Round i: In all subsequent rounds, Duplicator plays in the same manner. After round i, we have elements $a_1, \ldots, a_i \in A$ and $b_1, \ldots, b_i \in B$ such that $\mathcal{A}^{a_1,\ldots,a_i} \equiv^m_{t(r-i)} \mathcal{B}^{b_1,\ldots,b_i}$. (One small point in generalizing from 2 to larger i: in reply to Spoiler playing $a_i \in A$, Duplicator consults his winning strategy for $\mathcal{A}^{a_1,\ldots,a_{i-1}} \equiv^m_{t(r-i+1)} \mathcal{B}^{b_1,\ldots,b_{i-1}}$ to obtain $b_i \in B \cup \{b_1^*, \ldots, b_{i-1}^*\}$ such that $(\mathcal{A}^{a_1,\ldots,a_{i-1}}, a_i) \equiv^m_{\gamma(t(r-i))} (\mathcal{B}^{b_1,\ldots,b_{i-1}}, b_i)$. Note that we are guaranteed that $b_i \notin \{b_1^*, \ldots, b_{i-1}^*\}$, since $a_i \notin \{a_1^*, \ldots, a_{i-1}^*\}$ and it takes only 2 ($\leqslant m$) variables and $i - 1$ ($\leqslant \gamma(t(r - i))$) quantifiers to express that a given element is among the $i - 1$ maximal elements in a finite linear order.)

At the end of the game (after r rounds) we have $\mathcal{A}^{a_1,\ldots,a_r} \equiv^m_{2r} \mathcal{B}^{b_1,\ldots,b_r}$. It follows that $\mathcal{A} \equiv^m_r \mathcal{B}$ by Lemma 6.10(2). However, this yields a contradiction, since $\mathcal{A}$ and $\mathcal{B}$ disagree on the sentence φ which has variable complexity m and quantifier rank r. □

As an immediate consequence of Corollary 6.3 (the variable hierarchy is infinite on finite ordered graphs) and Theorem 6.11 ($L^{m-1} = L^m \implies L^m = L^{m+1}$), we have:

Corollary 6.12 (Variable Hierarchy Theorem). *The variable hierarchy is strict on finite ordered graphs.* □

6.3 Average-case definability of k-Clique without order

This section studies the average-case first-order definability of k-CLIQUE on finite graphs without a linear order. We first show that $k/2$ variables are necessary in §6.3.1 (improving the $k/4$ variable lower bound which we proved in the ordered case). We then prove two upper bounds: a warm-up of $k/2 + \log k + O(1)$ in §6.3.2, before our final upper bound of $k/2 + O(1)$ in §6.3.3.

All formulas in this section are assumed to be formulas in the language of graphs (with adjacency relation $\sim$ only).

6.3.1 $k/2$ variables are necessary

The following notion of ℓ-extendibility shows up in the first-order theory of the infinite random graph, as well as the 0-1 law for first-order logic (see [4, 25, 50, 70]).

Definition 6.13. *A graph G is ℓ-extendible if it has $\geqslant \ell$ vertices and for all distinct vertices $v_1, \ldots, v_\ell$ and every $I \subseteq \{1, \ldots, \ell\}$, there exists a vertex w distinct from $v_1, \ldots, v_\ell$ such that w is adjacent to v_i for all $i \in I$ and non-adjacent to v_j for all $j \in \{1, \ldots, \ell\} \setminus I$.*

The next lemma gives a key property of ℓ-extendible graphs.

Lemma 6.14. *Every two ℓ-extendible graphs (finite or infinite) satisfy exactly the same sentences of $L^{\ell+1}$.*

Lemma 6.14 is easily argued using the $\ell + 1$-pebble game. Duplicator has a particularly simple winning strategy in the r-round $\ell + 1$-pebble game (for any r, indeed even $r = \infty$): so long as the current configuration of pebbles describes a partial isomorphism, the ℓ-extension property guarantees that, for any move of the Spoiler, there is a suitable reply which extends the previous configuration to a new partial isomorphism.

The following lemma is standard (see [4, 70]).

Lemma 6.15. *If $\min\{p(n), 1 - p(n)\} = \omega(n^{-1/\ell} \log^\ell n)$, then w.h.p. $G(n, p)$ is ℓ-extendible.*

Lemma 6.15 is essentially the fact that $\Theta(n^{-1/\ell} \log^\ell n)$ is the threshold for the monotone graph property that every ℓ vertices have a common neighbor.

Theorem 6.16. *No sentence with $\leqslant k/2$ variables solves k-CLIQUE w.h.p. on $G(n, p)$.*

Proof. Let $G \sim G(n, p)$ and let $\ell = \lfloor k/2 \rfloor - 1$. We have $\min\{p, 1 - p\} = p \geqslant \omega(n^{-1/\ell} \log^\ell n)$. Therefore, w.h.p. G is ℓ-extendible by Lemma 6.14. It follows from Lemma 6.15 that $\lim_{n \to \infty} \Pr[G \models \varphi] \in \{0, 1\}$ for every first-order sentence φ with $\ell + 1$ variables. But $0 < \lim_{n \to \infty} \Pr[G \text{ has a } k\text{-clique}] < 1$ since $p(n) = \Theta(n^{-2/(k-1)})$ is a threshold function for k-CLIQUE (Lemma 2.3). Therefore, $\ell + 1 = \lfloor k/2 \rfloor$ variables are insufficient. $\square$

6.3.2 $k/2 + \log k + O(1)$ **variables suffice**

The following lemma shows that ℓ-extendibility allows counting up to ℓ using only $\lceil \log \ell \rceil$ extra variables.

Lemma 6.17. *For every formula $\varphi(\vec{x}, y)$ and $\ell \in \mathbb{N}$, there is a formula $\psi(\vec{x})$ such that*

- *$\psi(\vec{x})$ is logically equivalent to $\exists^{\geqslant \ell} y\ \varphi(\vec{x}, y)$ on the class of ℓ-extendible graphs, and*

- *$\psi(\vec{x})$ contains $\lceil \log \ell \rceil$ more variables than $\varphi(\vec{x}, y)$.*

Proof. Let

$$\psi(\vec{x}) \equiv \exists z_1 \exists z_2 \ldots \exists z_{\lceil \log \ell \rceil} \bigwedge_{i \in \{0, \ldots, \ell - 1\}} \exists y\ \varphi(\vec{x}, y) \wedge \theta_i(y, \vec{z})$$

where

$$\theta_i(y, \vec{z}) \equiv \bigwedge_{j \in \{1, \ldots, \lceil \log \ell \rceil\}} \begin{cases} \neg(y \sim z_j) & \text{if the jth binary digit of i is 0,} \\ y \sim z_j & \text{if the jth binary digit of i is 1.} \end{cases}$$

It is easy to see that ψ has the correct semantics, using the fact that ℓ-extendibility implies that for all distinct $a_1, \ldots, a_\ell$ there exist $b_1, \ldots, b_{\lceil \log \ell \rceil}$ such that $\bigwedge_{i \in \{0, \ldots, \ell - 1\}} \theta_i(a_i, \vec{b})$. $\square$

Proposition 6.18. *$k/2 + \log k + O(1)$ variable suffice to define k-CLIQUE w.h.p. on $G(n, p)$.*

70

Proof. Let $\ell = \lfloor \frac{k}{2} \rfloor - 1$. Define formula $\psi(x_1, \ldots, x_{k-\ell}, y)$ by

$$\psi(\vec{x}, y) \equiv \bigwedge_{i \in \{1, \ldots, k-\ell\}} x_i \sim y.$$

$\psi(\vec{x}, y)$ expresses that y is a common neighbor of $x_1, \ldots, x_{k-\ell}$. By Lemma 6.17, there exists a sentence $\chi(\vec{x})$ such that

- on ℓ-extendible graphs, $\chi(x_1, \ldots, x_{k-\ell})$ is logically equivalent to "$x_1, \ldots, x_{k-\ell}$ has $\geqslant \ell$ common neighbors", and

- $\chi(x_1, \ldots, x_{k-\ell})$ has $k - \ell + 1 + \lceil \log \ell \rceil = \frac{k}{2} + \log k + O(1)$ variables.

Define the sentence φ by

$$\begin{aligned}
\varphi \equiv \exists x_1 \exists x_2 \ldots \exists x_{k-\ell} \ &\wedge \ \bigwedge_{1 \leqslant i < j \leqslant k-\ell} (x_i \sim x_j) \\
&\wedge \ \forall y \forall z \ \big((y \neq z) \wedge \psi(\vec{x}, y) \wedge \psi(\vec{x}, z)\big) \rightarrow (y \sim z) \\
&\wedge \ \chi(\vec{x}).
\end{aligned}$$

In plain language, φ says that there exist vertices $x_1, \ldots, x_{k-\ell}$ such that

- $x_1, \ldots, x_{k-\ell}$ form a clique,

- there is an edge between every two distinct common neighbors of $x_1, \ldots, x_{k-\ell}$, and

- $\chi(\vec{x})$ holds.

We will show that φ defines k-CLIQUE w.h.p. on $G \sim G(n, p)$. We first claim that w.h.p., G is ℓ-extendible and contains no P-subgraph for a particular pattern P.

- Note that $p(n) = \min\{p(n), 1 - p(n)\} = \omega(n^{-1/\ell} \log^\ell n)$. Therefore, w.h.p. G is ℓ-extendible by Lemma 6.15.

- Let P be the pattern with $V_P = \{1, \ldots, k+1\}$ and

$$E_P = \binom{\{1, \ldots, k\}}{2} \cup \Big\{\{i, k+1\} : i \in \{1, \ldots, k - \ell\}\Big\}$$

(that is, P is a complete pattern on k vertices plus one additional vertex of degree $k - \ell$). We have

$$\theta(P) \leqslant \frac{|V_P|}{|E_P|} = \frac{k+1}{\binom{k}{2} + k - \ell} < \frac{2}{k-1}.$$

Therefore, $n^{-2/(k-1)} = o(n^{-\theta(P)})$. By Lemma 2.3, it follows that w.h.p. G has no P-subgraph.

Let G be an arbitrary ℓ-extendible graph with no P-subgraph. It suffices to show that $G \models \varphi$ iff G has a k-clique.

For one direction, assume that $G \models \varphi$. Fix witnesses $x_1, \ldots, x_{k-\ell}$ for the existential quantifiers in φ. Since $G \models \chi(\vec{x})$ and G is ℓ-extendible, $x_1, \ldots, x_{k-\ell}$ have at least ℓ common neighbors (by definition of the formula $\chi(\vec{x})$). Since $x_1, \ldots, x_{k-\ell}$ form a clique and every two distinct common neighbors of $x_1, \ldots, x_{k-\ell}$ are adjacent (by definition of φ), it follows that G contains a k-clique.

71

For the other direction, assume that G contains a k-clique. Let $x_1, \ldots, x_{k-\ell}$ be any vertices in G which belong to a k-clique. Since G is ℓ-extendible and $x_1, \ldots, x_{k-\ell}$ have at least ℓ common neighbors, we have $G \models \chi(\vec{x})$. To show that $G \models \varphi$, it suffices to show that every two common neighbors of $x_1, \ldots, x_{k-\ell}$ are adjacent. This holds thanks to the fact that G has no P-subgraph: fix any $y_1, \ldots, y_\ell$ which extend $x_1, \ldots, x_{k-\ell}$ to a k-clique and note that if $x_1, \ldots, x_{k-\ell}$ have any other common neighbor $z \notin \{y_1, \ldots, y_\ell\}$, then G contains a P-subgraph on vertices $\vec{x}$, $\vec{y}$ and z. $\qquad\square$

6.3.3 $k/2 + O(1)$ variables suffice

In this section we shave $\log k$ off the previous $k/2 + \log k + O(1)$ upper bound.

Definition 6.19. *For an integer $\ell \geqslant 0$, let A_ℓ denote the following graph property: for all distinct vertices $x_1, \ldots, x_{\ell-2}, y_1, \ldots, y_\ell$, there exists a vertex z such that*

$$x_1, \ldots, x_{\ell-2}, y_i, y_j, z \text{ have a common neighbor} \iff |i - j| \leqslant 1$$

for all $i, j \in \{1, \ldots, \ell\}$.

We wish to establish that A_ℓ holds w.h.p. on $G(n, p)$ for a certain range of $p(n)$. First, we need a result of Shelah and Spencer [67] (from their work on a 0-1 law for random graphs $G(n, n^{-\alpha})$ for irrational $\alpha > 0$). First, a definition:

Definition 6.20. *A rooted graph is a pair (R, H) where $H = (V_H, E_H)$ is a graph and R is a subset of V_H. Let $|V_{R,H}| = |V_H \setminus R|$ and $|E_{R,H}| = |E_H \setminus \binom{R}{2}|$. For $\alpha \in (0, 1)$, we say that (R, H) is:*

- *α-dense*   *if $|V_{R,H}| - \alpha|E_{R,H}| < 0$,*

- *α-sparse*   *if $|V_{R,H}| - \alpha|E_{R,H}| > 0$,*

- *α-rigid*   *if for all S with $R \subseteq S \subset V_H$, (S, H) is α-dense,*

- *α-safe*   *if for all S with $R \subseteq S \subseteq V_H$, $(R, H|_S)$ is α-sparse.*

Now the required lemma of Spencer of Shelah (see Ch. 5 of [70] and Ch. 10 of [4] for a discussion of this result).

Lemma 6.21 (Generic Extension Theorem [67]). *Fix $\alpha > 0$ and integers $r, s, t \geqslant 0$. W.h.p., $G \sim G(n, n^{-\alpha})$ has the property that for all vertices $x_1, \ldots, x_r$ and every α-safe rooted graph (R, H) with $V_H = \{X_1, \ldots, X_r, Y_1, \ldots, Y_s\}$ and $R = \{X_1, \ldots, X_r\}$, there exist vertices $y_1, \ldots, y_s$ such that*

- *x_i, y_j (resp. y_i, y_j) are adjacent in G iff X_i, Y_j (resp. Y_i, Y_j) are adjacent in H (<u>note:</u> we don't care if adjacencies match up between x's and between X's), and*

- *for all vertices $z_1, \ldots, z_t$ such that $(\{x_1, \ldots, x_r, y_1, \ldots, y_s\}, G|_{\{x_1, \ldots, x_r, y_1, \ldots, y_s, z_1, \ldots, z_t\}})$ is α-rigid, there are no adjacencies between any pair y_i, z_j.*

Using Lemma 6.21, we now show the following:

Lemma 6.22. *If $n^{-1/\ell} \leqslant p(n) \leqslant n^{-1/(\ell+(1/2))}$, then w.h.p. $G(n, p)$ has property A_ℓ.*

Proof. Fix α in the range $1/(\ell + (1/2)) \leqslant \alpha \leqslant 1/\ell$ and let $G \sim G(n, n^{-\alpha})$. W.h.p., G satisfies the conclusion of Lemma 6.21 for $r = 2\ell - 2$, $s = \ell$ and $t = 1$. We will assume that this is the case and show that G has property A_ℓ.

Let $x_1, \ldots, x_{\ell-2}, y_1, \ldots, y_\ell$ be any distinct vertices of G. Consider the rooted graph (R, H) where $R = \{X_1, \ldots, X_{\ell-2}, Y_1, \ldots, Y_\ell\}$ and $H = (V_H, E_H)$ is given by

$$V_H = \{X_1, \ldots, X_{\ell-2}, Y_1, \ldots, Y_\ell, Z, W_1, \ldots, W_{\ell-1}\},$$

$$E_H = \bigcup_{i \in \{1, \ldots, \ell-1\}} \{\{X_1, W_i\}, \ldots, \{X_{\ell-2}, W_i\}, \{Y_i, W_i\}, \{Y_{i+1}, W_i\}, \{Z, W_i\}\}.$$

Note that $|V_{R,H}| = \ell$ and $|E_{R,H}| = \ell^2 - 1$ and hence

$$|V_{R,H}| - \alpha|E_{R,H}| = \ell - \frac{\ell^2 - 1}{\ell} = \frac{1}{\ell} > 0.$$

Thus (R, H) is α-sparse, and indeed is α-safe. By the generic extension property of Lemma 6.21, there exist vertices $w_1, \ldots, w_{\ell-1}, z$ in G such that

(i) not counting edges among $x_1, \ldots, x_{\ell-2}, y_1, \ldots, y_\ell$, the graph

$$G|_{\{x_1, \ldots, x_{\ell-2}, y_1, \ldots, y_\ell, w_1, \ldots, w_{\ell-1}, z\}}$$

(i.e., the induced subgraph of G on vertices $x_1, \ldots, x_{\ell-2}, y_1, \ldots, y_\ell, w_1, \ldots, w_{\ell-1}, z$) is isomorphic to H via the obvious bijection (x_i goes to X_i, etc.), and

(ii) for every neighbor v of z, the rooted graph

$$(\{x_1, \ldots, x_{\ell-2}, y_1, \ldots, y_\ell, w_1, \ldots, w_{\ell-1}, z\}, G|_{\{x_1, \ldots, x_{\ell-2}, y_1, \ldots, y_\ell, w_1, \ldots, w_{\ell-1}, z, v\}})$$

is not α-rigid.

By (i), vertices $x_1, \ldots, x_{\ell-2}, y_i, y_{i+1}, z$ have a common neighbor (namely, w_i) for every $i \in \{1, \ldots, \ell - 1\}$. To show that G has property A_ℓ, we must show that $x_1, \ldots, x_{\ell-2}, y_i, y_j, z$ do not have a common neighbor whenever $|i - j| \geqslant 2$. This follows from (ii). Toward a contradiction, assume that for some i, j with $|i - j| \geqslant 2$, $x_1, \ldots, x_{\ell-2}, y_i, y_j, z$ have a common neighbor v. Writing (R', H') for the rooted graph in (ii), note that $|V_{R',H'}| = 1$ and $|E_{R',H'}| \geqslant \ell + 1$ and hence

$$|V_{R',H'}| - \alpha|E_{R',H'}| < 1 - \frac{\ell + 1}{\ell - (1/2)} < 0.$$

But this means that (R', H') is α-dense and hence α-rigid, which contradicts (ii). $\qquad\square$

Theorem 6.23. *$k/2 + O(1)$ variable suffice to define k-CLIQUE w.h.p. on $G(n, p)$.*

Proof. Let $\ell = \lfloor \frac{k-1}{2} \rfloor$ and note that $n^{-1/(\ell+(1/2))} \leqslant n^{-2/(k-1)} \leqslant n^{-1/\ell}$. By Lemma 6.22, w.h.p. $G \sim G(n, n^{-2/(k-1)})$ has property A_ℓ. The sentence φ defining k-CLIQUE w.h.p. on G says the following: "There exist $x_1, \ldots, x_{k-\ell}$ and z such that $x_1, \ldots, x_{k-\ell}$ form a clique and, letting N be the set of common neighbors of $x_1, \ldots, x_{k-\ell}$ and letting R be the binary relation on N defined by

$$(y, y') \in R \iff y \neq y' \text{ and } x_1, \ldots, x_{\ell-2}, y, y', z \text{ have a common neighbor,}$$

73

the graph (N, R) is a path of length ℓ." Note that

1. the set N is defined by a formula with $k - \ell + 1$ variables (that is, one free variable in addition to $x_1, \ldots, x_{k-\ell}$),

2. the binary relation R is defined by a formula with $k - \ell + O(1)$ variables (that is, $O(1)$ extra variables—including two free variables y, y'—in addition to $x_1, \ldots, x_{\ell-2}$ and z),

3. the statement "(N, R) is isomorphic to a path of length ℓ" requires only $O(1)$ extra variables, in addition to the $k - \ell + O(1)$ total variables required to define N and R.

To see this last claim, note that a graph H is isomorphic to a path of length ℓ if, and only if, H has maximum degree 2, no isolated vertices, exactly two vertices of degree 1, and

$$\exists x \exists y \ (x \sim y) \wedge \exists z \ (y \sim z) \wedge (z \neq x) \wedge \exists x \ (z \sim x) \wedge (x \neq y) \wedge \exists y \ (x \sim y) \wedge (y \neq z) \wedge \exists z \ \cdots .$$

Thus, the entire sentence φ can be expressed using only $k - \ell + O(1) = k/2 + O(1)$ variables. Finally, it is easy to see that φ is equivalent to "there exists a k-clique" on graphs with property A_ℓ. $\qquad \square$

Chapter 7

Extensions and Open Problems

In this final chapter, we describe some extensions of our results and discuss some further questions raised by our work.

7.1 Extensions of our results

In this section we mention some extensions of our results and techniques:

- (§7.1.1) We identify the size-depth and size-gaps tradeoffs at the boundaries of our lower bounds.

- (§7.1.2) We give a version of our lower bounds with larger planted cliques.

- (§7.1.3) We mention a lower bound of Amano [6], using our techniques, for the subgraph isomorphism problem.

7.1.1 Size-depth and size-gap tradeoffs

Our lower bound of $\Omega(n^k)$ for bounded-depth circuits (Theorem 3.1) exhibits no size-depth tradeoff up to depth $k^{-2}\log n/\log\log n$. Above depth $k^{-2}\log n/\log\log n$, however, our bound begins to exhibit a tradeoff. We state this tradeoff below, followed by a brief explanation of the changes to the proof required for this different range of parameters. As usual, $p(n) = \Theta(n^{-2/(k-1)})$.

Theorem 7.1. *For all* $\lambda \in (0,1)$, *Boolean circuits of size* $O(n^{(1-\lambda)k/4})$ *and depth* $\lambda k^{-1}\log n/\log\log n$ *cannot solve* k-Clique *w.h.p. on* $G(n,p)$.

For example, we get a lower bound of $\Omega(n^{k/8})$ for circuits of depth $\frac{1}{2}k^{-1}\log n/\log\log n$. To prove Theorem 7.1, we need only to modify single lemma (Lemma 3.15) in the proof of Theorem 3.1. The modified lemma is as follows:

Lemma 7.2. *Let* P *be a fixed small or medium pattern and let* $p(n) = n^{-2/(k-1)}$. *Suppose* $f : \mathscr{G}^n \to \{0,1\}^{n^{o(1)}}$ *is computed by circuits of size* $n^{O(1)}$ *and depth* $\lambda k^{-1}\log n/\log\log n + O(1)$. *Then*

$$\Pr_{\substack{G\sim G(n,p) \\ H\sim \mathrm{Plant}(n,P)}} \left[H \text{ is an } f^{\mathbf{G}}\text{-core}\right] = \begin{cases} o(n^{-1}) & \text{if } P \text{ is nonempty,} \\ o(n^{-(1-\lambda)\frac{k}{4}-\frac{1}{4}}) & \text{if } P \text{ is medium.} \end{cases}$$

The proof of Lemma 7.2 is the same as the proof of Lemma 3.15, except that we set $q(n)$ to $n^{-(\lambda k^{-1}+k^{-3})}$ instead of $n^{-(k^{-2}+k^{-3})}$.

We get a similar "size-gap tradeoff" for our monotone circuit lower bound:

Theorem 7.3. *For all* $\lambda \in (0,1)$*, monotone circuits of size* $O(n^{(1-\lambda)k/4})$ *cannot solve* k-CLIQUE *w.h.p. on both* $G(n,p)$ *and* $G(n,p+p^{1+\lambda k^{-1}})$.

Compared with Theorem 4.1, the "gap" between the two thresholds is reduced from $p^{1+k^{-2}}$ to $p^{1+\lambda k^{-1}}$. The proof of this theorem modifies the argument of Theorem 4.1 simply by setting δ equal to λk^{-1} instead of k^{-2}.

7.1.2 Planting a larger clique

Our two primary lower bounds (Theorems 3.2 and 4.3) can be strengthened in the direction of planting a larger clique.

Theorem 7.4. *Let* f *be a Boolean graph function and let* $\boldsymbol{G} \sim G(n,p)$ *and* $\boldsymbol{G}^- \sim G(n,p^{1+k^{-2}})$ *and* $\boldsymbol{K}_{\kappa(n)} \sim \mathrm{Plant}(n,K_{\kappa(n)})$ *where* $\kappa(n) = \lceil n^{k^{-2}} \rceil$.

1. *If* f *is computed by Boolean circuits of size* $O(n^{k/4})$ *and depth* $k^{-2}\log n/\log\log n$, *then w.h.p.* $f(\boldsymbol{G}) = f(\boldsymbol{G} \cup \boldsymbol{K}_{\kappa(n)})$.

2. *If* f *is computed by monotone circuits of size* $O(n^{k/4})$ *and* $\mathrm{E}[f(\boldsymbol{K}_{\kappa(n)})] = 1 - o(1)$, *then* $\mathrm{E}[f(\boldsymbol{G}^-)] = 1 - \exp(-\Omega(n^{k^{-2}}))$.

Here k is still fixed, but we are considering the random planted $\kappa(n)$-clique $\boldsymbol{K}_{\kappa(n)}$ instead of the random planted k-clique $\boldsymbol{K}_k$ (as in Theorems 3.2 and 4.3). To prove Theorem 7.4 we observe that the only relevant fact in the proofs of Theorems 3.2 and 4.1 is that $\boldsymbol{K}_k$ has $o(n^{1/k})$ medium subgraphs (indeed, $\boldsymbol{K}_k$ has only $2^{\binom{k}{2}}$ subgraphs). These proofs remain valid for Theorem 7.4 since $\boldsymbol{K}_{\kappa(n)}$ has at most $2^{\binom{k-1}{2}}\binom{\kappa(n)}{k-1} = o(n^{1/k})$ medium subgraphs (note that medium graphs have at most $k-1$ non-isolated vertices).

7.1.3 Subgraph isomorphism problem

For a fixed graph H, the *H-subgraph isomorphism problem* is the problem of determining whether a graph contains a subgraph isomorphic to H. Using our technique on bounded-depth circuits (as presented in the paper [64]), Amano [6] proved a lower bound of $\Omega(n^{c(H)})$ for the average-case complexity on AC^0 circuits of the H-subgraph isomorphism problem for every fixed H. Here $c(H)$ is a constant depending on H (for instance, $c(K_k) = k/4$). The generalization is achieved essentially by picking an appropriate class of "small" patterns P relative to the graph H (for instance, as in the case $H = K_k$, one could define "small" to mean "has $< |V_H|$ non-isolated vertices"). Additionally, Amano proved a lower bound of $\Omega(n^{k(1-(\ln\ell+2)/(\ell-1))})$ for the average-case complexity on AC^0 circuits of the k-clique problem on ℓ-uniform hypergraphs for all $k > \ell \geqslant 2$. This lower bound leads to a stronger Size Hierarchy Theorem for uniform AC^0: whereas our result implies that

$$\text{uniform } \mathsf{AC}^0(\text{size } O(n^{k/4})) \neq \text{uniform } \mathsf{AC}^0(\text{size } O(n^k))$$

for all $k \in \mathbb{N}$ (hence the size hierarchy is infinite), Amano obtains the sharper result that

$$\text{uniform } \mathsf{AC}^0(\text{size } O(n^\alpha)) \neq \text{uniform } \mathsf{AC}^0(\text{size } O(n^\beta))$$

76

for all $\beta > \alpha > 0$ (hence the size hierarchy is strict).

7.2 Open problems

Finally, we discuss a few questions raised by this work.

- (§7.2.1) What is the complexity of finding a clique of size $(\frac{1}{2} + \varepsilon)k$ in the random graph $G(n, n^{-2/(k-1)})$?

- (§7.2.2) Does our lower bound for monotone circuits hold at a single threshold?

We ask some additional questions in §7.2.3.

7.2.1 Karp's question at the k-clique threshold

We recall a question stated earlier in §1.2.2:

Question 7.5. *Is there an $O(n^{(1-\delta)\varepsilon^2 k})$ algorithm which w.h.p. finds a clique of size $(\frac{1}{2}+\varepsilon)k$ in $G(n, n^{-2/(k-1)})$, for any constants $\delta > 0$ and $\varepsilon \in (0, \frac{1}{2})$?*

This question is a "scaling" to $G(n, n^{-2/(k-1)})$ of Karp's question [46] about finding cliques of size $(1 + \varepsilon) \log n$ in $G(n, \frac{1}{2})$ in polynomial time (see §1.2.1). One direction of future research is to show, using the techniques of this thesis, that the answer to Question 7.5 is "no" for bounded-depth circuits.

7.2.2 Monotone lower bound at a single threshold

We showed that monotone circuits of size $O(n^{k/4})$ cannot solve k-CLIQUE w.h.p. on both $G(n, p_1)$ and $G(n, p_2)$ for two sufficiently separated threshold functions (Theorem 4.1). It is open whether this lower bound can be extended to monotone circuits which solve k-CLIQUE at a single threshold (as in our bounded-depth circuit result). We conjecture that it can.

Conjecture 7.6. *Monotone circuits of size $O(n^{k/4})$ cannot solve k-CLIQUE w.h.p. on $G(n, n^{-2/(k-1)})$.*

We feel that Conjecture 7.6 may be a hard problem. In particular, it seems to call for an approach outside the framework of Razborov's approximation method, which seems to break down when the distributions on positive and negative inputs are brought too close together.

7.2.3 Additional questions

We briefly mention a few additional questions and directions for future work.

1. What is the worst-case complexity of k-CLIQUE on constant-depth circuits? Similarly, how many variables are needed to express "there exists a k-clique" on finite ordered graphs, in the usual (non-average-case) sense? Despite our results for the average case, upper bounds of $n^{k-\Omega(1)}$ or $< k$ variables would be surprising.

2. We showed that $k/4$ first-order variables are necessary to define k-CLIQUE in the average case on finite linearly ordered graphs. Are $k/4 + O(1)$ variables sufficient? We suspect that the answer is "yes". This can perhaps be shown using ideas along the lines of our $k/2+O(1)$ variable upper bound for unordered graphs (but we have not worked out the details). We remark that $k/4+O(1)$ variables are sufficient on graphs with built-in arithmetic operations $+$ and $\times$. (This follows from the observation that the circuits in our upper bound, Theorem 5.1, can be made uniform, together with a descriptive complexity characterization of uniform AC^0 [37].)

3. What is the minimum *quantifier rank* required to define k-CLIQUE in the average case on graphs (with or with a linear order)? Joel Spencer (personal communication) observed that, for all k in the range of a certain "busy beaver" function BB : $\mathbb{N} \to \mathbb{N}$, quantifier rank $k/2 + \mathrm{BB}^{-1}(k) + O(1)$ is sufficient. It would be interesting to see a lower bound of $k/2 + \omega_k(1)$.

4. Finally, in §4.3 we introduced a new notion of "quasi-sunflowers" (Definition 4.9) and gave a "quasi-sunflower lemma" (Theorem 4.11) along the lines of the Erdős-Rado sunflower lemma [26], which played a key role in our monotone circuit lower bound. Since quasi-sunflowers seem to be a natural relaxation of sunflowers, it would be interesting to see other applications.

Bibliography

[1] Miklós Ajtai. Σ_1^1 formulae on finite structures. *Annals of Pure and Applied Logic*, 24:1–48, 1983.

[2] Noga Alon and Ravi B. Boppana. The monotone circuit complexity of boolean functions. *Combinatorica*, 7(1):1–22, 1987.

[3] Noga Alon, Michael Krivelevich, and Benny Sudakov. Finding a large hidden clique in a random graph. *Random Structures and Algorithms*, 13(3-4):457–466, 1998.

[4] Noga Alon and Joel Spencer. *The Probabilistic Method, 3rd Edition*. John Wiley, 2008.

[5] Noga Alon, Raphael Yuster, and Uri Zwick. Color-coding. *Journal of the ACM*, 42(4):844–856, 1995.

[6] Kazuyuki Amano. k-Subgraph isomorphism on AC^0 circuits. *Computational Complexity*, 19(2):183–210, 2010.

[7] Kazuyuki Amano and Akira Maruoka. A superpolynomial lower bound for a circuit computing the clique function with at most $(1/6)\log\log n$ negation gates. *SIAM Journal on Computing*, 35(1):201–215, 2005.

[8] Alexander E. Andreev. On a method for obtaining lower bounds for the complexity of individual monotone functions. *Doklady Akademii Nauk SSSR*, 282(5):1033–1037 (in Russian), 1985.

[9] David Mix Barrington, Jonathan Buss, and Neil Immerman. Number of variables is equivalent to space. *Journal of Symbolic Logic*, 66:1217–1230, 2001.

[10] Jon Barwise. On Moschovakis closure ordinals. *Journal of Symbolic Logic*, 42(2):292–296, 1977.

[11] Paul Beame. Lower bounds for recognizing small cliques on CRCW PRAM's. *Discrete Appl. Math.*, 29(1):3–20, 1990.

[12] Paul Beame. A switching lemma primer. Technical Report UW-CSE-95-07-01, Department of Computer Science and Engineering, University of Washington, November 1994.

[13] Béla Bollobás. Complete subgraphs are elusive. *Journal of Combinatorial Theory, Series B*, 21(1):1–7, 1976.

[14] Béla Bollobás. Threshold functions for small subgraphs. *Math. Proc. Camb. Phil. Soc.*, 90:197–206, 1981.

[15] Béla Bollobás. *Random Graphs (2nd Edition)*. Cambridge University Press, 2001.

[16] Ravi B. Boppana. The average sensitivity of bounded-depth circuits. *Inf. Process. Lett.*, 63(5):257–261, 1997.

[17] Ravi B. Boppana and Michael Sipser. The complexity of finite functions. In *Handbook of theoretical computer science (Vol. A): Algorithms and complexity*, pages 757–804. MIT Press, Cambridge, MA, USA, 1990.

[18] Coenraad Bron and Joep Kerbosch. Finding all cliques of an undirected graph (algorithm 457). *Commun. ACM*, 16(9):575–576, 1973.

[19] Shiva Chaudhuri and Jaikumar Radhakrishnan. Deterministic restrictions in circuit complexity. In *STOC '96: Proceedings of the 28th Annual ACM Symposium on Theory of Computing*, pages 30–36, 1996.

[20] Jianer Chen, Xiuzhen Huang, Iyad A. Kanj, and Ge Xia. Strong computational lower bounds via parameterized complexity. *J. Comput. Syst. Sci.*, 72(8):1346–1367, 2006.

[21] Don Coppersmith and Shmuel Winograd. Matrix multiplication via arithmetic progressions. *J. Symb. Comput.*, 9(3):251–280, 1990.

[22] Anuj 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.

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

[24] Rod G. Downey and Michael R. Fellows. Fixed-parameter tractability and completeness I: Basic results. *SIAM Journal on Computing*, 24(4):873–921, 1995.

[25] Heinz-Dieter Ebbinghaus and Jorg Flum. *Finite Model Theory*. Springer-Verlag, 1996.

[26] Paul Erdős and Richard Rado. Intersection theorems for systems of sets. *J. London Math. Soc.*, 35:85–90, 1960.

[27] Paul Erdős and Alfred Rényi. On the evolution of random graphs. *Mat. Kutató Int. Közl*, 5:17–60, 1960.

[28] Merrick L. Furst, James B. Saxe, and Michael Sipser. Parity, circuits, and the polynomial-time hierarchy. *Mathematical Systems Theory*, 17:13–27, 1984.

[29] Mikael Goldmann and Johan 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.

[30] E. Grädel, P.G. Kolaitis, L. Libkin, M. Marx, J. Spencer, M.Y. Vardi, Y. Venema, and S. Weinstein. *Finite Model Theory and its Applications*. Springer, 2007.

[31] Martin Grohe. Finite variable logics in descriptive complexity theory. *Bulletin of Symbolic Logic*, 4(4):345–398, 1998.

[32] Yuri Gurevich and Harry R. Lewis. A logic for constant-depth circuits. *Information and Control*, 61(1):65–74, 1984.

[33] 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.

[34] Johan Håstad. Clique is hard to approximate within $n^{1-\epsilon}$. *Acta Mathematica*, 182(1):105–142, 1999.

[35] Neil Immerman. Upper and lower bounds for first order expressibility. *J. Comput. Syst. Sci.*, 25(1):76–98, 1982.

[36] Neil Immerman. Languages that capture complexity classes. *SIAM Journal of Computing*, 16:760–778, 1987.

[37] Neil Immerman. Expressibility and parallel complexity. *SIAM Journal on Computing*, 18(3):625–638, 1989.

[38] Neil Immerman. *Descriptive Complexity*. Graduate Texts in Computer Science. Springer-Verlag, New York, 1999.

[39] Svante Janson. Poisson approximation for large deviations. *Random Structures and Algorithms*, 1(2):221–230, 1990.

[40] Svante Janson, Tomasz Luczak, and Andrzej 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.

[41] Svante Janson, Tomasz Luczak, and Andrzej Ruciński. *Random Graphs*. John Wiley, 2000.

[42] Mark Jerrum. Large cliques elude the metropolis process. *Random Structures and Algorithms*, 3(4):347–359, 1992.

[43] Ari Juels and Marcus Peinado. Hiding cliques for cryptographic security. *Des. Codes Cryptography*, 20(3):269–280, 2000.

[44] Stasys Jukna. *Extremal Combinatorics with Applications in Computer Science*. Springer, Heidelberg, 2001.

[45] Richard M. Karp. Reducibility among combinatorial problems. In R. E. Miller and J. W. Thatcher, editors, *Complexity of Computer Computations*, pages 85–103. Plenum Press, 1972.

[46] Richard M. Karp. Probabilistic analysis of some combinatorial search problems. In J. F. Traub, editor, *Algorithms and Complexity: New Directions and Recent Results*, pages 1–19. Academic Press, 1976.

[47] Michal Koucky, Clemens Lautemann, Sebastian Poloczek, and Denis Therien. Circuit lower bounds via Ehrenfeucht-Fraisse games. In *CCC '06: Proceedings of the 21st Annual IEEE Conference on Computational Complexity*, pages 190–201, 2006.

[48] Luděk Kučera. Expected complexity of graph partitioning problems. *Discrete Appl. Math.*, 57(2-3):193–212, 1995.

[49] Leonid Levin. Average case complete problems. *SIAM Journal on Computing*, 15(1):285–286, 1986.

[50] Leonid Libkin. *Elements of Finite Model Theory*. Springer-Verlag, 2004.

[51] Nathan Linial, Yishay Mansour, and Noam Nisan. Constant depth circuits, fourier transform, and learnability. *Journal of the ACM*, 40(3):607–620, 1993.

[52] James F. Lynch. A depth-size tradeoff for boolean circuits with unbounded fan-in. In *Structure in Complexity Theory Conference*, pages 234–248, 1986.

[53] John W. Moon and Leo Moser. On cliques in graphs. *Israel J. Math.*, 3:23–28, 1965.

[54] Jaroslav Nešetřil and Svatopluk Poljak. On the complexity of the subgraph problem. *Comment. Math. Univ. Carolinae.*, 26(2):415–419, 1985.

[55] Martin Otto. *Bounded variable logics and counting: a study in finite models*. Springer-Verlag, 1997.

[56] Marcus Peinado. Hard graphs for the randomized Boppana-Halldórsson algorithm for MAXCLIQUE. *Nordic J. of Computing*, 1(4):493–515, 1994.

[57] Marcus Peinado. Go with the winners algorithms for cliques in random graphs. In Peter Eades and Tadao Takaoka, editors, *Algorithms and Computation*, volume 2223 of *Lecture Notes in Computer Science*, pages 525–537. Springer Berlin / Heidelberg, 2001.

[58] Bruno Poizat. Deux ou trois choses que je sais de L_n. *Journal of Symbolic Logic*, 47(3):641–658, 1982.

[59] Prabhakar Ragde and Avi Wigderson. Linear-size constant-depth polylog-threshold circuits. *Inf. Process. Lett.*, 39(3):143–146, 1991.

[60] Alexander 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.

[61] Alexander A. Razborov. Lower bounds on the size of bounded depth circuits over a complete basis with logical addition. *Math. Notes*, 41:333–338, 1987.

[62] Alexander A. Razborov. On the method of approximations. In *Proceedings of the 21st Annual ACM Symposium on the Theory of Computing*, pages 167–176, 1989.

[63] Alexander A. Razborov and Steven Rudich. Natural proofs. *J. Comput. Syst. Sci.*, 55(1):24–35, 1997.

[64] Benjamin Rossman. On the constant-depth complexity of k-clique. In *STOC '08: Proceedings of the 40th Annual ACM Symposium on Theory of Computing*, pages 721–730, 2008.

[65] Benjamin Rossman. Ehrenfeucht-Fraïssé games on random structures. In *WoLLIC '09: Proceedings of 15th Workshop on Logic, Language, Information and Computation*, pages 350–364, 2009.

[66] Benjamin Rossman. The monotone complexity of k-clique on random graphs. In *FOCS '10: Proceedings of 51st Annual IEEE Symposium on Foundations of Computer Science*, 2010 (to appear).

[67] Saharon Shelah and Joel Spencer. Zero-one laws for sparse random graphs. *J. Amer. Math. Soc.*, 1:97–115, 1988.

[68] Michael Sipser. Borel sets and circuit complexity. In *STOC '83: Proceedings of the Fifteenth Annual ACM Symposium on Theory of Computing*, pages 61–69, New York, NY, USA, 1983. ACM Press.

[69] Roman Smolensky. Algebraic methods in the theory of lower bounds for boolean circuit complexity. In *STOC '87: Proceedings of the 19th Annual ACM Symposium on Theory of Computing*, pages 77–82, 1987.

[70] Joel Spencer. *The Strange Logic of Random Graphs*. Springer, 2001.

[71] Éva Tardos. The gap between monotone and non-monotone circuit complexity is exponential. *Combinatorica*, 8(1):141–142, 1988.

[72] Ingo Wegener. On the complexity of branching programs and decision trees for clique functions. *Journal of the ACM*, 35(2):461–471, 1988.

[73] Ingo Wegener. *The Complexity of Boolean Functions*. Wiley, 1991.

[74] Andrew C-C. Yao. Separating the polynomial-time hierarchy by oracles. In *Proc. 26th Annual Symposium on Foundations of Computer Science*, pages 1–10, 1985.

[75] David Zuckerman. Linear degree extractors and the inapproximability of max clique and chromatic number. *Theory of Computing*, 3(1):103–128, 2007.