COMPUTATIONAL COMPLEXITY AND EXTREMAL COMBINATORICS

BRUNO PASQUALOTTO CAVALAR

Contents

1. Introduction

Recently there has been a growing interest in *average-case* lower bounds, in which one tries to prove not only that a certain complexity class cannot compute a particular Boolean function, but also that it cannot even approximate it under a given distribution. Average-case lower bounds are interesting in their own right, as they provide sharper separation between complexity classes than a worst-case lower bound, but they have also found applications in other branches of complexity theory, such as pseudorandomness [16] and learning theory [15]. Moreover, product distributions are thought to be a source of hard instances for monotone Boolean functions [22], so that proving average-case lower bounds under a product distribution for monotone circuits is a major question regarding monotone computation. In this project, we are interested in studying average-case complexity questions for restricted but natural classes of circuits, such as bounded-depth circuits, formulas and monotone circuits.

In this line of research, Rossman has recently obtained many new results of major importance. First, in a series of works [18, 21, 22] along with a joint work with Li and Razborov [14], Rossman obtained average-case lower bounds for subgraph isomorphism in various settings. We remark

Date: 2019/03/07, 9:24pm

that [21] implied the first *size-hierarchy theorem* for bounded-depth circuits. Secondly, Håstad, Rossman, Servedio and Tan [11] obtained the first average-case depth hierarchy theorem for bounded-depth circuits, thus solving a longstanding open problem. Finally, Rossman developed in [20] a new technique called *pathset complexity*, allowing him to prove an average-case lower bound for $\mathbf{AC}^0$ formulas. Using the same technique, Rossman proved the first correlation bound against a monotone class of circuits (monotone formulas) under a product distribution [22].

In this project, we aim at developing some of these results, by seeking a stronger average-case size-hierarchy theorem (Problem 3), new applications of the pathset complexity technique (Problem 5), an average-case lower bound against k-CLIQUE for monotone formulas under an Erdős–Rényi random graph (Problem 2), and by understanding the behaviour of monotone circuits under two-sliced distributions (Problem 6).

2. A review of basic results of circuit complexity

Boolean circuits are arguably the most important combinatorial model of computation studied in circuit complexity. Since any polynomial-time algorithm can be implemented by a sequence of polynomial-size circuits (one for each input length), obtaining a superpolynomial lower bound on the minimum circuit size of any problem in **NP** is enough to separate **P** from **NP**.

Definition 1. For every $n \in \mathbb{N}$, an *n-input, single-output Boolean circuit* is a directed acyclic graph with n *sources* and one *sink*. All nonsource vertices are called *gates* and are labeled with one of $\{\vee, \wedge, \neg\}$. The vertices labeled with $\vee$ and $\wedge$ usually have *fan-in* (that is, in-degree) equal to 2, but sometimes we allow unbounded fan-in. When we do, we will mention it explicitly. The vertices labeled with $\neg$ always have fan-in 1. The *size* of a circuit C is the number of gates it contains, and the *depth* is the maximum number of gates on a path from an input to the output. Finally, the *alternation-depth* of a circuit is the maximum number of alternations between $\wedge$ and $\vee$ gates in a path from an input to the output.

One reason for studying the computational model of Boolean circuits is the hope that combinatorial methods might be succesful in proving unconditional lower bounds. Unfortunately, until now no superlinear lower bound on the circuit size is known for any explicit Boolean function. This motivates the search for lower bounds in restricted classes of Boolean circuits, as we explain next.

One of the most basic complexity classes is $\mathbf{AC}^0$, the class of polynomial-size circuits with constant alternation-depth or, equivalently, polynomial-size circuits with unbounded fan-in and constant depth. It is one of the complexity classes on which we have had the most success in proving lower bounds. For example, a celebrated result of Furst, Saxe and Sipser [7] and Ajtai [1] is that no $\mathbf{AC}^0$ circuit can compute the function PARITY : $\{0,1\}^n \to \{0,1\}$, which returns the parity of $\sum_{i\in[n]} x_i$ for $x \in \{0,1\}^n$. This result was further optimized by Håstad [9] through his *switching lemma*, proving that constant-depth circuits computing PARITY must have exponential size.

Another model of computation on which we have had considerable success in proving lower bounds is that of *monotone circuits*. Monotone circuits are circuits without a $\neg$ gate. It is easy to see that monotone circuits only compute monotone Boolean functions, and that every monotone Boolean function can be computed by a monotone circuit. The first superpolynomial lower bound on the monotone complexity of a Boolean function was obtained by Razborov [17]. The result

goes as follows. Let k-CLIQUE : $\{0,1\}^{\binom{n}{2}} \to \{0,1\}$ be the function that, given an adjacency matrix of a graph G on n vertices, outputs 1 if and only if G contains a k-clique. The results of Razborov [17], together with the improvements of Andreev [3] and Alon and Boppana [2], imply that the monotone complexity of k-CLIQUE is $n^{\Omega(\sqrt{k})}$ for all $k \leqslant n^{1/4}$.

2.1. **Average-case complexity.** A probability distribution μ over $\{0,1\}^n$ is said to be a *product distribution* if there exists $\mu_1, \mu_2, \ldots, \mu_n$ such that $\mathbb{P}_{x\sim\mu}[x_i = 1] = \mu_i$, and every *bit* of $x \sim \mu$ is independent of the others. Thus, for every $a \in \{0,1\}^n$ we have $\mathbb{P}_{x\sim\mu}[x = a] = \left(\prod_{a_i=1} \mu_i\right)\left(\prod_{a_i=0}(1-\mu_i)\right)$. Observe that the Erdős–Rényi random graph $G(n,p)$ can be interpreted as a product distribution on $\{0,1\}^{\binom{n}{2}}$. A boolean function f is said to be γ-*hard* for a class of circuits $\mathcal{C}$ under a distribution μ if $\mathbb{P}_{x\sim\mu}[f(x) = C(x)] \leqslant \gamma$ for every circuit $C \in \mathcal{C}$. By default, μ is the uniform distribution and γ is typically written as $1/2 + \delta$ or $1 - \delta$, where $\delta = \delta(n) \to 0$. We moreover say that a Boolean function g has *correlation* γ (or is γ-*correlated*) with f under μ if $\mathbb{P}_{x\sim\mu}[f(x) = C(x)] \geqslant \gamma$, and we use the same term for any family of circuits $\mathcal{C}$ such that there exists a function g computable by a circuit $C \in \mathcal{C}$ that is γ-correlated with f.

Understanding the (average-case) circuit complexity of natural computational problems under interesting input distributions is an important research question in the theory of computation. Indeed, proving that a Boolean function f is γ-hard for a family of Boolean circuits $\mathcal{C}$ (i.e., an average-case lower bound or correlation bound) is a stronger result than proving that no circuit of $\mathcal{C}$ can compute f exactly (i.e., a worst-case lower bound).

Moreover, recently some worst-case lower bounds have been matched by average-case lower bounds under product distributions. For example, Håstad [12] recently proved that PARITY is $(1/2+2^{-\Omega(n/(\log S)^{d-1})})$-hard for depth-$d$ circuits of size S under the uniform distribution, through a stronger version of the above mentioned switching lemma, called *multi-switching lemma*. Let k-CYCLE : $\{0,1\}^{\binom{n}{2}} \to \{0,1\}$ be the function that, given an adjacency matrix of a graph G on n vertices, outputs 1 if and only if G contains a k-cycle. Using his new technique called *pathset complexity* (see Section 3.3), Rossman [22] also proved that k-CYCLE is $(1/2 + n^{-1/2+o(1)})$-hard under $G(n, p_c^{C_k})$ for monotone polynomial-size formulas (also known as $\mathbf{mNC}^1$), where $p_c^{C_k} \asymp 1/n$ is such that $\mathbb{P}[G(n, p_c^{C_k}) \text{ contains a } k\text{-cycle}] = 1/2$.

In this project we are interested in studying circuit complexity problems in the average-case setting for important classes of circuits, such as bounded-depth circuits and monotone formulas, as we explain in the next sections.

3. Research questions

3.1. **Lower bounds for subgraph isomorphism.** Subgraph isomorphism is one of the most important problems in computational complexity. For example, a lower bound of $n^{\Omega(k)}$ for k-CLIQUE for every fixed k would be enough to separate $\mathbf{P}$ from $\mathbf{NP}$.[1] This motivates the study of the circuit complexity of subgraph isomorphism, not only for cliques but for other fixed graphs.

One breakthrough of the area was a result due to Rossman [18], who proved that every $\mathbf{AC}^0$ circuit of size $O(n^{k/4})$ has correlation $1/2 + n^{-\Omega(k)}$ with k-CLIQUE under $G(n, p_c^{K_k})$, where $p_c^{K_k}$ is such that $\mathbb{P}[G(n, p_c^{K_k}) \text{ contains a } k\text{-clique}] = 1/2$. This result implied the first *size-hierarchy*

[1] This holds because, in parametrized complexity, k-CLIQUE is $\mathbf{W}[1]$-complete. Proving, for every constant k, a lower bound of the form $\Omega(n^{c_k})$, where $c_k \to \infty$ when $k \to \infty$, would imply that k-CLIQUE $\notin \mathbf{FPT}$, thus proving $\mathbf{FPT} \neq \mathbf{W}[1]$, which implies $\mathbf{P} \neq \mathbf{NP}$.

theorem for bounded-depth circuits (see Section 3.2 for more details). Moreover, this result was extended by Li, Razborov and Rossman [14] to SUB(G), the function computing whether a given graph contains a fixed graph G as a subgraph, tying the average-case complexity of SUB(G) to a new graph-theoretical measure closely related to tree-width [14].

In the same framework, Rossman proved a lower bound of $\Omega(n^{k/4})$ on the size of monotone circuits approximating k-CLIQUE under a distribution which is half the time $G(n, p_c^{K_k})$ with a planted k-CLIQUE and the other half is $G(n, p_c^{K_k})$ conditioned on being k-clique free. Unfortunately, this distribution is not a product distribution, like a "pure" Erdős–Rényi random graph. The importance of product distributions for monotone computation is explained in Section 3.4.

The first (and, so far, only) average-case lower bound for a monotone problem under a product distribution came in a recent work of Rossman [22], applying the pathset complexity technique explained in Section 3.3. The correlation bound of [22] shows that k-CYCLE is hard for $\mathbf{mNC}^1$ under $G(n, p_c^{C_k})$. A natural complexity question along these lines would be the following.

Problem 2. Decide if k-CLIQUE is hard for $\mathbf{mNC}^1$ under $G(n, p_c^{K_k})$.

It is conjectured in [21] that k-CLIQUE is hard on average under $G(n, p_c^{K_k})$ for monotone circuits in general. We see Problem 2 as a first step in this direction. Rossman's work on formula lower bounds (see Section 3.3), along with his work for subgraph isomorphism in general [14, 18, 21], provides good initial ideas for attacking this problem [20, 22].

3.2. **Stronger average-case size-hierarchy theorem for bounded-depth circuits.** One of the main aims of computational complexity is to understand the amount of computational resources needed to perform certain computational tasks. These resources can be, for example, time and space for algorithms in Turing machines, or size and depth in Boolean circuits. A natural question regarding any computational resource is whether increasing the access to this resource also increases the power of the computational model. This is known to hold with respect to time (Time Hierarchy Theorem [8]) and space (Space Hierarchy Theorem [24]) in Turing machines. The same question could be asked with respect to size and depth in Boolean circuits.

Analogous results have been proved in the worst-case for $\mathbf{AC}^0$ circuits. Sipser [23] was the first to prove (in 1983) a depth hierarchy theorem for small-depth circuits. He showed that, for every $d \in \mathbb{N}$, there exists a Boolean function $F_d : \{0,1\}^n \to \{0,1\}$ such that F_d is computable by a linear-size depth-d circuit but any depth-$(d-1)$ circuit requires superpolynomial size to compute F_d.

A size-hierarchy theorem for bounded-depth circuits was harder to come by, appearing only in a result of 2008 due to Rossman [18]. The result goes as follows. Denote by $\mathbf{AC}^0(\text{size } O(n^k))$ the class of constant-depth Boolean circuits with size at most $O(n^k)$. Rossman showed that, even though the k-CLIQUE function is computed by a depth-2 circuit of size $O(n^k)$, there exists no constant-depth circuit of size $O(n^{k/4})$ that computes this function. This implies that the hierarchy of complexity classes $\mathbf{AC}^0(\text{size } O(n^k))$, parametrized by k, is infinite.

An *average-case* depth hierarchy theorem was recently obtained in a breakthrough result by Håstad, Rossman, Servedio and Tan [11]. They proved that, for every $d \in \mathbb{N}$, there exists a monotone Boolean function $F_d : \{0,1\}^n \to \{0,1\}$ such that F_d is computable by a depth-d monotone formula, but any depth-$(d-1)$ circuit of subexponential size has correlation at most $1/2 + n^{-\Omega(1/d)}$ with F_d. This can be seen as a strenghtening of the result of Sipser [23]. Obtaining an average-case version of this result was an open problem posed by Håstad in 1986 [10].

The size-hierarchy theorem due to Rossman [18] mentioned above is actually an average-case result. Rossman proved not only that constant-depth circuits of size $O(n^{k/4})$ cannot compute k-CLIQUE, but also that such circuits have correlation at most $1/2 + n^{-\Omega(k)}$ with k-CLIQUE under $G(n, p_c^{K_k})$ (see Section 3.1). This provides a sharper separation between the computational power of $\mathbf{AC}^0$(size $O(n^k)$) and $\mathbf{AC}^0$(size $O(n^{k/4})$) than a worst-case lower bound.

The average-case result of Rossman shows that the correlation of $\mathbf{AC}^0$(size $O(n^{k/4})$) with k-CLIQUE under $G(n, p_c^{K_k})$ is only a polynomial fraction better than a constant function. Given that k-CLIQUE is monotone, in a sense this is best possible up to polynomial factors, since a well-known result [13] implies that any monotone function has $1/2 + \Omega(n^{-1})$ correlation with a trivial depth-1 circuit (that is, a circuit that outputs a constant or an input x_i) under the uniform distribution. However, it is reasonable to suppose that we can prove an even sharper average-case size-hierarchy theorem by considering a non-monotone function and proving a correlation bound that decreases exponentially.

Problem 3. Prove an exponentially decreasing average-case lower bound separating the hierarchy $\mathbf{AC}^0$(size $O(n^k)$) parametrized by k.

A result of Håstad [12], which shows that the correlation of PARITY with depth-d circuits is $1/2+2^{-\Omega(n)}$, might be a place to start. We are confident that a combination of this technique with the techniques of Rossman for proving lower bounds for subgraph isomorphism (see Section 3.1) can be successful in tackling Problem 3.

3.3. **Formula lower bounds via pathset complexity.** Understanding the relative power of Boolean formulas vs. circuits is a central challenge in complexity theory. Circuits are a powerful model of computation, capable of efficiently simulating Turing machines. On the other hand, formulas are thought to be a much weaker model of computation. Many natural problems solvable by small circuits, such as st-connectivity, are believed to require large formulas. However, no super-polynomial gap between the formula complexity and circuit complexity of any Boolean function has ever been shown. The existence of such a gap is a major open question.

Question 4. Are polynomial-size Boolean circuits strictly more powerful than polynomial-size Boolean formulas? Equivalently, is $\mathbf{NC}^1$ (polynomial-size formulas) strictly contained in $\mathbf{P/poly}$ (polynomial-size circuits)?

Let Distance-$k(n)$-Connectivity denote the problem of deciding whether there exists a path of length at most $k(n)$ between two given vertices s and t in a given graph. Working towards Question 4, Rossman [20] recently proved a $n^{\Omega(\log k)}$ lower bound on the $\mathbf{AC}^0$ formula size of Distance-$k(n)$-Connectivity for all $k(n) \leqslant \log\log n$ and depth up to $\log n/(\log\log n)^{O(1)}$. Since circuits of depth $O(\log k)$ already solve this problem, the result of Rossman shows a sharp separation between the power of bounded-depth circuits vs. bounded-depth formulas.

This groundbreaking result was proved through an innovative framework developed in [20], called *pathset complexity*, which provides a combinatorial explanation of why Boolean formulas fail to detect long paths in Erdős–Rényi random graphs. Roughly, his technique consists in reducing formula size lower bounds to a purely combinatorial lower bound on the minimum cost of constructing a set of paths via the operations of union and relational join, subject to certain "density constraints."

The pathset complexity is a new technique, and we expect many new formula lower bounds can come out of it. For example, there are many other "formula-like" computational models which may be malleable to pathset complexity arguments. We are interested in pursuing new lower bounds in these other models through this technique.

Problem 5. Obtain new lower bounds in other "formula-like" computational models, such as $\mathbf{AC}^0[\oplus]$ formulas[2] and multi-linear arithmetic formulas.[3]

For multi-linear arithmetic formulas in particular, it would be interesting to obtain a lower bound for the *iterated matrix multiplication* problem, which is an algebraic variety of the connectivity problem in graphs. Since the pathset framework has been succesful in the Boolean setting for st-connectivity, we can expect it to work also on the algebraic setting. We note that the iterated matrix multiplication has recently been studied in the multilinear setting by Chillara, Limaye and Srinivasan [6].

3.4. **Monotone circuits.** Given $k \in [n]$, a probability distribution μ over $\{0,1\}^n$ is called *k-slice* if, for every $x \in \mathrm{supp}(\mu)$, we have $|x|_1 = k$, where $|x|_1 := \sum_{i\in[n]} x_i$. A famous result due to Berkowitz [4] states that the monotone complexity of any monotone Boolean function is polynomially equivalent to its non-monotone complexity under a k-slice distribution. Moreover, by the Chernoff inequality product distributions are concentrated around a few slices. For this reason, product distributions are believed to be a source of hard instances for monotone functions. This motivates the study of average-case lower bounds in the monotone setting under product distributions.

The theorem of Berkowitz mentioned above gives possible evidence that the gap under product distributions between monotone circuits and non-monotone circuits is at most polynomial. However, nothing is known about this gap. If this gap were at most polynomial, then the lower bound for k-CYCLE [22] mentioned above would be extended to the non-monotone setting, thus separating $\mathbf{NC}^1$ from $\mathbf{P/poly}$ and thereby solving Question 4. Therefore, investigating this gap is an important quest for circuit complexity. We also remark that an exponential gap between monotone complexity and non-monotone complexity is well-known in the worst-case [25].

Given $k \in [n-1]$, a probability distribution μ over $\{0,1\}^n$ is called *$(k, k+1)$-slice* if, for every $x \in \mathrm{supp}(\mu)$, we have $|x|_1 \in \{k, k+1\}$. An important first step to understand the aforementioned gap between non-monotone and monotone circuits under product distributions would be the following question.

Problem 6. Decide if the gap under a $(k, k+1)$-slice distribution between monotone and non-monotone circuits is at most polynomial.

Surprisingly, nothing is known so far about this question.

4. Technical overview

In this section we give an overview of some of the results discussed in the previous sections, explaining in more detail some of the technical aspects of them.

[2] Bounded-depth formulas with parity gates.
[3] Algebraic formulas computing only multilinear polynomials.

4.1. **Definitions and preliminaries.** For a circuit C, we denote by $\mathrm{Gates}(C)$ the set of gates of C, and for each $\nu \in \mathrm{Gates}(C)$ we let $\mathrm{Children}(\nu)$ denote the set of the child gates of ν in C. The *height* of a gate ν is its maximum distance to an input gate. Furthermore, we denote by $\mathrm{size}(C)$ and $\mathrm{depth}(C)$ the size and the depth of the circuit, respectively. Finally, we write $\mathrm{wires}(C)$ to denote the number of wires in C (that is, we have $\mathrm{wires}(C) = \sum_{\nu \in \mathrm{Gates}(C)} |\mathrm{Children}(\nu)|$.)

For a Boolean function $f : \{0,1\}^n \to \{0,1\}$, we denote by $\mathrm{DT}_{\mathrm{depth}}(f)$ the minimum depth of a decision-tree computing f.

Let $\mathcal{G}^n$ denote the set of all graphs on n vertices. We usually identify $\mathcal{G}^n$ with $\{0,1\}^{\binom{n}{2}}$, and treat a graph function $f : \mathcal{G}^n \to \{0,1\}$ as a Boolean function with $\binom{n}{2}$ inputs. For a graph H, we define

$$m(H) := \max_{\substack{J \subseteq H, \\ |V(J)|>0}} \frac{|E(J)|}{|V(J)|}.$$

Observe that any $p_H(n) = \Theta(n^{-1/m(H)})$ is a threshold function for the containment of H in an Erdös-Rényi random graph.

For graphs H and G, we let $\mathrm{sub}(H,G)$ denote the number of copies of the graph H in G. By applying Janson's inequality, one is able to prove the following lemma.

Lemma 1. Let P be a fixed graph and let $G \sim G(n,p)$, where $p(n) = n^{\Omega(1)-1/m(P)}$. We have

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

4.2. **Background on bounded-depth complexity.** We begin by reviewing a few key results in bounded-depth complexity.

Definition 2. For $q,p \in [0,1]$ and $n \in \mathbb{N}$, we denote by $\mathcal{R}^n(q,p)$ the distribution of random restrictions $\rho : [n] \to \{0,1,*\}$ such that the $\rho(i)$ are independent and

$$\begin{aligned}\mathbb{P}[\rho(i) = *] &= q,\\ \mathbb{P}[\rho(i) = 1] &= (1-q)p,\\ \mathbb{P}[\rho(i) = 0] &= (1-q)(1-p).\end{aligned}$$

When n is clear from the context, we will omit it and write simply $\mathcal{R}(q,p)$.

Lemma 3 (Switching Lemma [9])**.** Suppose a Boolean function $f : \{0,1\}^n \to \{0,1\}$ is an **AND** or an **OR** of depth-r decision trees. Then for all $q \in [0,1]$ and $s \in \mathbb{N}$, we have

$$\mathbb{P}_{\rho \sim \mathcal{R}(q,1/2)}[\mathrm{DT}_{\mathrm{depth}}(f \restriction_\rho) > s] \leqslant (5pr)^s.$$

Remark 4. The lemma above is a stronger version of the original switching lemma due to Håstad, since every function computable by a depth-r decision tree can be written as both a r-CNF and a r-DNF.

An important consequence of the switching lemma, obtained by Boppana [5], goes as follows. Let $f : \{0,1\}^n \to \{0,1\}$ be a Boolean function and $x \in \{0,1\}^n$, and define

$$S(f,x) := \{i \in [n] : f(x) \neq f(x \oplus e_i)\}.$$

Define also the *average-sensitivity* $\mathrm{as}(f)$ of f as

$$\mathrm{as}(f) = 2^{-n} \sum_{x \in \{0,1\}^n} |S(f,x)|.$$

Theorem 5 (Bounded-depth circuits have low average-sensitivity [5])**.** If $f : \{0,1\}^n \to \{0,1\}$ is computable by a depth-d circuit of size S, then

$$\mathrm{as}(f) = O(\log S)^{d-1}.$$

Equivalently, we have

$$\mathbb{P}_{\substack{x \in_U \{0,1\}^n \\ i \in_U [n]}}[i \in S(f,x)] \leqslant \frac{O(\log S)^{d-1}}{n 2^n}.$$

Remark 6. It is not hard to see that $\mathrm{as}(\mathrm{PARITY}_n) = n$. Theorem 5 then implies the optimal $2^{\Omega(n^{1/(d-1)})}$ lower bound on the size of depth-d circuits computing PARITY_n, thus proving that $\mathrm{PARITY} \notin \mathbf{AC}^0$.

4.3. **A lower bound for k-CLIQUE against $\mathbf{AC}^0$.** In this section we explain in more detail the main technical contributions that led to the $\omega(n^{k/4})$ lower bound on the size of $\mathbf{AC}^0$ circuits approximating k-CLIQUE for fixed k, a result due to Rossman [18] mentioned in Section 3.1 and Section 3.2. However, for the sake of simplicity, here we prove a weaker version version of the result whose proof is clearer from the technical point of view, though the full result can be obtained with basically the same proof, only with a bit more of care. Our presentation follows that of Rossman [19].

Definition 7. We say that a sequence of circuits $(C_n)_{n\in\mathbb{N}}$ *computes* a sequence of Boolean functions $(f_n)_{n\in\mathbb{N}}$ on graphs of n vertices *with high probability under* $G(n,p)$ if

$$\lim_{n\to\infty} \mathbb{P}_{G\sim G(n,p)}[C_n(G) = f_n(G)] = 1.$$

Observe that, for instance, computing k-CLIQUE with high probability is a stronger condition than having correlation $(1/2 + \varepsilon)$ with the same function. Therefore, a lower bound against the former class of circuits is weaker result than a lower bound against the latter. In this section we will prove the following theorem.

Theorem 8. Boolean circuits of size $O(n^{k/4})$ and depth at most $k^{-2}\log n/\log\log n$ cannot compute k-CLIQUE with high probability on $G(n, n^{-2/(k-1)})$[4].

Before we proceed with the proof, a few observations and definitions are in order. Let us call *pattern* a constant-size graph $P \subseteq K_k$ with no isolated vertices. Let $\mathrm{Poisson}(\lambda)$ denote the Poisson distribution with mean λ and, given a pattern P, let $\mathrm{Plant}(n,P)$ denote the random graph with n vertices and edge set $\{\alpha(v)\alpha(w) : vw \in E(P)\}$, where α is uniformly chosen among all injective functions from $V(P)$ to $[n]$. Let also $d_{\mathrm{TV}}(\cdot,\cdot)$ denote the *total variation distance* between two distributions or random variables. Finally, let $\kappa(G)$ denote the number of k-cliques of a given graph G. The following is a technical lemma that can be proved by means of simple but long calculations.

Lemma 9. Let $G \sim G(n, cn^{-2/(k-1)})$ and $\mathbb{K}_k \sim \mathrm{Plant}(n, K_k)$. Let also $G^+ \sim G(n, n^{-2/(k-1)})$ conditioned on $\kappa(G) > 0$. We have

$$d_{\mathrm{TV}}(G \cup \mathbb{K}_k, G^+) = o(1).$$

[4]The probability function could be any $p(n) = \Theta(n^{-2/(k-1)})$. Since we are considering circuits computing k-CLIQUE with high probability, any probability function $p(n)$ such that $\mathbb{P}[k\text{-CLIQUE}(G(n,p(n))) = 1]$ is bounded away from both 0 and 1 would work. However, if we were aiming at a correlation bound result, we would have to take $p(n) = p_c^{K_k}$, as defined in the survey above.

We may now proceed with the proof of Theorem 8. The result will folow from the following theorem, which we will prove in the course of this section.

Theorem 10. Suppose $f : \mathcal{G}^n \to \{0,1\}$ is computed by a circuit 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 $\mathbb{K}_k \sim \mathrm{Plant}(n, K_k)$. Then $f(G) = f(G \cup \mathbb{K}_k)$ w.h.p.

Proof of Theorem 8 from Theorem 10. Suppose that $f : \{0,1\}^n \to \{0,1\}$ is computed by a Boolean circuit of size $O(n^{k/4})$ and depth at most $k^{-2}\log n/\log\log n$. Suppose moreover that f agrees with k-CLIQUE w.h.p on $G(n, n^{-2/(k-1)})$. We will derive a contradiction.

First, observe that, since f agrees with k-CLIQUE w.h.p and $n^{-2/(k-1)}$ is a threshold function for k-CLIQUE, we have

$$\mathop{\mathbb{P}}_{G\sim G(n,n^{-2/(k-1)})}[f(G)=1] = \mathop{\mathbb{P}}_{G\sim G(n,n^{-2/(k-1)})}[k\text{-CLIQUE}(G)=1] + o(1) = c + o(1), \tag{1}$$

for some constant $c \in (0,1)$.

Secondly, one notes that Lemma 9 and (1) imply

$$\begin{aligned}
&\mathop{\mathbb{P}}_{\substack{G\sim G(n,n^{-2/(k-1)})\\ \mathbb{K}_k\sim\mathrm{Plant}(n,K_k)}}[f(G\cup\mathbb{K}_k)=1] \\
&\qquad = o(1) + \mathop{\mathbb{P}}_{G\sim G(n,n^{-2/(k-1)})}[f(G)=1 \,|\, G \text{ contains a } k\text{-clique}] \\
&\qquad = o(1) + \mathop{\mathbb{P}}_{G\sim G(n,n^{-2/(k-1)})}[k\text{-CLIQUE}(G)=1 \,|\, G \text{ contains a } k\text{-clique}] \\
&\qquad = 1 + o(1).
\end{aligned}$$

Therefore, by Theorem 10 we get

$$\begin{aligned}
c + o(1) &= \mathop{\mathbb{P}}_{G\sim G(n,n^{-2/(k-1)})}[f(G)=1] \\
&= o(1) + \mathop{\mathbb{P}}_{\substack{G\sim G(n,n^{-2/(k-1)})\\ \mathbb{K}_k\sim\mathrm{Plant}(n,K_k)}}[f(G\cup\mathbb{K}_k)=1] \\
&= 1 + o(1),
\end{aligned}$$

a desired contradiction. □

We now proceed to prove Theorem 10. **We will first present the main technical lemmas without proving, so as not to stop the flow of text. The interested reader is pointed Section 4.3.4, where the remaining proofs are presented.**

4.3.1. *A consequence of the switching lemma.* The first ingredient of the proof is a consequence of the switching lemma, which allows us to consider "biased" random restrictions.

Definition 11. For a Boolean function $f : \{0,1\}^n \to \{0,1\}$, we say that f depends on $i \in [n]$ if there exists $x \in \{0,1\}^n$ such that $f(x) \neq f(x \oplus e_i)$. We define $\mathrm{Live}(f)$ as the set of coordinates $i \in [n]$ such that f depends on i. We say that a coordinate i is *live* if $i \in \mathrm{Live}(f)$.

Remark 12. Observe that a Boolean function f with decision-tree depth d depends on at most 2^d variables, i.e.: satisfies $|\mathrm{Live}(f)| \leqslant 2^d$.

Lemma 13 (Consequence of the Switching Lemma)**.** Let $p \in [0,1/2]$, $q \in [0,1]$, $c \geqslant 5$ and $t \geqslant 1$. If f is a Boolean function computed by a circuit C of size S and depth d such that $d \leqslant \log(1/q)/(tc + \log\log S)$, then

$$\mathbb{P}_{\rho\sim\mathcal{R}(pq,p)}\left[\mathrm{DT}_{\mathrm{depth}}(f\restriction_\rho) > \frac{\log S}{c}\right] \leqslant S^{1-t}.$$

Consequently, by Remark 12, we have

$$\mathbb{P}_{\rho\sim\mathcal{R}(pq,p)}\left[|\mathrm{Live}(f\restriction_\rho)| > S^{1/c}\right] \leqslant S^{1-t}.$$

4.3.2. *H-shaped average sensitivity.* We now present the main technical lemma of Rossman [18,19], which generalizes the low-average sensitivity result of Boppana (see Theorem 5). The proof goes by applying the consequence of the switching lemma described in the previous section (Lemma 13) and a consequence of Janson's inequality (Lemma 1). In what follows, for a graph $G \in \mathcal{G}^n$ and a Boolean function $f : \mathcal{G}^n \to \{0,1\}$, we write f^G to denote the function such that $f^G(H) = f(G \cup H)$ for all $H \in \mathcal{G}^n$.

Lemma 14. For every function $f : \mathcal{G}^n \to \{0,1\}$ and graph $H \in \mathcal{G}^n$, there exists a unique minimal graph T such that $f(H') = f(H' \cap T)$ for every $H' \subseteq H$.

Proof. Let $\mathcal{T}$ be the set of all graphs T such that $f(H') = f(H' \cap T)$ for every $H' \subseteq H$. It suffices to observe that $\mathcal{T}$ is not empty and that $\mathcal{T}$ is closed under intersection. □

Definition 15. For a function $f : \mathcal{G}^n \to \{0,1\}$ and a graph $H \in \mathcal{G}^n$, we denote by $\mathrm{Sens}(f,H)$ the unique minimal subgraph of H guaranteed to exist by Lemma 14. We call $\mathrm{Sens}(f,H)$ the *f-sensitive subgraph of H*. Moreover, when $\mathrm{Sens}(f,H) = H$, we say that f is *sensitive* over H.

Remark 16. Observe that the edges of $\mathrm{Sens}(f,H)$ are precisely

$$\left\{e \in E(H) : \text{ exists } H', H'' \subseteq H \text{ such that } E(H') = E(H'') \setminus \{e\} \text{ and } f(H') \neq f(H'')\right\}.$$

Definition 17. Let $H, G \in \mathcal{G}^n$. We define the restriction $\rho[G,H] : \binom{[n]}{2} \to \{0,1,*\}$ as follows:

$$\rho[G,H](e) = \begin{cases} * & \text{if } e \in E(H), \\ 1 & \text{if } e \in E(G) \setminus E(H), \\ 0 & \text{otherwise.} \end{cases}$$

Remark 18. Let $P \subseteq K_k$ be a pattern and let $G \sim G(n,p)$ and $H \sim \mathrm{Plant}(n,P)$. Remark 16 implies that the event $\{f^G \text{ is sensitive over } H\}$ is equivalent to the event $\{f\restriction_{\rho[G,H]} \text{ depends on all variables}\}$ (i.e.: the event that $\left|\mathrm{Live}(f\restriction_{\rho[G,H]})\right| = |E(H)|$).

Proposition 19 (Main technical lemma)**.** Let $P \subseteq K_k$ be a pattern and $p, q : \mathbb{N} \to [0,1/2]$ functions such that $p(n)q(n) = n^{\Omega(1)-1/m(P)}$. Suppose that $f : \mathcal{G}^n \to \{0,1\}$ is computed by circuits of size $n^{O(1)}$ and depth at most $\log(1/q)/(\omega(1) + \log\log n)$. Then

$$\mathop{\mathbb{P}}_{\substack{G\sim G(n,p)\\ H\sim\mathrm{Plant}(n,P)}}[f^G \text{ is sensitive over } H] \leqslant \frac{n^{o(1)}}{\mathbb{E}[\mathrm{sub}(P, G(n,pq))]}.$$

Remark 20. Observe that, by Remark 18, Proposition 19 is similar to the average sensitivity result of Boppana (Theorem 5). Indeed, when P is a single edge, Proposition 19 is almost the same result, with the only difference that the input is not chosen uniformly at random but from $G(n,p)$.

We now apply Proposition 19 to specific types of graphs, which we now define.

Definition 21. We say that a pattern P is *small* if $|V(P)| \leqslant k/2$; *medium*, if $|V(P)| > k/2$ and P is a union of two small patterns; *large* otherwise. Accordingly, we say that a graph is *small* if it contains at most $k/2$ non-isolated vertices; *medium*, if it contains more than $k/2$ non-isolated vertices and is a union of two small graphs; *large*, otherwise.

Applying Proposition 19, we obtain a more useful bound when P is a medium or a nonempty small pattern and $p(n)$ is at threshold for the containment of a k-clique.

Lemma 22 (Consequence of Proposition 19)**.** Let $P \subseteq K_k$ be a pattern and $p(n) = \Theta(n^{-2/(k-1)})$. Suppose that $f : \mathcal{G}^n \to \{0,1\}$ is computed by circuits of size $n^{O(1)}$ and depth at most $k^{-2}\log n/\log\log n + O(1)$. Then

$$\mathop{\mathbb{P}}_{\substack{G\sim G(n,p)\\ H\sim \mathrm{Plant}(n,P)}} [f^G \text{ is sensitive over } H] = \begin{cases} O(n^{-1}) & \text{if } P \text{ is nonempty and small,} \\ O(n^{-k/4-1/k}) & \text{if } P \text{ is medium.} \end{cases}$$

4.3.3. *Final inductive argument.* We are now ready to (almost) prove Theorem 10. The lower bound we will give here is only against circuits C with $\mathrm{wires}(C) = O(n^{k/4})$, not $\mathrm{size}(C) = O(n^{k/4})$, as promised. However, since $\mathrm{wires}(C) \leqslant \mathrm{size}(C)^2$, this yields a $\omega(n^{k/8})$ lower bound for the size. The full proof of Theorem 10 follows almost the same line of argument we give, but it is slightly more involved. We observe that the change from a lower bound against wires to a lower bound against size is equivalent to the change from a lower bound against a fan-in 2 circuit with bounded alternation-depth to an unbounded fan-in circuit with bounded depth. Moreover, note that $\mathrm{wires}(C) = O(n^{k/4})$ is a stronger assumption since $\mathrm{size}(C) \leqslant \mathrm{wires}(C)$. Finally, since this section wraps up everything together, we give full proofs here.

Proposition 23 (Theorem 10 for wires)**.** Suppose $f : \mathcal{G}^n \to \{0,1\}$ is computed by a circuit with $O(n^{k/4})$ wires and depth at most $k^{-2}\log n/\log\log n$. Let $G \sim G(n, n^{-2/(k-1)})$ and $\mathbb{K}_k \sim \mathrm{Plant}(n, K_k)$. Then $f(G) = f(G \cup \mathbb{K}_k)$ w.h.p.

The proof of Proposition 23 uses an inductive argument expressed in the following lemma.

Lemma 24. Let C be a circuit of fan-in 2 computing a Boolean function f and H be a graph such that, for every $\nu \in \mathrm{Gates}(C)$, the graph $\mathrm{Sens}(\nu, H)$ is not medium. Then $\mathrm{Sens}(f, H)$ is small.

Proof. We argue by inducion on the gates that $\mathrm{Sens}(\nu, H)$ is small for every gate $\nu \in \mathrm{Gates}(C)$. Indeed, for the base case it suffices to see that $\mathrm{Sens}(\nu, H)$ is clearly small for every input gate ν. Moreover, for the induction step, supposing that ν has children gates μ_1 and μ_2, we have that $\mathrm{Sens}(\nu, H) \subseteq \mathrm{Sens}(\nu, \mu_1) \cup \mathrm{Sens}(\nu, \mu_2)$. Since both $\mathrm{Sens}(\mu_1, H)$ and $\mathrm{Sens}(\mu_2, H)$ are small by the induction hypothesis, we have that $\mathrm{Sens}(\nu, H)$ is either small or medium, by definition. Therefore, it follows by the Lemma assumption that $\mathrm{Sens}(\nu, H)$ is small. □

Proof of Proposition 23. Consider a circuit computing f with $O(n^{k/4})$ wires and depth at most $k^{-2}\log n/\log\log n$. Let C be the modification of this circuit that replaces each gate by a fan-in 2 gate. Circuit C still satisfies $\mathrm{size}(C) \leqslant \mathrm{wires}(C) = O(n^{k/4})$, but the depth is unbounded. However, each gate $\nu \in \mathrm{Gates}(C)$ can be computed by a circuit of size $O(n^{k/4})$ wires and depth

at most $k^{-2}\log n/\log\log n$, by collapsing gates of the same kind. For this reason, we are able to apply Lemma 22, as we will do below.

Let $\nu \in \mathrm{Gates}(C)$. We have

$$\begin{aligned}
&\mathbb{P}[\mathrm{Sens}(\nu^G, \mathbb{K}_k) \text{ is medium}] \\
&\qquad \leqslant \sum_{\text{medium patterns } P} \mathbb{P}\left[\mathrm{Sens}(\nu^G, \mathbb{K}_k) \text{ is } P\text{-subgraph of } \mathbb{K}_k\right] \\
&\qquad \leqslant \sum_{\text{medium patterns } P} \mathbb{E}\left[\left|\left\{H : H \text{ is a } \nu^G\text{-sensitive } P\text{-subgraph of } G\right\}\right|\right] \\
&\qquad = \sum_{\text{medium patterns } P} \mathrm{sub}(P, K_k) \mathop{\mathbb{P}}_{\substack{G\sim G(n,p)\\ \mathbb{K}_k\sim \mathrm{Plant}(n,K_k)\\ H\sim \mathrm{Plant}(n,P)}} \left[\nu^G \text{ is sensitive over } H \,\middle|\, H \subseteq \mathbb{K}_k\right] \\
&\qquad = \sum_{\text{medium patterns } P} \mathrm{sub}(P, K_k) \mathop{\mathbb{P}}_{\substack{G\sim G(n,p)\\ \mathbb{K}_k\sim \mathrm{Plant}(n,K_k)\\ H\sim \mathrm{Plant}(n,P)}} \left[\nu^G \text{ is sensitive over } H\right],
\end{aligned}$$

since $\left\{\nu^G \text{ is sensitive over } H\right\}$ is independent of $\{H \subseteq \mathbb{K}_k\}$. Moreover, note that

$$\sum_{\text{medium patterns } P} \mathrm{sub}(P, K_k) \leqslant 2^{k^2} = O(1).$$

By Lemma 22, we obtain

$$\mathop{\mathbb{P}}_{\substack{G\sim G(n,p)\\ \mathbb{K}_k\sim \mathrm{Plant}(n,K_k)\\ H\sim \mathrm{Plant}(n,P)}} \left[\nu^G \text{ is sensitive over } H\right] = O(n^{-k/4-1/k})$$

Therefore, since $\mathrm{size}(C) \leqslant n^{k/4}$, we obtain by an union bound that, with high probability, the graph $\mathrm{Sens}(\nu^G, \mathbb{K}_k)$ is not medium for all gates $\nu \in \mathrm{Gates}(C)$. Applying Lemma 24, we obtain that $\mathrm{Sens}(f^G, \mathbb{K}_k)$ is small w.h.p. By a similar calculation, one is able to prove that

$$\mathbb{P}[\mathrm{Sens}(\nu^G, \mathbb{K}_k) \text{ is non-empty and small}] = o(1).$$

Hence, it follows that, with high probability, $\mathrm{Sens}(f^G, \mathbb{K}_k)$ is empty. We conclude that, with high probability, we have

$$f(G) = f^G(\emptyset) = f^G(\mathrm{Sens}(f^G, \mathbb{K}_k)) = f^G(\mathbb{K}_k) = f(G \cup \mathbb{K}_k). \qquad \square$$

4.3.4. *Proofs of the technical lemmas.* In this section, we give proofs for all the important technical innovations used in the proof of Theorem 8 which we did not prove above.

We begin by proving Lemma 13, a consequence of the switching lemma (Lemma 3).

Proof of 13. We generate $\rho \sim \mathcal{R}(pq, p)$ as a sequence of random restrictions in the following way: first, let $\rho_0 \sim \mathcal{R}(p, \lambda)$ be a random restriction of the variables of C, where

$$\lambda := \frac{p(1-pq) - p(1-q)/2}{1-p}.$$

One may check that indeed $\lambda \in [0,1]$. For $i \in [d]$, we let $\rho_i \sim \mathcal{R}(q^{1/d}, 1/2)$ be a restriction applied to the variables which were left unrestricted by the previous restrictions $\rho_0, \ldots, \rho_{i-1}$. For $i \in \{0, \ldots, d\}$, we define ρ^i as the composition of the restrictions $\rho_0, \rho_1, \ldots, \rho_i$, and we are able to check that $\rho^d \sim \mathcal{R}(pq, p)$.

For a gate $\nu \in \mathrm{Gates}(C)$ with height h, let E_ν be the event that $\mathrm{DT}_{\mathrm{depth}}(\nu \restriction_{\rho^h}) \leqslant (\log S)/c$. If $h = 0$, then $\mathbb{P}(E_\nu) = 1$. Otherwise, supposing $h \geqslant 1$, we have by the Switching Lemma (Lemma 3) that

$$\begin{aligned}
&\mathbb{P}\left[\neg X_\nu \,\middle|\, \bigwedge_{\mu \in \mathrm{Children}(\nu)} X_\mu\right] \\
&\quad = \mathbb{P}\left[\mathrm{DT}_{\mathrm{depth}}\left((\nu \restriction_{\rho^{h-1}}) \restriction_{\rho^h}\right) > \frac{\log S}{c} \,\middle|\, \bigcap_{\mu \in \mathrm{Children}(\nu)} \mathrm{DT}_{\mathrm{depth}}(\nu \restriction_{\rho^{h-1}}) \leqslant \frac{\log S}{c}\right] \\
&\quad \leqslant \left(\frac{5q^{1/d}\log S}{c}\right)^{(\log S)/c} \\
&\quad \leqslant S^{-t}.
\end{aligned}$$

We may therefore conclude

$$\begin{aligned}
\mathbb{P}\left[\mathrm{DT}_{\mathrm{depth}}(f \restriction_\rho) > \frac{\log S}{c}\right] &= \mathbb{P}\left[\neg X_{\text{output gate}}\right] \\
&\leqslant \mathbb{P}\left[\bigcup_{\nu \in \mathrm{Gates}(C)} \neg X_\nu\right] \\
&\leqslant \sum_{\nu \in \mathrm{Gates}(C)} \mathbb{P}\left[\neg X_\nu \cap \bigcap_{\mu \in \mathrm{Children}(\nu)} X_\mu\right] \\
&\leqslant \sum_{\nu \in \mathrm{Gates}(C)} \mathbb{P}\left[\neg X_\nu \,\middle|\, \bigcap_{\mu \in \mathrm{Children}(\nu)} X_\mu\right] \\
&\leqslant S^{1-t}.
\end{aligned}$$

□

We may now prove the main technical lemma of Rossman [18, 19], described in Proposition 19. Our proof applies Lemma 13 which we proved above and a consequence of Janson's inequality (Lemma 1).

Proof of Proposition 19. Fix $\varepsilon > 0$. It suffices to prove that

$$\mathop{\mathbb{P}}_{\substack{G \sim G(n,p) \\ H \sim \mathrm{Plant}(n,P)}} [f^G \text{ is sensitive over } H] \leqslant \frac{n^\varepsilon}{\mathbb{E}[\mathrm{sub}(P, G(n, pq))]}.$$

Towards the proof, we consider the random graph $Q \sim G(n, pq)$, independent of both G and H. Observe that

$$\begin{aligned}
\mathrm{Live}(f \restriction_{\rho[G,Q]}) &= E(\mathrm{Sens}(f^G, Q)) \\
&= \left\{e \in E(Q) : \exists Q', Q'' \subseteq Q \text{ s.t. } E(Q') = E(Q'') \setminus \{e\} \text{ and } f(Q') \neq f(Q'')\right\}.
\end{aligned}$$

We now consider two events:

1. $E_1 = \left[\mathrm{sub}(P, Q) \geqslant (1/2)\mathbb{E}[\mathrm{sub}(P, Q)]\right]$,
2. $E_2 = \left[\left|\mathrm{Live}(f \restriction_{\rho[G,Q]})\right| \leqslant n^{\varepsilon/2|E(P)|}\right]$.

Let us also consider the random graph $\widetilde{H}$ defined as follows: if E_1 holds, then $\widetilde{H}$ is a P-subgraph of Q chosen uniformly at random; otherwise, $\widetilde{H}$ is $H \sim \mathrm{Plant}(n, P)$. Observe that $\widetilde{H}$ is indepedent of both G and H and that (G, H) and $(G, \widetilde{H})$ have the same joint distribution.

We now obtain

$$\begin{aligned}\mathbb{P}\left[f^G \text{ is sensitive over } H\right] &= \mathbb{P}\left[f^G \text{ is sensitive over } \widetilde{H}\right] \\ &\leqslant \mathbb{P}\left[f^G \text{ is sensitive over } \widetilde{H} \,\middle|\, E_1, E_2\right] + \mathbb{P}[\neg E_1] + \mathbb{P}[\neg E_2].\end{aligned} \tag{2}$$

Let us first prove

$$\mathbb{P}\left[f^G \text{ is sensitive over } \widetilde{H} \,\middle|\, E_1, E_2\right] \leqslant \frac{1}{2}\frac{n^{\varepsilon}}{\mathbb{E}[\mathrm{sub}(P, G(n, pq))]}.$$

Indeed, if E_1 holds, then $\widetilde{H}$ is uniformly distributed among the P-subgraphs of Q and, for this reason, if f^G is sensitive over $\widetilde{H}$, then $E(\widetilde{H}) \subseteq \mathrm{Live}(f \restriction_{\rho[G,Q]})$. Moreover, if E_2 holds, then there are at most $\binom{n^{\varepsilon/2|E(P)|}}{|E(P)|} \leqslant n^{\varepsilon/2}$ subgraphs of Q satisfying $E(\widetilde{H}) \subseteq \mathrm{Live}(f \restriction_{\rho[G,Q]})$. Therefore, we have

$$\begin{aligned}\mathbb{P}\left[f^G \text{ is sensitive over } \widetilde{H} \,\middle|\, E_1, E_2\right] &\leqslant \frac{\#\left\{P\text{-subgraphs } \widetilde{P} \text{ of } Q \text{ s.t. } E(\widetilde{P}) \subseteq \mathrm{Live}(f \restriction_{\rho[G,Q]})\right\}}{\mathrm{sub}(P, Q)} \\ &\leqslant \frac{n^{\varepsilon/2}}{(1/2)\mathbb{E}[\mathrm{sub}(P, Q)]} \\ &\leqslant \frac{1}{2}\frac{n^{\varepsilon}}{\mathbb{E}[\mathrm{sub}(P, G(n, pq))]}.\end{aligned}$$

Observe now that $\mathbb{E}[\mathrm{sub}(P, G(n, pq))] = n^{O(1)}$. Therefore, because of (2), it suffices to prove that $\mathbb{P}[\neg E_1] = n^{-\omega(1)}$ and $\mathbb{P}[\neg E_2] = n^{-\omega(1)}$.

Since $\mathbb{P}[\neg E_1] = n^{-\omega(1)}$ follows easily from Lemma 1, it only remains to prove $\mathbb{P}[\neg E_2] = n^{-\omega(1)}$, which we now do.

Let C be a circuit computing f with size and depth as in the hypothesis. Let $S := \mathrm{size}(C)$ and $d := \mathrm{depth}(C)$. By hypothesis we have $n^{O(1)}$, and we may assume without loss of generality that $S = n^{\Theta(1)}$. Let also

$$c := \frac{\log S}{(\varepsilon/2\,|E(P)|)\log n} + 5,$$

and it is easily checkable that $5 \leqslant c = O(1)$. Furthermore, one may also check that

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

Observe moreover that the random restriction $\rho[G, H]$ follows the distribution $\mathcal{R}(pq, p)$, since G and Q are independent. Finally, by the choice of c we have

$$\mathbb{P}\Big[\,\big|\mathrm{Live}(f \restriction_{\rho[G,Q]})\big| > n^{\varepsilon/2|E(P)|}\Big] \leqslant \mathbb{P}\Big[\,\big|\mathrm{Live}(f \restriction_{\rho[G,Q]})\big| > S^{1/c}\Big].$$

We are now ready to bound $\mathbb{P}[\neg E_2]$. By Lemma 13, we have

$$\begin{aligned}\mathbb{P}[\neg E_2] &= \mathbb{P}\Big[\big|\mathrm{Live}(f\restriction_{\rho[G,Q]})\big| > n^{\varepsilon/2|E(P)|}\Big]\\ &\leqslant \mathbb{P}\Big[\big|\mathrm{Live}(f\restriction_{\rho[G,Q]})\big| > S^{1/c}\Big]\\ &\leqslant S^{-\omega(1)}\\ &= n^{-\omega(1)},\end{aligned}$$

since $S = n^{\Theta(1)}$. This finishes the proof. □

We can now apply Proposition 19 and obtain Lemma 22. First, we will need a technical lemma about medium patterns which we do not prove here.

Lemma 25. For every medium pattern P, we have

$$|V(P)| - \frac{2}{k-1}|E(P)| \geqslant \frac{k+1}{4} + \frac{2}{k-1}.$$

Proof of Lemma 22. Let $q(n) = n^{-k^{-2}+k^{-3}}$. The following observations can be easily checked.

1. $p(n)q(n) = n^{\Omega(1)-1/m(P)}$;
2. $k^{-2}\log n/\log\log n + O(1) \leqslant \log(1/q)/(\omega(1)+\log\log n)$;

Observe now that

$$\begin{aligned}\mathbb{E}[\mathrm{sub}(P, G(n,pq))] &\geqslant n^{|V(P)|}(pq)^{|E(P)|}\\ &= n^{|V(P)|-(2/(k-1)+k^{-2}+k^{-3})|E(P)|+o(1)}\\ &\geqslant n^{|V(P)|-2/(k-1)|E(P)|-1/4-1/4k+o(1)} \qquad (\text{since } |E(P)| \leqslant k^2/4)\\ &> \begin{cases} n^{1+o(1)} & \text{if } P \text{ is nonempty and small,}\\ n^{k/4+1/k+o(1)} & \text{if } P \text{ is medium, by Lemma 25.}\end{cases}\end{aligned}$$

The lemma now follows by applying Proposition 19. □

4.4. **A quick overview of pathset complexity.** The main result of Rossman [20], as mentioned in Section 3.3, is the following:

Theorem 26. Formulas of depth $\log n/(\log\log n)^6$ solving Distance-$k(n)$-Connectivity require size $n^{\Omega(\log k)}$ for all $k(n) \leqslant \log\log n$.

The same technique employed in [20] was also employed in [22], as explained in Section 3. Here we give a brief overview of the main technical innovations that led to these results.

We denote by $P_k = (V_k, E_k)$ the directed path with k edges, where

$$V_k = \{v_0, v_1, \ldots, v_k\}, \quad E_k = \{v_i v_{i+1} : 0 \leqslant i \leqslant k-1\}.$$

We denote by $\mathcal{G}(k,n)$ the set of k-layered graphs Γ, satisfying

$$V(\Gamma) = \{v^i : v \in V_k, i \in [n]\}, \quad E(\Gamma) \subseteq \{v^i w^j : vw \in E_k, i, j \in [n]\}.$$

We will consider Boolean functions of the form $f : \mathcal{G}(k,n) \to \{0,1\}$, by identifying $\mathcal{G}(k,n)$ with $\{0,1\}^{kn^2}$. We identify the vertex s with v_0^1, and t with v_k^1. The variation of the Distance-$k(n)$-Connectivity problem we here consider consists of answering if there exists a path between s and t in a given k-layered graph Γ. We will also consider the graphs Γ as coming from the

random graph distribution where each potential edge appers with probability $1/n$, independent of each other edge. Observe that $1/n$ is below the threshold for the containment of a st-path.

A graph G is called a *pattern* if it is a subgraph of P_k with no isolated vertices (note the analogy with the patterns in the k-CLIQUE lower bound in Definition 21). For a pattern G, a *G-pathset* is a relation $\mathcal{A} \subseteq [n]^{V_G}$. We define the *density* $\mu(\mathcal{A})$ of a G-pathset $\mathcal{A}$ as $\mu(\mathcal{A}) := |\mathcal{A}| / n^{|V(G)|}$. We say that a G-pathset is G-small if it obeys a series of constraints on its density, which we omit here for the sake of brevity. Finally, we define (informally) the *pathset complexity* $\chi(\mathcal{A})$ of a G-pathset $\mathcal{A}$ as the minimum number of operations required to construct $\mathcal{A}$ via unions and relational joins, with the restriction that the relational joins can be made only between small pathsets. This restriction serves as bottleneck, which is explored in [20] to obtain a lower bound against pathset complexity. This lower bound is the combinatorial heart of the paper.

Theorem 27. For every P_k-pathset $\mathcal{A}$, we have

$$\chi(\mathcal{A}) \geqslant \frac{n^{\Omega(\log k)}}{2^{O(2^k)}} \mu(\mathcal{A}).$$

We then associate, for every $x \in [n]^{V(G)}$, every pattern graph G and random graph Γ, a random restriction $\rho^{\Gamma}_{G,x}$, whose definition is quite similar to Definition 17. Finally, we associate for every Boolean function $f : \mathcal{G}(k,n) \to \{0,1\}$ a random pathset $\mathcal{A}^{\Gamma}_{f,G}$ defined as follows

$$\mathcal{A}^{\Gamma}_{f,G} = \left\{ x \in [n]^{V(G)} : f \restriction \rho^{\Gamma}_{G,x} \text{ depends on all variables} \right\}.$$

By employing the Switching Lemma (Lemma 3) and Janson's inequality, Rossman proved, in a very similar fashion to the proof of Proposition 19 and Lemma 22, an average-sensitivity type lemma, as follows.

Lemma 28. Suppose $f : \mathcal{G}(k,n) \to \{0,1\}$ is computable by a circuit of size n^k and depth $\log n/(\log\log n)^6$. Then, for all patterns G, we have

$$\mathop{\mathbb{P}}_{\Gamma}[\mathcal{A}^{\Gamma}_{f,G} \text{ is not } G\text{-small}] \leqslant O(n^{-2k}).$$

Lemma 22 is used in a union bound, in a similar manner to what we do in the proof of Proposition 23.

By employing a simple but new top-down argument, one is also able to prove the following lemma, which is the essential bridge connecting pathset complexity with formula size. Observe that this is the only part where the fact that F is a formula is used; elsewhere, F could also be a circuit.

Lemma 29. Let F be any fan-in 2 formula and let $\Gamma \in \mathcal{G}(k,n)$. If $\mathcal{A}^{\Gamma}_{f,G}$ is G-small for all gates $f \in F$ and patterns G, then

$$\chi(\mathcal{A}^{\Gamma}_{f_{\text{out}},G}) \leqslant 2^{O(k^2)} \operatorname{depth}(F) \operatorname{size}(F).$$

Finally, by employing another simple but crucial argument, it is proved in [20] the following lower bound on the density of the random pathset.

Lemma 30. Suppose that a circuit F computes Distance-$k(n)$-Connectivity w.h.p. on random graphs Γ. Then, w.h.p, we have $\mu(\mathcal{A}^{\Gamma}_{f_{\text{out}},P_k}) \geqslant 0.99n^{-2}$.

Combining all these ingredients, one is now able to prove Theorem 26.

Proof of Theorem 26. By Lemma 28, with high probability $\mathcal{A}^{\Gamma}_{f,G}$ is G-small for every $f \in F$ and pattern G. Moreover, by Lemma 30, we have $\mu(\mathcal{A}^{\Gamma}_{f_{\text{out}},P_k}) \geqslant 0.99n^{-2}$ w.h.p. Therefore, we have

$$\begin{aligned}
\operatorname{size}(F) &\geqslant 2^{-O(k^2)} \operatorname{depth}(f)^{-k} \chi(\mathcal{A}^{\Gamma}_{f_{\text{out}},P_k}) && \text{(by Lemma 29)} \\
&\geqslant 2^{-O(k^2)} \operatorname{depth}(f)^{-k} \frac{n^{\Omega(\log k)}}{2^{O(2^k)}} \mu(\mathcal{A}) && \text{(by Theorem 27)} \\
&\geqslant n^{\Omega(\log k)} && \text{(by Lemma 30).}
\end{aligned}$$

□

References

[1] M. Ajtai, Σ^1_1-*formulae on finite structures*, Ann. Pure Appl. Logic **24** (1983), no. 1, 1–48. MR706289 ↑2

[2] N. Alon and R. B. Boppana, *The monotone circuit complexity of Boolean functions*, Combinatorica **7** (1987), no. 1, 1–22. MR905147 ↑2

[3] A. E. Andreev, *A method for obtaining lower bounds on the complexity of individual monotone functions*, Dokl. Akad. Nauk SSSR **282** (1985), no. 5, 1033–1037. MR796937 ↑2

[4] S. J. Berkowitz, *On some relationships between monotone and non-monotone circuit complexity*, 1982. ↑3.4

[5] R. B. Boppana, *The average sensitivity of bounded-depth circuits*, Inform. Process. Lett. **63** (1997), no. 5, 257–261. MR1475339 ↑4.2, 5

[6] S. Chillara, N. Limaye, and S. Srinivasan, *Small-depth multilinear formula lower bounds for iterated matrix multiplication, with applications*, 35th Symposium on Theoretical Aspects of Computer Science, 2018, pp. Art. No. 21, 15. MR3779302 ↑3.3

[7] M. Furst, J. B. Saxe, and M. Sipser, *Parity, circuits, and the polynomial-time hierarchy*, Math. Systems Theory **17** (1984), no. 1, 13–27. MR738749 ↑2

[8] J. Hartmanis and R. E. Stearns, *On the computational complexity of algorithms*, Trans. Amer. Math. Soc. **117** (1965), 285–306. MR0170805 ↑3.2

[9] J Håstad, *Almost optimal lower bounds for small depth circuits*, Proceedings of the eighteenth annual acm symposium on theory of computing, 1986, pp. 6–20. ↑2, 3

[10] J. Håstad, *Computational limitations of small-depth circuits*, MIT Press, Cambridge, MA, USA, 1987. ↑3.2

[11] J. Håstad, B. Rossman, R. A. Servedio, and L.-Y. Tan, *An average-case depth hierarchy theorem for Boolean circuits*, J. ACM **64** (2017), no. 5, Art. 35, 27. MR3716890 ↑1, 3.2

[12] J. Håstad, *On the correlation of parity and small-depth circuits*, SIAM J. Comput. **43** (2014), no. 5, 1699–1708. MR3262612 ↑2.1, 3.2

[13] M. Kearns, M. Li, and L. Valiant, *Learning Boolean formulas*, J. Assoc. Comput. Mach. **41** (1994), no. 6, 1298–1328. MR1371501 ↑3.2

[14] Y. Li, A. Razborov, and B. Rossman, *On the* AC^0 *complexity of subgraph isomorphism*, 55th Annual IEEE Symposium on Foundations of Computer Science—FOCS 2014, 2014, pp. 344–353. MR3344884 ↑1, 3.1, 3.1

[15] N. Linial, Y. Mansour, and N. Nisan, *Constant depth circuits, Fourier transform, and learnability*, J. Assoc. Comput. Mach. **40** (1993), no. 3, 607–620. MR1370363 ↑1

[16] N. Nisan and A. Wigderson, *Hardness vs. randomness*, J. Comput. System Sci. **49** (1994), no. 2, 149–167. MR1293639 ↑1

[17] A. A. Razborov, *Lower bounds on the monotone complexity of some Boolean functions*, Dokl. Akad. Nauk SSSR **281** (1985), no. 4, 798–801. MR785629 ↑2

[18] B. Rossman, *On the constant-depth complexity of k-clique*, STOC '08, 2008, pp. 721–730. MR2582693 ↑1, 3.1, 3.1, 3.2, 4.3, 4.3.2, 4.3.4

[19] ———, *Average-Case Complexity of Detecting Cliques*, ProQuest LLC, Ann Arbor, MI, 2010. Thesis (Ph.D.)–Massachusetts Institute of Technology (MIT). MR2873600 ↑4.3, 4.3.2, 4.3.4

[20] ———, *Formulas vs. circuits for small distance connectivity*, Proceedings of the forty-sixth annual acm symposium on theory of computing, 2014, pp. 203–212. ↑1, 3.1, 3.3, 4.4, 4.4, 4.4

[21] ———, *The monotone complexity of k-clique on random graphs*, SIAM J. Comput. **43** (2014), no. 1, 256–279. MR3166976 ↑1, 3.1

[22] ———, *Correlation bounds against monotone* NC^1, 30th Conference on Computational Complexity, 2015, pp. 392–411. MR3441814 ↑1, 2.1, 3.1, 3.1, 3.4, 4.4

[23] M. Sipser, *Borel sets and circuit complexity*, Proceedings of the fifteenth annual acm symposium on theory of computing, 1983, pp. 61–69. ↑3.2

[24] R. E. Stearns, J. Hartmanis, and P. M. Lewis, *Hierarchies of memory limited computations*, Proceedings of the 6th annual symposium on switching circuit theory and logical design (swct 1965), 1965, pp. 179–190. ↑3.2

[25] É. Tardos, *The gap between monotone and nonmonotone circuit complexity is exponential*, Combinatorica **8** (1988), no. 1, 141–142. MR952004 ↑3.4

Instituto de Matemática e Estatística, Universidade de São Paulo, Rua do Matão 1010, 05508–090 São Paulo, SP

Endereços eletrônicos: `brunopc@ime.usp.br`