Lower Bounds for Subgraph Isomorphism

Benjamin Rossman

May 3, 2018

Abstract

We consider the problem of determining whether an Erdős-Rényi random graph contains a subgraph isomorphic to a fixed pattern, such as a clique or cycle of constant size. The computational complexity of this problem is tied to fundamental open questions including P vs. NP and NC^1 vs. L. We give an overview of unconditional average-case lower bounds for this problem (and its colored variant) in a few important restricted classes of Boolean circuits.

1 Background and preliminaries

The *subgraph isomorphism problem* is the computational task of determining whether a "host" graph H contains a subgraph isomorphic to a "pattern" graph G. When both G and H are given as input, this is a classic NP-complete problem which generalizes both MAXIMUM CLIQUE and HAMILTONIAN CYCLE [20]. We refer to the *G-subgraph isomorphism problem* in the setting where the pattern G is fixed and H alone is given as input. As special cases, this includes the k-CLIQUE and k-CYCLE problems when G is a complete graph or cycle of order k.

For patterns G of order k, the G-subgraph isomorphism problem is solvable in time $O(n^k)$ by the obvious exhaustive search.[1] This upper bound can be improved to $O(n^{\alpha\lceil k/3 \rceil})$ using any $O(n^\alpha)$ time algorithm for fast matrix multiplication [28] (the current record has $\alpha < 2.38$ [23]). Additional upper bounds are tied to structural parameters of G, such as an $O(n^{w+1})$ time algorithm for patterns G of tree-width w [30]. (See [26] for a survey on upper bounds.)

[1] Throughout this article, asymptotic notation ($O(\cdot)$, $\Omega(\cdot)$, etc.), whenever bounding a function of n, hides constants that may depend on G.

The focus of this article are *lower bounds* which show that the G-subgraph isomorphism problem cannot be solved with insufficient computational resources. It is conjectured that k-CLIQUE requires time $n^{\Omega(k)}$ and that a colored version of G-subgraph isomorphism (described in §2) requires time $n^{\Omega(w/\log w)}$ for patterns G of tree-width w. Conditionally, these lower bounds are known to follow from the Exponential Time Hypothesis [9, 25]. Proving such lower bounds unconditionally would separate P from NP in a very strong way. Since that goal is a long way off, we shall restrict attention to complexity measures much weaker than sequential time; specifically, we focus on restricted classes of Boolean circuits (described in §1.2).

1.1 The average-case setting

The lower bounds for the G-subgraph isomorphism problem described in this article are obtained in the natural average-case setting where the input is an Erdős-Rényi graph $\boldsymbol{G}_{n,p}$ (or G-colored version thereof). This is the random n-vertex graph in which each potential edge is included independently with probability p. For many patterns of interest including cliques and cycles, $\boldsymbol{G}_{n,p}$ is conjectured to be a source of hard-on-average instances at an appropriate threshold p. These conjectures are natural targets for the combinatorial and probabilistic approach of circuit complexity. Strong enough lower bounds for the average-case G-subgraph isomorphism problem would resolve P vs. NP and other fundamental questions, as we explain next.

In the average-case version of the k-CLIQUE problem, we are given an Erdős-Rényi graph $\boldsymbol{G}_{n,p}$ at the critical threshold $p = \Theta(n^{-2/(k-1)})$ (where the existence of a k-clique occurs with probability bounded away from 0 and 1). Our task is to determine, asymptotically almost surely[2] correctly, whether or not the given graph contains a k-clique. One natural approach is to make several independent runs of the following *randomized greedy algorithm*: start with a uniform random vertex v_1, then select a vertex v_2 uniformly at random from among the neighbors of v_1, next select a vertex v_3 uniformly at random from among the common neighbors v_1 and v_2, and so on until reaching a maximal (though not necessarily maximum) clique in the given graph. It is easy to show that a single run of the greedy algorithm on $\boldsymbol{G}_{n,p}$, which only requires linear time with very high probability, almost surely produces a clique of size $\lfloor \frac{k}{2} \rfloor$ or $\lceil \frac{k}{2} \rceil$. To find a clique of size $\lfloor \frac{(1+\varepsilon)k}{2} \rfloor$

[2] Throughout this article, *asymptotically almost surely* (abbreviated as *a.a.s.*) means with probability $1 - o(1)$, that is, with probability that tends to 1 as $n \to \infty$.

where $\varepsilon < 1$, it suffices to repeat the greedy algorithm $n^{\varepsilon^2 k/4}$ times, while $n^{k/4+O(1/k)}$ iterations suffice to find a k-clique in $\boldsymbol{G}_{n,p}$ if any exists. The average-case k-CLIQUE problem is thus solvable in time $n^{k/4+O(1)}$.

It is unknown whether this iterated greedy algorithm is optimal. In other words, is $\Omega(n^{k/4})$ a lower bound on the complexity of the average-case k-CLIQUE problem? This question may be seen a scaled-down version of a famous open question of Karp [21] concerning the uniform random graph $\boldsymbol{G}_{n,1/2}$. It is well-known that $\boldsymbol{G}_{n,1/2}$ has expected maximum clique size $\approx 2\log n$, while the randomized greedy algorithm almost surely finds a clique of size $\approx \log n$. Karp asked whether any polynomial-time algorithm a.a.s. succeeds in finding a clique of size $(1+\varepsilon)\log n$ for any constant $\varepsilon > 0$. Karp's question, together with a variant where $\boldsymbol{G}_{n,1/2}$ is augmented by a very large planted clique, have stimulated a great deal of research in theoretical computer science. The hardness of detecting planted cliques is used as a cryptographic assumption [18], while lower bounds have been shown against specific algorithms such as the metropolis process [17], the sum-of-squares semidefinite programming hierarchy [4], and a class of statistical query algorithms [10].

The k-CYCLE problem is another instance where $\boldsymbol{G}_{n,p}$ at the critical threshold $p = \Theta(1/n)$ is thought to be a source of hard-on-average instances. Compared to the k-CLIQUE problem, the average-case k-CYCLE problem has relatively low complexity: it is solvable in just $n^{2+o(1)}$ time and moreover in logarithmic space. Nevertheless, $\boldsymbol{G}_{n,p}$ is believed to be hard-on-average with respect to formula size (a combinatorial complexity measure which we shall discuss shortly). The smallest known formulas solving k-CYCLE have size $n^{O(\log k)}$ and this upper bound is conjectured to be optimal even in the average-case. Proving such a lower bound unconditionally would separate complexity classes NC^1 and L.

1.2 Circuit complexity

Circuit complexity is the quest for unconditional lower bounds in combinatorial models of computation. Among such models, Boolean circuits (acyclic networks of $\wedge$, $\vee$ and $\neg$ gates) are the most basic and important. Every polynomial-time algorithm can be implemented by a sequence of polynomial-size Boolean circuits, one for each input length n. To separate P from NP, it therefore suffices to prove a *super-polynomial lower bound* on the minimum circuit size of any problem in NP, as represented by a sequence of Boolean functions $\{0,1\}^n \to \{0,1\}$.

Claude Shannon in 1949 showed that *almost all* Boolean functions require circuits of exponential size [39]. Yet after nearly 70 years of efforts, no one has yet proved a super-linear lower bound on the circuit size of any *explicit* Boolean function. In the meantime, the majority of research in circuit complexity has focused on restricted classes of Boolean circuits and other combinatorial models with the aim of developing sharper insights and techniques. Below, we describe three natural and important restricted settings: formulas (tree-like circuits), the AC^0 setting (bounded alternation), and the monotone setting (the absence of negations).

Definitions. A *circuit* is a finite directed acyclic graph in which every node of in-degree 0 ("input") is labeled by a literal (i.e., a variable x_i or its negation $\neg x_i$), there is a unique node of out-degree 0 (the "output"), and each non-input ("gate") has in-degree 2 and is labeled by $\wedge$ or $\vee$. Every n-variable circuit computes a Boolean function $\{0,1\}^n \to \{0,1\}$ in the obvious way.

The *size* of a circuit is the number of gates it contains. The complexity class *P/poly* consists of sequences of Boolean functions $\{0,1\}^n \to \{0,1\}$ computable by n-variable circuits of polynomial size (i.e., $O(n^c)$ for any constant c). (The more familiar class P is obtained by imposing a uniformity condition on the sequence of n-variable circuits.)

The *depth* of a circuit is the maximum number of gates on an input-to-output path. The class NC^1 consists of Boolean functions computable by circuits of depth $O(\log n)$. Note that $\mathrm{size}(C) \leq 2^{\mathrm{depth}(C)}$ for all circuits C, hence $NC^1 \subseteq P/poly$. This containment is believed but not known to be proper.

The *alternation-depth* of a circuit is the maximum number of alternations between $\wedge$ and $\vee$ gates on an input-to-output path. The complexity class AC^0 consists of Boolean functions computed by circuits of polynomial size and constant alternation-depth.[3] Breakthrough lower bounds of the 1980's showed that AC^0 is a proper subclass of NC^1 [1, 11]. Quantitatively, the strongest of these lower bounds shows that circuits with alternation-depth d require size $2^{\Omega(n^{1/(d-1)})}$ to compute the n-variable PARITY function [14].

Another important restricted class of circuits are *formulas*: circuits with the structure of a binary tree (i.e., in which every non-output node has out-degree 1). In the context of formulas, *size* and *depth* are closely related complexity measures, as every formula of size s is equivalent to a formula of

[3] AC^0 is usually defined in terms of constant depth circuits with AND and OR gates of unbounded in-degree. In this article, we adopt the equivalent definition in terms of alternation-depth, since the simplest version of our lower bounds naturally applies to binary $\wedge$ and $\vee$ gates.

depth $O(\log s)$ [40]. As a corollary, NC^1 is equivalent to the class of Boolean functions computed by polynomial-size formulas.

In contrast to circuits, formulas are memoryless in the sense that the result of each sub-computation is only used once. However, despite this obvious weakness, the strongest lower bound on the formula size of an explicit Boolean function is only $n^{3-o(1)}$ [15, 41]. The challenge of proving a *super-polynomial formula-size lower bound* (i.e., showing that any explicit Boolean function is not in NC^1) is one of the major frontiers in circuit complexity.

1.3 The monotone setting

Monotonicity is both a property of circuits and a property of Boolean functions. A circuit is *monotone* if it has no negations (i.e., inputs are labeled by positive literals only). A Boolean function $f : \{0,1\}^n \to \{0,1\}$ is *monotone* if $f(x) \leq f(y)$ whenever $x_i \leq y_i$ for all coordinates i. Note that the G-subgraph isomorphism problem is monotone when viewed as a sequence of functions $\{0,1\}^{\binom{n}{2}} \to \{0,1\}$.

It is natural to study the *monotone complexity* of monotone functions f (i.e., the minimum size of a monotone circuit or formula which computes f). This has been an extremely fruitful restricted setting in circuit complexity beginning with celebrated results in the 1980's. In a groundbreaking paper which introduced the sunflower-plucking approximation method, Razborov [33] showed that the k-CLIQUE problem requires monotone circuits of size $\Omega(n^k/(\log n)^{2k})$ for any constant k.[4] By an entirely different technique based on communication complexity, Karchmer and Wigderson [19] proved an $n^{\Omega(\log k)}$ lower bound on the size of monotone formulas solving DISTANCE-k ST-CONNECTIVITY, a problem which is equivalent to k-CYCLE up to a polynomial factor. These results and several others [12, 29, 31, 32] imply essentially all separations $AC^0 \subset TC^0 \subset NC^1 \subset L \subset NL \subset P \subset NP$ in the monotone world (i.e., for the monotone versions of these classes), whereas in the non-monotone world it is open whether TC^0 (the class of constant-depth threshold circuits) is equal to NP.

Unfortunately, it is unclear if any of the lower bound techniques developed in the monotone setting have the potential to extend to non-monotone classes. A "barrier" emerges from the observation that essentially all monotone lower bounds in the literature are obtained by pitting a class of *sparse 1-inputs* (e.g., isolated k-cliques or st-paths) against a class of *dense 0-inputs*

[4]Note that this monotone lower bound is quantitatively stronger than the non-monotone $O(n^{2.73\lceil k/3\rceil})$ upper bound from fast matrix multiplication. This reveals a gap between monotone vs. non-monotone complexity (see [42]).

(complete $k-1$-partite graphs or st-cuts). In this circumstance, note that the sets of relevant 0- and 1-inputs are separable (in the anti-monotone direction) by a mere threshold function. No monotone lower bound with this property can therefore extend to TC^0.

This observation motivates the challenge of proving *average-case lower bounds under product distributions* in the monotone setting, in particular for problems like k-CLIQUE and k-CYCLE on Erdős-Rényi graphs. This challenge may be seen as a step toward non-monotone lower bounds insofar as product distributions like $\boldsymbol{G}_{n,p}$ resemble slice distributions like $\boldsymbol{G}_{n,m}$ (the random graph with exactly m edges), due to the fact that monotone and non-monotone complexity measures coincide on slice distributions up to a polynomial factor [6].

1.4 Outline of the article

In the rest of this article, we give an overview of lower bounds which characterize the circuit size, as well as the formula size, of the average-case G-subgraph isomorphism problem in both the AC^0 and monotone settings. The basic technique originated in work of the author [35] where it is shown that AC^0 circuits solving the average-case k-CLIQUE problem require size $\Omega(n^{k/4})$, matching the upper bound from the greedy algorithm. This result improved the previous $\Omega(n^{k/89d^2})$ lower bound of Beame [5] for circuits of alternation-depth d. This is significant for eliminating the dependence on d in the exponent of n up to $O(\log n/k^2 \log\log n)$, at which point the technique breaks down (though the lower bound is conjectured to hold for unbounded d).

Amano [3] generalized the technique to the G-subgraph isomorphism problem for arbitrary patterns G and also gave an extension to hypergraphs. Subsequent work of Li, Razborov and the author [24] further generalized the technique to a colored variant of the G-subgraph isomorphism problem, obtaining an $n^{\Omega(w/\log w)}$ lower bound for patterns of tree-width w. This result is presented in §4.

The challenge of proving stronger lower bounds for formulas was addressed by the author in [36] where it is shown that AC^0 formulas solving the average-case k-CYCLE problem require size $n^{\Omega(\log k)}$. This result sharply separates the power of formulas vs. circuits in the AC^0 setting, as k-CYCLE is solvable by AC^0 circuits of size $n^{O(1)}$. A lower bound for arbitrary patterns G in terms of tree-depth (a graph invariant akin to tree-width) was subsequently shown using recent results in graph minor theory [22] (joint work with Kawarabayashi). These results are described in §5.

These lower bounds in the AC^0 setting apply more generally to any Boolean circuit (or formula) all of whose subcircuits (subformulas) have "low sensitivity with respect to planted subgraphs of G" in a certain sense made precise in §3. By considering a different notion of "sensitivity", quantitatively similar lower bounds for *monotone* circuits and formulas are obtained in [37, 38]. For most patterns G, these lower bounds are merely average-case with respect to a non-product distribution (a convex combination of $\boldsymbol{G}_{n,p}$ and $\boldsymbol{G}_{n,p+o(p)}$). However, in the special case of the k-CYCLE problem, the technique produces an average-case lower bound under $\boldsymbol{G}_{n,p}$. This is significant for being the first super-polynomial lower bound against monotone formulas under any product distribution.

It is hoped that the framework behind these lower bounds might eventually offer an approach to proving super-polynomial lower bounds for unrestricted Boolean formulas and circuits.

2 Colored G-subgraph isomorphism

The main target problem for our lower bounds is actually a colored version of the G-subgraph isomorphism problem, which we denote by $\mathrm{SUB}(G)$. In this problem, the input is a G-colored graph X with vertex set $V(G) \times \{1, \dots, n\}$ and the task to determine whether X contains a copy of the pattern G that involves one vertex from each color class. Compared with the previously discussed *uncolored* G-subgraph isomorphism problem, which we denote by $\mathrm{SUB}_{\mathrm{uncol}}(G)$, the colored variant turns out to be better structured and admits a richer class of threshold distributions. All average-case lower bounds for $\mathrm{SUB}(G)$ in this article extend to the average-case $\mathrm{SUB}_{\mathrm{uncol}}(G)$ as a special case (as we explain in Example 2.6).

Definitions. All *graphs* in this article are finite simple graphs without isolated vertices. Formally, a graph G consists of a set $V(G)$ of vertices and a set $E(G) \subseteq \binom{V(G)}{2}$ of unordered edges such that $V(G) = \bigcup_{\{v,w\} \in E(G)} \{v, w\}$. A *subgraph* of G is a graph H such that $E(H) \subseteq E(G)$ (we simply write $H \subseteq G$). A graph G thus has $2^{|E(G)|}$ subgraphs, which are naturally identified with points in the hypercube $\{0,1\}^{|E(G)|}$. An *isomorphism* between graphs G and G' is a bijection $\pi : V(G) \to V(G')$ such that $\{v, w\} \in E(G) \Leftrightarrow \{\pi(v), \pi(w)\} \in E(G')$ for all distinct vertices v, w of G.

The n-*blowup* of a graph G, denoted $G^{\uparrow n}$, has vertices $v^{(1)}, \dots, v^{(n)}$ for each $v \in V(G)$ and edges $\{v^{(a)}, w^{(b)}\}$ for each $\{v, w\} \in E(G)$ and $a, b \in [n]$ $(:= \{1, \dots, n\})$. We view $G^{\uparrow n}$ and its subgraphs as "G-colored graphs" under the vertex-coloring $v^{(a)} \mapsto v$.

The *colored G-subgraph isomorphism problem*, denoted $\mathrm{SUB}(G)$ for short, is the computational task, given a G-colored graph $X \subseteq G^{\uparrow n}$ as input, of determining whether X contains a subgraph that is isomorphic G via the map $v^{(a)} \mapsto v$. Formally, this problem is represented by a sequence of Boolean functions $\{0,1\}^{kn^2} \to \{0,1\}$ where $k = |E(G)|$ and $kn^2 = |E(G^{\uparrow n})|$.

Henceforth, H is always a subgraph of G, while X is a subgraph of $G^{\uparrow n}$. For an element $\alpha \in [n]^{V(H)}$, let $H^{(\alpha)}$ denote the copy of H in $G^{\uparrow n}$ with vertices $v^{(\alpha_v)}$ for $v \in V(H)$ and edges $\{v^{(\alpha_v)}, w^{(\alpha_w)}\}$ for $\{v,w\} \in E(H)$. We refer to subgraphs of X of the form $H^{(\alpha)}$ as *H-subgraphs* of X. Let $\mathrm{sub}_H(X)$ denote the number of H-subgraphs of X, that is, $\mathrm{sub}_H(X) := |\{\alpha \in [n]^{V(H)} : H^{(\alpha)} \subseteq X\}|$.

On the relationship between $\mathrm{SUB}_{\mathrm{uncol}}(G)$ and $\mathrm{SUB}(G)$. For every pattern G, the color-coding method of [2] provides an efficient many-one reduction from $\mathrm{SUB}_{\mathrm{uncol}}(G)$ to $\mathrm{SUB}(G)$. The colored version of G-subgraph isomorphism is therefore the harder problem in general. However, for many graphs G of interest such as cliques, these two problems are in fact equivalent. Namely, if G is a *core* (meaning every homomorphism $G \to G$ is an isomorphism), then there is a trivial reduction from $\mathrm{SUB}(G)$ to $\mathrm{SUB}_{\mathrm{uncol}}(G)$, as the only subgraphs of $G^{\uparrow n}$ that are isomorphic to G are those of the form $G^{(\alpha)}$.

2.1 Threshold random graphs

For the average-case analysis of the problem $\mathrm{SUB}(G)$, it is natural to study a G-colored version of the Erdős-Rényi random graph. For a vector $\vec{p} \in [0,1]^{E(G)}$ of edge probabilities (one $p_e \in [0,1]$ for each $e \in E(G)$), let $\boldsymbol{G}_{n,\vec{p}}$ denote the random subgraph of $G^{\uparrow n}$ which includes each potential edge $\{v^{(a)}, w^{(b)}\}$ independently with probability $p_{\{v,w\}}$. The class of "threshold vectors" for the existence of G-subgraphs in $\boldsymbol{G}_{n,\vec{p}}$ has a characterization in terms of certain edge-weightings on G.

Definition 2.1 (Threshold weighting)**.** Let G be a graph, let θ be a function $E(G) \to [0,2]$, and let Δ_θ be the function $\{\text{subgraphs of } G\} \to \mathbb{R}_{\geq 0}$ defined by

$$\Delta_\theta(H) := |V(H)| - \textstyle\sum_{e \in E(H)} \theta(e).$$

We say that θ is a *threshold weighting* on G if $\Delta_\theta(G) = 0$ and $\Delta_\theta(H) \geq 0$ for all $H \subseteq G$. We say that θ is *strict* if, moreover, $\Delta_\theta(H) > 0$ for all proper subgraphs $\emptyset \subset H \subset G$.

The set of threshold weightings on G forms a convex polytope in $[0,2]^{E(G)}$. For connected graphs G, the strict threshold weightings form the interior of this polytope. Note that only connected graphs admit strict threshold weightings, as it follows from the definition that $\Delta_\theta(H) = 0$ whenever H is a union of connected components of G.

Example 2.2. For every graph G, the function $\theta : E(G) \to [0,2]$ defined by $\theta(\{v,w\}) := \frac{1}{\deg(v)} + \frac{1}{\deg(w)}$ is a threshold weighting. In particular, if G is r-regular, then the constant function $\theta = \frac{2}{r}$ is a threshold weighting. (Two additional constructions of threshold weightings are described at the end of this section.)

Definition 2.3 (The random graph $\boldsymbol{X}_\theta$)**.** Every threshold weighting θ on G gives rise to a sequence of random graphs $\boldsymbol{X}_{n,\theta}$, defined as the G-colored Erdős-Rényi graph $\boldsymbol{G}_{n,\vec{p}}$ where $\vec{p} \in [0,1]^{E(G)}$ is the vector of edge probabilities $p_e = n^{-\theta(e)}$. That is, $\boldsymbol{X}_{n,\theta}$ is the random subgraph of $G^{\uparrow n}$ which includes each potential edge $\{v^{(a)}, w^{(b)}\}$ independently with probability $n^{-\theta(\{v,w\})}$. To simplify notation, we will generally omit the parameter n and simply write $\boldsymbol{X}_\theta$.

Observe that the function Δ_θ characterizes the expected number of H-subgraphs in $\boldsymbol{X}_\theta$: for every $H \subseteq G$, we have $\mathrm{E}[\ \mathrm{sub}_H(\boldsymbol{X}_\theta)\ ] = n^{\Delta_\theta(H)}$ by linearity of expectation. In particular, $\mathrm{sub}_G(\boldsymbol{X}_\theta)$ has expectation 1 (since $\Delta_\theta(G) = 0$). Moreover, when θ is strict, $\mathrm{sub}_G(\boldsymbol{X}_\theta)$ is asymptotically Poisson and $\mathrm{sub}_H(\boldsymbol{X}_\theta)$ is highly concentrated around its mean for all proper subgraphs $H \subset G$.

Proposition 2.4. *For every graph G and threshold weighting θ, the probability that $\boldsymbol{X}_\theta$ contains a G-subgraph converges to a limit in* $(0,1)$*. When θ is strict, this limit is* $1 - \frac{1}{e}$*.*

In light of Proposition 2.4, it makes sense to study the average-case complexity of $\mathrm{SUB}(G)$ on $\boldsymbol{X}_\theta$, that is, the complexity of functions $f : \{$subgraphs of $G^{\uparrow n}\} \to \{0,1\}$ such that $f(\boldsymbol{X}_\theta) = 1 \Leftrightarrow \mathrm{sub}_G(\boldsymbol{X}_\theta) \geq 1$ holds asymptotically almost surely. We conclude this section with two constructions of threshold weightings.

Example 2.5 (Threshold weightings from Markov chains)**.** Let G be any graph and let $M : V(G) \times V(G) \to [0,1]$ be a *Markov chain* on G satisfying

- $M(v,w) > 0 \Rightarrow \{v,w\} \in E(G)$ and
- $\sum_w M(v,w) = 1$ for every v.

Then the function $E(G) \to [0,2]$ given by $\{v,w\} \mapsto M(v,w) + M(w,v)$ is a threshold weighting on G. (This construction generalizes Example 2.2, which corresponds to the Markov chain where $M(v,w) = \frac{1}{\deg(v)}$ for every $\{v,w\} \in E(G)$.) The associated function Δ_M has the property that $\Delta_M(H)$ equals the amount of M-flow leaving the subgraph H (i.e., $\sum_{v,w} M(v,w)$ over pairs v,w with $v \in V(H)$ and $\{v,w\} \in E(G) \setminus E(H)$). In §4.1 we use this construction of threshold weightings to bound the AC^0 circuit size of $\mathrm{SUB}(G)$ in terms of the tree-width of G.

Example 2.6 (The uncolored setting)**.** The threshold for the existence of G-subgraphs in the Erdős-Rényi random graph $\boldsymbol{G}_{n,p}$ is well-known to be $p = \Theta(n^{-c})$ where c is the constant $\min_{H \subseteq G} \frac{|V(H)|}{|E(H)|}$ [7]. For all intents and purposes, the average-case analysis of $\mathrm{SUB}_{\mathrm{uncol}}(G)$ on $\boldsymbol{G}_{n,p}$ is equivalent to the average-case analysis of $\mathrm{SUB}(G)$ on $\boldsymbol{X}_{\theta_{\mathrm{uncol}}}$ where $\theta_{\mathrm{uncol}} : E(G) \to \{0,c\}$ is the threshold weighting defined by $\theta_{\mathrm{uncol}}(e) = c \Leftrightarrow$ there exists $H \subseteq G$ such that $e \in E(H)$ and $\frac{|V(H)|}{|E(H)|} = c$. All lower and upper bounds described in this article translate easily between these two average-case settings, modulo insignificant constant factors as between $n^{|V(G)|}$ and $\binom{n}{|V(G)|}$.

3 H-subgraph sensitivity

AC^0 functions are known to have *low average sensitivity* in the following sense [8]: for any AC^0 function $f : \{0,1\}^n \to \{0,1\}$ and independent uniform random $\boldsymbol{x} \in \{0,1\}^n$ and $\boldsymbol{i} \in [n]$, it holds that

$$\Pr_{\boldsymbol{x},\boldsymbol{i}}[\ f(\boldsymbol{x}) \neq f(\boldsymbol{x} \text{ with its } \boldsymbol{i}^{\text{th}} \text{ coordinate flipped})\ ] \leq n^{-1+o(1)}.$$

Analogously, a key lemma in our lower bounds shows that AC^0 functions $f : \{\text{subgraphs of } G^{\uparrow n}\} \to \{0,1\}$ have what might be termed "low average H-subgraph sensitivity on $\boldsymbol{X}_\theta$".

Definition 3.1. For any graph F, let $\mathbb{B}(F)$ denote the set of functions $\{\text{subgraphs of } F\} \to \{0,1\}$.

We say that a function $f \in \mathbb{B}(F)$ *depend on all coordinates* if for every $e \in E(F)$, there exists a subgraph $F' \subseteq F$ such that $f(F') \neq f(F' - e)$ where $F' - e$ is the graph with edge set $E(F') \setminus \{e\}$ (in other words, if f depends on all coordinates when viewed as a Boolean function $\{0,1\}^{|E(F)|} \to \{0,1\}$).

For a function $f \in \mathbb{B}(F)$ and graphs $X, H \subseteq F$,

- let $f^{\cup X} \in \mathbb{B}(F)$ denote the function $f^{\cup X}(F') := f(X \cup F')$ and

- let $f\restriction_H \in \mathbb{B}(H)$ denote the restriction of f to domain $\{$subgraphs of $H\}$.

Note that the function $f^{\cup X}\restriction_H \in \mathbb{B}(H)$ depends on all coordinates if, and only if, for every $e \in E(H)$, there exists a subgraph $H' \subseteq H$ such that $f(X \cup H') \neq f(X \cup (H' - e))$.

Fix any graph G and threshold weighting θ. Consider any subgraph $H \subseteq G$ and let $\boldsymbol{\alpha}$ be a uniform random element of $[n]^{V(H)}$, independent of $\boldsymbol{X}_\theta$. For a function $f \in \mathbb{B}(G^{\uparrow n})$, we consider the randomly restricted function $f^{\cup \boldsymbol{X}_\theta}\restriction_{H^{(\boldsymbol{\alpha})}} \in \mathbb{B}(H^{(\boldsymbol{\alpha})})$. When f is AC^0-computable, the following lemma bounds the probability that $f^{\cup \boldsymbol{X}_\theta}\restriction_{H^{(\boldsymbol{\alpha})}}$ depends on all coordinates.

Lemma 3.2 (H-subgraph sensitivity of AC^0 functions [24])**.** *Suppose $f \in \mathbb{B}(G^{\uparrow n})$ is an AC^0-computable sequence of functions. Then for every subgraph $H \subseteq G$,*

$$\Pr_{\boldsymbol{X}_\theta,\, \boldsymbol{\alpha} \in [n]^{V(H)}}[\ f^{\cup \boldsymbol{X}_\theta}\restriction_{H^{(\boldsymbol{\alpha})}} \textit{ depends on all coordinates }] \leq n^{-\Delta_\theta(H)+o(1)}.$$

When $\Delta_\theta(H) > 0$, the $n^{-\Delta_\theta(H)+o(1)}$ bound of Lemma 3.2 is nontrivial and moreover tight.[5] However, note that this lemma says nothing when $\Delta_\theta(H) = 0$, in particular when $H = G$. The main tools in the proof are Håstad's Switching Lemma [14], which shows that random restrictions simplify AC^0 circuits, and Janson's Inequality [16], which implies lower tail bounds for random variables $\mathrm{sub}_H(\boldsymbol{X}_\theta)$. The assumption that f is AC^0-computable is necessary, as for instance if f is the PARITY function (mapping $X \subseteq G^{\uparrow n}$ to $|E(X)| \bmod 2$), then the restricted function $f^{\cup \boldsymbol{X}_\theta}\restriction_{H^{(\boldsymbol{\alpha})}}$ depends on all coordinates with probability 1. (In the case that H is a single-edge subgraph of G, Lemma 3.2 essentially equivalent to aforementioned bound on the average sensitivity of AC^0 functions, only with respect to a product distribution rather than the uniform distribution.)

The next lemma is an analogue of Lemma 3.2 in the monotone setting. It shows that every monotone function, irrespective of its monotone circuit complexity, has "low average H-subgraph sensitivity of f on $\boldsymbol{X}_\theta$" in a different sense. Namely, we consider the event that $H^{(\boldsymbol{\alpha})}$ is a common minterm of f and $f^{\cup \boldsymbol{X}_\theta}$ (i.e., $f(H^{(\boldsymbol{\alpha})}) = 1$ and $f^{\cup \boldsymbol{X}_\theta}(H^{(\boldsymbol{\alpha})} - e) = 0$ for every $e \in E(H^{(\boldsymbol{\alpha})})$).

[5] Let f be the function $f(X) = 1 \Leftrightarrow \bigvee_{\alpha \in \mathcal{A}} (H^{(\alpha)} \subseteq X)$ where $\mathcal{A}$ is a generic (i.e., almost any choice of) subset of $[n]^{V(H)}$ of size $|\mathcal{A}| = n^{|V(H)| - \Delta_\theta(H)}$. Then f is AC^0-computable and $\Pr[\ f^{\cup \boldsymbol{X}_\theta}\restriction_{H^{(\boldsymbol{\alpha})}}$ depends on all coordinates $] = \Omega(n^{-\Delta_\theta(H)})$.

Lemma 3.3 (H-subgraph sensitivity of monotone functions [38])**.** *For every monotone function $f \in \mathbb{B}(G^{\uparrow n})$ and subgraph $H \subseteq G$,*

$$\Pr_{\boldsymbol{X}_\theta,\, \boldsymbol{\alpha} \in [n]^{V(H)}}[\ H^{(\boldsymbol{\alpha})} \textit{ is a common minterm } f \textit{ and } f^{\cup \boldsymbol{X}_\theta}\ ] \leq n^{-\Delta_\theta(H)+o(1)}.$$

In §5.3 we explain how Lemma 3.3 is used in place of Lemma 3.2 to derive lower bounds for monotone circuits and formulas using the same framework as our AC^0 lower bounds.

4 The AC^0 circuit size of $\mathrm{SUB}(G)$

This section presents results of Li, Razborov and the author [24], which characterize the average-case AC^0 circuit size of $\mathrm{SUB}(G)$ on $\boldsymbol{X}_\theta$ for any G and θ in terms of a combinatorial invariant $\kappa_\theta(G)$. This invariant is defined by dual min-max and max-min expressions.

Definition 4.1. A *union family* for a graph G is a set $\mathcal{F}$ of subgraphs of G such that $G \in \mathcal{F}$ and every $F \in \mathcal{F}$ with at least two edges is the union of two proper subgraphs which both belong to $\mathcal{F}$ (i.e., there exist proper subgraphs $F_1, F_2 \subset F$ with $F_1 \cup F_2 = F$ and $F_1, F_2 \in \mathcal{F}$). Intuitively, $\mathcal{F}$ is a blueprint for constructing G out of individual edges by taking pairwise unions of subgraphs.

A *hitting family* for G is a set $\mathcal{H}$ of subgraphs of G such that $\mathcal{F} \cap \mathcal{H} \neq \emptyset$ for every union family $\mathcal{F}$ for G.

For any threshold weighting θ on G, the invariant $\kappa_\theta(G)$ is defined by the pair of dual expressions

$$\kappa_\theta(G) := \min_{\text{union families } \mathcal{F}} \ \max_{F \in \mathcal{F}} \ \Delta_\theta(F) = \max_{\text{hitting families } \mathcal{H}} \ \min_{H \in \mathcal{H}} \ \Delta_\theta(H).$$

Example 4.2. We illustrate these definitions by working through an example. Let K_k be the k-clique graph (i.e., the complete graph of order $k \geq 2$) and let θ be the constant threshold weighting $\frac{2}{k-1}$. We will show that $\kappa_\theta(K_k) = \frac{k}{4} + O(\frac{1}{k})$ by constructing a union family $\mathcal{F}$ and a hitting family $\mathcal{H}$ that witness matching upper and lower bounds for $\kappa_\theta(K_k)$.

Let $\mathcal{F}$ be the set of subgraphs $F \subseteq K_k$ such that F is either a clique (i.e., a complete subgraph $K_I \subseteq K_k$ where $I \subseteq [k]$ with $|I| \geq 2$) or a clique minus a single edge. Note that $\mathcal{F}$ is a union family for K_k, as $K_k \in \mathcal{F}$ and every graph in $\mathcal{F}$ with at least two edges is the union of two proper subgraphs in $\mathcal{F}$ (e.g., $K_{\{1,\dots,j\}}$ minus the edge $\{1,j\}$ is the union of $K_{\{1,\dots,j-1\}}$ and $K_{\{2,\dots,j\}}$). A straightforward calculation shows $\kappa_\theta(K_k) \leq \max_{F \in \mathcal{F}} \Delta_\theta(F) =$

$\max_{F\in\mathcal{F}} |V(F)| - \frac{2}{k-1}|E(F)| = \frac{k}{4} + O(\frac{1}{k})$, where this maximum over $F \in \mathcal{F}$ is attained by a clique of size $\lceil \frac{k}{2} \rceil$ minus a single edge.

To obtain a matching lower bound on $\kappa_\theta(K_k)$, we consider the hitting family $\mathcal{H}$ consisting of subgraphs $H \subseteq K_k$ such that $|V(H)| \geq \frac{k}{2}$ and $H = H_1 \cup H_2$ for some H_1, H_2 satisfying $|V(H_1)|, |V(H_2)| < \frac{k}{2}$. The minimum of $\Delta_\theta(H)$ over $H \in \mathcal{H}$ is again attained by a clique of size $\lceil \frac{k}{2} \rceil$ minus a single edge. This shows that the $\frac{k}{4} + O(\frac{1}{k})$ upper bound coming from $\mathcal{F}$ is tight.

Example 4.3. If G is an r-regular expander and $\theta = \frac{2}{r}$, then we obtain a lower bound $\kappa_\theta(G) = \Omega(|V(G)|)$ (for a constant depending on the edge-expansion of G) by considering the hitting family $\{H \subseteq G : \frac{1}{3} \leq \frac{|V(H)|}{|V(G)|} < \frac{2}{3}\}$.

We next state the main theorem of [24] and outline its proof.

Theorem 4.4. *For every graph G and threshold weighting θ, the average-case AC^0 circuit size of* $\mathrm{SUB}(G)$ *on $\boldsymbol{X}_\theta$ is at least $n^{\kappa_\theta(G)-o(1)}$ and at most $n^{2\kappa_\theta(G)+O(1)}$.*

Theorem 4.4 together with Examples 2.6 and 4.2 imply a lower bound of $\Omega(n^{k/4})$ on the AC^0 circuit size of the average-case k-CLIQUE problem on $\boldsymbol{G}_{n,p}$ at the threshold $p = \Theta(n^{-2/(k-1)})$.

The upper bound. We give a high-level description of an algorithm that solves $\mathrm{SUB}(G)$ a.a.s. correctly on $\boldsymbol{X}_\theta$ in time $n^{2\kappa_\theta(G)+O(1)}$, omitting details of the implementation by AC^0 circuits. We use the fact that, with high probability, $\mathrm{sub}_H(\boldsymbol{X}_\theta)$ is at most $n^{\Delta_\theta(H)+o(1)}$ for all $H \subseteq G$ (by Markov's inequality). Fix an optimal union family $\mathcal{F}$ such that $\kappa_\theta(G) = \max_{F\in\mathcal{F}} \Delta_\theta(F)$. Also fix an enumeration $F_1, \ldots, F_m$ of graphs in $\mathcal{F}$ such that $F_m = G$ and each F_i is either a single edge or the union of two previous graphs in the sequence. In order for $k = 1, \ldots, m$, the algorithm will compile a list of all F_k-subgraphs in $\boldsymbol{X}_\theta$. When F_k is a single edge, this takes time $O(n^2)$. When $F_k = F_i \cup F_j$ for $i, j < k$, this is done by examining each pair of subgraphs $F_i^{(\alpha)} \subseteq \boldsymbol{X}_\theta$ and $F_j^{(\beta)} \subseteq \boldsymbol{X}_\theta$ from the previously compiled lists: if $\alpha_v = \beta_v$ for all $v \in V(F_i) \cap V(F_j)$, then $F_k^{(\alpha\cup\beta)}$ is added to the list of F_k-subgraphs. Compiling this list therefore takes time $O(\mathrm{sub}_{F_i}(\boldsymbol{X}_\theta) \cdot \mathrm{sub}_{F_j}(\boldsymbol{X}_\theta))$, which with high probability is at most $n^{\Delta_\theta(F_i)+\Delta_\theta(F_j)+o(1)} \leq n^{2\kappa_\theta(G)+o(1)}$. Since there are only $O(1)$ (at most $2^{|E(G)|}$) lists to compute and nonemptiness of the final list determines whether $\boldsymbol{X}_\theta$ contains a G-subgraph, this algorithm has expected time $n^{2\kappa_\theta(G)+O(1)}$.

The lower bound. Let C be a sequence of AC^0 circuits of size $n^{\kappa_\theta(G)-\Omega(1)}$ which compute functions $f \in \mathbb{B}(G^{\uparrow n})$. Our goal is to show that f does not agree with SUB(G) a.a.s. on $\boldsymbol{X}_\theta$. We consider the randomly restricted function $f^{\cup \boldsymbol{X}_\theta}\restriction_{G(\boldsymbol{\alpha})}$ where $\boldsymbol{\alpha}$ is a uniform random element of $[n]^{V(G)}$ independent of $\boldsymbol{X}_\theta$. We will show that

$$\Pr[\ f^{\cup \boldsymbol{X}_\theta}\restriction_{G(\boldsymbol{\alpha})} \text{ depends on all coordinates }] = o(1). \tag{4.1}$$

Inequality (4.1) uses Lemma 3.2 on the "H-subgraph sensitivity" of AC^0 functions. However, (4.1) does not follow by directly applying Lemma 3.2 to f with $H = G$ (as the $n^{-\Delta_\theta(H)+o(1)}$ bound of Lemma 3.2 is trivial when $H = G$). Rather, we apply Lemma 3.2 to all functions g computed by subcircuits of C with respect to all subgraphs $H \subseteq G$ which come from an optimal hitting family for G. We present the argument in detail in a moment.

On the other hand, we show that every function $f \in \mathbb{B}(G^{\uparrow n})$ which agrees with SUB(G) a.a.s. on $\boldsymbol{X}_\theta$ satisfies

$$\Pr[\ f^{\cup \boldsymbol{X}_\theta}\restriction_{G(\boldsymbol{\alpha})} \text{ depends on all coordinates }] = \Omega(1). \tag{4.2}$$

Since (4.1) and (4.2) are contradictory for sufficiently large n, we conclude that functions f computed by C do not solve SUB(G) on $\boldsymbol{X}_\theta$.

We first justify (4.2), which is the more straightforward inequality. To illustrate the general idea, we make the stronger assumption that f coincides with SUB(G) on all inputs and we further assume that θ is strict. In this case, Proposition 2.4 implies that $\boldsymbol{X}_\theta$ has no G-subgraph with probability $\frac{1}{e} - o(1)$. A straightforward union bound shows that, a.a.s., if $\boldsymbol{X}_\theta$ has no G-subgraph, then neither does $\boldsymbol{X}_\theta \cup H^{(\boldsymbol{\alpha})}$ for any proper subgraph $H \subset G$. (By "$H^{(\boldsymbol{\alpha})}$" we mean $H^{(\boldsymbol{\alpha}_{V(H)})}$, which is a uniform random H-subgraph of $G^{\uparrow n}$ independent of $\boldsymbol{X}_\theta$.) It follows that, with probability $\frac{1}{e} - o(1)$, the randomly restricted function $f^{\cup \boldsymbol{X}_\theta}\restriction_{G(\boldsymbol{\alpha})} \in \mathbb{B}(G^{(\boldsymbol{\alpha})})$ outputs 1 on G and 0 on every $H \subset G$ (i.e., $f^{\cup \boldsymbol{X}_\theta}\restriction_{G(\boldsymbol{\alpha})}$ is the AND function over coordinates $G^{(\boldsymbol{\alpha})}$). Since this function depends on all coordinates, inequality 4.2 follows. (When we only assume that f agrees with SUB(G) a.a.s. on $\boldsymbol{X}_\theta$, this argument additionally requires showing that the total variation distance between random graphs $\boldsymbol{X}_\theta$ and $\boldsymbol{X}_\theta \cup H^{(\boldsymbol{\alpha})}$ is $1 - \Omega(1)$ for every $H \subseteq G$.)

Onto the more interesting inequality (4.1), showing that a.a.s. $f^{\boldsymbol{X}_\theta}\restriction_{G(\boldsymbol{\alpha})}$ does not depend on all coordinates. Let $\mathcal{G} \subseteq \mathbb{B}(G^{\uparrow n})$ be the set of functions computed by subcircuits of C. For every $g \in \mathcal{G}$ and $H \subseteq G$, Lemma 3.2 implies that the randomly restricted function $g^{\cup \boldsymbol{X}_\theta}\restriction_{H(\boldsymbol{\alpha})}$ depends on all coordinates with probability at most $n^{-\Delta_\theta(H)+o(1)}$. Let us now fix an optimal hitting family $\mathcal{H} \subseteq \{\text{subgraphs of } G\}$ such that $\kappa_\theta(G) = \min_{H \in \mathcal{H}} \Delta_\theta(H)$. Taking a union bound over $g \in \mathcal{G}$ and $H \in \mathcal{H}$, we have

$$\text{(4.3)} \quad \Pr[\ (\exists g \in \mathcal{G})(\exists H \in \mathcal{H})\ g^{\cup \boldsymbol{X}_\theta}\restriction_{H(\boldsymbol{\alpha})} \text{ depends on all coordinates }] \\ \leq |\mathcal{G}| \cdot |\mathcal{H}| \cdot n^{-\kappa_\theta(G)+o(1)} = o(1)$$

since $|\mathcal{G}| \leq \text{size}(C) = n^{\kappa_\theta(G)-\Omega(1)}$ and $|\mathcal{H}| \leq 2^{|E(G)|} = O(1)$. Inequality (4.1) now follows by combining (4.3) with the following non-probabilistic claim.

Claim 4.5. *For any $X \subseteq G^{\uparrow n}$ and $\alpha \in [n]^{V(G)}$, if $f^{\cup X}\restriction_{G(\alpha)}$ depends on all coordinates, then there exist $g \in \mathcal{G}$ and $H \in \mathcal{H}$ such that $g^{\cup X}\restriction_{H(\alpha)}$ depends on all coordinates.*

To prove Claim 4.5, assume $f^{\cup X}\restriction_{G(\alpha)}$ depends on all coordinates. Let $\mathcal{F}$ be the family of subgraphs $F \subseteq G$ for which there exists $g \in \mathcal{G}$ such that $g^{\cup X}\restriction_{F(\alpha)}$ depends on all coordinates. It suffices to show that $\mathcal{F}$ is a union family for G. The claim then follows from the fact that $\mathcal{F} \cap \mathcal{H}$ is nonempty (since $\mathcal{H}$ is a hitting family for G). To show that $\mathcal{F}$ is a union family, we first note that $G \in \mathcal{F}$ (by the assumption that $f^{\cup X}\restriction_{G(\alpha)}$ depends on all coordinates).

Now consider any $F \in \mathcal{F}$ with ≥ 2 edges. It remains to show that F is the union of two proper subgraphs which belong to $\mathcal{F}$. By definition of $\mathcal{F}$ and $\mathcal{G}$, there exists a function $g \in \mathbb{B}(G^{\uparrow n})$ computed by a subcircuit of C such that $g^{\cup X}\restriction_{F(\alpha)}$ depends on all coordinates. Fix a choice of g computed by a subcircuit of minimal depth in C. Since $g^{\cup X}\restriction_{F(\alpha)}$ depends on ≥ 2 coordinates (namely all edges of $F^{(\alpha)}$), it cannot correspond to an input of C and must therefore come from a gate of C. Let g_1 and g_2 be the functions computed by the two subcircuits feeding into this gate. The function g is thus either $g_1 \wedge g_2$ or $g_1 \vee g_2$.

For $i = 1, 2$, let F_i be the graph consisting of edges $\{v, w\} \in E(F)$ such that $g_i^{\cup X}\restriction_{F(\alpha)}$ depends on the corresponding edge $\{v^{(\alpha_v)}, w^{(\alpha_w)}\} \in E(F^{(\alpha)})$. Observe that the function $g_i^{\cup X}\restriction_{F_i^{(\alpha)}} \in \mathbb{B}(F_i^{(\alpha)})$ depends on all coordinates. Therefore, $F_i \in \mathcal{F}$. Next, note that F_i must be proper subgraph of F by the minimality in our choice of g. Finally, observe that $F = F_1 \cup F_2$ (since if g depends on a given coordinate in $E(F)$, then so must one or both of g_1 and g_2, and the same is true after applying the restriction ${}^{\cup X}\restriction_{F(\alpha)}$ to all three functions). As we have shown that F is the union of two proper subgraphs which belong to $\mathcal{F}$, this completes the proof. $\square$

By a similar argument, we obtain a similar $n^{\kappa_\theta(G)-o(1)}$ lower bound on the *monotone* circuit size of SUB(G). In this argument, Lemma 3.3 plays the role of Lemma 3.2 in bounding the "H-subgraph sensitivity" of each subcircuit. However, as we explain in §5.3, for most patterns G, the lower bound we obtain in the monotone setting is only worst-case, or average-case under a non-product distribution.

4.1 Tree-width

Tree-width, denoted $\mathsf{tw}(G)$, is an important invariant that arises frequently in parameterized complexity and several areas of graph theory. Roughly speaking, it measures the extent to which a graph is "tree-like": trees and forests have tree-width 1, while the complete graph of order k has tree-width $k-1$.

In the introduction, it was mentioned that the G-subgraph isomorphism problem is solvable in time $O(n^{\mathsf{tw}(G)+1})$ for all patterns G. In fact, SUB(G) is solvable by monotone AC^0 circuits of size $O(n^{\mathsf{tw}(G)+1})$. If we compare this upper bound to the lower bound of Theorem 4.4, we see that $\max_\theta \kappa_\theta(G) \leq \mathsf{tw}(G)+1$. The next proposition shows that this inequality is nearly tight.

Proposition 4.6. *Every graph G admits a threshold weighting θ such that $\kappa_\theta(G) = \Omega(\mathsf{tw}(G)/\log \mathsf{tw}(G))$.*

Proof sketch. We will use a lemma of Grohe and Marx [13] which states that, for every G with tree-width k, there exists a set $W \subseteq V(G)$ of size $|W| = \Omega(k)$ together with a concurrent flow on G with vertex-capacity 1 which routes $\Omega(\frac{1}{k \log k})$ units of flow between every pair of vertices in W. This concurrent flow is easily transformed to a Markov chain M on G (in the sense of Example 2.5) with the property that $\Delta_M(H) = \Omega(\frac{|V(H)\cap W|\cdot|V(H)\setminus W|}{k \log k})$ for all $H \subseteq G$. We now consider the hitting family $\mathcal{H}$ consisting of subgraphs $H \subseteq G$ such that $\frac{1}{3} \leq \frac{|V(H)\cap W|}{|W|} < \frac{2}{3}$ (similar to Example 4.3). This gives the bound $\kappa_M(G) \geq \min_{H\in\mathcal{H}} \Delta_M(H) = \Omega(\frac{k}{\log k})$ with respect to the threshold weighting $\{v,w\} \mapsto M(v,w) + M(w,v)$ induced by M. $\square$

We remark that the upper bound $\max_\theta \kappa_\theta(G) \leq \mathsf{tw}(G)+1$ has a direct proof that does not appeal to Theorem 4.4. In fact, the next proposition shows that $\max_\theta \kappa_\theta(G)$ is at most the *branch-width* of G, an invariant that is related to tree-width by $\mathsf{bw}(G) \leq \mathsf{tw}(G)+1 \leq \frac{3}{2}\mathsf{bw}(G)$ [34].

Proposition 4.7. $\kappa_\theta(G) \leq \mathsf{bw}(G)$ *for every threshold weighting θ on G.*

Proof. Branch-width admits a simple characterization in terms of union families:

$$\mathsf{bw}(G) = \min_{\text{complement-closed union families } \mathcal{F}} \; \max_{F\in\mathcal{F}} |V(F) \cap V(\overline{F})|.$$

Here *complement-closed* means $F \in \mathcal{F} \Rightarrow \overline{F} \in \mathcal{F}$ where $\overline{F}$ is the graph with $E(\overline{F}) = E(G) \setminus E(F)$. It follows from Def. 2.1 threshold weighting that $\Delta_\theta(F) \leq \Delta_\theta(F) + \Delta_\theta(\overline{F}) = |V(F) \cap V(\overline{F})|$ for every threshold weighting θ and subgraph $F \subseteq G$. Therefore, $\kappa_\theta(G) = \min_{\text{union families } \mathcal{F}} \max_{F \in \mathcal{F}} \Delta_\theta(F) \leq \mathsf{bw}(G)$. $\square$

5 The restricted formula size of $\mathrm{SUB}(G)$

In this section we sketch an extension the lower bound technique that yields quantitatively stronger lower bounds for formulas vis-à-vis circuits in both the AC^0 and monotone settings. The improvement is significant for patterns of constant tree-width such as paths and cycles where $\mathrm{SUB}(G)$ is computable by polynomial-size circuits but is conjecture to require super-polynomial size formulas.

An outline of this section is as follows. In §5.1 we introduce the key notion of *pathsets* (relations $\mathcal{A} \subseteq [n]^{V(H)}$ that satisfy certain density constraints related to the bounds on "H-subgraph sensitivity" given by Lemmas 3.2 and 3.3). We next define *pathset formulas*, which are a tree-like model for constructing pathsets. In §5.2 we describe a randomized reduction which transforms any AC^0 formula that solves average-case $\mathrm{SUB}(G)$ on $\boldsymbol{X}_\theta$ into a pathset formula that computes a dense subset of $[n]^{V(G)}$. In §5.3 we outline a similar transformation for monotone formulas.

In §5.4 we arrive at the combinatorial heart of the technique: an $n^{\tau_\theta(G)-o(1)}$ lower bound on the size of pathset formulas that compute a dense subset of $[n]^{V(G)}$. Here $\tau_\theta(G)$ is an invariant of the threshold-weighted graphs, which plays an analogous role to $\kappa_\theta(G)$ in the context of formulas. Although $\tau_\theta(G)$ turns out to be much harder to compute, we are able to bound $\tau_\theta(G)$ in a few special cases of interest, such as when G is a cycle, path, or complete binary tree. Finally, in §5.5 we discuss a relationship between $\max_\theta \tau_\theta(G)$ and the tree-depth of G.

5.1 Pathset formulas

In what follows, we fix a graph G and a threshold weighting θ, as well as $n \in \mathbb{N}$ and an arbitrary "density parameter" $\varepsilon \in [0, 1]$. (In our applications, we take ε to be $n^{1-o(1)}$ and later $n^{1/2-o(1)}$.)

Definition 5.1. Let $\mathcal{A} \subseteq [n]^V$ where V is any finite set. (We regard $\mathcal{A}$ as a "V-ary relation with universe $[n]$".) The *density* of $\mathcal{A}$ is defined by

$$\mu(\mathcal{A}) := \Pr_{\boldsymbol{\alpha}\in[n]^V}[\ \boldsymbol{\alpha} \in \mathcal{A}\ ]\ (= |\mathcal{A}|/n^{|V|}).$$

For $S \subseteq V$ and $\beta \in [n]^S$, the *conditional density* of $\mathcal{A}$ on β is defined by

$$\mu(\mathcal{A} \,|\, \beta) := \Pr_{\boldsymbol{\alpha}\in[n]^V}[\ \boldsymbol{\alpha} \in \mathcal{A} \mid \boldsymbol{\alpha}_S = \beta\ ].$$

The *join* of relations $\mathcal{A} \subseteq [n]^V$ and $\mathcal{B} \subseteq [n]^W$ is the relation $\mathcal{A} \bowtie \mathcal{B} \subseteq [n]^{V\cup W}$ consisting of $\gamma \in [n]^{V\cup W}$ such that $\gamma_V \in \mathcal{A}$ and $\gamma_W \in \mathcal{B}$.

Definition 5.2. Let H be a subgraph of G. An *H-pathset* (with respect to $G, \theta, n, \varepsilon$) is a relation $\mathcal{A} \subseteq [n]^{V(H)}$ satisfying density constraints

$$\mu(\mathcal{A} \,|\, \beta) \leq \varepsilon^{\Delta_\theta(H_1)} \quad \text{for all } H_1 \uplus H_2 = H \text{ and } \beta \in [n]^{V(H_2)}. \tag{5.1}$$

Here the pair H_1, H_2 range over vertex-disjoint partitions of H (such that $H_1 \cup H_2 = H$ and $V(H_1) \cap V(H_2) = \emptyset$). Thus, if H has t connected components, then (5.1) includes 2^t separate inequalities. Note that the inequality corresponding to $H_1 = H$ and $H_2 = \emptyset$ (the empty graph) is $\mu(\mathcal{A}) \leq \varepsilon^{\Delta_\theta(H)}$, while the inequality corresponding to $H_1 = \emptyset$ and $H_2 = H$ is vacuous since $\Delta_\theta(\emptyset) = 0$. If H is connected, it follows that a relation $\mathcal{A} \subseteq [n]^{V(H)}$ is an H-pathset if and only if $\mu(\mathcal{A}) \leq \varepsilon^{\Delta_\theta(H)}$. Finally, note that every relation $\mathcal{A} \subseteq [n]^{V(G)}$ is a G-pathset since $\Delta_\theta(G_1) = 0$ whenever G_1 is a union of connected components of G.

Definition 5.3. A *pathset formula* (with respect to $G, \theta, n, \varepsilon$) is a rooted binary tree F together with an indexed family of relations $\{\mathcal{A}_{f,H} \subseteq [n]^{V(H)}\}_{f\in V(F),\, H\subseteq G}$ subject to three conditions:

(i) $\mathcal{A}_{f,H}$ is a H-pathset,

(ii) if f is a leaf and $|E(H)| \geq 2$, then $\mathcal{A}_{f,H} = \emptyset$,

(iii) if f is a non-leaf with children f_1 and f_2, then

$$\mathcal{A}_{f,H} \subseteq \textstyle\bigcup_{H_1,H_2\subseteq H\,:\,H_1\cup H_2=H} (\mathcal{A}_{f_1,H_1} \bowtie \mathcal{A}_{f_2,H_2}).$$

We view F as "computing" the family of pathsets $\{\mathcal{A}_{f_{\mathrm{out}},H}\}_{H\subseteq G}$ (and in particular the G-pathset $\mathcal{A}_{f_{\mathrm{out}},G}$) where f_{out} is the root of F.

5.2 Transforming AC^0 formulas to pathset formulas

For any Boolean function $f \in \mathbb{B}(G^{\uparrow n})$ and a subgraph $H \subseteq G$, let $\mathcal{A}^{\boldsymbol{X}_\theta}_{f,H} \subseteq [n]^{V(H)}$ be the random relation defined by

$$\mathcal{A}^{\boldsymbol{X}_\theta}_{f,H} := \{\alpha \in [n]^{V(H)} : f^{\cup \boldsymbol{X}_\theta}\restriction_{H^{(\alpha)}} \text{ depends on all coordinates}\}.$$

When f is AC^0-computable, Lemma 3.2 is equivalent to the expectation bound $\mathrm{E}[\ \mu(\mathcal{A}^{\boldsymbol{X}_\theta}_{f,H})\ ] \leq n^{-\Delta_\theta(H)+o(1)}$. This can be extended to show that $\mu(\mathcal{A}^{\boldsymbol{X}_\theta}_{f,H}) > n^{-(1-\delta)\Delta_\theta(H)}$ with exponentially small probability for any constant $\delta > 0$ (i.e., with probability $\exp(-\Omega(n^c))$ where $c > 0$ depends on δ and the minimum nonzero value of Δ_θ). It is a small additional step to show that $\mathcal{A}^{\boldsymbol{X}_\theta}_{f,H}$ fails to be an H-pathset (with respect to G, θ, n and $\varepsilon = n^{-1+\delta}$) with exponentially small probability.

If F is an AC^0 formula, it follows that the family of relations $\mathcal{A}^{\boldsymbol{X}_\theta}_{f,H} \subseteq [n]^{V(H)}$ (indexed by subformulas f of F and subgraphs $H \subseteq G$) a.a.s. constitutes a pathset formula. Condition (i) of Def. 5.3 is established by taking a union bound, over the $n^{O(1)}$ pairs of f and H, of the exponentially small probability that $\mathcal{A}^{\boldsymbol{X}_\theta}_{f,H}$ fails to be an H-pathset. Conditions (ii) and (iii) both hold with probability 1 (by observations which appeared earlier in the proof of Claim 4.5). Finally, if F solves SUB(G) a.a.s. correctly on $\boldsymbol{X}_\theta$, it follows that the G-pathset computed by F is .99-dense with constant probability by an argument similar to inequality (4.2).

5.3 Transforming monotone formulas to pathset formulas

Let F be a monotone formula of polynomial size. As a first attempt to transform F to a pathset formula, for each subformula f of F and subgraph $H \subseteq G$, let $\mathcal{M}^{\boldsymbol{X}_\theta}_{f,H} \subseteq [n]^{V(H)}$ be the relation consisting of $\alpha \in [n]^{V(H)}$ such that $H^{(\alpha)}$ is a minterm of $f^{\cup \boldsymbol{X}_\theta}$. This family of relations satisfies conditions (ii) and (iii) of Def. 5.3 with probability 1. (Condition (iii) follows from the elementary fact that every minterm of $f_1 \vee f_2$ is a minterm of f_1 or f_2, while every minterm of $f_1 \wedge f_2$ is the union of a minterm of f_1 and a minterm of f_2.) However, the relation $\mathcal{M}^{\boldsymbol{X}_\theta}_{f,H}$ can fail to be an H-pathsets with probability $\Omega(1/n)$ (e.g., if f is the monotone threshold function $f(X) = 1 \Leftrightarrow |E(X)| \geq \sum_{e \in E(G)} n^{2-\theta(e)}$). This failure probability is too large for us to establish condition (i) by taking a union bound over pairs f and H.

To get around this issue, we consider different relations defined in terms of an increasing sequence $\vec{\boldsymbol{X}}_\theta$ of random graphs $\boldsymbol{X}^0_\theta \subseteq \cdots \subseteq \boldsymbol{X}^m_\theta$ where $m = n^{o(1)}$. This sequence is generated as $\boldsymbol{X}^0_\theta := \boldsymbol{X}_\theta$ and $\boldsymbol{X}^i_\theta := \boldsymbol{X}^{i-1}_\theta \cup \boldsymbol{Y}^i$

where $\boldsymbol{Y}^i$ is an independent copy of the G-colored Erdős-Rényi graph $\boldsymbol{G}_{n,\vec{p}}$ where $p_e := n^{-(1-\delta)\theta(e)}$ for a small constant $\delta > 0$ (i.e., each $\boldsymbol{Y}^i$ is a sparse version of $\boldsymbol{X}_\theta$). If f is a depth-d subformula of F, we say that $H^{(\alpha)}$ is a *persistent minterm* of $f^{\cup \vec{\boldsymbol{X}}_\theta}$ if it is a common minterm of $f^{\cup \boldsymbol{X}_\theta^i}$ and $f^{\cup \boldsymbol{X}_\theta^j}$ for some $0 \le i < j \le m$ with $j - i = \binom{d+|E(H)|}{|E(H)|}$. Finally, we consider relations

$$\mathcal{P}_{f,H}^{\vec{\boldsymbol{X}}_\theta} := \{\alpha \in [n]^{V(H)} : H^{(\alpha)} \text{ is a persistent minterm of } f^{\cup \vec{\boldsymbol{X}}_\theta}\}.$$

The definition of persistent minterms ensures that, just like $\mathcal{M}_{f,H}^{\boldsymbol{X}_\theta}$, this family of relations satisfies conditions (ii) and (iii) of Def. 5.3 with probability 1. An extension of Lemma 3.3 shows that $\mathcal{P}_{f,H}^{\vec{\boldsymbol{X}}_\theta}$ fails to be an H-pathset (with respect to $\varepsilon = n^{-1+2\delta}$) with exponentially small probability. A union bound now shows that this family of relations a.a.s. satisfies condition (i), thus transforming F to a pathset formula.

In order for this pathset formula to compute a .99-dense G-pathset with constant probability, we require two additional assumptions: first, that F has depth $O(\log n)$ so that $\binom{\mathrm{depth}(F)+|E(G)|}{|E(G)|} \le m = n^{o(1)}$ (this is without loss of generality by balancing F via Spira's theorem [40]); and second, that F solves $\mathrm{SUB}(G)$ a.a.s. on both $\boldsymbol{X}_\theta$ and $\boldsymbol{X}_\theta^m$ ($= \boldsymbol{X}_\theta \cup \boldsymbol{Y}^1 \cup \cdots \cup \boldsymbol{Y}^m$). This is akin to solving $\mathrm{SUB}_{\mathrm{uncol}}(G)$ a.a.s. correctly on both $\boldsymbol{G}_{n,p}$ and $\boldsymbol{G}_{n,p+p^{1+\delta}}$, or alternatively on a convex combination of these random graphs. The lower bounds that we obtain in the monotone setting are therefore merely worst-case, or average-case under a non-product distribution.

However, in the special case of $G = C_k$ and $\theta = 1$ (corresponding to the average-case k-CYCLE problem on $\boldsymbol{G}_{n,p}$ at the threshold $p = \Theta(1/n)$), we may take each $\boldsymbol{Y}^i$ to be the union of $n^{1/2-\delta}$ random paths of length k. In this case we are able to show that relations $\mathcal{P}_{f,H}^{\vec{\boldsymbol{X}}_\theta}$ are pathsets with respect to density parameter $\varepsilon = n^{1/2-2\delta}$. Moreover, random graphs $\boldsymbol{X}_\theta$ and $\boldsymbol{X}_\theta^m$ have total variation distance $o(1)$. As a result, we obtain an average-case lower bound for $\mathrm{SUB}(G)$ on $\boldsymbol{X}_\theta$ alone.

5.4 Pathset complexity

At this point, we are left with the task of proving lower bounds on the size of pathset formulas computing dense G-pathsets. This is by far the hardest part of the overall technique. Here we present only a brief outline. We introduce a family of complexity measures, each associated with different union family. However, rather than viewing a union family as a set of subgraphs of G, we explicitly consider the underlying tree structure.

Definition 5.4. A *union tree* A is a rooted binary tree whose leaves are labeled by edges of G. We denote by G_A the subgraph of G formed by the edges that label the leafs in A. We say that A is an H-*union tree* if $G_A = H$. For union trees A and B, let $\langle A, B\rangle$ denote the union tree consisting of a root attached to A and B (with $G_{\langle A,B\rangle} = G_A \cup G_B$). Notation $A \preceq B$ denotes that A is a subtree of B formed by a node of B together with all of its descendants.

Definition 5.5. *Pathset complexity* (with respect to $G, \theta, n, \varepsilon$) is the unique pointwise maximal family of functions $\chi_A : \{G_A\text{-pathsets}\} \to \mathbb{N}$, one for each union tree A, subject to the following inequalities:

- $\chi_A(\mathcal{A}) \le 1$ whenever A is a union tree of size 1,
- $\chi_A(\mathcal{A}) \le \sum_i \chi_A(\mathcal{A}_i)$ whenever $\mathcal{A} \subseteq \bigcup_i \mathcal{A}_i$,
- $\chi_A(\mathcal{A}) \le \max\{\chi_B(\mathcal{B}), \chi_C(\mathcal{C})\}$ whenever $A = \langle B, C\rangle$ and $\mathcal{A} \subseteq \mathcal{B} \bowtie \mathcal{C}$.

Pathset complexity gives lower bounds on pathset formula size (and by extension lower bounds on AC^0 formula size and monotone formula size). We describe the relationship between pathset formula size and pathset complexity in terms of a parameter $\tau_\theta(G)$, which plays an analogous role to $\kappa_\theta(G)$ in our formula lower bounds.

Definition 5.6. For each union tree A, let Φ_A be the maximum constant (depending on G and θ alone) such that the inequality $\chi_A(\mathcal{A}) \ge (1/\varepsilon)^{\Phi_A} \cdot \mu(\mathcal{A})$ holds for every G_A-pathset $\mathcal{A}$ and every setting of parameters n and ε. The invariant $\tau_\theta(G)$ is defined as the minimum value of Φ_A over G-union trees A.

For comparison, note that the invariant $\kappa_\theta(G)$ equals the minimum value of $\max_{A' \preceq A} \Delta_{A'}$ over G-union trees A, writing $\Delta_{A'}$ to abbreviate $\Delta_\theta(G_{A'})$. The constant Φ_A thus plays a similar role in our formula lower bounds as $\max_{A' \preceq A} \Delta_{A'}$ in our circuit lower bounds.

It follows from the above definitions, though not entirely straightforwardly, that any pathset formula F computing a .99-dense G-pathset (i.e., such that $\mu(\mathcal{A}_{f_{\mathrm{out}},G}) \ge .99$) must have size $\Omega((1/\varepsilon)^{\tau_\theta(G)})$. (This $\Omega(\cdot)$ hides a factor of $(1/2)^{2^{|E(G)|}}$, which arises from partitioning $\mathcal{A}_{f_{\mathrm{out}},G}$ according to a union tree that accounts for the construction of each of its elements in F.) Combined with the reduction outlined in §5.2, this implies the following lower bound, which is a version of Theorem 4.4 for AC^0 formulas.

Theorem 5.7 ([36])**.** *The average-case AC^0 formula size of* $\mathrm{SUB}(G)$ *on* $\boldsymbol{X}_\theta$ *is at least* $n^{\tau_\theta(G)-o(1)}$.

Using the reduction outlined in §5.3, we get the following lower bounds in the monotone setting.

Theorem 5.8 ([38])**.** *For all G and θ, the worst-case monotone formula (resp. circuit) size* $\mathrm{SUB}(G)$ *is at least* $n^{\tau_\theta(G)-o(1)}$ *(resp.* $n^{\kappa_\theta(G)-o(1)}$*). In the case of $G = C_k$ and $\theta = 1$, the average-case monotone formula size of* $\mathrm{SUB}(G)$ *on* $\boldsymbol{X}_\theta$ *is at least* $n^{\frac{1}{2}\tau_\theta(G)-o(1)}$.

It remains to prove lower bounds on $\tau_\theta(G)$, especially in cases of interest like $G = C_k$ and $\theta = 1$. This requires us to prove lower bounds on constants Φ_A for every possible G-union tree A. In principle, this is a problem in the realm of graph theory, since Φ_A depends on G and θ alone. Unfortunately, we do not have any nice expression for Φ_A, nor even an efficient method of computing these constants. Nevertheless, we are able to deduce some useful inequalities. For starters, it is simple to show that $\Phi_A \geq \Delta_A$ and moreover $\Phi_A \geq \Delta_{A'}$ for every $A' \preceq A$. However, this merely amounts to the inequality $\tau_\theta(G) \geq \kappa_\theta(G)$, which is the unsurprising fact that our formula lower bounds are not weaker than our circuit lower bounds.

To derive stronger lower bounds on Φ_A, we make use of structural properties of pathset complexity:

- **(projection lemma)** $\chi_{A'}(\mathrm{proj}_{A'}(\mathcal{A})) \leq \chi_A(\mathcal{A})$ for all union trees $A' \preceq A$ and every G_A-pathset $\mathcal{A}$, where $\mathrm{proj}_{A'}(\mathcal{A}) \subseteq [n]^{V(G_{A'})}$ is the projection of $\mathcal{A}$ to coordinates in $V(G_{A'})$,
- **(restriction lemma)** $\chi_{A\restriction H_1}(\mathcal{A} \,|\, \beta) \leq \chi_A(\mathcal{A})$ for every vertex-disjoint partition $G_A = H_1 \uplus H_2$ and $\beta \in [n]^{V(H_2)}$, where $A\restriction H_1$ is the union tree obtained from A by deleting every leaf that is labeled by an edge of H_2.

These lemmas allow us to derive two useful inequalities on constants Φ_A: for all union trees $A = \langle B, C\rangle$ and $B' \preceq B$ and $C' \preceq C$,

$$\Phi_A \geq \Phi_{B'} + \Delta_C + \Delta_{A\ominus C}, \tag{5.2}$$

$$\Phi_A \geq \tfrac{1}{2}(\Phi_{B'} + \Phi_{C'\ominus B'} + \Delta_A + \Delta_{A\ominus\langle B',C'\rangle}). \tag{5.3}$$

Here $\ominus$ is the following operation on union trees: $A \ominus B$ is the union tree obtained from A by deleting every leaf that is labeled an edge whose connected component in G_A contains any vertex of G_B.

In the case of $G = C_k$ and $\theta = 1$, inequalities (5.2) and (5.3) can be used to show that $\kappa_\theta(G) \geq \frac{1}{6}\log_2(k)$. This yields the following corollary of Theorems 5.7 and 5.8.

Corollary 5.9. *AC^0 formulas, as well as monotone formulas, which solve the average-case k-CYCLE problem on $\boldsymbol{G}_{n,p}$ at the threshold $p = \Theta(1/n)$ require size $n^{\Omega(\log k)}$.*

In unpublished work in progress, we explore an additional inequality on constant Φ_A. Consider *any* root-to-leaf branch in a union tree A, and let $A_1, \dots, A_m$ enumerate the union trees hanging off this branch in *any* order. For example, we might have $A = \langle A_3, \langle\langle A_1, \langle A_5, A_2\rangle\rangle, A_4\rangle\rangle$. For all such A and $A_1, \dots, A_m$, there is an inequality

$$\Phi_A \geq \Delta_{A_1} + \Delta_{A_2 \ominus A_1} + \Delta_{A_3 \ominus (A_1 \cup A_2)} + \cdots + \Delta_{A_m \ominus (A_1 \cup \cdots \cup A_{m-1})}. \tag{5.4}$$

Again in the case $G = C_k$ and $\theta = 1$, using (5.4) we can show that if A is a G-union tree with *left-depth* d (i.e., no root-to-leaf branch in A descends to the left more than d times), then $\Phi_A \geq \Omega(dk^{1/d}) - O(d)$. This in turn leads to nearly tight tradeoffs between the size and alternation-depth of AC^0 formulas solving the average-case k-CYCLE problem. Inequality (5.4) is also useful in bounding $\tau_\theta(G)$ for additional patterns of interest, such as complete binary trees.

5.5 Tree-depth

The *tree-depth* of a graph G, denoted $\mathsf{td}(G)$, is the minimum height of a forest F with the property that every edge of G connects a pair of vertices that have an ancestor-descendant relationship to each other in F (see [27]). Analogous to the relationship between tree-width and the circuit size, it turns out that SUB(G) is solvable by monotone AC^0 formulas of size $O(n^{\mathsf{td}(G)})$. Comparing this upper bound to the lower bound of Theorem 5.7, it follows that $\max_\theta \tau_\theta(G) \leq \mathsf{td}(G)$.

Using a recent result in graph minor theory of Kawarabayashi and the author [22], we are able to show that $\max_\theta \tau_\theta(G) > \mathsf{td}(G)^c$ for all patterns G where $c > 0$ is an absolute constant. The result of [22] reduces this inequality to three special cases when the pattern G is a grid, a path, or a complete binary tree. By bounding $\max_\theta \tau_\theta(G)$ in these three cases, we obtain an $\Omega(n^{\mathsf{td}(G)^c})$ lower bound on both the AC^0 and monotone formula size of SUB(G) for arbitrary patterns G.

Acknowledgements

The author's work is supported by NSERC and a Sloan Research Fellowship.

References

[1] Miklós Ajtai. Σ_1^1 formulae on finite structures. *Annals of Pure and Applied Logic*, 24:1–48, 1983.

[2] Noga Alon, Raphael Yuster, and Uri Zwick. Color-coding. *Journal of the ACM*, 42(4):844–856, 1995.

[3] Kazuyuki Amano. k-Subgraph isomorphism on AC^0 circuits. *Computational Complexity*, 19(2):183–210, 2010.

[4] Boaz Barak, Samuel B Hopkins, Jonathan Kelner, Pravesh Kothari, Ankur Moitra, and Aaron Potechin. A nearly tight sum-of-squares lower bound for the planted clique problem. In *57th IEEE Symposium on Foundations of Computer Science*, pages 428–437, 2016.

[5] Paul Beame. Lower bounds for recognizing small cliques on CRCW PRAM's. *Discrete Appl. Math.*, 29(1):3–20, 1990.

[6] Stuart J. Berkowitz. On some relationships between monotone and nonmonotone circuit complexity. Technical report, Department of Computer Science, University of Toronto, 1982.

[7] Béla Bollobás. Threshold functions for small subgraphs. *Math. Proc. Camb. Phil. Soc.*, 90:197–206, 1981.

[8] Ravi B Boppana. The average sensitivity of bounded-depth circuits. *Information Processing Letters*, 63(5):257–261, 1997.

[9] Jianer Chen, Benny Chor, Mike Fellows, Xiuzhen Huang, David Juedes, Iyad A Kanj, and Ge Xia. Tight lower bounds for certain parameterized np-hard problems. *Information and Computation*, 201(2):216–231, 2005.

[10] Vitaly Feldman, Elena Grigorescu, Lev Reyzin, Santosh Vempala, and Ying Xiao. Statistical algorithms and a lower bound for detecting planted cliques. In *45th ACM Symposium on Theory of Computing*, pages 655–664, 2013.

[11] Merrick L. Furst, James B. Saxe, and Michael Sipser. Parity, circuits, and the polynomial-time hierarchy. *Mathematical Systems Theory*, 17:13–27, 1984.

[12] Michelangelo Grigni and Michael Sipser. Monotone complexity. *Boolean function complexity*, 169:57–75, 1992.

[13] Martin Grohe and Dániel Marx. On tree width, bramble size, and expansion. *Journal of Combinatorial Theory, Series B*, 99(1):218–228, 2009.

[14] Johan Håstad. Almost optimal lower bounds for small depth circuits. In *18th ACM Symposium on Theory of Computing*, pages 6–20, 1986.

[15] Johan Håstad. The shrinkage exponent of de Morgan formulas is 2. *SIAM Journal on Computing*, 27(1):48–64, 1998.

[16] Svante Janson. Poisson approximation for large deviations. *Random Structures and Algorithms*, 1(2):221–229, 1990.

[17] Mark Jerrum. Large cliques elude the metropolis process. *Random Structures and Algorithms*, 3(4):347–359, 1992.

[18] Ari Juels and Marcus Peinado. Hiding cliques for cryptographic security. *Des. Codes Cryptography*, 20(3):269–280, 2000.

[19] Mauricio Karchmer and Avi Wigderson. Monotone circuits for connectivity require super-logarithmic depth. *SIAM Journal on Discrete Mathematics*, 3(2):255–265, 1990.

[20] Richard M. Karp. Reducibility among combinatorial problems. In *Complexity of computer computations*, pages 85–103. Springer, 1972.

[21] Richard M. Karp. The probabilistic analysis of some combinatorial search algorithms. *Algorithms and complexity: New directions and recent results*, 1:19, 1976.

[22] Ken-ichi Kawarabayashi and Benjamin Rossman. A polynomial excluded-minor characterization of treedepth. In *29th Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 234–246, 2018.

[23] François Le Gall. Powers of tensors and fast matrix multiplication. In *Proceedings of the 39th International Symposium on Symbolic and Algebraic Computation*, pages 296–303. ACM, 2014.

[24] Yuan Li, Alexander A. Razborov, and Benjamin Rossman. On the AC^0 complexity of subgraph isomorphism. In *55th IEEE Symposium on Foundations of Computer Science*, pages 344–353, 2014.

[25] Dániel Marx. Can you beat treewidth? *Theory of Computing*, 6:85–112, 2010.

[26] Dániel Marx and Michał Pilipczuk. Everything you always wanted to know about the parameterized complexity of subgraph isomorphism (but were afraid to ask). In *31st International Symposium on Theoretical Aspects of Computer Science*, page 542, 2014.

[27] Jaroslav Nešetřil and Patrice Ossona de Mendez. Tree depth, subgraph coloring and homomorphism bounds. *European J. Combin.*, 27(6):1022–1041, 2006.

[28] Jaroslav Nešetřil and Svatopluk Poljak. On the complexity of the subgraph problem. *Comment. Math. Univ. Carolinae.*, 26(2):415–419, 1985.

[29] Toniann Pitassi and Robert Robere. Strongly exponential lower bounds for monotone computation. In *Proceedings of the 49th Annual ACM SIGACT Symposium on Theory of Computing*, pages 1246–1255. ACM, 2017.

[30] Jürgen Plehn and Bernd Voigt. Finding minimally weighted subgraphs. In *International Workshop on Graph-Theoretic Concepts in Computer Science*, pages 18–29. Springer, 1990.

[31] Aaron Potechin. Bounds on monotone switching networks for directed connectivity. In *51st IEEE Symposium on Foundations of Computer Science*, pages 553–562, 2010.

[32] Ran Raz and Pierre McKenzie. Separation of the monotone NC hierarchy. In *38th Annual Symposium on Foundations of Computer Science*, pages 234–243. IEEE, 1997.

[33] Alexander A Razborov. Lower bounds on the monotone complexity of some boolean functions. *Dokl. Akad. Nauk SSSR*, 281(4):798–801, 1985.

[34] Neil Robertson and Paul D Seymour. Graph minors. X. Obstructions to tree-decomposition. *Journal of Combinatorial Theory, Series B*, 52(2):153–190, 1991.

[35] Benjamin Rossman. On the constant-depth complexity of k-clique. In *40th ACM Symposium on Theory of Computing*, pages 721–730, 2008.

[36] Benjamin Rossman. Formulas vs. circuits for small distance connectivity. In *46th ACM Symposium on Theory of Computing*, pages 203–212, 2014.

[37] Benjamin Rossman. The monotone complexity of k-clique on random graphs. *SIAM Journal on Computing*, 43(1):256–279, 2014.

[38] Benjamin Rossman. Correlation bounds against monotone NC^1. In *LIPIcs-Leibniz International Proceedings in Informatics*, volume 33. Schloss Dagstuhl-Leibniz-Zentrum fuer Informatik, 2015.

[39] Claude Shannon et al. The synthesis of two-terminal switching circuits. *Bell System Technical Journal*, 28(1):59–98, 1949.

[40] P.M. Spira. On time-hardware complexity tradeoffs for boolean functions. In *4th Hawaii Symposium on System Sciences*, pages 525–527, 1971.

[41] Avishay Tal. Shrinkage of De Morgan formulae by spectral techniques. In *55th IEEE Foundations of Computer Science*, pages 551–560, 2014.

[42] Éva Tardos. The gap between monotone and non-monotone circuit complexity is exponential. *Combinatorica*, 8(1):141–142, 1988.