Lower Bounds for Subgraph Isomorphism

Benjamin Rossman

January 7, 2018

Abstract

We consider the computational complexity 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. We give an overview of unconditional lower bounds for this problem, focusing on circuits and formulas in the $AC^0$ and monotone settings.

1 Background and preliminaries

The subgraph isomorphism problem is the computational task of determining whether a “host” graph $H$ contains a copy of a “pattern” $G$ as a subgraph. When both $G$ and $H$ are given as input, this is a classic $NP$-complete problem which generalizes both MAXIMUM CLIQUE and HAMILTONIAN CYCLE [21]. We refer to the $G$-subgraph isomorphism problem in the setting whether the pattern $G$ is fixed and $H$ alone is given as input. This family of problems, which includes the $k$-CLIQUE and $k$-CYCLE problems for fixed $k$, are each solvable in polynomial time $O(n^{|V(G)|})$ by the obvious exhaustive search.\(^1\) Faster algorithms are known in many cases. For example, $k$-CYCLE has an $O(n^\omega)$ time algorithm where $\omega < 2.38$ is the exponent of matrix multiplication [2], while $k$-CLIQUE is solvable in $O(n^{\omega[k/3]})$ time [27]. Additional upper bounds are based on invariants of $G$ and $H$ (see [25] for a survey), such as an $O(n^{k+1})$ time algorithm for patterns of tree-width $k$ [28].

The focus of this article is lower bounds: impossibility results which show, unconditionally, that the $G$-subgraph isomorphism problem cannot be solved with insufficient computational resources, such as time or space. Conditionally, it is known that the Exponential Time Hypothesis implies that $k$-CLIQUE requires time $n^{\Omega(k)}$ [9] and that $G$-subgraph isomorphism

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

1
for patterns of tree-width \( k \) requires time \( n^{\Omega(k/\log k)} \) [24]. Proving such lower bounds unconditionally would separate \( P \) from \( NP \), moreover in a very strong way. We shall therefore restrict our attention to complexity measures weaker than sequential time (namely, circuit and formula size in the \( \mathcal{AC^0} \) and monotone settings, to be defined in §1.1).

The average-case setting. The lower bounds for \( G \)-subgraph isomorphism described in this article are obtained in the natural average-case setting where the input is an Erdős-Rényi graph \( G_{n,p} \) (or \( G \)-colored version thereof). Recall that \( G_{n,p} \) is the random \( n \)-vertex graph that includes each potential edge independently with probability \( p \). For many patterns of interest such as cliques and cycles, \( G_{n,p} \) 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 often 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.

Consider the average-case \( k \)-clique problem of determining, asymptotically almost surely\(^2\) correctly, whether or not \( G_{n,p} \) at the threshold \( p = \Theta(n^{-2/(k-1)}) \) contains a \( k \)-clique (i.e., complete subgraph of order \( k \)). One approach to this problem involves the following randomized greedy algorithm: start with a uniform random vertex \( v_1 \), then select a uniform random neighbor \( v_2 \) of \( v_1 \), then a common neighbor \( v_3 \) of \( v_1, v_2 \), and so on until reaching a maximal (though not necessarily maximum) clique of the input graph. A single run of the greedy algorithm on \( G_{n,p} \) a.a.s. produces a clique of size \( \lfloor \frac{k}{2} \rfloor \) or \( \lceil \frac{k}{2} \rceil \). To find a clique of size \( \lfloor (1+\varepsilon)\frac{k}{2} \rfloor \) where \( \varepsilon < 1 \), it suffices to run the greedy algorithm \( n^{\varepsilon^2 k/4} \) times, while \( n^{k/4+O(1/k)} \) runs suffice to find a \( k \)-clique in \( G_{n,p} \) if any exists. Since each iteration of the greedy algorithm takes linear time, this approach solves the average-case \( k \)-clique problem in \( n^{k/4+O(1)} \) time.

It is an open question whether this greedy approach is optimal, that is, whether \( \Omega(n^{k/4}) \) is a lower bound on the average-case complexity of \( k \)-clique. This may be seen a scaled-down version of a famous open question of Karp [22] concerning the uniform random graph \( G_{n,1/2} \), which has expected maximum clique size \( \approx 2 \log n \), whereas the greedy algorithm a.a.s. outputs 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

---

\(^2\)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 \( n \to \infty \).
constant $\varepsilon > 0$. Karp’s question, as well as a variant where $G_{n,1/2}$ is augmented by a large planted clique, have stimulated a great deal of research in theoretical computer science. The hardness of detecting planted cliques has been used as a cryptographic assumption [19], while lower bounds have been shown against specific algorithms, such as the metropolis process [18], the sum-of-squares semi-definite programming hierarchy [4] and statistical query algorithms [10].

The $k$-cycle problem is another example where $G_{n,p}$ at the threshold $p = \Theta(1/n)$ is believed to be a source of hard-on-average instances. The average-case $k$-cycle problem has relatively low complexity: it is solvable a.a.s. on $G_{n,p}$ in deterministic $O(\log n)$ space (independent of $k$). Nevertheless, $G_{n,p}$ is believed to be hard-on-average with respect to a weaker complexity measure: Boolean formula size. The smallest known formulas solving $k$-cycle have size $n^{O(\log k)}$ and this is believed to be optimal even in the average-case. Proving such a lower bound unconditionally would separate complexity classes $NC^1$ and $L$.

1.1 Circuit complexity

Circuit complexity is the quest for unconditional lower bounds in combinatorial models of computation. Among such models, Boolean circuits (acyclic networks of $\land$, $\lor$ 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 decision problem in $NP$, as represented by a sequence of Boolean functions $\{0,1\}^n \rightarrow \{0,1\}$.

Claude Shannon in 1949 showed that almost all Boolean functions require circuits of exponential size [36]. 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 $\land$ or $\lor$. Every $n$-variable circuit computes a Boolean function $\{0,1\}^n \rightarrow \{0,1\}$ 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 $\{0,1\}^n \rightarrow \{0,1\}$ computed by $n$-variable circuits of polynomial size (i.e., $O(n^c)$ for any constant $c$). (Imposing a uniformity condition on the sequence of $n$-variable circuits results in the more familiar class $P$.)

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

The alternation-depth of a circuit is the maximum number of alternations between $\land$ and $\lor$ 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]. 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 and natural restricted class of circuits are formulas, that is, circuits with the structure of a tree (i.e., in which every non-output node has out-degree 1). In the context of formulas, size and depth are essentially the same complexity measure, as it is known that every formula of size $s$ can be transformed to an equivalent formula of depth $O(\log s)$ [37]. As a corollary, $NC^1$ is equivalent to the class of Boolean functions computed by polynomial-size formulas.

Intuitively, 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, 38]. 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 a major frontier in circuit complexity.

The monotone setting. A Boolean function $f$ is said to be monotone if $f(x) \leq f(y)$ whenever $x_i \leq y_i$ for all coordinates $i$. The $G$-subgraph isomorphism problem is monotone when viewed as a sequence of functions

\(^3\) $AC^0$ is equivalently defined in terms of constant depth circuits with AND and OR gates of unbounded in-degree. In this article, we adopt the nonstandard definition in terms of alternation-depth, since simplest version of our lower bounds naturally applies to binary $\land$ and $\lor$ gates.
\(\{0,1\}_2^G \rightarrow \{0,1\}\). Monotonicity is also a syntactic property of circuits: a circuit is said to be monotone if it has no negations (i.e., inputs are labeled by positive literals only).

Monotonicity has been an extremely fruitful restriction in circuit complexity, beginning with celebrated results in the 1980’s. In the paper [30] which introduced the sunflower-plucking approximation method, Razborov showed that \(k\text{-CLIQUE}\) has monotone circuit size \(\Omega(n^k/(\log n)^{2k})\) for any constant \(k\). By an entirely different technique based on communication complexity, Karchmer and Wigderson [20] proved an \(n^{\Omega(\log k)}\) lower bound on the monotone formula size of \(\text{DISTANCE-}k\text{ ST-CONNECTIVITY}\), a problem which is equivalent to \(k\text{-CYCLE}\) up to a polynomial factor. These results and many others (such as [12, 29] to name just two) 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 it is open whether \(TC^0\) (the class of constant-depth threshold circuits) equals \(NP\) in the non-monotone world.

However, despite the extensive knowledge of lower bounds, it is unclear if any of the techniques developed in the monotone setting has the potential to extend to the non-monotone world. Pretty much every monotone lower bound technique works by pitting a class of sparse 1-inputs (e.g., isolated \(k\)-cliques or st-paths) against a class of dense 0-inputs (complete \(k - 1\)-partite graphs or st-cuts). Note that any technique subject to this “Hamming weight gap” is useless against non-monotone classes as weak as \(TC^0\), since such 0- and 1-inputs are separable by the function \(x \mapsto 1_{|x| < t}\) for a suitable threshold \(t\).

This observation motivates the challenge of proving average-case lower bounds under product distributions in the monotone setting, in particular for natural problems like \(k\text{-CLIQUE}\) and \(k\text{-CYCLE}\) on \(G_{n,p}\) where the “Hamming weight gap” does not apply. This challenge can be seen as a step toward non-monotone lower bounds. We remark that monotone and non-monotone complexity coincide (up to a polynomial factor) on slice distributions [6], such as the random graph \(G_{n,m}\) with exactly \(m\) edges.

Outline of the article. This article gives an overview of lower bounds (and a few upper bounds) from papers [32, 33, 34, 35] of the author, [3] of Kazuyuki Amano, and joint work [23] of the author with Yuan Li and Alexander Razborov and [16] with Ken-ichi Kawarabayashi. These results 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 lower bound technique originated in [32], where it is shown that $AC^0$ circuits solving the average-case $k$-CLIQUE problem require size $\Omega(n^{k/4})$, establishing optimality of the greedy algorithm in the $AC^0$ setting. This result improved the previous $\Omega(n^{k/89d^2})$ lower bound of Beame [5] for circuits of alternation-depth $d$, in a significant way, by removing dependence on the parameter $d$ up to $O(\log n/\log \log n)$. (Beyond this point, a key lemma in the proof (Lemma 3.2) breaks down, although the $\Omega(n^{k/4})$ lower bound is conjectured to hold for arbitrary $d$.) This $k$-CLIQUE lower bound was generalized by Amano [3] to arbitrary patterns, as well as hypergraphs. Subsequent work of Li, Razborov and the author [23] further extended the technique to a colored variant of the $G$-subgraph isomorphism problem, obtaining an $n^{\Omega(k/\log k)}$ lower bound for patterns of tree-width $k$. The results of [23] are presented in §2-4 of this article.

The challenge of proving stronger lower bounds for formulas vis-à-vis circuits was addressed in [33], 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 has $AC^0$ 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) follows from a recent result in graph minor theory [16].

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 [34, 35]. For most patterns $G$, these lower bounds are only average-case with respect to a non-product distribution (a convex combination of $G_{n,p}$ and $G_{n,p+O(p)}$). However, in the special case of the $k$-CYCLE problem, the technique produces an average-case lower bound under $G_{n,p}$. This is significant for being the first super-polynomial lower bound against monotone formulas under any product distribution. In §5 we describe the pathset complexity framework underlying this result. It is hoped that this framework might be a viable approach to eventually proving a super-polynomial lower bound for unrestricted Boolean formulas.

2 Colored $G$-subgraph isomorphism

The actual target problem for our lower bounds will be a colored variant of the $G$-subgraph isomorphism problem, denoted SUB($G$). Here the input
is a “G-colored graph” $X$ on vertices $V(G) \times \{1, \ldots, n\}$ and the task to determine whether $X$ contains a copy of the pattern $G$ involving one vertex from each color class. Compared with the standard uncolored $G$-subgraph isomorphism problem, which we denote by $\text{SUB}_{\text{uncol}}(G)$, the colored variant is better structured and admits a richer class of threshold distributions. All average-case lower bounds for $\text{SUB}(G)$ in this article extend to the average-case $\text{SUB}_{\text{uncol}}(G)$ 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 $E(G) \subseteq (V(G))^2$ 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 $\{v,w\} \in E(G) \iff \{\pi(v),\pi(w)\} \in E(G')$ for all distinct vertices $v, w$ of $G$.

The $n$-blowup of a graph $G$, denoted $G^{\uparrow n}$, has vertices $v^{(1)}, \ldots, v^{(n)}$ for each $v \in V(G)$ and edges $\{v^{(a)}, w^{(b)}\}$ for each $\{v, w\} \in E(G)$ and $a, b \in [n]$ ($:= \{1, \ldots, n\}$). 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 $\text{SUB}(G)$ 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 to $G$ 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 $\text{sub}_H(X)$ denote the number of $H$-subgraphs of $X$, that is, $\text{sub}_H(X) := |\{\alpha \in [n]^{V(H)} : H^{(\alpha)} \subseteq X\}|$.

**On the relationship between $\text{SUB}_{\text{uncol}}(G)$ and $\text{SUB}(G)$.** For every pattern $G$, the color-coding method of [2] provides an efficient many-one reduction from $\text{SUB}_{\text{uncol}}(G)$ to $\text{SUB}(G)$. 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 $\text{SUB}(G)$ to $\text{SUB}_{\text{uncol}}(G)$,
as the only subgraphs of $G^r_n$ 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 $\text{SUB}(G)$, 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 $G_{n,\vec{p}}$ denote the random subgraph of $G^r_n$ 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 $G_{n,\vec{p}}$ has a characterization in terms of certain edge-weightings on $G$.

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

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

We say that $\theta$ 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 $\theta$ 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 $\theta : E(G) \rightarrow [0,2]$ defined by $\theta(\{v,w\}) := \frac{1}{\deg(v)} + \frac{1}{\deg(w)}$ is a threshold weighting. In particular, in $G$ is $r$-regular, then the constant function $\theta \equiv \frac{2}{r}$ is a threshold weighting. (Two additional constructions of threshold weightings are described at the end of this section.)

**Definition 2.3 (The random graph $X_\theta$).** Every threshold weighting $\theta$ on $G$ gives rise to a sequence of random graphs $X_{n,\theta}$, defined as the $G$-colored Erdős-Rényi graph $G_{n,\vec{p}}$ where $\vec{p} \in [0,1]^{E(G)}$ is the vector of edge probabilities $p_e = n^{-\theta(e)}$. That is, $X_{n,\theta}$ is the random subgraph of $G^r_n$ 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 $X_\theta$. 
Observe that the function $\Delta_\theta$ characterizes the expected number of $H$-subgraphs in $X_\theta$: for every $H \subseteq G$, we have $E[\text{sub}_H(X_\theta)] = n^{\Delta_\theta(H)}$ by linearity of expectation. In particular, $\text{sub}_G(X_\theta)$ has expectation 1 (since $\Delta_\theta(G) = 0$). Moreover, when $\theta$ is strict, $\text{sub}_G(X_\theta)$ is asymptotically Poisson and $\text{sub}_H(X_\theta)$ is highly concentrated around its mean for all proper subgraphs $H \subset G$.

**Proposition 2.4.** For every graph $G$ and threshold weighting $\theta$, the probability that $X_\theta$ contains a $G$-subgraph converges to a limit in $(0,1)$. When $\theta$ 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 $\text{SUB}(G)$ on $X_\theta$, that is, the complexity of functions $f : \{\text{subgraphs of } G \} \rightarrow \{0,1\}$ such that $f(X_\theta) = 1 \iff \text{sub}_G(X_\theta) \geq 1$ 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 $M : V(G) \times V(G) \rightarrow [0,1]$ 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 $v \in V(G)$.

Then the function $E(G) \rightarrow [0,2]$ 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 $M(v,w) = \frac{1}{\text{deg}(v)}$ for every $\{v,w\} \in E(G)$.) The associated function $\Delta_M$ 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 $\text{SUB}(G)$ 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 $G_{n,p}$ 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 $\text{SUB}_{\text{ancol}}(G)$ on $G_{n,p}$ is equivalent to the average-case analysis of $\text{SUB}(G)$ on $X_{\theta_{\text{ancol}}}$ where $\theta_{\text{ancol}} : E(G) \rightarrow \{0,c\}$ is the threshold weighting defined by $\theta_{\text{ancol}}(e) = c \iff$ 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 $\left(\frac{n}{|V(G)|}\right)$.
3 \( H \)-subgraph sensitivity

The fundamental weakness of \( AC^0 \) functions is low average sensitivity. More specifically, for an \( AC^0 \) function \( f : \{0,1\}^n \to \{0,1\} \) and independent uniform random \( x \in \{0,1\}^n \) and \( i \in [n] \), it is known that \( \Pr_{x,x'}[f(x) \neq f(x \text{ with its } i^{th} \text{ bit flipped})] \leq n^{-1+o(1)} \) [8]. By analogy, a key lemma in our lower bounds shows that \( AC^0 \)-computable functions \( f : \{\text{subgraphs of } G^{\uparrow n}\} \to \{0,1\} \) have what might be termed “low average \( H \)-subgraph sensitivity of \( f \) on \( X_\theta \)”. This notion involves the randomly restricted function \( f^{\cup X_\theta}|_{H(\alpha)} \) (defined below) where the uniform random \( \alpha \in [n]^{V(H)} \) plays the role of \( i \in [n] \).

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 \( f|_H \in \mathbb{B}(H) \) denote the restriction of \( f \) to domain \( \{\text{subgraphs of } H\} \).

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

Lemma 3.2 (\( H \)-subgraph sensitivity of \( AC^0 \) functions [23]). Suppose \( f \in \mathbb{B}(G^{\uparrow n}) \) is an \( AC^0 \)-computable sequence of functions. Then for every subgraph \( H \subseteq G \),

\[
\Pr_{X_\theta, \alpha \in [n]^{V(H)}}[f^{\cup X_\theta}|_{H(\alpha)} \text{ depends on all coordinates}] \leq n^{-\Delta_\theta(H)+o(1)}.
\]
When $\Delta_\theta(H) > 0$, the $n^{-\Delta_\theta(H)+o(1)}$ bound of Lemma 3.2 is nontrivial and moreover tight. However, note that the 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 [17], which gives lower tail bounds for random variables $\text{sub}_H(X_\theta)$. The assumption that $f$ is $AC^0$-computable is necessary, as for instance if $f$ is the parity function (mapping $X \subseteq G^{tn}$ to $|E(X)| \mod 2$), then the restricted function $f \cup X_\theta |_{H(\alpha)}$ depends on all coordinates with probability 1. We remark that in the case where $H$ is a single-edge subgraph of $G$, Lemma 3.2 essentially equivalent to aforementioned bounds 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 $X_\theta$” in a different sense. Namely, we consider the event that $H(\alpha)$ is a common minterm of $f$ and $f \cup X_\theta$ (i.e., $f(H(\alpha)) = 1$ and $f \cup X_\theta(H(\alpha) - e) = 0$ for every $e \in E(H(\alpha))$).

**Lemma 3.3** ($H$-subgraph sensitivity of monotone functions [35]). For every monotone function $f \in \mathbb{B}(G^{tn})$ and subgraph $H \subseteq G$,

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

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 which match our lower bounds in $AC^0$ setting.

### 4 The $AC^0$ circuit size of SUB$(G)$

This section presents results of Li, Razborov and the author [23], which characterize the average-case $AC^0$ circuit size of $\text{SUB}(G)$ on $X_\theta$ for any $G$ and $\theta$ 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 $F_1, F_2 \subseteq F$ 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 $\theta$ 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 $\theta$ be the constant threshold weighting $\frac{k}{k-1}$. 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., $K_{\{1,\ldots,j\}}$ minus the edge $\{1,j\}$ is the union of $K_{\{1,\ldots,j-1\}}$ and $K_{\{2,\ldots,j\}}$). 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 $\left\lceil \frac{k}{2} \right\rceil$ 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 $\left\lceil \frac{k}{2} \right\rceil$ 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 $\theta \equiv \frac{2}{r}$, 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 [23] and outline its proof.

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

Note that the lower and upper bounds of Theorem 4.4 are separated by a quadratic gap (modulo the $n^{O(1)}$ factor). It is an open question to close this gap, for example, by improving the upper bound to $n^{\kappa(G) + O(1)}$. By Examples 2.6 and 4.2, Theorem 4.4 implies a lower bound of $\Omega(n^{k/4})$ on the $AC^0$ circuit size of the average-case $k$-CLIQUE problem on $G_{n,p}$ at the threshold $p = \Theta(n^{-2/(k-1)})$. 

12
The upper bound. We give a high-level description of an algorithm that solves \( \text{SUB}(G) \) a.a.s. on \( X_\theta \) 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, \( \text{sub}_H(X_\theta) \) is at most \( n^{\Delta_\theta(H)+o(1)} \) for all \( H \subseteq G \) (by Markov’s inequality).

Fix an optimal union family \( F \) such that \( \kappa_\theta(G) = \max_{F \in F} \Delta_\theta(F) \). Also fix an enumeration \( F_1, \ldots, F_m \) of graphs in \( F \) 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 \( X_\theta \). 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 \( F_\alpha \subseteq X_\theta \) and \( F_\beta \subseteq X_\theta \) from the previously compiled lists: if \( \alpha_v = \beta_v \) for all \( v \in V(F_i) \cap V(F_j) \), then \( F_k \) is added to the list of \( F_k \)-subgraphs. Compiling this list therefore takes time \( O(\text{sub}_{F_i}(X_\theta) \cdot \text{sub}_{F_j}(X_\theta)) \), 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 \( X_\theta \) 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 solve \( \text{SUB}(G) \) a.a.s. on \( X_\theta \). We consider the randomly restricted function \( f \cup X_\theta |_{G(\alpha)} \) where \( \alpha \) is a uniform random element of \( [n]^V(G) \) independent of \( X_\theta \). We will show that

\[
\text{Pr}[ f \cup X_\theta |_{G(\alpha)} \text{ depends on all coordinates } ] = o(1).
\]

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 solves \( \text{SUB}(G) \) a.a.s. on \( X_\theta \) satisfies

\[
\text{Pr}[ f \cup X_\theta |_{G(\alpha)} \text{ depends on all coordinates } ] = \Omega(1).
\]

Since (4.1) and (4.2) are contradictory for sufficiently large \( n \), we conclude that functions \( f \) computed by \( C \) do not solve \( \text{SUB}(G) \) a.a.s. on \( X_\theta \).
solves SUB( variation distance between random graphs $X_f$ on all coordinates, inequality 4.2 follows. (When we merely assume that $X$ and $H$ is the $H$ which is a uniform random $X$ that is nonempty $X \subseteq \theta(X)$ does not depend on all coordinates. Let $f$ for every $H$ which is a hitting family $G \subseteq \theta \cup \alpha$. The claim then follows from the fact that $F \cap H = \emptyset$ since $\kappa_\theta(G) = \min_{H \in H} \Delta_\theta(H)$. Taking a union bound over $g \in G$ and $H \in H$, we have

\begin{equation}
\Pr[ (\exists g \in G)(\exists H \in H) \text{ } g^{\cup X}_\theta \downarrow_{H^{(\alpha)}} \text{ depends on all coordinates } ] \leq |G| \cdot |H| \cdot n^{-\kappa_\theta(G) + o(1)} = o(1)
\end{equation}

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

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

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

Now consider any $F \in F$ with $\geq 2$ edges. It remains to show that $F$ is the union of two proper subgraphs which belong to $F$. By definition of $F$
and \( G \), there exists a function \( g \in \mathbb{B}(G^n) \) computed by a subcircuit of \( C \) such that \( g^{\cup X} | _{F^{(a)}} \) depends on all coordinates. Fix a choice of \( g \) computed by a subcircuit of minimal depth in \( C \). Since \( g^{\cup X} | _{F^{(a)}} \) depends on \( \geq 2 \) coordinates (namely all edges of \( F^{(a)} \)), 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 \( g_1 \land g_2 \) or \( g_1 \lor g_2 \).

For \( i = 1, 2 \), let \( F_i \) be the graph consisting of edges \{\( v, w \)\} \( \in E(F) \) such that \( g_i^{\cup X} | _{F^{(a)}} \) depends on the corresponding edge \{\( v^{(\alpha_v)}, w^{(\alpha_w)} \)\} \( \in E(F^{(a)}) \). Observe that the function \( g_i^{\cup X} | _{F^{(a)}} \in \mathbb{B}(F^{(a)}) \) 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 \( ^\cup X | _{F^{(a)}} \) 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.

By essentially the same argument, we obtain a similar \( n^{\kappa_{\theta}(G) - o(1)} \) lower bound on the monotone circuit size of \( \text{SUB}(G) \). 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, the lower bound we obtain in the monotone setting is only worst-case (or average-case with respect to a non-product distribution).

### 4.1 Tree-width

For every pattern \( G \), there exist monotone \( AC^0 \) circuits of size \( O(n^{\text{tw}(G)} + 1) \) which solve \( \text{SUB}(G) \) in the worst-case, where \( \text{tw}(\cdot) \) denotes tree-width. Comparing this upper bound to the lower bound of Theorem 4.4, we see that \( \max_{\theta} \kappa_{\theta}(G) \leq \text{tw}(G) + 1 \). The next proposition shows that this inequality is nearly tight.

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

**Proof.** We rely on a result of [13] showing 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 \( \Delta_M(H) = \Omega(\frac{|W|}{k\log k}) \) 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 \( \kappa_M(G) \geq \min_{H \in \mathcal{H}} \Delta_M(H) = \Omega \left( \frac{k}{\log k} \right) \) with respect to the threshold weighting \( \{v, w\} \mapsto M(v, w) + M(w, v) \) induced by \( M \).

It is conjectured in [23] that \( \max_\theta \kappa_\theta(G) = \Omega(\text{tw}(G)) \) for all graphs \( G \).

As for the upper bound \( \max_\theta \kappa_\theta(G) \leq \text{tw}(G) + 1 \), it can be directly shown (without appealing to Theorem 4.4) that \( \max_\theta \kappa_\theta(G) \) is at most the branch-width of \( G \), an invariant that is related to tree-width by \( \text{bw}(G) \leq \text{tw}(G) + 1 \leq \frac{3}{2} \text{bw}(G) \) [31].

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

**Proof.** Branch-width has a characterization in terms of union families:

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

Here \textit{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 easily from the definition of threshold weighting that \( \Delta_\theta(F) \leq \Delta_\theta(F) + \Delta_\theta(\overline{F}) = |V(F) \cap V(\overline{F})| \) for every threshold weighting \( \theta \) and subgraph \( F \subseteq G \). Therefore, \( \kappa_\theta(G) = \min_{\text{union families } \mathcal{F}} \max_{F \in \mathcal{F}} \Delta_\theta(F) \leq \text{bw}(G) \). \( \square \)

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

In this section we sketch an extension the lower bound technique which produces stronger lower bounds for formulas vis-à-vis circuits in both the \( AC^0 \) and monotone settings. This is significant for patterns with constant tree-width, such as paths and cycles, which are computable small circuits but are believed to require large formulas.

A brief outline of this section: §5.1 introduces the key notion of pathsets \( \mathcal{A} \subseteq [n]^{V(H)} \) which satisfy certain density constraints (related to the bounds on “\( H \)-subgraph sensitivity” given by Lemma 3.2 and 3.3). We then define pathset formulas, which are a tree-like model for constructing pathsets. §5.2 presents a reduction which transforms any \( AC^0 \) formula that solves average-case \( \text{SUB}(G) \) on \( X_\theta \) to a pathset formula that computes a dense subset of \( [n]^{V(G)} \). §5.3 outlines a similar transformation for monotone formulas that solve \( \text{SUB}(G) \) in a weaker average-case sense. §5.4 outlines the combinatorial heart of the technique: an \( n^{\tau_\theta(G) - o(1)} \) lower bound on the size of pathset formulas that 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 setting of formulas. Although \( \tau_\theta(G) \) is 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 \S 5.5 we describe 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 \( \theta \), as well as \( n \in \mathbb{N} \) and an arbitrary “density parameter” \( \varepsilon \in [0, 1] \). (In our applications, we take \( \varepsilon \) to be \( n^{1-o(1)} \) and later \( n^{1/2-o(1)} \).)

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

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

For \( S \subseteq V \) and \( \beta \in [n]^S \), the *conditional density* of \( A \) on \( \beta \) is defined by

\[
\mu(A|\beta) := \Pr_{\alpha \in [n]^V} [\alpha \in A | \alpha_S = \beta].
\]

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

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

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

Here the pair \( H_1, H_2 \) range over vertex-disjoint partitions of \( H \) (which satisfy \( H_1 \cup H_2 = H \) and \( V(H_1) \cap V(H_2) = \emptyset \)), of which there are \( 2^t \) possibilities if \( H \) has \( t \) connected components.

When \( H_1 = H \) and \( H_2 \) is the empty graph, note that (5.1) specializes to \( \mu(A) \leq \varepsilon \Delta_\theta(H) \). When \( H_1 \) is the empty graph, (5.1) is vacuous (since \( \Delta_\theta(\emptyset) = 0 \)). In the case that \( H \) is a connected subgraph of \( G \), we see that a relation \( A \subseteq [n]^{V(H)} \) is an \( H \)-pathset if and only if \( \mu(A) \leq \varepsilon \Delta_\theta(H) \). Note that every relation \( A \subseteq [n]^{V(G)} \) 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 (w.r.t. \( G, \theta, n, \varepsilon \)) is a rooted binary tree \( F \) together with an indexed family of relations \( \{A_{f,H} \subseteq [n]^{V(H)}\}_{f \in V(F), H \subseteq G} \) subject to three conditions:
(i) \( A_{f,H} \) is a \( H \)-pathset,

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

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

\[
A_{f,H} \subseteq \bigcup_{H_1, H_2 \subseteq H : H_1 \cup H_2 = H} (A_{f_1,H_1} \Join A_{f_2,H_2}).
\]

We view \( F \) as “computing” the family of pathsets \( \{ A_{f_\text{out},H} \}_{H \subseteq G} \) (and in particular the \( G \)-pathset \( A_{f_\text{out},G} \)) where \( f_\text{out} \) is the root of \( F \).

5.2 Transforming \( AC^0 \) formulas to pathset formulas

For any Boolean function \( f \in \mathbb{B}(G^n) \) and a subgraph \( H \subseteq G \), let \( A_{X_\theta}^{f,H} \subseteq [n]^{V(H)} \) be the random relation defined by

\[
A_{X_\theta}^{f,H} := \{ \alpha \in [n]^{V(H)} : f \cup X_\theta \mid_{H(\alpha)} \text{ depends on all coordinates} \}.
\]

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

If \( F \) is an \( AC^0 \) formula, it follows that the family of relations \( A_{X_\theta}^{f,H} \subseteq [n]^{V(H)} \) (indexed by subformulas \( f \) of \( F \) and subgraphs \( H \subseteq G \) a.a.s. constitutes a pathset formula. Condition (i) of Def. 5.3 holds a.a.s. by taking a union bound, over the \( n^{O(1)} \) pairs of \( f \) and \( H \), of the exponentially small probability that \( A_{X_\theta}^{f,H} \) fails to be an \( H \)-pathset. Conditions (ii) and (iii) hold with probability 1 (by observations which appeared earlier in the proof of Claim 4.5). Finally, if we assume that \( F \) solves \( \text{SUB}(G) \) a.a.s. on \( X_\theta \), then 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 \( M_{f,H}^{X_\theta} \subseteq [n]^{V(H)} \) be the relation consisting of \( \alpha \in [n]^{V(H)} \) such that \( H(\alpha) \) is a minterm of \( f \cup X_\theta \). 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 \( f_1 \lor f_2 \) is a minterm of \( f_1 \) or \( f_2 \), while every minterm of \( f_1 \land f_2 \) is the union of a minterm of \( f_1 \) and a minterm of \( f_2 \).

However, the relation \( \mathcal{M}_{f,H}^\theta \) can fail to be an \( H \)-pathset with probability \( \Omega(1/n) \) (e.g., if \( f \) is the monotone threshold function \( f(x) = 1 \iff |E(X)| \geq \sum_{e \in E(G)} n^{2-\theta(e)} \)). 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 \( \tilde{X}_\theta \) of random graphs \( X_\theta^0 \subseteq \cdots \subseteq X_\theta^m \) where \( m = n^o(1) \). This sequence is generated as \( X_\theta^0 := X_\theta \) and \( X_\theta^i := X_\theta^{i-1} \cup Y^i \) where \( Y^i \) is an independent copy of the \( G \)-colored Erdős-Rényi graph \( G_{n,\tilde{p}} \) where \( \tilde{p}_e := n^{-(1-\delta)\theta(e)} \) for a small constant \( \delta > 0 \) (i.e., each \( Y^i \) is a sparse version of \( X_\theta \)). If \( f \) is a depth-\( d \) subformula of \( F \), we say that \( H^{(\alpha)} \) is a persistent minterm of \( f \cdot \tilde{X}_\theta \) if it is a common minterm of \( f \cdot X_\theta^i \) and \( f \cdot Y^j \) for some \( 0 \leq i < j \leq m \) with \( j - i = (d+|E(H)|) \). Finally, we consider relations

\[
P_{f,H}^{\tilde{X}_\theta} := \{ \alpha \in [n]^{|V(H)|} : H^{(\alpha)} \text{ is a persistent minterm of } f \cdot \tilde{X}_\theta \}.
\]

The definition of persistent minterms ensures that, just like \( \mathcal{M}_{f,H}^X \), this family of relations satisfies conditions (ii) and (iii) of Def. 5.3 with probability 1. An extension of Lemma 3.3 shows that \( \mathcal{P}_{f,H}^{\tilde{X}_\theta} \) fails to be an \( H \)-pathset (w.r.t. \( \varepsilon = n^{-1+2\delta} \)) with exponentially small probability. Therefore, by a union bound, 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 assumptions: first, that \( F \) has depth \( O(\log n) \) so that \( \frac{(\text{depth}(F)+|E(G)|)}{|E(G)|} \leq m = n^o(1) \) (this is without loss of generality by balancing \( F \) via Spira’s theorem [37]); and second, that \( F \) solves \( \text{SUB}(G) \) a.a.s. on both \( X_\theta \) and \( X_\theta^m = X_\theta \cup Y^1 \cup \cdots \cup Y^m \). This is akin to solving \( \text{SUB}_{\text{uncat}}(G) \) a.a.s. on both \( G_{n,p} \) and \( G_{n,p+p^{1+\delta}} \), 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.

In the special case of \( G = C_k \) and \( \theta \equiv 1 \) (corresponding to the average-case \( k \)-CYCLE problem on \( G_{n,p} \) at the threshold \( p = \Theta(1/n) \)), we may take each \( Y^i \) to be the union of \( n^{1/2-\delta} \) random paths of length \( k \). In this case we are able to show that relations \( \mathcal{P}_{f,H}^{\tilde{X}_\theta} \) are pathsets w.r.t. density parameter \( \varepsilon = n^{1/2-2\delta} \). Furthermore, random graphs \( X_\theta \) and \( X_\theta^m \) have total variation distance \( o(1) \), which results in average-case lower bound on \( X_\theta \) 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 leaves 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 (w.r.t. $G, \theta, n, \varepsilon$) 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:

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

Pathset complexity gives lower bounds on pathset formula size (and by extension lower bounds on $AC^\theta$ 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 $\Phi_A$ be the maximum constant (depending on $G$ and $\theta$ alone) such that the inequality $\chi_A(A) \geq (1/\varepsilon)^{\Phi_A} \cdot \mu(A)$ holds for every $G_A$-pathset $A$ and every setting of parameters $n$ and $\varepsilon$. The invariant $\tau_\theta(G)$ is defined as the minimum value of $\Phi_A$ 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 $\Phi_A$ 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 $\mu(\mathcal{A}_{\text{out},G}) \geq .99$) must have size $\Omega((1/\varepsilon)^{\tau_\theta(G)})$. (This $\Omega(\cdot)$ hides a factor of $(1/2)^{2^{G}(E(G)}}$, which arises from partitioning $\mathcal{A}_{\text{out},G}$ 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 ([33]).** The average-case $AC^0$ formula size of $\text{SUB}(G)$ on $X_\theta$ 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 ([35]).** For all $G$ and $\theta$, the worst-case monotone formula (resp. circuit) size $\text{SUB}(G)$ is at least $n^{\tau_\theta(G)-o(1)}$ (resp. $n^{\kappa_\theta(G)-o(1)}$). In the case of $G = C_k$ and $\theta \equiv 1$, the average-case monotone formula size of $\text{SUB}(G)$ on $X_\theta$ 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 $\theta \equiv 1$. This requires us to prove lower bounds on constants $\Phi_A$ for every possible $G$-union tree $A$. In principle, this is a problem in the realm of graph theory, since $\Phi_A$ depends on $G$ and $\theta$ alone. Unfortunately, we do not know any nice expression for $\Phi_A$, nor any way of computing these constants exactly. 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 $A' \subseteq A$. 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 $\Phi_A$, we make use of structural properties of pathset complexity:

- **(projection lemma)** $\chi_{A'}(\text{proj}_{A'}(A)) \leq \chi_A(A)$ for all union trees $A' \subseteq A$ and every $G_{A'}$-pathset $A$, where $\text{proj}_{A'}(A) \subseteq [n]^{V(G_{A'})}$ is the projection of $A$ to coordinates in $V(G_{A'})$,

- **(restriction lemma)** $\chi_{A|H_1}(A|\beta) \leq \chi_A(A)$ for every vertex-disjoint partition $G_A = H_1 \uplus H_2$ and $\beta \in [n]^{V(H_2)}$, where $A|H_1$ is the union tree obtained from $A$ by deleting every leaf that is labeled by an edge of $H_2$. 

21
These lemmas allow us to derive two useful inequalities on constants \( \Phi_A \):
for all union trees \( A = \langle B, C \rangle \) and \( B' \preceq B \) and \( C' \preceq C \),

\[
\Phi_A \geq \Phi_{B'} + \Delta_C + \Delta_{A \ominus C},
\]

\[
\Phi_A \geq \frac{1}{2}(\Phi_{B'} + \Phi_{C' \ominus B'} + \Delta_A + \Delta_{A \ominus \langle B', C' \rangle}).
\]

(5.2)

(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 \( \theta \equiv 1 \), inequalities (5.2) and (5.3) can be used to show that \( \kappa_\theta(G) \geq \frac{1}{k} \log_2(k) \). This results in 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 \( G_{n,p} \) 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 \( \Phi_A \). Consider any root-to-leaf branch in a union tree \( A \), and let \( A_1, \ldots, A_m \) enumerate the union trees hanging off this branch in any order. For example, we might have \( A = \langle A_3, \langle A_1, \langle A_5, A_2 \rangle \rangle, A_4 \rangle \). For all such \( A \) and \( A_1, \ldots, A_m \), there is an inequality

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

(5.4)

Again in the case \( G = C_k \) and \( \theta \equiv 1 \), 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 \( \text{td}(G) \), 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 \( F \) [26]. For every pattern \( G \), there exist monotone \( AC^0 \) formulas of size \( O(n^{\text{td}(G)}) \) which solve \( \text{SUB}(G) \) in the worst-case. Comparing this upper bound to the lower bound of Theorem 5.7, it follows that \( \max_\theta \tau_\theta(G) \leq \text{td}(G) \). Using a recent result in graph minor theory of Kawarabayashi and the author [16], we are
able to show that \( \max_{\theta} \tau_{\theta}(G) \geq t_d(G)^c \) for all patterns \( G \) where \( c > 0 \) is an absolute constant. The result of [16] reduces this inequality to three special cases (grids, complete binary trees and paths) where we can bound \( \max_{\theta} \tau_{\theta}(G) \) by hand. As a corollary, this gives an \( \Omega(n^{t_d(G)^c - o(1)}) \) lower bound on the both \( AC^0 \) and monotone formula size of \( \text{SUB}(G) \). We conjecture that the true lower bound is \( n^{\Omega(t_d(G))} \), even for unrestricted Boolean formulas.

**Acknowledgements**

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

**References**


[25] 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.


