k-Subgraph Isomorphism on AC_0 Circuits

Kazuyuki Amano[*]

February 13, 2009

Abstract

Recently, Rossman [STOC '08] established a lower bound of $\omega(n^{k/4})$ on the size of constant-depth circuits for the k-clique function on n-vertex graphs, which is the first lower bound that does not depend on the depth of circuits in the exponent of n. He showed, in fact, a stronger statement: Suppose $f_n : \{0,1\}^{\binom{n}{2}} \to \{0,1\}$ is a sequence of functions computed by constant-depth circuits of size $O(n^t)$. For any positive integer k and $0 < \alpha \le 1/(2t-1)$, let $G = \mathbb{ER}(n, n^{-\alpha})$ be an Erdős-Rényi random graph with edge probability $n^{-\alpha}$ and let K_A be a k-clique on a uniformly chosen k vertices of G. Then $f_n(G) = f_n(G \cup K_A)$ asymptotically almost surely.

In this paper, we prove that this bound is essentially tight by showing that there *exists* a sequence of Boolean functions $f_n : \{0,1\}^{\binom{n}{2}} \to \{0,1\}$ that can be computed by constant-depth circuits of size $O(n^t)$ such that $f_n(G) \neq f_n(G \cup K_A)$ asymptotically almost surely for the same distributions with $\alpha = 1/(2t-5.5)$ and $k = 4t - c$ (where c is a small constant independent of k). This means that there are constant-depth circuits of size $O(n^{\frac{k}{4}+c})$ that correctly compute the k-clique function with high probability when the input is a random graph with independent edge probability around $n^{-2/(k-1)}$. Several extensions of his lower bound method to the problem of detecting general patterns as well as some upper bounds are also described. In addition, we provide an explicit construction of DNF formulas that are almost incompressible by any constant-depth circuits.

[*] Dept. of Computer Science, Gunma University, 1-5-1 Tenjin, Kiryu, Gunma 376-8515, Japan, Email: amano@cs.gunma-u.ac.jp

1 Introduction and Results

Proving a good lower bound on the size of Boolean circuits for explicit functions is one of the main challenges in theoretical computer science. The model of *constant-depth* circuits is one of two restricted models that has been a great success (see e.g., [1, 5, 12, 13, 23]). The other model is *monotone* circuits (see e.g., [2, 3, 19, 20]).

In this paper, we study the complexity of the k-subgraph isomorphism problem on constant-depth circuits. The k-subgraph isomorphism problem is, given a fixed "pattern" graph H on k vertices, to answer whether an input graph contains H as a subgraph. The problem has been widely investigated (e.g, [4, 9, 10, 17]). This work is strongly motivated by a recent work by Rossman [21], in which he established an $\omega(n^{k/4})$ lower bound on the size of constant-depth circuits for the k-clique problem on n-vertex graphs, which is the first lower bound that does not depend on the depth of circuits in the exponent of n. Throughout the paper we consider constant-depth circuits with $\wedge$ and $\vee$ gates of unbounded fan-in, and $\neg$ gates of fan-in one.

In the first part of the paper, we concentrate on the k-clique problem. We below briefly describe the approach taken by Rossman [21]. Consider an Erdős-Rényi random graph $\mathbb{ER}(n,p)$ with edge probability $p \sim n^{-(2/(k-1)+\varepsilon)}$. The exponent $2/(k-1)$ is called the *threshold exponent* for k-clique by the reasons (i) if $\alpha < 2/(k-1)$ then a graph $G \in \mathbb{ER}(n, n^{-\alpha})$ contains a k-clique asymptotically almost surely (a.a.s. for short), and (ii) if $\alpha > 2/(k-1)$ then a graph $G \in \mathbb{ER}(n, n^{-\alpha})$ does not contain a k-clique a.a.s..

Let $\alpha(t)$ be the maximal value such that for all $k \in \mathbb{N}$ and every sequence of functions $f_n : \{0,1\}^{\binom{n}{2}} \to \{0,1\}$ that computed by constant-depth circuits of size $O(n^t)$, it holds that $f_n(G) = f_n(G \cup K_A)$ a.a.s. where $G \in \mathbb{ER}(n, n^{-(\alpha+\varepsilon)})$ and K_A is a k-clique on a uniformly chosen k vertices of G. Rossman [21] showed that $\alpha(t) \geq 1/(2t-1)$ by a novel use of the famous Håstad's *switching lemma* [13] together with the result on the number of small subgraphs in the Erdős-Rényi random graph. This bound immediately implies a lower bound of $\omega(n^{k/4})$ on the size of constant-depth circuits for the k-clique problem since if we put $t = k/4$, the value of $\alpha = 1/(2t-1) = 2/(k-2)$ is greater than the threshold exponent of k-clique.

If we don't put any restrictions on the depth of circuits, it is known that the k-clique problem can be solved by circuits of size $O(n^{w\lceil k/3 \rceil})$ ([18] or see [22]), where $w \sim 2.376$ is the exponent of the fast matrix multiplication [8]. However, this method would not be applicable for constant-depth circuits. Thus, it is natural to expect that the lower bound would be improved to $\Theta(n^k)$. Apparently, a better lower bound on $\alpha(t)$ gives a higher lower bound on the constant-depth complexity for k-clique, and so the determination of the value of $\alpha(t)$ was stated as an open question in [21].

In this paper, we give an essentially tight upper bound of $\alpha(t) \leq 1/(2t-5.5)$ (Corollary 1). We show this by giving an explicit construction of constant-depth circuits of size $O(n^{\frac{k}{4}+c})$ (where c is a small constant independent of k) that correctly computes the k-clique function with high probability when the input is a random graph with independent edge probability around $n^{-2/(k-1)}$ (See Theorem 2 in Section 3 for more formal statement). Hence the lower bound of $\omega(n^{k/4})$ seems to almost reach the limit of the method based on the difficulties of the k-clique problem on random graphs around the threshold.

In the second part of the paper, we consider the k-subgraph isomorphism problem as a natural generalization of the k-clique problem. Suppose that the pattern H is a k-star. Then detecting H is equivalent to seeing whether there is a vertex of degree $\geq k$. Since it is known that the k-threshold function can be computed by depth-three circuits of size $O(n \log n)$ (for constant k, see e.g., [11] or [22]), k-star can be detected by depth-four circuits of size $O(n^2 \log n)$, here the exponent is independent of k. Hence it is interesting to investigate the relationship between the shape of the pattern H and the complexity of detecting it.

We first explain that the color-coding method introduced by Alon, Yuster and Zwick [4] is trivially applicable on constant-depth circuits, and thus obtain an upper bound of $O(n^{t+1} \log n)$ for a pattern of treewidth t. Then, we extend the Rossman's lower bound method to the problem of detecting a general pattern. We show intuitively that if the random graph $\mathbb{ER}(n, n^{-\alpha})$ with α being the threshold exponent of H contains a large number of copies of a certain induced subgraph of H, then the method can yield a good lower bound on the size of constant-depth circuits for detecting H (See Theorem 5 for more formal statement). As an example of applications our extended method, we show that the constant-depth circuit complexity of detecting a $k \times k$-grid is between $\omega(n^{(3\sqrt{2}-4)k-\varepsilon}) = \omega(n^{0.246k})$ and $O(n^{k+1} \log n)$ (Theorem 6).

Our extended method also reveals that sharper lower bounds can be obtained when we consider the k-clique problem on *hypergraphs*. Based on this, we give explicit DNFs that are almost incompressible by any constant-depth circuits. Precisely, for every integer $t \geq 2$ and $\varepsilon > 0$, we give a construction of an $O(n^t)$-term DNF formula with term length $O(1)$ that cannot be computed by any constant-depth circuits of size $O(n^{t-\varepsilon})$ (Corollary 2). Our construction is highly uniform, i.e., it is a naive DNF formula representing the k-hyperclique problem on ℓ-uniform hypergraphs (for a suitable choice of k and ℓ). As far as the author's knowledge, this is the first natural construction of functions with such a property. A simple counting argument would not be able to show even the existence of such DNFs. Note that an incompressible result in a sharper form has recently shown for monotone circuits of depth at most four [16].

1.1 Organization of the Paper

In Section 2, we give some notations and definitions. In Section 3, we give an explicit construction of constant-depth circuits of size $O(n^{k/4+c})$ for detecting a k-clique when the input is a random graph with independent edge probability around the threshold. In Section 4, we give a generalized version of the lower bound method developed by Rossman as well as some upper bounds. We also give an explicit construction of DNFs that are almost incompressible by any constant-depth circuits. Finally, we close the paper by giving some open problems in Section 5.

2 Notations and Definitions

A *circuit* C on $X = \{x_1, \ldots, x_n\}$ is an acyclic directed graph consisting of input nodes and gate nodes. Input nodes are labeled by one of the input variables, and gate nodes are labeled by one of the elements of $\{\neg, \wedge, \vee\}$. Gate nodes labeled by $\wedge$ and by $\vee$ have unbounded fan-in, and those labeled by $\neg$ have fan-in one. A subset of nodes in C are designated as output nodes. A circuit computes a Boolean function (or a set of Boolean functions) in an obvious way. The output of C for an input $x \in \{0,1\}^n$ is denoted by $C(x)$. A graph G on m vertices is naturally represented by $\binom{m}{2}$ Boolean variables. If an input represents a graph G, then the output of C for this input is written as $C(G)$. When C has t output nodes, then $C(x) \in \{0,1\}^t$. The *size* of a circuit C is the number of gate nodes in C, and the *depth* of C is the maximum number of gates on a path from an input node to an output node. The class of all Boolean functions that can be computed by constant-depth polynomial-size circuits is denoted by AC_0.

The *k-clique function* on n-vertex graphs, denoted by Clique_k^n, is a Boolean function on $\binom{n}{2}$ variables that outputs 1 iff the graph represented by an input contains a k-clique (denoted by K_k), i.e., a complete graph on k vertices. A complete graph whose support is A is denoted by K_A. For a graph $H = (V_H, E_H)$ on k vertices, the *H-detecting function* on n-vertex graphs is a Boolean function on $\binom{n}{2}$ variables that outputs 1 iff the graph represented by an input contains a copy of H as a subgraph.

Let $\mathbb{ER}(n,p)$ denote the Erdős-Rényi random graph on n vertices with independent edge probability p. For the properties of $\mathbb{ER}(n,p)$, see e.g., [7, 15]. For a graph G, $V(G)$ denotes the vertex set of G and $E(G)$ denotes the edge set of G. The *density* of a graph G is defined as $|E(G)|/|V(G)|$, and the *maximum density* is defined as $\max_{H \subseteq G} |E(H)|/|V(H)|$ where H ranges over all induced subgraphs of G. The *threshold exponent* of a graph G, denoted by $\mathsf{thre}(G)$, is the inverse of the maximum density of G, i.e., $\mathsf{thre}(G) = \min_{H \subseteq G} |V(H)|/|E(H)|$. For example, $\mathsf{thre}(K_k) = 2/(k-1)$. We say that an event E_n, describing a property of a random structure depending on a parameter n, holds *asymptotically almost surely* (abbreviated a.a.s.), if $\Pr[E_n] \to 1$ as $n \to \infty$.

For a natural number n, $[n]$ stands for the set $\{1, \ldots, n\}$.

3 Detecting k-Clique around the Threshold

Following Rossman [21, Sect. 6], we define $\alpha(t)$ to be the maximal value such that for all $k \in \mathbb{N}$ and every sequence of functions $f_n : \{0,1\}^{\binom{n}{2}} \to \{0,1\}$ that are computed by constant-depth circuits of size $O(n^t)$, it holds that $f_n(G) = f_n(G \cup K_A)$ asymptotically almost surely where $G \in \mathbb{ER}(n, n^{-(\alpha+\varepsilon)})$ and A is a uniform random set of k elements of $V(G)$. In [21], Rossman gave the following lower bound on $\alpha(t)$, and left the problem of determination of this value as an open problem.

Theorem 1 *[21]* $\alpha(t) \geq 1/(2t-1)$.

The lower bound of $\omega(n^{k/4})$ for the constant-depth circuit complexity of the k-clique problem immediately follows from this theorem by letting $t = k/4$. For this choice of t, $G \in \mathbb{ER}(n, n^{-\alpha})$ contains no clique a.a.s. since $\alpha = 2/(k-2) > 2/(k-1) = \mathsf{thre}(K_k)$.

A careful inspection of his proof reveals that the lower bound of $\omega(n^{k/4})$ comes from the maximum of the expected number of s-cliques among $2 \leq s \leq k$ in the random graph $\mathbb{ER}(n, n^{-\alpha})$ where $\alpha = \mathsf{thre}(K_k) = 2/(k-1)$. For $2 \leq s \leq k$, let X_s be a random variable that represents the number of s-cliques in $\mathbb{ER}(n, n^{-\alpha})$ where $\alpha = \mathsf{thre}(K_k) = 2/(k-1)$. The expectation of X_s is $\Theta(n^{s-\alpha\binom{s}{2}})$. Figure 1 shows the exponent of this expectation for $k = 100$. This takes a maximum of $\sim k/4$ at $s = k/2$.

We below show that Theorem 1 is essentially tight.

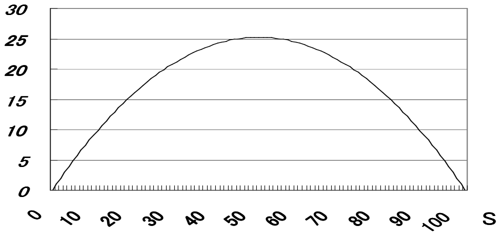

Figure 1: The exponent of the expected number of s-cliques in the Erdős-Renyi graph $\mathbb{ER}(n, n^{-\alpha})$ where $\alpha = \mathsf{thre}(K_k) = 2/(k-1)$ and $k = 100$.

Theorem 2 *Let α be a constant such that $\alpha \geq \mathsf{thre}(K_k) = 2/(k-1)$. Let d be an arbitrary constant. Then, there is a sequence of Boolean functions f_n on n-vertex graphs that can be computed by constant-depth circuits of size $O(n^{(k+10)/4})$ such that*

$$\Pr[f_n(G) = \mathsf{Clique}^n_k(G)] \quad \geq \quad 1 - o(n^{-d}).$$

This holds for both of two distributions (i) $G \in \mathbb{ER}(n, n^{-\alpha})$, (ii) $G = G' \cup K_A$ where $G' \in \mathbb{ER}(n, n^{-\alpha})$ and A is a uniform k-subset of $V(G)$.

By letting $k = 4t - 10$ for the above theorem, the following corollary is immediate.

Corollary 1 $\alpha(t) \leq 1/(2t - 5.5)$. □

In the rest of this section, we describe the proof of Theorem 2. In order to show this, we use the following result on the complexity of the k-threshold function, which is formally defined as the Boolean function on n variables $\{x_1, \ldots, x_n\}$ that outputs 1 iff $\sum_{i=1}^n x_i \geq k$ and is denoted by Th^n_k. See e.g., [22, Chap. 8.2] for the proof.

Theorem 3 Th^n_k *can be computed by a depth three formula (i.e., a special form of circuits in which every gate has fan-out one) of size $O(k^{k+1} n \log n)$. In particular, if k is a constant then Th^n_k can be computed by a depth three formula of size $O(n \log n)$.* □

Proof (of Theorem 2). We give a construction of a circuit for f_n that satisfies the condition of the theorem. For the sake of simplicity, we concentrate on the distribution (i) for a while. The proof for the distribution (ii) is analogous, and will be noted at the end of the proof.

The construction consists of $k-1$ stages, from stage 2 to stage k. Intuitively, at stage ℓ, we maintain all ℓ-cliques contained in an input graph. The number of ℓ-cliques in $G \in \mathbb{ER}(n, n^{-\alpha})$ is well concentrated around its expectation, which is shown as the line of the subgraph-count plot (See Fig. 1). If we can maintain these small cliques efficiently, then we can construct a circuit for the k-clique function whose size is close to the maximum number of ℓ-cliques among $2 \leq \ell \leq k$, which is shown as the peak of the subgraph-count plot. In the following, a variable that represents whether there is an edge between two vertices u and v is denoted by $x_{u,v}$. We consider that the vertex set V is ordered in a natural way (e.g., by the index of a vertex).

We begin with Stage 2. Let $\mathcal{S}_2$ be a partition of $\binom{[n]}{2}$ into blocks of equal size n^{b_2}, where the value of b_2 will be chosen later. Note that the number of sets is $\Theta(n^{2-b_2})$. For every

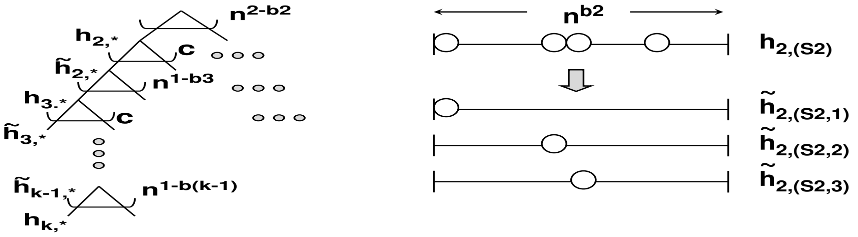

Figure 2: Left: Our circuit has a tree-like structure with alternating levels of h's and $\tilde{h}$'s. Right: The relationship between the functions h and $\tilde{h}$. A circle denotes the output bit that gets the value 1. If there are more than c 1's in h, then these will be discarded.

$S_2 \in \mathcal{S}_2$, define $h_{2,(S_2)} : \{0,1\}^{\binom{[n]}{2}} \to \{0,1\}^{n^{b_2}}$ such that the i-th bit of $h_{2,(S_2)}(G)$ is 1 if and only if the i-th element of S_2 is connected by an edge in G.

Let c be a sufficiently large constant. For every $c_2 \in [c]$, define $\tilde{h}_{2,(S_2,c_2)} : \{0,1\}^{\binom{[n]}{2}} \to \{0,1\}^{n^{b_2}}$ such that the i-th bit of $\tilde{h}_{2,(S_2,c_2)}(G)$ is 1 if and only if the i-th bit of $h_{2,(S_2)}(G)$ is 1 and it is the c_2-th 1 counting from the top. Formally, the i-th bit of $\tilde{h}_{2,(S_2,c_2)}$ is defined as

$$\mathsf{Th}^{i}_{c_2}(h^{[i]}_{2,(S_2)}) \wedge \overline{\mathsf{Th}^{i-1}_{c_2}(h^{[i-1]}_{2,(S_2)})}, \tag{1}$$

where $h^{[i]}$ denotes the first i bits of the outputs of h (see Fig. 2). Note that the number of outputs of $\tilde{h}_{2,(S_2,c_2)}$ that have the value 1 is at most one. By Theorem 3, each bit of $\tilde{h}_{2,(S_2,c_2)}$ can be computed by a depth five circuit of size $O(n^{b_2}\log n)$.

We say that a graph G is *bad on stage 2* if for some $S_2 \in \mathcal{S}_2$, the number of 1's in $h_{2,(S_2)}(G)$ exceeds c. If we pick G from $\mathbb{ER}(n, n^{-\alpha})$, the probability that a graph is bad on stage 2 is at most

$$O(n^{2-b_2})\binom{n^{b_2}}{c+1}n^{-\alpha(c+1)} = O(n^{2-b_2}n^{(b_2-\alpha)(c+1)}). \tag{2}$$

We put $b_2 = \alpha - \varepsilon$ for a sufficiently small constant $\varepsilon > 0$. Then the exponent in the above probability is strictly smaller than $-d$ when c is sufficiently large.

We now proceed to Stage 3. Let $\mathcal{S}_3$ be a partition of $[n]$ into blocks of equal size n^{b_3}, where the value of b_3 will be chosen later. Note that the number of sets is $\Theta(n^{1-b_3})$. For every $S_2 \in \mathcal{S}_2$, $c_2 \in [c]$ and $S_3 \in \mathcal{S}_3$, define $|S_3|$-output function $h_{3,(S_2,c_2,S_3)}$ such that (i) each output is corresponding to a vertex in v in S_3, (ii) for every $v_3 \in S_3$, the corresponding output is defined as

$$\bigvee_{\substack{(v_1,v_2)\in S_2\\ v_3 > v_1,v_2}} x_{v_1,v_3}x_{v_2,v_3}\tilde{h}^{(v_1,v_2)}_{2,(S_2,c_2)}(G),$$

where $\tilde{h}^{(v_1,v_2)}_{2,(S_2,c_2)}$ denotes the output bit of $\tilde{h}_{2,(S_2,c_2)}$ corresponding to $(v_1,v_2) \in S_2$. After computing $h_{3,(S_2,c_2,S_3)}$, we compute $\tilde{h}_{3,(S_2,c_2,S_3,c_3)}$ for each $c_3 \in [c]$ which is defined analogously to Stage 2, i.e., the i-th bit of this is 1 iff the i-th bit of the output of $h_{3,(S_2,c_2,S_3)}$ is the c_3-th 1 counting from the top. This can be done by using a similar circuit as to (1). A graph G is *bad on stage 3* if there is a function $h_{3,\star}$ such that the output of it contains more than c

one's. The probability that a graph in $\mathbb{ER}(n, n^{-\alpha})$ is bad on stage 3 is

$$O(n^{2-b_2}n^{1-b_3})\binom{n^{b_3}}{c+1}n^{-2\alpha(c+1)} = O(n^3 n^{(b_3-2\alpha)(c+1)}).$$

This probability is $o(n^{-d})$ for sufficiently large c if we put $b_3 = 2\alpha - \varepsilon$ for sufficiently small constant $\varepsilon > 0$.

We then continue the above construction. The Stage ℓ is as follows: Let $\mathcal{S}_\ell$ be a partition of $[n]$ into blocks of equal size n^{b_ℓ}, where the value of b_ℓ will be chosen later. For every choice of $(S_2, c_2, \ldots, S_\ell)$, define $|S_\ell|$-output function $h_{\ell,(S_2,c_2,\ldots,S_\ell)}$ such that (i) each output is corresponding to a vertex v_ℓ in S_ℓ, (ii) for every $v_\ell \in S_\ell$, the corresponding output is defined as

$$\bigwedge_{3\le m\le \ell-1}\left(\bigvee_{\substack{v_m\in S_m\\ v_\ell>v_m}} x_{v_m,v_\ell}\tilde{h}^{(v_m)}_{m,(S_2,c_2,\ldots,S_m,c_m)}(G)\right)\wedge\left(\bigvee_{\substack{(v_1,v_2)\in S_2\\ v_\ell>v_1,v_2}} x_{v_1,v_\ell}x_{v_2,v_\ell}\tilde{h}^{(v_1,v_2)}_{2,(S_2,c_2)}(G)\right),$$

where $\tilde{h}^{(v_m)}_{m,(\ldots,S_m,c_m)}$ denotes the output bit of $\tilde{h}_{m,(\ldots,S_m,c_m)}$ corresponding to $v_m \in S_m$. After computing $h_{\ell,(S_2,c_2,\ldots,S_\ell)}$, we compute $\tilde{h}_{\ell,(S_2,c_2,\ldots,S_\ell,c_\ell)}$ analogously to the lower stages. A graph G is *bad on stage* ℓ if there is a function $h_{\ell,\star}$ such that the output of it contains more than c one's. The probability that a graph in $\mathbb{ER}(n, n^{-\alpha})$ is bad on stage ℓ is

$$O(n^{2-b_2}n^{1-b_3}\cdots n^{1-b_\ell})\binom{n^{b_\ell}}{c+1}n^{-(\ell-1)\alpha(c+1)} = O(n^\ell n^{(b_\ell-(\ell-1)\alpha)(c+1)}).$$

This will be smaller than $o(n^{-d})$ if we put $b_\ell = \min\{1, (\ell-1)\alpha - \varepsilon\}$ for some $\varepsilon > 0$.

The final output of the circuit is the OR of the all $h_{k,\star}$'s. A graph G is said to be *good* if G is not bad for every stage. The probability that a graph $G \in \mathbb{ER}(n, n^{-\alpha})$ is good is at least $1 - o(n^{-d})$. If an input graph G is good, then the output of our circuit is equal to $\mathsf{Clique}^n_k(G)$.

It is obvious that the depth of a circuit is linear in k, which is a constant when k is a constant. The size of a circuit is dominated by the final stage. The number of blocks at stage k is

$$O\left(n^{\sum_{i=2}^{k}(1-b_i)}\right),$$

and each block has size $O(n^{b_k})$. Hence the total number of elements $h_{k,\star}$'s at stage k is

$$O\left(n^{b_k}\cdot n^{\sum_{i=2}^{k}(1-b_i)}\right) \quad = \quad O\left(n^{1+\sum_{i=2}^{k-1}(1-b_i)}\right).$$

The exponent of the above equation is

$$\begin{aligned}1+\sum_{i=2}^{k-1}(1-b_i) &= 1+\sum_{i=2}^{\lceil\frac{k-1}{2}\rceil}\left(1-\frac{2(i-1)}{k-1}\right)+\varepsilon' = \lceil\frac{k-1}{2}\rceil - \frac{2}{k-1}\sum_{i=1}^{\lceil\frac{k-1}{2}\rceil-1} i+\varepsilon' \\ &= \lceil\frac{k-1}{2}\rceil - \frac{\lceil\frac{k-1}{2}\rceil(\lceil\frac{k-1}{2}\rceil-1)}{k-1}+\varepsilon' \le \frac{k}{4}+\frac{1}{3}+\varepsilon'.\end{aligned}$$

for some sufficiently small constant $\varepsilon' > 0$. Here the first equality follows from the fact that $b_\ell \geq 1-\varepsilon$ for $\ell > \lceil \frac{k-1}{2} \rceil$ (and in fact $b_\ell = 1$ for $\ell > \lceil \frac{k-1}{2} \rceil + 1$). Thus the total number of output bits in all $\tilde{h}_{k-1,\star}$ and $h_{k,\star}$ is $O(n^{\frac{k}{4}+\frac{4}{3}+\varepsilon'})$. Each of these bits can be computed by a circuit of size at most $O(n \log n)$. Hence the total size of our circuit is $O(n^{\frac{k}{4}+\frac{7}{3}+\varepsilon''}) = O(n^{\frac{k}{4}+\frac{5}{2}})$.

We now show the correctness of the theorem for the distribution (ii). We can upper bound the probability that $f(G) \neq \mathsf{Clique}_k^n(G)$ when G is chosen according to the distribution (ii) in an analogs way to the case for the distribution (i). Since the addition of K_A will affect only on k vertices, for example, the probability of the failure at stage 2 is upper bounded by

$$O(n^{2-b_2}) \binom{n^{b_2}}{c+1} n^{-\alpha(c+1-k)} = O(n^{2-b_2} n^{(b_2-\alpha)(c+1-k)}).$$

This will be smaller than $o(n^{-d})$ when c is sufficiently large. The failure probability for any other stages can be bounded in an analogous way. □

4 Detecting General Patterns

4.1 Color-Coding on $\mathbf{AC}^0$

The color-coding method introduced by Alon, Yuster and Zwick [4] is implementable on constant-depth circuits, and thus we can obtain an upper bound on the constant-depth circuit complexity for detecting a pattern H (of a fixed size k) of $O(n^{t+1} \log n)$ where t is the treewidth of H. See e.g., [6] for a good survey on the notion of treewidth.

Definition 1 *A tree-decomposition of a graph $G = (V, E)$ is a pair $(\mathcal{X}, T)$ with $T = (I, F)$ a tree, and $\mathcal{X} = \{X_i \mid i \in I\}$ a family of subsets of V, one for each node of T, such that (i) $\bigcup_{i \in I} X_i = V$, (ii) for all edges $\{u, v\} \in E$ there exists an $i \in I$ with $u \in X_i$ and $v \in X_i$, and (iii) for all $i, j, k \in I$, if j is on the path from i to k in T, then $X_i \cap X_k \subseteq X_j$. The treewidth of a tree-decomposition $(\mathcal{X}, T)$ is $\max_{i \in I} |X_i| - 1$. The treewidth of a graph G is the minimum treewidth over all possible tree-decomposition of G.* □

Theorem 4 *Let $H = (V_H, E_H)$ be a graph on k vertices. If the tree-width of H is t, then there is a constant depth circuit of size $O(n^{t+1} \log n)$ that detects H in a graph G on n vertices, if one exists.* □

The theorem can easily be shown by applying the derandomized version of the color-coding method [4]. See Appendix (Section 6.1) for some more details.

4.2 Lower Bound for Detecting General Patterns

The lower bound method developed by Rossman [21] can naturally be extended to the problem of detecting a general pattern.

Throughout this section, $H = (V_H, E_H)$ denotes a pattern graph on k vertices, and $G = (V, E)$ denotes an input graph on n vertices. We first extend the notion of the *clique-sensitive core* introduced by Rossman [21] to be able to handle more general graphs.

Let A be a mapping from V_H to V. For a subset $V' \subseteq V_H$, let $\mathsf{Im}_A(V') = \{A(v') \mid v' \in V'\}$. For a subset $V' \subseteq \mathsf{Im}_A(V_H)$, let $H_{A|V'}$ denote a graph on V whose edge set is

$\{(A(u), A(v)) \mid (u,v) \in E_H \ \& \ u, v \in A^{-1}(V')\}$. In particular, $H_{A|\mathsf{Im}_A(V_H)}$ is simply denoted by H_A, i.e., H_A consists of a single copy of H on the vertex set $\mathsf{Im}_A(V_H)$.

Definition 2 *Let f be an arbitrary (not necessary Boolean) function on n-vertex graphs. Let $H = (V_H, E_H)$ be a pattern graph on k vertices. For $A : V_H \to V$, $V' \subseteq \mathsf{Im}_A(V_H)$ and $s \in \mathbb{N}$, we define*

$$\begin{aligned} T^{f,G}(A, V') &= \left\{a \in V' \mid \exists U \subseteq V' s.t. f(G \cup H_{A|U}) \neq f(G \cup H_{A|(U \setminus \{a\})})\right\} \\ T^{f,G}_{\langle s\rangle}(A) &= \bigcup_{V' \subseteq \mathsf{Im}_A(V_H) : |V'| \leq s} T^{f,G}(A, V'). \end{aligned}$$

If $T^{f,G}(A, V') = V'$ then we say that $T^{f,G}(A, V')$ is full. □

It can be verified that the all desired properties described in Section 3 in [21] are valid under these extensions.

For a graph H on k vertices and $2 \leq s \leq k$, let $m(H, s)$ denote the maximum density of an induced subgraph of H with size s, i.e.,

$$m(H, s) = \max\{|E(H')|/|V(H')| \mid H' \subseteq H, |V(H')| = s\}.$$

Note that the threshold exponent of a graph H on k vertices is equal to $\min_{2 \leq s \leq k} m(H, s)^{-1}$.

We can show the following lemma, whose proof is omitted in this version, that is analogous to Lemma 4.8 in [21].

Lemma 1 *Let $\beta \geq 0$ be a constant. Let $f_n : \{0,1\}^{\binom{n}{2}} \to \{0,1\}^{n^\beta}$ be Boolean functions that can be computed by AC_0 circuits (with n^β output gates). Let H be a graph on k vertices, and let H' be an induced subgraph of H on s vertices where $2 \leq s \leq k$. Suppose that $0 < \alpha < \mathsf{thre}(H')$. Then,*

$$\Pr[T^{f,G}(A', \mathsf{Im}_{A'}(V(H'))) \text{ is full }] \leq n^{\alpha|E(H')| + (\beta-1)s + o(1)},$$

where the probability is taken over a random graph $G \in \mathbb{ER}(n, n^{-\alpha})$ and A' is a uniformly chosen mapping from $V(H') \to V$. □

Let C be a circuit on $\binom{n}{2}$ inputs, and let g be an arbitrary gate in C whose fan-in is denoted by $\mathsf{fanin}(g)$. Let $\tilde{g} : \{0,1\}^{\binom{n}{2}} \to \{0,1\}^{\mathsf{fanin}(g)+1}$ be the function that outputs the value of g and all of its children. Let $T^{\tilde{g},G}_{\langle s\rangle}(A)$ be defined as

$$T^{\tilde{g},G}_{\langle s\rangle}(A) \quad = \quad T^{g,G}_{\langle s\rangle}(A) \cup \bigcup_{h \text{ is a child of } g} T^{h,G}_{\langle s\rangle}(A).$$

We will simply write $T(g)$ instead of $T^{g,G}_{\langle s\rangle}(A)$, and $T(\tilde{g})$ instead of $T^{\tilde{g},G}_{\langle s\rangle}(A)$. We identify a circuit C with a function that computed by C.

Lemma 2 *Let $H = (V_H, E_H)$ be a graph on k vertices and $A : V_H \to V$ be a mapping. If $C(G) \neq C(G \cup H_A)$ and $T(C) = \emptyset$, then there is a gate g in C such that $|T(\tilde{g})| > s$.* □

The proof of the above lemma is analogous to the proof of Lemma 3.6 in [21], which is described in Appendix. If we consider that a pair of inputs $(G, G \cup H_A)$ is assigned to a gate with $|T(\tilde{g})| > s$, then the role of Lemma 2 may be clarified.

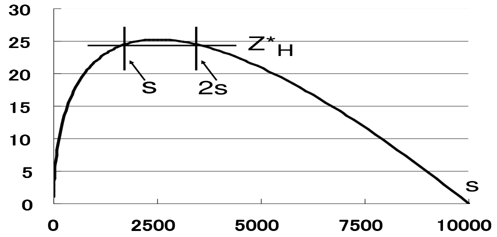

Figure 3: The subgraph-count function $Z_H(s)$ for 100×100 grid graph H. The threshold $\mathsf{thre}(H)$ is $\frac{k}{2(k-1)}$ and the value of Z^*_H is ~ 24.6.

The subgraph-count function $Z_H(s)$ is defined to be $Z_H(s) := s\{1 - \mathsf{thre}(H) \cdot m(H, s)\}$. Note that for every induced subgraph $H' \subseteq H$ of size s, $\mathbb{ER}(n, n^{-\mathsf{thre}(H)})$ contains $\Omega(n^{Z_H(s)})$ copies of H' in expectation. Let $\hat{Z}_H(s) := \min_{2\leq s'\leq s} Z_H(s)$ and let

$$s^\star(H) := \min\{s \mid m(H, s)^{-1} = \mathsf{thre}(H)\}.$$

In other words, $s^\star(H)$ is the smallest s satisfying $Z_H(s) = 0$, and so $\hat{Z}_H(s) > 0$ iff $s < s^\star(H)$. Finally, we define $\tilde{Z}_H(s) := \min_{s+1\leq\tilde{s}\leq 2s} Z_H(\tilde{s})$, and $Z^\star_H := \max_{s\geq 1:\hat{Z}_H(s)>0} \tilde{Z}_H(s)$ (or equivalently $Z^\star_H := \max_{1\leq s\leq s^\star(H)} \tilde{Z}_H(s)$). See Fig. 3 for an example.

Lemma 3 *For every $\epsilon > 0$, there are constants $\alpha > \mathsf{thre}(H)$ and $\beta > 0$ such that the following holds. For every AC_0 computable function $f : \{0,1\}^{\binom{n}{2}} \to \{0,1\}^{n^\beta}$,*

$$\Pr_{G,A}[|T^{f,G}_{\langle s\rangle}(A)| > s] \quad \leq \quad n^{-\tilde{Z}_H(s)+\varepsilon}, \tag{3}$$

and

$$\Pr_{G,A}[T^{f,G}_{\langle s\rangle}(A) \neq \emptyset] \quad \leq \quad n^{-\hat{Z}_H(s)+\varepsilon}, \tag{4}$$

where $G \in \mathbb{ER}(n, n^{-\alpha})$ and A is a random mapping $V_H \to V$. □

The proof of the above lemma is in Appendix (Section 6.3). By putting them together, we have a generalized version of the lower bound results by Rossman [21].

Theorem 5 *Let $H = (V_H, E_H)$ be a graph on k vertices. Suppose that $(f_n)_{n\in\mathbb{N}}$ is a sequence of Boolean functions on $\binom{n}{2}$ variables that computed by constant-depth circuits $C = (C_n)_{n\in\mathbb{N}}$. Suppose also that f_n computes the H-detecting function on n-vertex graphs. Then the size of C_n is $\Omega(n^{Z^\star_H-\varepsilon})$ for every $\varepsilon > 0$.*

Proof (sketch). Fix a sufficiently small constant ε. We now pick $\alpha > \mathsf{thre}(H)$ and $\beta > 0$ as in Lemma 3. Let s be the value such that $Z^\star_H = \tilde{Z}_H(s)$. Without loss of generality, we can assume that all gates in C have fan-in $n^\beta - 1$ (without increasing the size or depth by more than constant factors).

Let $G \in \mathbb{ER}(n, n^{-\alpha})$ and A be a uniformly chosen random mapping from V_H to V. Then, $C_n(G) = 0$ a.a.s., and $C_n(G \cup H_A) = 1$ with probability 1. Moreover, $T(C) = 0$ a.a.s. (here we use Eq. (4) of Lemma 3). By Lemma 2, for almost all G and H_A, there is a gate g in C such that $|T(\tilde{g})| > s$. However, this contradicts Eq. (3) of Lemma 3. □

It should be mentioned that $Z^{\star}_{K_k} = \frac{2k}{9} + o(1)$, and the above theorem generalizes the weaker $\omega(n^{2k/9})$ lower bound of Section 3.3 of [21], rather than the stronger $\omega(n^{k/4})$ lower bound of Section 4 of that paper.

In the rest of this subsection, we derive a lower bound on the size of constant-depth circuits that detect a $k \times k$-grid, as an illustrative example of our generalized method.

Let $H = (V_H, E_H)$ be the $k \times k$ grid, i.e., $V_H = \{\{i,j\} \mid i,j \in [k]\}$ and $E_H = \{(\{i_1,j_1\},\{i_2,j_2\}) \mid |i_1 - i_2| + |j_1 - j_2| = 1\}$. It is easy to check that $|E_H| = 2k(k-1)$. The following is a simple exercise (whose proof is in Appendix (Section 6.4)).

Fact 1 *For $k \times k$ grid H, the threshold exponent of H is $\frac{k}{2(k-1)}$ and $Z^{\star}_H \geq (3\sqrt{2}-4)k$.* □

Since it is folklore that the treewidth of the $k \times k$ grid is k, we have the following:

Theorem 6 *The constant-depth circuit complexity of the problem for detecting a $k \times k$-grid is $\omega(n^{(3\sqrt{2}-4)k-\varepsilon}) = \omega(n^{0.246k})$ and $O(n^{k+1}\log n)$ for every constant $\varepsilon > 0$.* □

The peak of the subgraph-count plot for the $k \times k$-grid is around $k/4$ (see Fig. 3), and so it would be possible to improve the above lower bound to $\Omega(n^{k/4})$ by more detailed analysis as in the proof of Theorem 1.1 in [21]. However, it would be also possible to construct constant-depth circuits of size $O(n^{k/4+c})$ that correctly separates two sets of test inputs with high probability as in the proof of Theorem 2. Thus, again, this lower bound seems to (almost) reach the limits of the method.

4.3 Hypergraphs and Incompressible DNFs

As we have seen in the previous sections, the method yields a sharper lower bound when the slope of a subgraph-count plot is steep. Based on this, we can show that DNF formulas representing the k-clique problems on ℓ-uniform hypergraphs are almost incompressible for constant-depth circuits when ℓ is large.

Let K^{ℓ}_k denote the complete ℓ-uniform hypergraph with k vertices. The ℓ-uniform k-hyperclique function, denoted by $\mathsf{HypClique}^n_{\ell,k}$, is a Boolean function on $\binom{n}{\ell}$ variables that outputs 1 iff the ℓ-uniform hypergraph represented by the input contains K^{ℓ}_k. A random hypergraph $\mathcal{H}(\ell, n, p)$ is a hypergraph on the vertex set $[n]$ where each of $\binom{n}{\ell}$ ℓ-element subsets of $[n]$ is an edge with probability p independently of the other ℓ-element subsets of $[n]$.

We can obtain a lower bound on the constant-depth circuit size for $\mathsf{HypClique}^n_{\ell,k}$ by a proof along the lines of the proof for k-clique. The sketch of the proof is described in Appendix (Section 6.5).

Theorem 7 *For every $k > \ell > 2$, every constant-depth circuit that computes $\mathsf{HypClique}^n_{\ell,k}$ contains at least $\Omega(n^{k(1-(\ln \ell + 2)/(\ell - 1))})$ gates.* □

Corollary 2 *For every integer $t \geq 2$ and every constant $\varepsilon > 0$, there is a uniform sequence of Boolean functions f_n on n variables such that (i) f_n can be represented by an $O(n^t)$-term DNF formula with term length $O(1)$, and (ii) f_n cannot be computed by constant-depth circuits of size $O(n^{t-\varepsilon})$.*

Proof. Let ℓ be the smallest integer such that $(\ln \ell + 2)/(\ell - 1) < \varepsilon/t$, and let $k = \ell t$. Then the corollary immediately follows from Theorem 7. □

5 Discussions

There remains a considerable gap between $O(n^k)$ upper bounds and $\omega(n^{k/4})$ lower bounds on the constant-depth circuit complexity of the k-clique function. The proof of Theorem 2 suggests that the k-clique function may not be the hardest for graphs with $\Theta(n^{2-2/(k-1)})$ edges in average. To find distributions on positive and negative inputs that are unlikely to be distinguished by constant-depth circuits of size $O(n^{k/4+c})$ seems an important step for improving the lower bounds.

Some other open problems are listed below.

- Is there an algorithm for the k-subgraph isomorphism that beats the upper bound based on the tree-decomposition described in Section 4.1? Note that fast algorithms are known when input graphs are restricted to planar graphs or some more general graphs (see e.g., [9, 10, 17]).
- Can we give an explicit construction of DNFs of size $O(n^t)$ that cannot be computed by any constant-depth circuits of size $o(n^t)$?

Acknowledgment

The author would like to thank an anonymous reviewer on an earlier version of this paper for providing a huge number of constructive suggestions and comments and in particular for pointing out an error in the proof of Theorem 5 and fixing it. The author also would like to Koichi Yamazaki for enjoyable discussions.

References

[1] M. Ajtai, "Σ_1^1 Formula on Finite Structures", Ann. of Pure Appl. Logic, **24**, 1–44 (1983)

[2] K. Amano and A. Maruoka, "The Potential of the Approximation Method", SIAM J. Comput.,**33(2)**, 433-447 (2004) (Preliminary version in the Proc. of the 37th FOCS, 431–440 (1996))

[3] A. Andreev, "On a Method for Obtaining Lower Bounds for the Complexity of Individual Monotone Functions", Dolk. Akad. Nauk. SSSR, **282(5)**, 1033–1037 (1985) (in Russian) English Translation: Soviet Math. Dokl., **31(3)**, 530–534 (1985)

[4] N. Alon, R. Yuster and U. Zwick, "Color-Coding", J. ACM, **42(4)**, 844-856 (1995)

[5] P. Beame, "Lower Bounds for Recognizing Small Cliques on CRCW PRAM's", Disc. Appl. Math. **29(1)**, 3–20 (1990)

[6] H.L. Bodlaender, "Discovering Treewidth", Proc. of the 31st SOFSEM, LNCS 3381, 1–16 (2005)

[7] B. Bollobás, Random Graphs, 2nd Eds., Cambridge Univ. Press (2001)

[8] D. Coppersmith and S. Winograd, "Matrix Multiplication via Arithmetic Progressions", J. Symbolic Comput., **9**, 251–280 (1990)

[9] D. Eppstein, "Subgraph Isomorphism in Planar Graphs and Related Problems", J. Graph Algorithms and Applications, **3(3)**, 1–27 (1999)

[10] D. Eppstein, "Diameter and Treewidth in Minor-Closed Graph Families", Algorithmica, **27**, 275–291 (2000)

[11] J. Friedman, "Constructing $O(n \log n)$ Size Monotone Formulae for the k-th Threshold Function of n Boolean Variables", SIAM J. Comput., **15**, 641–654 (1986)

[12] M.L. Furst, J.B. Saxe and M. Sipser, "Parity, Circuits and the Polynomial-Time Hierarchy", Math. Syst. Theory, **17**, 13–27 (1984)

[13] J. Håstad, "Almost Optimal Lower Bounds for Small Depth Circuits", Proc of the 18th STOC, 6–20 (1986)

[14] S. Janson, T. Luczak and A. Ruciński, "An Exponential Bound for the Probability of Nonexistence of a Specified Subgraph in a Random Graph", Random Graph '87, 73–87 (1990)

[15] S. Janson, T. Luczak and A. Ruciński, Random Graphs, John Wiley & Sons (2000)

[16] M. Krieger, "On the Incompressibility of Monotone DNFs", Theory of Comput. Sys., **41**, 211–231 (2007)

[17] J. Nešetřil and P. O de Mendez, "Linear Time Low Tree-Width Partitions and Algorithmic Consequences", Proc. of the 38th STOC, 391–400 (2006)

[18] J. Nešetřil and S. Poljak, "Complexity of the Subgraph Problem", Comment. Math. Univ. Carol, **26(2)**, 415–420 (1985)

[19] A. Razborov, "Lower Bounds on the Monotone Complexity of Some Boolean Function", Dolk. Akad. Nauk. SSSR, **281(4)**, 598–607 (1985) (in Russian) English Translation in Soviet Math. Dokl. **31**, 354–357 (1985)

[20] A. Razborov, "On the Method of Approximation", Proc. of the 21st STOC, 167–176 (1989)

[21] B. Rossman, "On the Constant-Depth Complexity of k-Clique", Proc. of the 40th STOC, 721–730 (2008)

[22] I. Wegener, The Complexity of Boolean Functions, Wiley-Teubner (1987)

[23] A.C.C. Yao, "Separating the Polynomial-Time Hierarchy by Oracles", Proc. of the 26th FOCS, 1–10 (1985)

6 Appendix

6.1 Proof Sketch of Theorem 4

Proof (sketch). We use the derandomized version of the color-coding method introduced by Alon, Yuster and Zwick [4]. Let $\mathcal{C}$ be a class of colorings $c : V \to \{1, \dots, k\}$ such that every sequence of k vertices $v_1, \ldots, v_k$ chosen from V is colored consecutively by $1, \ldots, k$ in some coloring in $\mathcal{C}$. It is known that such a class of size $|\mathcal{C}| = k^{O(k)} \log |V|$ exists (see e.g., [4, Section 4]). For a graph $G = (V, E)$ that contains an H and a coloring c, an H in G is said to be *properly-colored* under c if the vertices on it are consecutively colored by $1, 2, \ldots, k$.

The output of our function is expressed as

$$\bigvee_{c \in \mathcal{C}} \text{“}G \text{ contains a properly-colored copy of } H \text{ under } c\text{”}.$$

Given a tree-decomposition of H and a coloring c, we can easily construct a constant-depth circuit of size $O(n^{t+1})$ for detecting a properly-colored copy of H in a way as described in [4, Theorem 5.2]. The construction is in the bottom-up way and see [4] for the details. □

6.2 Proof of Lemma 2

We use the following fact that is obvious from the definition of the generalized sensitive core (Definition 2).

Fact 2 *Let $T = T^{f,G}_{\langle s \rangle}(A)$ and suppose that B is a set with $T \subseteq B \subseteq \mathsf{Im}_A(V_H)$ and $|B| \leq s$. Then $f(G \cup H_{A|T}) = f(G \cup H_{A|B})$.* □

Proof (of Lemma 2). Suppose for the contrary that $T(\tilde{g}) \leq s$ for every gate g in C. We first show the following claim.

Claim 1 *For every gate g in C, $g(G \cup H_{A|T(\tilde{g})}) = g(G \cup H_A)$.*

Proof. The claim is shown by the induction on the depth of g. Let g be an input, i.e., $g = x_{u,v}$. If $(u, v) \in G$, then $g(G \cup H_{A|T(\tilde{g})}) = g(G \cup H_A) = 1$. Suppose that $(u, v) \notin G$. If $(u, v) \notin H_A$, then $g(G \cup H_{A|T(\tilde{g})}) = g(G \cup H_A) = 0$. Suppose now that $(u, v) \in H_A$. Then $T(\tilde{g}) = \{u, v\}$, and so $g(G \cup H_{A|T(\tilde{g})}) = g(G \cup H_A) = 1$.

For the induction step, let g be a non-input node and suppose that $h(G \cup H_{A|T(\tilde{h})}) = h(G \cup H_A)$ for every child h of g. Let h be an arbitrary fixed child of g. Since $T(h) \subseteq T(\tilde{g}) \subseteq A$ and $|T(\tilde{g})| \leq s$, we have $h(G \cup H_{A|T(h)}) = h(G \cup H_{A|T(\tilde{g})})$ (by Fact 2). Since $T(h) \subseteq T(\tilde{h}) \subseteq A$ and $|T(\tilde{h})| \leq s$, we also have $h(G \cup H_{A|T(h)}) = h(G \cup H_{A|T(\tilde{h})})$ (again by Fact 2). The induction hypothesis implies that $h(G \cup H_{A|T(\tilde{h})}) = h(G \cup H_A)$. Putting them together we have $h(G \cup H_{A|T(\tilde{g})}) = h(G \cup H_A)$. Since this holds for every child h of g, we can conclude that $g(G \cup H_{A|T(\tilde{g})}) = g(G \cup H_A)$. □

We now go back to the proof of Lemma 2. By applying the claim to the output of C, we have $C(G \cup H_{A|T(\tilde{C})}) = C(G \cup H|_A)$. Since $\emptyset = T(C) \subseteq T(\tilde{C}) \subseteq A$ and $|T(\tilde{C})| \leq s$, we have $C(G \cup H_{A|\emptyset}) = C(G) = C(G \cup H_A)$, which completes the proof. □

6.3 Proof of Lemma 3

The key properties of our extended sensitive core in Definition 2 is the following:

$$\begin{aligned} T^{f,G}_{\langle s \rangle}(A) \neq \emptyset \quad &\Longleftrightarrow \quad \exists V' \subseteq \mathsf{Im}_A(V_H) \text{ with } 2 \leq |V'| \leq s \text{ s.t. } T^{f,G}_{\langle s \rangle}(A, V') \text{ is full,} \\ |T^{f,G}_{\langle s \rangle}(A)| > s \quad &\Longrightarrow \quad \exists V' \subseteq \mathsf{Im}_A(V_H) \text{ with } s+1 \leq |V'| \leq 2s \text{ s.t. } T^{f,G}_{\langle s \rangle}(A, V') \text{ is full,} \end{aligned}$$

which corresponding to statements 1 and 2 of Lemma 3.5 in [21], and can be verified analogously. We can then prove Lemma 3.

Proof (of Lemma 3). The first inequality of the lemma is proved by the following series of inequalities.

$$\begin{aligned}
\Pr_{G,A}[|T^{f,G}_{\langle s\rangle}(A)| > s] &\leq \Pr_{G,A}[\exists V' \subseteq \mathsf{Im}_A(V_H) \text{ with } s+1 \leq |V'| \leq 2s \quad T^{f,G}(A,V') \text{ is full }] \\
&= \Pr_{G,A}[\exists H' \subseteq H \text{ with } s+1 \leq |V(H')| \leq 2s \quad T^{f,G}(A,\mathsf{Im}_A(V(H'))) \text{ is full }] \\
&= \sum_{\substack{H'\subseteq H\\ s+1\leq V(H')\leq 2s}} \Pr_{G,A}[T^{f,G}(A,\mathsf{Im}_A(V(H'))) \text{ is full }] \\
&= \sum_{\substack{H'\subseteq H\\ s+1\leq V(H')\leq 2s}} \Pr_{G,A':V(H')\to V}[T^{f,G}(A',\mathsf{Im}_{A'}(V(H'))) \text{ is full }] \\
&= \sum_{\tilde{s}=s+1}^{2s} \sum_{\substack{H'\subseteq H\\ |V(H')|=\tilde{s}}} \Pr_{G,A':V(H')\to V}[T^{f,G}(A',\mathsf{Im}_{A'}(V(H'))) \text{ is full }] \\
&\leq \sum_{\tilde{s}=s+1}^{2s} \sum_{\substack{H'\subseteq H\\ |V(H')|=\tilde{s}}} n^{\alpha|E(H')|+(\beta-1)s+o(1)} \quad \text{(By Lemma 1)} \\
&\leq \sum_{\tilde{s}=s+1}^{2s} O(1)\cdot n^{-Z_H(\tilde{s})+\varepsilon} \leq n^{-\min_{s+1\leq\tilde{s}\leq 2s} Z_H(\tilde{s})+\varepsilon} = n^{-\tilde{Z}_H(s)+\varepsilon}.
\end{aligned}$$

Similarly, the second inequality of the lemma is proved as follows.

$$\begin{aligned}
\Pr_{G,A}[T^{f,G}_{\langle s\rangle}(A) \neq \emptyset] &= \Pr_{G,A}[\exists V' \subseteq \mathsf{Im}_A(V_H) \text{ with } 2 \leq |V'| \leq s \quad T^{f,G}(A,V') \text{ is full }] \\
&\leq \sum_{\tilde{s}=2}^{s} \sum_{\substack{H'\subseteq H\\ |V(H')|=\tilde{s}}} \Pr_{G,A':V(H')\to V}[T^{f,G}(A',\mathsf{Im}_{A'}(V(H'))) \text{ is full }] \\
&\leq \sum_{\tilde{s}=2}^{s} \sum_{\substack{H'\subseteq H\\ |V(H')|=\tilde{s}}} n^{\alpha|E(H')|+(\beta-1)s+o(1)} \quad \text{(By Lemma 1)} \\
&\leq \sum_{\tilde{s}=2}^{s} O(1)\cdot n^{-Z_H(\tilde{s})+\varepsilon} \leq n^{-\min 2\leq\tilde{s}\leq s Z_H(\tilde{s})+\varepsilon} \leq n^{-\hat{Z}(s)+\varepsilon}.
\end{aligned}$$

This completes the proof of the lemma. □

6.4 Proof of Fact 1

We first show the following fact.

Fact 3 *Let $H' = (V_{H'}, E_{H'})$ be any induced subgraph of $k \times k$ grid. Then $|E_{H'}| \leq 2(|V_{H'}| - \sqrt{|V_{H'}|})$.* □

Proof. Consider the value $\sum_{v\in V_{H'}} \deg(v)$, where $\deg(v)$ denotes the degree of v. Each vertex has four possible neighbors; however, the most northern and the southern vertices in each column *lose* one neighbor each. Similarly, the most eastern and the western vertices in each row lose one neighbor each. Hence we have

$$\sum_{v\in V_{H'}} \deg(v) \quad \leq \quad 4|V_{H'}| - 2(\sharp\text{row} + \sharp\text{column}),$$

where $\sharp$row and $\sharp$column denote the number of rows and of columns spanned by $V_{H'}$. The value in the bracket is minimized when $\sharp\text{row} = \sharp\text{column} = \sqrt{V_{H'}}$, which immediately implies the fact. □

Proof (of Fact 1). The above fact immediately implies that the threshold exponent of $k \times k$ grid is $\frac{k}{2(k-1)}$. The second statement of Fact 1 is shown as follows: Let $F(s) := s^2 - \frac{ks(s-1)}{k-1}$. Then $F(s) \geq Z_H(s^2)$ for every s by the above fact. Let $s = (\sqrt{2}-1)k$. Then $F(s) = F(\sqrt{2}s) = (3\sqrt{2}-4)k$. It is easy to verify that $F(\tilde{s}) \geq (3\sqrt{2}-4)k$ for every $s \leq \tilde{s} \leq \sqrt{2}s$, which implies $Z_H(s') \geq (3\sqrt{2}-4)k$ for every $s^2 \leq s' \leq 2s^2$. □

6.5 Proof for $\mathsf{HypClique}^n_{\ell,k}$

In this subsection, we show the following lower bound described in Section 4.3.

Theorem 7 *For every $k > \ell \geq 2$, every constant-depth circuit that computes $\mathsf{HypClique}^n_{\ell,k}$ contains at least $\Omega(n^{k(1-(\ln \ell+2)/(\ell-1))})$ gates.* □

The theorem easily follows from a more general form of the result below.

Theorem 8 *Let $k > \ell \geq 2$. For every $s = \gamma k$ where $\gamma \leq \left(\frac{1}{2}\right)^{\frac{1}{\ell-1}}$, every constant-depth circuit that computes $\mathsf{HypClique}^n_{\ell,k}$ contains at least $n^{s-\frac{2k\binom{s}{\ell}}{\binom{k}{\ell}}-o(1)}$ gates.*

Proof. (From Theorem 8 to Theorem 7) Put $s = \gamma k$. From Theorem 8, the exponent of the size of circuits for $\mathsf{HypClique}^n_{\ell,k}$ is at least

$$\begin{aligned} s - \frac{2k\binom{s}{\ell}}{\binom{k}{\ell}} - o(1) \quad &= \quad s - 2k \cdot \frac{s}{k} \cdot \frac{s-1}{k-1} \cdots \frac{s-(\ell-1)}{k-(\ell-1)} - o(1) \\ &\geq \quad k\gamma - 2k\gamma^\ell = k\gamma(1-2\gamma^{\ell-1}). \end{aligned}$$

By putting $\gamma = \left(\frac{1}{\ell}\right)^{\frac{1}{\ell-1}}$, the value of the above exponent is

$$\begin{aligned} k\left(\frac{1}{\ell}\right)^{\frac{1}{\ell-1}}\left(1-\frac{2}{\ell}\right) \quad &\geq \quad k \cdot \left(\left(\frac{1}{\ell}\right)^{\frac{1}{\ell}}\right)^{\frac{\ell}{\ell-1}}\left(1-\frac{2}{\ell}\right) \\ &> \quad k\left(1-\frac{\ln \ell}{\ell}\right)^{\frac{\ell}{\ell-1}}\left(1-\frac{2}{\ell-1}\right) \\ &\geq \quad k\left(1-\frac{\ln \ell + 2}{\ell-1}\right) \end{aligned}$$

This completes the proof of Theorem 7. □

It remains to show Theorem 8. The outline of the proof is similar to the proof of the lower bound for the k-clique by Rossman [21], and so we assume some familiarity with [21].

In order to show Theorem 8, we need two technical lemmas.

The first lemma we use is the following which is an alteration of Lemma 5.2 in [21]. Let K^ℓ_A denote the complete ℓ-uniform hypergraph supported on a vertex set A. For $a, b, c \in \mathbb{N}$ such that $a \le \min(b, c)$ and $b + c - a \ge 1$, let $H^\ell_{a,b,c}$ denote the ℓ-uniform hypergraph whose vertex set is $B \cup C$ with $|B| = b, |C| = c$ and $|B \cap C| = a$ and consists of two ℓ-uniform hypercliques supported on B and on C. For $s \in \mathbb{N}$, let W_s denote the set of triples of integers (a, b, c) such that $a \le \min(b, c)$, $\max(b, c) \le s$ and $b + c - a \ge s + 1$.

Lemma 4 *Suppose that $b, c \le s$ and that $\gamma \le \left(\frac{1}{2}\right)^{\ell-1}$. Then $\mathsf{thre}(H^\ell_{a,b,c}) \ge \mathsf{thre}(K^\ell_k)$.*

Proof. By the definition of the threshold exponent, we have

$$\mathsf{thre}(H^\ell_{a,b,c}) \quad \ge \quad \min \frac{b' + c' - a'}{\binom{b'}{\ell} + \binom{c'}{\ell} - \binom{a'}{\ell}},$$

where the minimization ranges over all $a', b', c' \in \mathbb{N}$ with $a' \le b', c'$, $b' \le b$, $c' \le c$ and $b' + c' - a' \ge 1$. Let $s' = \max(b, c)$. Then we have

$$\mathsf{thre}(H^\ell_{a,b,c}) \ge \min_{a'} \frac{2s' - a'}{2\binom{s'}{\ell} - \binom{a'}{\ell}} \ge \frac{s'}{2\binom{s'}{\ell}} \ge \frac{s}{2\binom{s}{\ell}},$$

since $a \le s' \le s$. In order to show the lemma, it is sufficient to show

$$\frac{s}{2\binom{s}{\ell}} \ge \frac{k}{\binom{k}{\ell}},$$

or equivalently

$$\frac{1}{2} \ge \frac{s-1}{k-1} \cdot \frac{s-2}{k-2} \cdots \frac{s-(\ell-1)}{k-(\ell-1)}.$$

This holds if $\frac{s}{k} \le \left(\frac{1}{2}\right)^{\ell-1}$. □

The second lemma we need is Lemma 5, which is a natural extension of the corresponding lemma for the case $\ell = 2$ in [21, Lemma 5.4]. Here we need to extend the definition of the *clique-sensitive core* by Rossman [21, Definition 3.1] to be able to handle hypergraphs. This can be done in an obvious way.

Definition 3 *Let f be a function on ℓ-uniform hypergraphs. Let G be an ℓ-uniform hypergraph, $A \subseteq V(G)$ and $s \in \mathbb{N}$. Define*

$$\begin{aligned} T^{\ell,f,G}(A) &= \left\{a \in A \mid \exists B \subseteq A \text{ s.t. } f(G \cup K^\ell_B) \ne f(G \cup K^\ell_{B\setminus\{a\}})\right\}. \\ T^{\ell,f,G}_{\langle s\rangle}(A) &= \bigcup_{B \subseteq A : |B| \le s} T^{\ell,f,G}(B). \end{aligned}$$

This extension preserves all desired properties of the clique-sensitive core and thus we can show the following lemma. Recall that $\mathcal{H}(\ell, n, p)$ denotes a random ℓ-uniform hypergraph on the vertex set $[n]$ with independent edge probability p.

Lemma 5 *Let ℓ be a constant such that $\ell \geq 2$. Suppose $f = (f_n)_{n\in\mathbb{N}}$ is an sequence of AC_0 functions $f_n : \{0,1\}^{\binom{n}{\ell}} \to \{0,1\}^{n^\beta}$ for some constant $\beta \geq 0$. Let $s \in \mathbb{N}$ and $(a,b,c) \in W_s$ and $\alpha < \mathsf{thre}(H_{a,b,c})$. Then for a random hypergraph $G \in \mathcal{H}(\ell, n, n^{-\alpha})$ and uniform random sets $B \in \binom{[n]}{b}$ and $C \in \binom{[n]}{c}$ with $|B \cap C| = a$,*

$$\Pr[T^{\ell,f,G}(B) = B \textit{ and } T^{\ell,f,G}(C) = C] = n^{\alpha|E_{H_{a,b,c}}|+(\beta-1)|V_{H_{a,b,c}}|+o(1)}.$$

□

The proof of Lemma 5 is analogous to the proof of Lemma 5.4 in [21] and is omitted. Here we only note that we use the following properties of the number of subgraphs in $\mathcal{H}(\ell, n, n^{-\alpha})$ in the proof of Lemma 5. Let X_H denote a random variable that represents the number of copies of H in $G \in \mathcal{H}(\ell, n, n^{-\alpha})$. Note that the expectation of X_H is $\Theta(n^{|V(H)|-\alpha|E(H)|})$. The following theorem can be proved via an obvious extension of a method for deriving the case $\ell = 2$ (e.g., [14]). One can easily show this by applying the method described in [15, Chap 3.1].

Theorem 9 *For every ℓ-uniform hypergraph H and $\alpha > 0$, the following holds for $G \in \mathcal{H}(\ell, n, n^{-\alpha})$ as $n \to \infty$.*

- *if $\alpha > \mathsf{thre}(H)$ then $\Pr[X_H \neq 0] = o(1)$,*
- *if $\alpha < \mathsf{thre}(H)$, then for all $\varepsilon > 0$ $\Pr[X_H < n^{|V(H)|-\alpha|E(H)|-\varepsilon}] = \exp(-n^{\Omega(1)})$.* □

Proof sketch of Theorem 8. Put $s = \gamma k$ with $\gamma \leq \left(\frac{1}{2}\right)^{\frac{1}{\ell-1}}$. Consider a random hypergraph $\mathcal{H}(\ell, n, n^{-\alpha})$ for $\alpha = \mathsf{thre}(K_k^\ell) + \varepsilon = k/\binom{k}{\ell} + \varepsilon$. Suppose that constant-depth circuits of size $O(n^t)$ compute $\mathsf{HypClique}_{\ell,k}^n$, where $t = s - 2k\binom{s}{\ell}/\binom{k}{\ell}$. Without loss of generality, we can assume that all gates in C have fan-in $n^\beta - 1$ for sufficiently small constant $\beta > 0$ (without increasing the size or depth by more than constant factors). Then there is a gate g in C such that

$$\Pr[T^{\ell,f,G}(B) = B \text{ and } T^{\ell,f,G}(C) = C] = \Omega(n^{-t}), \tag{5}$$

where $f : \{0,1\}^{\binom{n}{\ell}} \to \{0,1\}^{n^\beta}$ is a list of outputs of all of g's children and g itself. By Lemma 4, we have $\alpha < \mathsf{thre}(H_{a,b,c})$. Hence by Lemma 5, we have

$$\Pr[T^{\ell,f,G}(B) = B \text{ and } T^{\ell,f,G}(C) = C] = n^{\alpha|E_{H_{a,b,c}}|+(\beta-1)|V_{H_{a,b,c}}|+o(1)}. \tag{6}$$

Since $\beta > 0$ is sufficiently small, the exponent of the RHS in Eq. (6) is at most

$$\frac{k|E_{H_{a,b,c}}|}{\binom{k}{\ell}} - |V_{H_{a,b,c}}|.$$

Since $|E_{H_{a,b,c}}| \leq 2\binom{s}{\ell}$ and $|V_{H_{a,b,c}}| > s$, this is strictly smaller than $-t$. This contradicts Eq. (5), and hence completes the proof. □