Formulas vs. Circuits for Small Distance Connectivity*

Benjamin Rossman†
University of Toronto

June 18, 2018

Abstract

We prove an $n^{\Omega(\log k)}$ lower bound on the AC^0 formula size of Distance $k(n)$ Connectivity for all $k(n) \leq \log\log n$ and formulas up to depth $\log n/(\log\log n)^{O(1)}$. This lower bound strongly separates the power of bounded-depth formulas vs. circuits, since Distance $k(n)$ Connectivity is solvable by polynomial-size AC^0 circuits of depth $O(\log k)$. For all $d(n) \leq \log\log\log n$, it follows that polynomial-size depth-d circuits—which are a semantic subclass of $n^{O(d)}$-size depth-d formulas—are *not* a semantic subclass of $n^{o(d)}$-size formulas of much higher depth $\log n/(\log\log n)^{O(1)}$.

Our lower bound technique probabilistically associates each gate in an AC^0 formula with an object called a *pathset*. We show that with high probability these random pathsets satisfy a family of density constraints called *smallness*, a property akin to low average sensitivity. We then study a complexity measure on small pathsets, which lower bounds the AC^0 formula size of Distance $k(n)$ Connectivity. The heart of our technique is an $n^{\Omega(\log k)}$ lower bound on this pathset complexity measure.

*Journal version of STOC 2014 paper
†Supported by NSERC and the JST ERATO Kawarabayashi Large Graph Project

Contents

1 Introduction

Understanding the relative power of Boolean formulas vs. circuits is a central challenge in complexity theory. Circuits are a powerful model of computation, capable of efficiently simulating Turing machines. On the other hand, formulas (that is, tree-like circuits with fan-out 1) are thought to be a much weaker model of computation. Many natural problems solvable by small circuits, such as st-connectivity, are believed to require large formulas. However, no super-polynomial gap between the formula complexity and circuit complexity of any Boolean function has ever been shown. The existence of such a gap is a major open question.

Question 1.1. *Are polynomial-size Boolean circuits strictly more powerful than polynomial-size Boolean formulas?*

Here we consider non-uniform sequences of circuits and formulas.[1] In terms of complexity classes, this is the question whether NC^1 is a proper subclass of $\mathsf{P/poly}$. (Recall that NC^1 is equivalent to the class of languages computable by polynomial-size Boolean formulas [25].) It is known that Question 1.1 cannot be resolved by simple counting arguments, as Savický and Woods [22] have shown that, for every constant $c > 1$, almost all Boolean functions with formula complexity $\leq n^c$ have circuit complexity $\geq n^c/c$.

The uniform version of Question 1.1 (i.e. whether $\mathsf{uniform}\text{-}\mathsf{NC}^1$ is a proper subclass of P) is also wide open.[2] To answer this question, of course we first need a super-polynomial lower bound on the formula size of any *explicit* Boolean function (say, in the class NP). However, despite the fact that *almost all* Boolean functions have formula complexity $\Omega(2^n/\log n)$ [17], the best lower bound for an explicit function is only $\Omega(n^{3-o(1)})$ [8].

While Question 1.1 remains intriguingly open, in the meantime we can hope to gain insight by studying the question of formulas vs. circuits in restricted settings where powerful lower bound techniques are available. In particular, there are natural analogues of Question 1.1 in both the *monotone setting* and the *bounded-depth setting*.

Monotone Formulas vs. Circuits. The separation of monotone formulas vs. circuits was shown in a classic paper of Karchmer and Wigderson [11] via a lower bound for directed st-connectivity (STCONN).

Theorem 1.2. *Monotone formulas solving* STCONN *require size* $n^{\Omega(\log n)}$.

As it was already known that STCONN has polynomial-size monotone circuits, Theorem 1.2 implies the separation of monotone classes mNC^1 and mP (in fact, it shows $\mathsf{mNC}^1 \neq \mathsf{mAC}^1$). In a notable recent development, Potechin [14] showed that monotone switching networks for STCONN require size $n^{\Omega(\log n)}$. Potechin's result strengthens Theorem 1.2 and implies the sharper separation $\mathsf{mL} \neq \mathsf{mNL}$.

[1] Whenever we speak of a circuits or formulas in this paper, this is understood to mean a sequence $(C_n)_{n=1}^{\infty}$ of circuits, one for each input size n. In the uniform setting, there is an underlying algorithm which, given 1^n as input, outputs a description of the circuit C_n. In the *non-uniform* setting, C_n are arbitrary. All bounds mentioned in this paper may be interpreted in the stronger sense: uniform upper bounds and non-uniform lower bounds.

[2] A separation of NC^1 from $\mathsf{P/poly}$ implies the separation of uniform classes $\mathsf{uniform}\text{-}\mathsf{NC}^1$ from P, as a consequence of the CIRCUIT EVALUATION problem being in P.

Bounded-Depth Formulas vs. Circuits. In the bounded-depth setting, we consider AC^0 circuits and formulas consisting of unbounded fan-in AND and OR gates and NOT gates. Let $\mathsf{Circuits}(s,d)$ (resp. $\mathsf{Formulas}(s,d)$) denote the class of languages computable by AC^0 circuits (resp. formulas) of size at most s and depth at most d. (We measure *circuit size* by the number of gates and *formula size* by the number of leaves.) Every depth-d circuits of size s can be converted into a depth-d formula of size at most s^d by repeatedly replacing overlapping subcircuits with disjoint copies. This brute-force conversion of circuits into formulas implies the following relationship of class for all functions $d(n)$:

$$\mathsf{Formulas}(\mathrm{poly}(n),d) \subseteq \mathsf{Circuits}(\mathrm{poly}(n),d) \subseteq \mathsf{Formulas}(n^{O(d)},d). \tag{1}$$

It is natural to ask whether this relationship of classes is best possible.

Question 1.3. *For which functions $d(n)$ does the containment (1) have a converse of the form*

$$\mathsf{Formulas}(\mathrm{poly}(n),d) \neq \mathsf{Circuits}(\mathrm{poly}(n),d) \textit{ or} \tag{2}$$

$$\mathsf{Circuits}(\mathrm{poly}(n),d) \nsubseteq \mathsf{Formulas}(n^{o(d)},d). \tag{3}$$

For constant $d(n) = O(1)$, the containment (1) implies $\mathsf{Formulas}(\mathrm{poly}(n),d) = \mathsf{Circuits}(\mathrm{poly}(n),d)$. However, there are reasons to believe that (2)—and even the stronger (3)—hold for all super-constant $d(n) = \omega(1)$. In particular, (2) for $d(n) = \log n$ is equivalent to conjecture that $\mathsf{NC}^1 \neq \mathsf{AC}^1$ (which separates NC^1 from $\mathsf{P/poly}$ in a strong way).

As a corollary of our main theorem (Corollary 2.3), we are able to show that (3) holds for all $d(n) \leq \log\log\log n$. (Prior to this paper, even the weaker separation (2) was not known to hold for any super-constant $d(n)$.) In fact, we show something even stronger:

$$\mathsf{Circuits}(\mathrm{poly}(n),d) \nsubseteq \mathsf{Formulas}(n^{o(d)}, \log n/(\log\log n)^6) \tag{4}$$

for the same range of $d(n) \leq \log\log\log n$. In other words, polynomial-size depth-d circuits cannot be simulated by AC^0 formulas of size $n^{o(d)}$ even allowing much greater depth.

In recent work [20] (subsequent to the initial conference publication of this paper), we show that (3) holds for all $d(n) \leq O(\log n/\log\log n)$ and (2) holds for all $d(n) \leq o(\log n)$ using a completely different technique. Similar results for $\mathsf{AC}^0[\oplus]$ formulas vs. circuits were shown by Rossman and Srinivasan [21]. However, in contrast to (4), the techniques in [20, 21] do not imply any separation between circuits of depth d and formulas of depth $d+1$ for any d.

Distance $k(n)$ Connectivity. As with the separation of monotone formulas vs. circuits in [11], our separation of bounded-depth formulas vs. circuits comes by way of a lower bound for (a parameterized version of) st-connectivity. In his survey on graph connectivity [27], Avi Wigderson wrote "Of all computational problems, graph connectivity is the one that has been studied on the largest variety of computational models, such as Turing machines, PRAMs, Boolean circuits, decision trees and communication complexity. It has proven a fertile test case for comparing basic resources such as time vs. space, nondeterminism vs. randomness vs. determinism, and sequential vs. parallel computation." There has been some significant progress in the 20 years since [27]. Notably, Reingold [16] showed that USTCONN (undirected st-connectivity) is in $\mathsf{DSPACE}(\log n)$. However, many basic questions remain open. Chief among these is the space complexity of STCONN. Savitch's theorem [23] that STCONN $\in \mathsf{DSPACE}(\log^2 n)$ is still the best known upper bound.

As for lower bounds, in addition to various results in monotone models of computation [11, 14, 15, 24, 26], there are lower bounds in structured models of computations whose basic operations manipulate pebblings on graphs. One result of this type, due to Edmonds, Poon and Achlioptas [5], gives a tight space lower bound of $\Omega(\log^2 n)$ on the NNJAG model. In arithmetic circuit complexity, lower bounds on the formula size of iterated matrix multiplication (the algebraic cousin of STCONN) were shown in restricted settings by Nisan and Wigderson [13] and Kumar and Saraf [12] among others.

In this paper, we consider a version of STCONN parameterized by distance. For a function $k : \mathbb{N} \to \mathbb{N}$ with $k(n) \le n$, *distance* $k(n)$ *connectivity*, denoted STCONN$(k(n))$, is the following problem: given a directed graph with n vertices and specified vertices s and t, determine whether or not there is a path of length at most $k(n)$ from s to t. In contrast to STCONN and USTCONN, the directed and undirected versions of distance $k(n)$ connectivity are essentially equivalent.[3] It is easy to show that STCONN$(k(n))$ has circuits (moreover, semi-unbounded monotone circuits) of size $O(kn^3)$ and depth $2\log k$ using the recursive doubling (a.k.a. repeated squared) method of Savitch [23]. (At the expense of logarithmic depth, one gets smaller circuits of size $O(kn^{2.38})$ using fast matrix multiplication.)

An important relationship between STCONN and its parameterized version STCONN$(k(n))$ is the fact every algorithm for STCONN$(k(n))$ "scales up" to an algorithm for STCONN by recursive kth powering. Conversely, every lower bound for STCONN "scales down" to a lower bound for STCONN$(k(n))$; in particular, Theorem 1.2 implies that monotone formulas solving STCONN$(k(n))$ require size $n^{\Omega(\log k)}$. This "scaling up" can be stated as the implication

$$\textsc{stconn}(k(n)) \in \mathsf{Circuits}(s,d) \implies \textsc{stconn} \in \mathsf{Circuits}(n^{O(1)} \cdot s, \tfrac{\log n}{\log k} \cdot d).$$

As noted in [27], if STCONN$(k(n))$ has polynomial-size circuits of depth $o(\log k)$, then STCONN has polynomial-size circuits of depth $o(\log n)$ and hence STCONN $\in \mathsf{DSPACE}(o(\log^2 n))$. This observation motivates the following question.

Question 1.4. *What is the minimum depth of polynomial-size circuits solving* STCONN$(k(n))$*?*

Furst, Saxe and Sipser [6] showed that STCONN $\notin \mathsf{AC}^0$ via the reduction from PARITY to STCONN. Via the same reduction, it follows from the PARITY lower bound of Håstad [7] that STCONN$(k(n)) \notin \mathsf{AC}^0$ for $k(n) = (\log n)^{\omega(1)}$. However, this implies nothing when $k(n) = (\log n)^{O(1)}$.

Ajtai [1] proved the first lower bound for small distances $k(n)$, showing that STCONN$(k(n)) \notin \mathsf{AC}^0$ for all super-constant $k(n) = \omega(1)$. By a careful analysis of Ajtai's proof, Bellantoni, Pitassi and Urquhart [3] extracted a lower bound of $\Omega(\log^* k)$ on the depth of polynomial-size circuits solving STCONN$(k(n))$. This was subsequently improved to $\Omega(\log\log k)$ for $k(n) = (\log n)^{O(1)}$ by Beame, Impagliazzo and Pitassi [2], using a special-purpose "connectivity switching lemma" tailored to STCONN$(k(n))$. It was left as an open problem to further narrow the gap between the $O(\log k)$ and $\Omega(\log\log k)$ upper and lower bounds. In this paper, we completely close this gap by proving a lower bound of $\Omega(\log k)$ for all $k(n) \le \log\log n$ (Corollary 2.2). For small but super-constant $k(n)$, we thereby rule out the possibility of showing that STCONN $\in \mathsf{DSPACE}(o(\log^2 n))$ by constructing polynomial-size circuits for STCONN$(k(n))$ of depth $o(\log k)$.

Recently Chen, Oliveira, Servedio and Tan [4] showed a nearly tight $n^{\Omega(k^{1/d}/d)}$ size-depth tradeoff for AC^0 circuits computing STCONN$(k(n))$ via a clever reduction from STCONN$(k(n))$ to an

[3] The reduction from STCONN$(k(n))$ to USTCONN$(k'(n'))$ converts a directed graph on n vertices into a layered undirected graph on $n' = (k+1)n$ vertices where $k'(n') = k(n)$.

unbalanced Sipser function. Their result implies an $\Omega(\log k/\log\log k)$ lower bound on the depth of polynomial-size circuits solving STCONN$(k(n))$, which moreover extends to all $k(n)$. For small depths $d \le \log k/\log\log k$ where the bound of [4] is nontrivial, the $n^{\Omega(k^{1/d}/d)}$ lower bound of [4] is quantitatively stronger than the $n^{\Omega(\log k)}$ bound in the present paper. However, the significance of our result is that it applies to formulas of depths much larger than $\log k$, at which point the circuit size of STCONN$(k(n))$ is only $n^{O(1)}$.

2 Our Results

Our main theorem is a tight lower bound on the AC^0 formula size solving distance $k(n)$ connectivity.

Theorem 2.1 (Main Result)**.** *Formulas of depth* $\log n/(\log\log n)^6$ *solving* STCONN$(k(n))$ *require size* $n^{\Omega(\log k)}$ *for all* $k(n) \le \log\log n$*.*

This lower bound is not only worst-case, it even applies to formulas which STCONN$(k(n))$ in the average-case (see §12). The following two corollaries of Theorem 2.1 were previously mentioned in the introduction. As discussed, these corollaries answer Questions 1.3 and 1.4 for a limited range of $d(n)$ and $k(n)$.

Corollary 2.2. *Polynomial-size circuits solving* STCONN$(k(n))$ *require depth* $\Omega(\log k)$ *for all* $k(n) \le \log\log n$*.*

Proof. For contradiction, assume C is a circuit of size $n^{O(1)}$ and depth $d(n) = o(\log k)$ solving STCONN$(k(n))$ for some $k(n) \le \log\log n$. By the naïve simulation of circuits by formulas, C is equivalent to a depth-d formula of size at most $n^{o(\log k)}$. But since $d = o(\log\log\log n) \le \log n/(\log\log n)^6$, we get a contradiction with Theorem 2.1. □

Corollary 2.3. *For all* $d(n) \le \log\log\log n$*, the class* $\mathsf{Circuits}(\mathrm{poly}(n), d)$ *of functions computable by polynomial-size depth-*d *circuits is not contained in the class* $\mathsf{Formulas}(n^{o(d)}, \log n/(\log\log n)^6)$ *of functions computable by formulas of size* $n^{o(d)}$ *and depth* $\log n/(\log\log n)^6$*.*

Proof. The separating language is STCONN$(k(n))$ where $k(n) = 2^{d/2}$. This is computable by polynomial-size circuits of depth $d(n) = 2\log k$ which implement the recursive-doubling algorithm. However, Theorem 2.1 implies that STCONN$(k(n))$ is not computable by formulas of size $n^{o(d)}$ and depth $\log n/(\log\log n)^6$, noting that $k(n) \le \log\log n$. □

Though we omit the analysis from this paper, we remark that the depth $\log n/(\log\log n)^6$ in our results can be extended to $\frac{c}{k\log k}\log n$ for an absolute constant $c > 0$. We state our results with $\log n/(\log\log n)^6$ since this makes calculations simple and allows us to use Håstad's switching lemma [7] in the standard way (rather than rely on a result of [20], which applies the switching lemma more efficiently to formulas).

3 Proof Overview

Our proof technique is centered on a new notion of *pathset complexity*. Informally, a *pathset* is a subset $\mathcal{A} \subseteq [n]^{k+1}$ whose elements represent potential paths of length k in a graph of size n. The *pathset complexity* of $\mathcal{A}$, denoted $\chi(\mathcal{A})$, measures of the minimum number of operations required to construct $\mathcal{A}$ via unions ($\cup$) and relational join ($\bowtie$), subject to certain density constraints. (The formal definition of $\chi(\mathcal{A})$, given in §5, is not important for this overview.)

The proof of Theorem 2.1 has two parts. Part 1 shows that every bounded-depth formula F solving STCONN($k(n)$) implies an upper bound on the pathset complexity of a certain (random) pathset $\mathcal{A}^\Gamma$. Part 2 is a general lower bound on $\chi(\mathcal{A})$ for arbitrary pathsets $\mathcal{A}$. Combining these two parts, we get the desired $n^{\Omega(\log k)}$ lower bound on the size of F.

Before explaining Parts 1 and 2 in more detail, we state the key property of STCONN($k(n)$) which our proof exploits. Instances for STCONN($k(n)$) are directed graphs with vertex set $[n]$ and distinguished vertices s and t (without loss of generality $s = 1$ and $t = 2$). An *st-path* is a sequence $(x_0, \dots, x_k) \in [n]^{k+1}$ such that $x_0 = s$ and $x_k = t$ and $x_i \neq x_j$ for all $i \neq j$.

Denote by Γ the random directed graph with edge probability $1/n$. (Note that $1/n$ is below the threshold for STCONN($k(n)$), that is, almost surely Γ contains no st-path of length k.) Define $\mathcal{A}^\Gamma$ as the set of st-paths $(x_0, \dots, x_k) \in [n]^{k+1}$ such that

- $(x_0, x_1), \dots, (x_{k-1}, x_k)$ are non-edges of Γ,
- $\Gamma \cup \{(x_0, x_1), \dots, (x_{k-1}, x_k)\}$ contains a unique st-path of length k (namely, $(x_0, \dots, x_k)$).

Then the (average-case) property of STCONN($k(n)$) that our proof exploits is:

Key Property (§6.3): *Almost surely, $\mathcal{A}^\Gamma$ contains 99% of st-paths of length k.*

We now state Parts 1 and 2 of the proof of Theorem 2.1 in more detail.

Part 1 (§6–7): *Suppose F is a formula of depth $\log n/(\log\log n)^{O(1)}$ solving* STCONN($k(n)$)*. Then, almost surely (with respect to Γ),*

$$\text{size}(F) \geq 2^{-O(k^2)} \cdot n^{-O(1)} \cdot \chi(\mathcal{A}^\Gamma). \tag{5}$$

Part 2 (§8–11): *For all pathsets $\mathcal{A} \subseteq [n]^{k+1}$, writing $\delta(\mathcal{A}) := |\mathcal{A}|/n^{k+1}$ for the density of $\mathcal{A}$,*

$$\chi(\mathcal{A}) \geq 2^{-O(2^k)} \cdot n^{\Omega(\log k)} \cdot \delta(\mathcal{A}). \tag{6}$$

Combining (5) and (6) with $\delta(\mathcal{A}^\Gamma) \geq 0.99n^{-2}$ (by the key property), we get the lower bound $\text{size}(F) \geq 2^{-O(2^k)} \cdot n^{\Omega(\log k)}$. Since $2^{-O(2^k)}$ is $n^{-O(1)}$ for $k(n) \leq \log\log n$, Theorem 2.1 is proved.

Part 1 builds on the technique of [18]. An essential new ingredient, which distinguishes formulas from circuits, is a top-down argument (Lemma 6.7) relating formula size to pathset complexity.

For Part 2, we develop a combinatorial framework for studying pathset complexity. This involves analyzing the *union tree* of joins which predominates the construction of a given pathset $\mathcal{A}$. In §8 we define an auxiliary notion of *pathset complexity with respect to a union tree*, denoted $\bar{\chi}(\mathcal{A})$. Part 2 then consists of 2a and 2b:

Part 2a (§9): *For every pathset $\mathcal{A}$, there exists $\mathcal{A}' \subseteq \mathcal{A}$ such that $\chi(\mathcal{A}) \geq \bar{\chi}(\mathcal{A}')$ and $\delta(\mathcal{A}') \geq 2^{-O(2^k)} \cdot \delta(\mathcal{A})$.*

Part 2b (§11): *For all pathsets $\mathcal{A}'$, $\bar{\chi}(\mathcal{A}') \geq n^{\Omega(\log k)} \cdot \delta(\mathcal{A}')$.*

Part 2a is relatively straightforward. This move from χ to $\bar{\chi}$ is precisely where we lose the factor of $2^{O(2^k)}$, which is the reason that our main theorem is limited to $k(n) \leq \log\log n$. Part 2b is the combinatorial heart of the paper. The proof involves an intricate induction on objects called union trees.

Organization of the Paper. Section 4 sets out the basic terminology and notation for the paper. Section 5 introduces the key notion of *pathset complexity.* Sections 6–7 contain Part 1 of the proof of Theorem 2.1. Sections 8–11 contain Part 2 of the proof. We state some conclusions and discuss future directions in Section 12. Three appendices (Sections A–C) contain supplementary material including key examples and relatively easier special cases of our main lower bound.

4 Preliminaries

Let n be an arbitrary positive integer (which we view as growing to infinity). Let $[n] := \{1, \ldots, n\}$. We note that, for all purposes in this paper, $[n]$ may be regarded as an arbitrary fixed set of size n. Let $k = k(n)$ and $d = d(n)$ be arbitrary functions of n. As parameters, k represents *distance* and d represents *depth.* No bound on k or d is assumed throughout the paper; assumptions like $k(n) \leq \log\log n$ are explicitly stated where needed. All constants in asymptotic notation ($O(\cdot)$, etc.) are universal (with no dependence on n, k, d).

Circuits and Formulas. The *circuits* and *formulas* considered in this paper are unbounded fan-in Boolean circuits and formulas with a single output node and NOT gates at the bottom level. Formally, a *circuit* is a finite acyclic directed graph with a unique output (node of out-degree 0) where each input (node of in-degree 0) is labeled by a literal (i.e. X_i or $\overline{X}_i$) and each gate (node of in-degree ≥ 1) is labeled by AND or OR. A *formula* is a tree-like circuit in which every node other than the output has out-degree 1. The *size* of a circuit is the number of gates, while the *size* of a formula is the number of leaves. (For a formula F, the circuit-size of F equals the formula-size of F minus 1.)

Graphs. All *graphs* in this paper are directed graph $G = (V_G, E_G)$ where V_G is a (possibly empty) set and $E_G \subseteq V_G \times V_G$. The edge from v to w is written simply as vw to cut down on unnecessary parentheses.

Two important graphs in this paper are P_k (the directed path of length k) and $P_{k,n}$ (the "complete k-layered graph" with $k+1$ layers of n vertices and kn^2 edges). Formally, let

$$P_k = (V_k, E_k) \text{ where } V_k = \{v_0, \ldots, v_k\} \text{ and } E_k = \{v_i v_{i+1} : 0 \leq i < k\}$$

where $v_0, \ldots, v_k$ are fixed abstract vertices. We will usually omit subscripts writing simply v and vw for arbitrary elements of V_k and E_k. To define $P_{k,n}$, we create $(k+1)n$ fresh vertices denoted v^i for each $v \in V_k$ and $i \in [n]$. Then

$$P_{k,n} = (V_{k,n}, E_{k,n}) \text{ where } V_{k,n} = \{v^i : v \in V_k,\ i \in [n]\} \text{ and } E_{k,n} = \{v^i w^j : vw \in E_k,\ i, j \in [n]\}.$$

We refer to subgraphs $\Gamma \subseteq P_{k,n}$ with $V_\Gamma = V_{k,n}$ as *k-layered graphs.* Throughout the paper, Γ consistently represents a (random) k-layered graph, while G, H, K are reserved for subgraphs of

P_k. We sometimes view Γ as the input to a circuit or formula; in this case, we identify the set of layered graphs with $\{0,1\}^N$ where N is a set of kn^2 variables indexed by elements of $E_{k,n}$.

Layered Distance $k(n)$ Connectivity. As with previous lower bounds for distance $k(n)$ connectivity [1, 2], we consider a variant of the problem on k-layered graphs. Let s, t denote vertices v_0^1, v_k^1 respectively. *Layered distance $k(n)$ connectivity* is the problem of determining whether a layered graph $\Gamma \in \{0,1\}^N$ contains a path from s to t. Following [2], we denote this problem by $\textsc{distconn}(k,n)$. The layered and unlayered versions of distance $k(n)$ connectivity are essentially equivalent.[4] This allows us to restate Theorem 2.1 as a lower bound on $\textsc{distconn}(k,n)$:

Theorem 2.1. (restated) *Formulas of depth* $\log n/(\log\log n)^6$ *solving* $\textsc{distconn}(k,n)$ *require size* $n^{\Omega(\log k)}$ *for all* $k(n) \le \log\log n$.

Boolean Functions and Restrictions. Let $f : \{0,1\}^I \to \{0,1\}$ be a Boolean function where I is an arbitrary finite set (of "variables"). We say that a variable $i \in I$ is *live* with respect to f if there exists $x \in \{0,1\}^I$ such that $f(x) \neq f(x')$ where x' equals x with its ith coordinate flipped. Let $\mathit{Live}(f) := \{i \in I : i \text{ is live w.r.t. } f\}$.

A *restriction* on I is any function $\theta : I \to \{0,1,*\}$. We denote by $f\lceil\theta : \{0,1\}^{\theta^{-1}(*)} \to \{0,1\}$ the function (over the "unrestricted" variables i such that $\theta(i) = *$) obtained from f by applying the restriction θ.

Probabilistic Notation. For a finite set I and $p, q \in [0,1]$, we write:

- $x \in \{0,1\}^I_p$ for the random tuple $x \in \{0,1\}^I$ where $\mathbb{P}[\, x_i = 1\,] = p$ independently for all $i \in I$ (in particular, we will consider the random layered graph $\Gamma \in \{0,1\}^N_{1/n}$),
- $R \subseteq_p I$ for the random subset R of I where $i \in R$ independently with probability p for all $i \in I$,
- $\theta \in \mathcal{R}(p,q)$ for the random restriction $\theta : I \to \{0,1,*\}$ where $\mathbb{P}[\,\theta(i) = *\,] = q$ and $\mathbb{P}[\,\theta(i) = 1\,] = (1-q)p$ for all $i \in I$.

Whenever we say *almost surely*, this is understood to mean *asymptotically almost surely* as $n \to \infty$ (i.e. with probability that goes to 1 as $n \to \infty$).

Tuples and Relations. The following notation pertains to "V-ary" tuples $x \in [n]^V$ and relations $\mathcal{A} \subseteq [n]^V$ where V is an arbitrary finite set.

Definition 4.1 (V-tuples)**.** For $x \in [n]^V$ and $S \subseteq V$, we denote by $x_S \in [n]^S$ the restriction of x to coordinates in S. For $x \in [n]^V$ and $y \in [n]^W$ where $V \cap W = \emptyset$, let $xy \in [n]^{V \cup W}$ denote the unique $z \in [n]^{V\cup W}$ such that $z_i = x_i$ for all $i \in V$ and $z_j = y_j$ for all $j \in W$; here $xy = yx$, as there is no intrinsic linear order on $V \cup W$. We adopt the convention $[n]^\emptyset = \{()\}$ where $()$ denotes the unique $\emptyset$-tuple.

[4] Since k-layered graphs are graphs with $(k+1)n$ vertices, there is a trivial reduction from $\textsc{distconn}(k,n)$ to $\textsc{stconn}(k'(n'))$ where $n' = (k+1)n$ and $k'(n') = k$. In the opposite direction, there is a simple reduction from $\textsc{stconn}(k(n))$ to $\textsc{distconn}(k,n)$ which converts graphs to k-layered graphs.

Definition 4.2 (Join)**.** For finite sets V and W and $\mathcal{A} \subseteq [n]^V$ and $\mathcal{B} \subseteq [n]^W$, the *join* of $\mathcal{A}$ and $\mathcal{B}$ is the set

$$\mathcal{A} \bowtie \mathcal{B} := \{x \in [n]^{V \cup W} : x_V \in \mathcal{A} \text{ and } x_W \in \mathcal{B}\}.$$

The join operation $\bowtie$ is a hybrid of intersection $\cap$ and cartesian product $\times$: if $V = W$ then $\mathcal{A} \bowtie \mathcal{B} = \mathcal{A} \cap \mathcal{B}$, and if $V \cap W = \emptyset$ then $\mathcal{A} \bowtie \mathcal{B}$ is the product $\mathcal{A} \times \mathcal{B}$. Note that $\mathcal{A} \bowtie \emptyset = \emptyset$ and $\mathcal{A} \bowtie \{()\} = \mathcal{A}$.

Definition 4.3 (Density, Projection, Restriction)**.** Let $\mathcal{A} \subseteq [n]^V$.

(i) The *density* of $\mathcal{A}$ is defined by $\delta(\mathcal{A}) := |\mathcal{A}| \,/\, n^{|V|}$.

(ii) For $S \subseteq V$, the *S-projection* and *S-projection density* of $\mathcal{A}$ are defined by

$$\mathrm{proj}_S(\mathcal{A}) := \{x_S : x \in \mathcal{A}\}, \qquad \pi_S(\mathcal{A}) := \delta(\mathrm{proj}_S(\mathcal{A})).$$

That is, $\pi_S(\mathcal{A}) = |\mathrm{proj}_S(\mathcal{A})| \,/\, n^{|S|}$, as δ here refers to the density of the S-ary relation $\mathrm{proj}_S(\mathcal{A}) \subseteq [n]^S$.

(iii) For $S \subseteq V$ and $z \in [n]^{V \setminus S}$, the *$S$-restriction* of $\mathcal{A}$ at z and *maximum S-restriction density* of $\mathcal{A}$ are defined by

$$\mathcal{A}|^z_S := \{y \in [n]^S : yz \in \mathcal{A}\}, \qquad \mu_S(\mathcal{A}) := \max_{z \in [n]^{V \setminus S}} \delta(\mathcal{A}|^z_S).$$

It will be convenient (later on in §10) to extend this notation as follows: for any sets S and $\overline{S}$ such that $S \cap \overline{S} = \emptyset$ and $V \subseteq S \cup \overline{S}$ and any $z \in [n]^{\overline{S}}$, let $\mathcal{A}|^z_S$ be understood as $\mathcal{A}|^{z'}_{V \cap S}$ where $z' = z_{V \cap \overline{S}}$.

The next lemma gives some basic inequalities on projection and restriction densities, which we will use throughout this paper.

Lemma 4.4. *For all $A \subseteq [n]^V$ and $S' \subseteq S \subseteq V$,*

$$\begin{array}{ccccccccc} \delta(\mathcal{A}) = \mu_V(\mathcal{A}) & \leq & \mu_S(\mathcal{A}) & \leq & \mu_{S'}(\mathcal{A}) & \leq & \mu_\emptyset(\mathcal{A}) & = & \begin{cases} 0 & \text{if } \mathcal{A} = \emptyset, \\ 1 & \text{if } \mathcal{A} \neq \emptyset. \end{cases} \\ \| & & \wedge\!| & & \wedge\!| & & \| & & \\ \pi_V(\mathcal{A}) & \leq & \pi_S(\mathcal{A}) & \leq & \pi_{S'}(\mathcal{A}) & \leq & \pi_\emptyset(\mathcal{A}) & & \end{array}$$

In addition, $\mu_{S'}(\mathcal{A}) \leq \mu_{S'}(\mathrm{proj}_S(\mathcal{A}))$.

Another basic inequality is given by the following lemma.

Lemma 4.5. *For all $A \subseteq [n]^V$ and $S \subseteq V$,*

$$\delta(\mathcal{A}) \leq \pi_S(\mathcal{A}) \cdot \mu_{V \setminus S}(\mathcal{A}).$$

Lemmas 4.4 and 4.5 follow easily from definitions once the notation is understood. We conclude this section by stating another basic inequality bounding the density of a join. The following lemma plays a key role later on (when it is reformulated as Lemma 10.9).

Lemma 4.6. *For all $\mathcal{A} \subseteq [n]^V$ and $\mathcal{B} \subseteq [n]^W$ and $S \subseteq V$ and $T \subseteq W$,*

$$\delta(\mathcal{A} \bowtie \mathcal{B}) \leq \pi_S(\mathcal{A}) \cdot \mu_{T \setminus S}(\mathrm{proj}_T(\mathcal{B})) \cdot \mu_{(V \cup W) \setminus (S \cup T)}(\mathcal{A} \bowtie \mathcal{B}).$$

Proof. This inequality is mainly derived by two applications of Lemma 4.5. We first project $\mathcal{A} \bowtie \mathcal{B}$ to $S \cup T$:

$$\delta(\mathcal{A} \bowtie \mathcal{B}) \leq \pi_{S \cup T}(\mathcal{A} \bowtie \mathcal{B}) \cdot \mu_{(V \cup W) \setminus (S \cup T)}(\mathcal{A} \bowtie \mathcal{B}).$$

We then project $\mathrm{proj}_{S \cup T}(\mathcal{A} \bowtie \mathcal{B})$ to S:

$$\pi_{S \cup T}(\mathcal{A} \bowtie \mathcal{B}) \leq \pi_S(\mathcal{A} \bowtie \mathcal{B}) \cdot \mu_{T \setminus S}(\mathrm{proj}_{S \cup T}(\mathcal{A} \bowtie \mathcal{B})).$$

Now note that $\pi_S(\mathcal{A} \bowtie \mathcal{B}) \leq \pi_S(\mathcal{A})$ and

$$\mu_{T \setminus S}(\mathrm{proj}_{S \cup T}(\mathcal{A} \bowtie \mathcal{B})) \leq \mu_{T \setminus S}(\mathrm{proj}_T(\mathcal{A} \bowtie \mathcal{B})) \leq \mu_{T \setminus S}(\mathrm{proj}_T(\mathcal{B})).$$

These inequalities combine to prove the lemma. □

5 Pathset Complexity

In this section, we define the key notion of pathset complexity, state our lower bound for pathset complexity (Theorem 5.8, to be proved in §8–11), and present a matching upper bound (Proposition 5.11).

Definition 5.1 (Pattern Graph)**.** Recall that $P_k = (V_k, E_k)$ is the directed path of length k where $V_k = \{v_i : 0 \leq i \leq k\}$ and $E_k = \{v_i v_{i+1} : 0 \leq i < k\}$. A *pattern graph* is a subgraph of P_k with no isolated vertices. That is, $G = (V_G, E_G)$ is a pattern graph if, and only if, $E_G \subseteq E_k$ and $V_G = \bigcup_{vw \in E_G} \{v, w\}$. We write $\wp_k$ for the set of pattern graphs. (We use the power set notation, since pattern graphs are in 1-1 correspondence with subsets of E_k.)

Note that every pattern graph is a (possibly empty) disjoint union of directed paths of length ≥ 1. We refer to maximal connected subsets of V_G simply as *components* of G. Two important parameters of pattern graphs are the number of components (= the number of maximal paths) and the length of the longest path (= the number of edges in the largest component). These are denoted by

$$\begin{aligned} \Delta_G &:= \#\text{ of components in } G\ (= |V_G| - |E_G|), \\ \ell_G &:= \text{length of the longest path in } G. \end{aligned}$$

Definition 5.2 (Pathset)**.** For a pattern graph G, a *G-pathset* is a set $\mathcal{A} \subseteq [n]^{V_G}$. Let $\mathscr{P}_G$ denote the set of all G-pathsets, i.e., $\mathscr{P}_G$ is the power set of $[n]^{V_G}$. We sometimes simply refer to *pathsets* when the pattern graph G is clear from context.

The intuition for pathsets is as follows. For a pattern graph G, we view each $x \in [n]^{V_G}$ as corresponding to a "lifting" of G inside the complete layered graph $P_{k,n}$, namely isomorphic copy of G with vertex set $\{v^i \in V_{k,n} : i = x_v\}$ and edge set $\{v^i w^j \in E_{k,n} : i = x_v \text{ and } j = x_w\}$. In this view, a pathset $\mathcal{A} \subseteq [n]^{V_G}$ corresponds to a set of liftings of G. The choice to define *pathsets* as relations (subsets of $[n]^{V_G}$) rather than sets of liftings of G (which better matches intuition) allows us more naturally to work with operations $\bowtie$ and proj_S and μ_S, etc.

We next introduce a notion called G-*smallness* of pathsets $\mathcal{A} \in \mathscr{P}_G$. Later on we will show that for any AC^0 computable function $f : \{0,1\}^N \to \{0,1\}$, a certain pathset $\mathcal{A}^{\Gamma}_{f,G}$ randomly associated with f is G-small with very high probability (Lemma 6.6). Roughly speaking, G-smallness of pathsets $\mathcal{A}^{\Gamma}_{f,G}$ is a property akin to low average sensitivity.

Definition 5.3 (G-small)**.**

(i) Let $\varepsilon \in [1/k, 1/2]$ be an arbitrary "smallness parameter", which we fix throughout the rest of this paper. Let $\tilde{n} := n^{1-\varepsilon}$.

(ii) A pathset $\mathcal{A} \in \mathscr{P}_G$ is G-*small* (we simply say *small* when G is understood from context) if, for all $1 \leq t \leq \Delta_G$ and $S \subseteq V_G$ such that S is the union of t components of G, $\mathcal{A}$ satisfies the density constraint

$$\mu_S(\mathcal{A}) \leq \tilde{n}^{-t}, \quad \text{ that is, } \quad \frac{|\{x \in \mathcal{A} : x_{V_G \setminus S} = y\}|}{n^{|S|}} \leq \tilde{n}^{-t} \text{ for all } y \in [n]^{V_G \setminus S}.$$

(iii) The set of G-small pathsets is denoted $\mathscr{P}^{\text{small}}_G$.

As the terminology suggests, G-smallness is a monotone decreasing property (i.e. if $\mathcal{A}$ is G-small, then so is every $\mathcal{A}' \subseteq \mathcal{A}$). Note that G-smallness consists of $2^{\Delta_G} - 1$ density constraints on $\mathcal{A}$, corresponding to the nonempty unions of the Δ_G components of G. Note that for $t = \Delta_G$ and $S = V_G$, the constraint $\mu_S(\mathcal{A}) \leq \tilde{n}^{-t}$ is equivalent to $\delta(\mathcal{A}) \leq \tilde{n}^{-\Delta_G}$. In the special case that G is connected (i.e. $\Delta_G = 1$), $\mathcal{A}$ is G-small $\iff$ $\delta(\mathcal{A}) \leq \tilde{n}^{-1}$.

Regarding the parameter ε, our lower bound for $\textsc{distconn}(k,n)$ will have the form $\tilde{n}^{\Omega(\log k)}$ for formulas of depth $\log n/(\log\log n)^6$. A tighter analysis (not included in this paper) extends this depth $O((\varepsilon/k \log k) \log n)$. Choosing $\varepsilon = 1/k$ or even $1/\log^* k$ produces an $n^{\Omega(\log k)}$ lower bound with the best $1/4.41$ constant, while choosing $\varepsilon = 1/2$ or even $1/100$ achieves the optimal depth $O((1/k \log k) \log n)$ with a slightly worse constant.

Example 5.4. Let G be the pattern graph with components $U = \{v_1, v_2, v_3\}$ and $U' = \{v_5, v_6\}$ (i.e. $V_G = \{v_1, v_2, v_3, v_5, v_6\}$ and $E_G = \{v_1v_2, v_2v_3, v_5v_6\}$). A union tree $\mathcal{A} \in \mathscr{P}_G$ is G-small if, and only if,

$$\delta(\mathcal{A}) \leq \tilde{n}^{-2}, \quad \mu_U(\mathcal{A}) \leq \tilde{n}^{-1}, \quad \mu_{U'}(\mathcal{A}) \leq \tilde{n}^{-1}.$$

For example, the pathset $\mathcal{A}_1 := \{x : x_1 = x_5 = 1\}$ is G-small (here x ranges over $[n]^{V_G}$ and we write x_i for x_{v_i}) since $\delta(\mathcal{A}_1) = n^{-2} < \tilde{n}^{-2}$ and $\mu_U(\mathcal{A}_1) = \mu_{U'}(\mathcal{A}_1) = n^{-1} < \tilde{n}^{-1}$. The pathset $\mathcal{A}_2 := \{x : x_1 = x_5 \text{ and } x_2 = x_6\}$ is G-small as well since $\delta(\mathcal{A}_2) = \mu_U(\mathcal{A}_2) = \mu_{U'}(\mathcal{A}_2) = n^{-2}$. However, pathsets

$$\mathcal{A}_3 := \{x : x_1 = x_2 = 1\}, \quad \mathcal{A}_4 := \{x : x_1 = x_5\}$$

are not G-small since $\mu_{U'}(\mathcal{A}_3) = 1 > \tilde{n}^{-1}$ and $\delta(\mathcal{A}_4) = n^{-1} > \tilde{n}^{-2}$.

The next lemma shows that smallness is preserved under joins. (Note to the reader: Although it is natural to state Lemma 5.5 now, we will not use this lemma until §11.)

Lemma 5.5. *If $\mathcal{A}$ is a small G-pathset and $\mathcal{B}$ is a small H-pathset, then $\mathcal{A} \bowtie \mathcal{B}$ is a small $G \cup H$-pathset.*

Proof. Assume $\mathcal{A}$ is a small G-pathset and $\mathcal{B}$ is a small H-pathset. To show that $\mathcal{A} \bowtie \mathcal{B}$ is a small $G \cup H$-pathset, consider any $1 \le t \le \Delta_{G\cup H}$ and $S \subseteq V_G \cup V_H$ such that S contains t distinct components $U_1, \dots, U_t$ of $G \cup H$. We must show that $\mu_S(\mathcal{A} \bowtie \mathcal{B}) \le \tilde{n}^{-t}$.

Without loss of generality, assume $U_1, \dots, U_t$ are ordered such that, for some $t' \le t$, we have $U_i \cap V_G \neq \emptyset$ for all $1 \le i \le t'$ and $U_j \cap V_G = \emptyset$ for all $t' < j \le t$. Let $S' = S \cap V_G$ and $S'' = U_{t'+1} \cup \dots \cup U_t$. Then S' contains $\ge t'$ components of G, since $U_i \cap V_G$ contains ≥ 1 component of G for all $1 \le i \le t'$. Next note that U_j is a component of H for all $t' < j \le t$, hence S'' is a union of $t - t'$ components of H. By G-smallness of $\mathcal{A}$ and H-smallness of $\mathcal{B}$, it follows that

$$\mu_{S'}(\mathcal{A}) \le \tilde{n}^{-t'} \quad \text{and} \quad \mu_{S''}(\mathcal{B}) \le \tilde{n}^{t'-t}.$$

Now fix $z \in [n]^{(V_G \cup V_H)\setminus S}$ which maximizes $\delta((\mathcal{A} \bowtie \mathcal{B})|^z_S)$. Using the basic properties of restrictions and joins (Lemmas 4.4 and 4.5), we have

$$\begin{aligned}\mu_S(\mathcal{A} \bowtie \mathcal{B}) = \delta((\mathcal{A} \bowtie \mathcal{B})|^z_S) &= \delta((\mathcal{A}|^z_{S'}) \bowtie (\mathcal{B}|^z_{S\cap V_H}))\\ &\le \delta(\mathcal{A}|^z_{S'}) \cdot \mu_{S\setminus V_G}(\mathcal{B}|^z_{S\cap V_H})\\ &\le \mu_{S'}(\mathcal{A}) \cdot \mu_{S''}(\mathcal{B}).\end{aligned}$$

It follows that $\mu_S(\mathcal{A} \bowtie \mathcal{B}) \le \tilde{n}^{-t}$, which completes the proof. $\square$

Definition 5.6 (Pathset Complexity)**.** For every pattern graph G and pathset $\mathcal{A} \in \mathscr{P}_G$, the *pathset complexity* $\chi_G(\mathcal{A})$ of $\mathcal{A}$ with respect to G is defined by the following induction:

(i) If G is the empty graph, then $\chi_G(\mathcal{A}) := 0$.

(ii) If G consists of a single edge, then $\chi_G(\mathcal{A}) := |\mathcal{A}|$.

(iii) If G has ≥ 2 edges, then

$$\chi_G(\mathcal{A}) := \min_{(H_i,K_i,\mathcal{B}_i,\mathcal{C}_i)_i} \sum_i \max\{\chi_{H_i}(\mathcal{B}_i),\, \chi_{K_i}(\mathcal{C}_i)\}$$

where $(H_i, K_i, \mathcal{B}_i, \mathcal{C}_i)_i$ ranges over sequences[5] where

$$H_i, K_i \subset G, \quad H_i \cup K_i = G, \quad \mathcal{B}_i \in \mathscr{P}^{\text{small}}_{H_i}, \quad \mathcal{C}_i \in \mathscr{P}^{\text{small}}_{K_i} \quad \text{and} \quad \mathcal{A} \subseteq \bigcup_i \mathcal{B}_i \bowtie \mathcal{C}_i.$$

In plain language, $(H_i, K_i, \mathcal{B}_i, \mathcal{C}_i)_i$ ranges over coverings of $\mathcal{A}$ by joins of small pathsets over proper subgraphs of G.

Note that pathset complexity satisfies the following inequalities:

(base case)	$\chi_\emptyset(\{()\}) \le 0$ and $\chi_G(\mathcal{A}) \le 1$	if $\lvert E_G\rvert = \lvert\mathcal{A}\rvert = 1$,
(monotonicity)	$\chi_G(\mathcal{A}') \le \chi_G(\mathcal{A})$	if $\mathcal{A}' \subseteq \mathcal{A}$,
(sub-additivity)	$\chi_G(\mathcal{A}_1 \cup \mathcal{A}_2) \le \chi_G(\mathcal{A}_1) + \chi_G(\mathcal{A}_2)$	for all $\mathcal{A}_1, \mathcal{A}_2$,
(join rule)	$\chi_{G\cup H}(\mathcal{A} \bowtie \mathcal{B}) \le \max\{\chi_G(\mathcal{A}),\, \chi_H(\mathcal{B})\}$	if $\mathcal{A} \in \mathscr{P}^{\text{small}}_G$, $\mathcal{B} \in \mathscr{P}^{\text{small}}_H$.

We will refer to these inequalities repeatedly throughout the paper.

[5] indexed by $i \in \{1, \dots, m\}$ for an arbitrary integer m (since the length of this sequence is arbitrary, we simplify notation by leaving the index set unspecified)

Remark 5.7. Pathset complexity has a *dual characterization* as the unique pointwise maximal function from pairs $(G, \mathcal{A})$ to $\mathbb{R}$ which satisfies (base case), (monotonicity), (sub-additivity) and (join rule). We will expand on this observation later in Remark 8.4.

We now state our lower bound on pathset complexity (to be proved in §8–11).

Theorem 5.8 (Pathset Complexity Lower Bound)**.** *For all* $\mathcal{A} \in \mathscr{P}_{P_k}$*,*

$$\chi_{P_k}(\mathcal{A}) \geq \frac{n^{(1/4.41)\log k}}{2^{O(2^k)}} \cdot \delta(\mathcal{A}).$$

In particular, for $k \leq \log\log n$ and non-negligible $\delta(\mathcal{A}) = n^{-O(1)}$, Theorem 5.8 implies $\chi_{P_k}(\mathcal{A}) \geq n^{(1/4.41)\log k - O(1)}$. In a moment, we will give an upper bound (Proposition 5.11) which shows that Theorem 5.8 is tight in the regime of $k \leq \log\log n$ and non-negligible $\delta(\mathcal{A})$. First, a couple of remarks which give a different perspective on the definition of $\chi_G(\mathcal{A})$:

Remark 5.9 (Pathset Complexity as Construction Cost)**.** Pathset complexity can be seen as a minimum construction cost. In this view, the goal is to construct a pathset $\mathcal{A} \in \mathscr{P}_G$ out of the fewest possible "atomic" pathsets (i.e., individual edges). The rules of construction are as follows:

(a) A single "atomic" pathset of the form $\mathcal{A} \in \mathscr{P}_G$ where $|E_G| = |\mathcal{A}| = 1$ may be bought for unit cost.

(b) Once a pathset $\mathcal{A}$ has been constructed, we may freely discard elements from $\mathcal{A}$ (i.e. replace $\mathcal{A}$ with any smaller $\mathcal{A}' \subseteq \mathcal{A}$).

(c) Having constructed two G-pathsets $\mathcal{A}$ and $\mathcal{A}'$, we may merge $\mathcal{A}$ and $\mathcal{A}'$ into a single G-pathset $\mathcal{A} \cup \mathcal{A}'$ (i.e. replace $\mathcal{A}$ and $\mathcal{A}'$ with $\mathcal{A} \cup \mathcal{A}'$) at no additional cost.

(d) Having constructed a G-pathset $\mathcal{A}$ and a H-pathset $\mathcal{B}$, provided both $\mathcal{A}$ and $\mathcal{B}$ are *small*, we may join $\mathcal{A}$ and $\mathcal{B}$ into a single $G \cup H$-pathset $\mathcal{A} \bowtie \mathcal{B}$ paying the maximum construction cost of $\mathcal{A}$ and $\mathcal{B}$.

For a pathset $\mathcal{A} \in \mathscr{P}_G$, $\chi_G(\mathcal{A})$ is equal to the minimum cost of constructing $\mathcal{A}$ according to these rules. Construction rules (a), (b), (c), (d) respectively correspond to inequalities (base case), (monotonicity), (sub-additivity), (join rule). Only applications of rule (a) increase cost (so minimum construction cost = fewest application of rule (a)). Rule (b) can be used to convert a non-small pathset into a small pathset (in order to use rule (d), for example). Note that only rule (c) can increase the density of pathsets.

Remark 5.10 (The Role of Smallness)**.** Suppose we modify construction rule (d) by dropping the smallness constraint on $\mathcal{A}$ and $\mathcal{B}$ (this is equivalent to substituting $\mathscr{P}_{H_i}$ and $\mathscr{P}_{K_i}$ for $\mathscr{P}^{\text{small}}_{H_i}$ and $\mathscr{P}^{\text{small}}_{K_i}$ in Definition 5.6(iii)). We could then construct the complete P_k-pathset $[n]^{V_k}$ at a total cost of kn^2 simply by joining pathsets $[n]^{\{v_i, v_{i+1}\}}$ for $0 \leq i < k$. This shows that the smallness constraint on joins is essential to Theorem 5.8. Intuitively, smallness is responsible for bottlenecks which drive up the cost of constructing sufficiently dense pathsets. However, smallness is not necessarily an obstacle for very sparse pathsets like $[\sqrt{n}]^{P_k}$: since $[\sqrt{n}]^{\{v_i, v_{i+1}\}}$ are small, we can take joins showing $\chi_{P_k}([\sqrt{n}]^{P_k}) \leq kn$.

We conclude this section with an upper bound.

Proposition 5.11 (Pathset Complexity Upper Bound)**.** *For all* $\mathcal{A} \in \mathscr{P}_{P_k}$*,*

$$\chi_{P_k}(\mathcal{A}) \le O(n^{(1/2)\lceil \log k \rceil + 2}).$$

For $k \le \log \log n$ and $\mathcal{A} \in \mathscr{P}_{P_k}$ with $\delta(\mathcal{A}) = n^{-O(1)}$, our lower and upper bounds show that $\chi_{P_k}(\mathcal{A}) = n^{\Theta(\log k)}$ where the constant in $\Theta(\log k)$ is between $\frac{1}{4.41}$ and $\frac{1}{2}$.

Notation 5.12. For a pattern graph G and an integer s, we denote by $G^{\triangleright s}$ the *s-shifted* pattern graph with vertex set $\{v_{i+s} : v_i \in V_G\}$ and edge set $\{v_{i+s}v_{i+s+1} : v_i v_{i+1} \in E_G\}$. For a pathset $\mathcal{A} \in \mathscr{P}_G$, we denote by $\mathcal{A}^{\triangleright s} \in \mathscr{P}_{G^{\triangleright s}}$ the corresponding *s-shifted* pathset. Note that pathset complexity is invariant under shifts (i.e. $\chi_G(\mathcal{A}) = \chi_{G^{\triangleright s}}(\mathcal{A}^{\triangleright s})$).

Proof of Proposition 5.11. For simplicity we assume $\sqrt{n}$ is an integer. For all $k \ge 1$, define $\mathcal{A}_k \in \mathscr{P}^{\text{small}}_{P_k}$ by

$$\mathcal{A}_k := \{x \in [n]^{\{0,\dots,k\}} : x_0, x_k \le \sqrt{n}\}.$$

(Note that $\delta(\mathcal{A}_k) = 1/n < 1/\tilde{n}$, so $\mathcal{A}_k$ is indeed P_k-small.)

Letting $j = \lceil k/2 \rceil$, we have

$$\mathcal{A}_j \bowtie \mathcal{A}^{\triangleright j}_{k-j} = \{x \in [n]^{\{0,\dots,k\}} : x_0, x_j, x_k \le \sqrt{n}\}.$$

Note that $\mathcal{A}_k$ is covered by $\sqrt{n}$ "copies" of $\mathcal{A}_j \bowtie \mathcal{A}^{\triangleright j}_{k-j}$ where, for $1 \le t \le \sqrt{n}$,

$$\text{Copy}_t(\mathcal{A}_j \bowtie \mathcal{A}^{\triangleright j}_{k-j}) := \{x \in [n]^{\{0,\dots,k\}} : x_0, x_k \le \sqrt{n} \text{ and } (t-1)\sqrt{n} < x_j \le t\sqrt{n}\}.$$

Note that pathset complexity is invariant under "copies" in this sense (i.e. χ_G is invariant under the action of coordinate-wise permutations of $[n]$ on $\mathscr{P}_G$):

$$\begin{aligned}
\chi_{P_k}(\text{Copy}_t(\mathcal{A}_j \bowtie \mathcal{A}^{\triangleright j}_{k-j})) &= \chi_{P_k}(\mathcal{A}_j \bowtie \mathcal{A}^{\triangleright j}_{k-j}) && \text{(invariance under "copies")}\\
&\le \max\{\chi_{P_j}(\mathcal{A}_j),\, \chi_{P^{\triangleright j}_{k-j}}(\mathcal{A}^{\triangleright j}_{k-j})\} && \text{(join rule)}\\
&\le \max\{\chi_{P_j}(\mathcal{A}_j),\, \chi_{P_{k-j}}(\mathcal{A}_{k-j})\} && \text{(invariance under shifts).}
\end{aligned}$$

Since $\mathcal{A}_k \subseteq \bigcup_{1 \le t \le \sqrt{n}} \text{Copy}_t(\mathcal{A}_j \bowtie \mathcal{A}^{\triangleright j}_{k-j})$, sub-additivity of χ implies

$$\chi_{P_k}(\mathcal{A}_k) \le \sum_{1 \le t \le \sqrt{n}} \chi_{P_k}(\text{Copy}_t(\mathcal{A}_j \bowtie \mathcal{A}^{\triangleright j}_{k-j})) \le \sqrt{n} \cdot \max\{\chi_{P_j}(\mathcal{A}_j),\, \chi_{P_{k-j}}(\mathcal{A}_{k-j})\}.$$

This recurrence implies

$$\chi_{P_k}(\mathcal{A}_k) \le (\sqrt{n})^{\lceil \log k \rceil} \cdot \chi_{P_1}(\mathcal{A}_1) = O(n^{(1/2)\lceil \log k \rceil + 1}).$$

Now note that the complete P_k-pathset $[n]^{V_k}$ is covered by n "copies" of P_k. Therefore, by a similar argument,

$$\chi_{P_k}([n]^{V_k}) \le n \cdot \chi_{P_k}(\mathcal{A}) = O(n^{(1/2)\lceil \log k \rceil + 2}).$$

Finally, monotonicity of χ implies that $\chi_{P_k}(\mathcal{A}) \le O(n^{(1/2)\lceil \log k \rceil + 2})$ for all $\mathcal{A} \in \mathscr{P}_{P_k}$. $\square$

6 From Formulas to Pathset Complexity

In this section we derive our main result (Theorem 2.1) from our lower bound on pathset complexity (Theorem 5.8). Let F_0 be a formula of depth $d(n)$ which solves $\textsc{distconn}(k,n)$ where $k(n) \le \log\log n$ and $d(n) \le \log n/(\log\log n)^6$. We must show that F_0 has size $n^{\Omega(\log k)}$.

As a first preliminary step: without loss of generality, we assume that F_0 has minimal size among all depth $d(n)$ formulas solving $\textsc{distconn}(k,n)$. In particular, we have $\text{size}(F_0) \le kn^{k-1}$ since $\textsc{distconn}(k,n)$ has DNFs of this size.

As a second preliminary step, we convert F_0 into a fan-in 2 formula F by replacing each unbounded fan-in AND/OR gate by a balanced binary tree of fan-in 2 AND/OR gates. We have

$$\text{size}(F) = \text{size}(F_0) \le n^k \quad \text{ and } \quad \text{depth}(F) \le \text{depth}(F_0) \cdot \log(\text{size}(F_0)) \le \log^2 n.$$

We write F_{in} for the set of inputs (i.e. leaves) in F, and F_{gate} for the set of gates in F, and f_{out} for the output gate in F. Note that each $f \in F$ is computed by an (unbounded fan-in) formula of size $\le n^k$ and depth $\le d(n)$ (by collapsing all adjacent AND/OR gates below f).

In order to lower bound size(F) in terms of pathset complexity, we define a family of pathsets $\mathcal{A}^{\Gamma}_{f,G}$ associated with each $f \in F$ and $G \in \wp_k$ and $\Gamma \in \{0,1\}^N$. Recall that we identify $\{0,1\}^N$ with the set of k-layered graphs where $N = E_{k,n} = \{v^i w^j : vw \in E_k,\ i,j \in [n]\}$.

Definition 6.1 (Pathsets $\mathcal{A}^{\Gamma}_{f,G}$)**.** For all $G \in \wp_k$ and $x \in [n]^{V_G}$ and $\Gamma \in \{0,1\}^N$ and $f \in F$:

(i) Let $N_{G,x} := \{v^i w^j \in N : i = x_v \text{ and } j = x_w\}$ $(= \{v^{x_v} w^{x_w} : vw \in E_G\})$.

(ii) Let $\rho^{\Gamma}_{G,x} : N \to \{0,1,*\}$ be the restriction which equals $*$ over $N_{G,x}$ and agrees with Γ over $N \setminus N_{G,x}$. In particular, applying $\rho^{\Gamma}_{G,x}$ to f, we get a function $f\lceil\rho^{\Gamma}_{G,x} : \{0,1\}^{N_{G,x}} \to \{0,1\}$ (whose variables correspond to edges of G via the bijection $N_{G,x} \cong E_G$).

(iii) Let $\mathcal{A}^{\Gamma}_{f,G}$ be the G-pathset defined by

$$\mathcal{A}^{\Gamma}_{f,G} := \{x \in [n]^{V_G} : \mathit{Live}(f\lceil\rho^{\Gamma}_{G,x}) = N_{G,x}\}.$$

That is, $\mathcal{A}^{\Gamma}_{f,G}$ is the set of $x \in [n]^{V_G}$ such that the restricted function $f\lceil\rho^{\Gamma}_{G,x}$ depends on all $|N_{G,x}|$ $(= |E_G|)$ of its variables.

In the next three subsections, we prove a sequence of claims about pathsets $\mathcal{A}^{\Gamma}_{f,G}$ in three cases where $f \in F_{\text{in}}$ and $f \in F_{\text{gate}}$ and $f = f_{\text{out}}$.

Remark 6.2. Claims 6.3, 6.4, 6.5 rely on few assumptions about F. In particular, these claims do not depend on the assumption that F_0 has bounded depth (i.e. F has bounded alternations), nor even that F is a formula as opposed to a circuit. In fact, these claims are valid if F is any B_2-circuit computing $\textsc{distconn}(k,n)$ where B_2 is the full binary basis.

Of course, we will eventually use both assumptions that (I) F_0 has bounded depth (i.e. F has bounded alternations), and (II) F is a formula as opposed to a circuit. Our main technical lemma (Lemma 6.6) relies on (I) but not (II) (not surprisingly, since the proof uses the Switching Lemma, which does not distinguish between circuits and formulas). A second key lemma (Lemma 6.7) relies on (II) but not (I) (using a top-down argument which only works for formulas).

6.1 Inputs of F

Suppose f is an input in F labeled by a literal (i.e. a variable or its negation) corresponding to some $v^i w^j \in N$. Then we have the following explicit description of $\mathcal{A}^{\Gamma}_{f,G}$:

- if G is the empty graph, then $\mathcal{A}^{\Gamma}_{f,G} = \{()\}$ (i.e. the singleton containing the 0-tuple),
- if $E_G = \{vw\}$, then $\mathcal{A}^{\Gamma}_{f,G} = \{x\}$ for the unique $x \in [n]^{\{v,w\}}$ with $x_v = i$ and $x_w = j$,
- otherwise (i.e. if $|E_G| \geq 2$ or E_G consists of an single edge other than vw), we have $\mathcal{A}^{\Gamma}_{f,G} = \emptyset$.

By the base case conditions (i) and (ii) in Definition 5.6 of pathset complexity, we have $\chi_\emptyset(\mathcal{A}) = 0$ and $\chi_G(\mathcal{A}) = |\mathcal{A}|$ if G has a single edge. The upshot of these observations is the following claim.

Claim 6.3 (Inputs of F)**.** *For all $f \in F_{\text{in}}$, $\sum_{G \in \wp_k} \chi_G(\mathcal{A}^{\Gamma}_{f,G}) = 1$.*

6.2 Gates of F

Suppose f is an AND or OR gate in F with children f_1 and f_2. Consider any $G \in \wp_k$ and $x \in \mathcal{A}^{\Gamma}_{f,G}$ (assuming $\mathcal{A}^{\Gamma}_{f,G}$ is nonempty). By definition of $\mathcal{A}^{\Gamma}_{f,G}$, the function $f\lceil\rho^{\Gamma}_{G,x} : \{0,1\}^{N_{G,x}} \to \{0,1\}$ depends on all variables in $N_{G,x}$. Since $f\lceil\rho^{\Gamma}_{G,x}$ is the AND or OR of functions $f_1\lceil\rho^{\Gamma}_{G,x}$ and $f_2\lceil\rho^{\Gamma}_{G,x}$, each variable in $N_{G,x}$ is a live variable for one or both $f_1\lceil\rho^{\Gamma}_{G,x}$ and $f_2\lceil\rho^{\Gamma}_{G,x}$.

Define sub-pattern graph $G_1 \subseteq G$ as follows: for each $vw \in E_G$, let vw be an edge in G_1 if and only if $v^{x_v} w^{x_w}$ ($\in N_{G,x}$) is a live variable for the function $f_1\lceil\rho^{\Gamma}_{G,x}$. Define $G_2 \subseteq G$ in the same way with respect to f_2. Since

$$\{v^{x_v} w^{x_w} : vw \in E_G\} = N_{G,x} = \mathit{Live}(f\lceil\rho^{\Gamma}_{G,x}) = \mathit{Live}(f_1\lceil\rho^{\Gamma}_{G,x}) \cup \mathit{Live}(f_2\lceil\rho^{\Gamma}_{G,x}),$$

it follows that $G_1 \cup G_2 = G$.

Let $y = x_{V_{G_1}}$ be the restriction of x ($\in [n]^{V_G}$) to coordinates in V_{G_1}. By definition of G_1, we have

- $v^{y_v} w^{y_w} = v^{x_v} w^{x_w} \in \mathit{Live}(f_1\lceil\rho^{\Gamma}_{G,x})$ for all $vw \in E_{G_1}$, and
- $v^{x_v} w^{x_w} \notin \mathit{Live}(f_1\lceil\rho^{\Gamma}_{G,x})$ for all $vw \in E_G \setminus E_{G_1}$.

It follows that $\mathit{Live}(f_1\lceil\rho^{\Gamma}_{G_1,y}) = \mathit{Live}(f_1\lceil\rho^{\Gamma}_{G,x}) = N_{G_1,y}$, hence $y \in \mathcal{A}^{\Gamma}_{f_1,G_1}$. Similarly, for $z = x_{V_{G_2}}$, we have $z \in \mathcal{A}^{\Gamma}_{f_2,G_2}$. This shows that $x \in \mathcal{A}^{\Gamma}_{f_1,G_1} \bowtie \mathcal{A}^{\Gamma}_{f_2,G_2}$.

The observation may be succinctly expressed as

$$\mathcal{A}^{\Gamma}_{f,G} \subseteq \bigcup_{G_1,G_2 \subseteq G \,:\, G_1 \cup G_2 = G} \mathcal{A}^{\Gamma}_{f_1,G_1} \bowtie \mathcal{A}^{\Gamma}_{f_2,G_2}.$$

Splitting this union into the cases that $G_1 = G$ or $G_2 = G$ or $G_1, G_2 \subset G$, we have proved:

Claim 6.4 (Gates of F)**.** *For every $f \in F_{\text{gates}}$ with children f_1, f_2 and every $G \in \wp_k$,*

$$\mathcal{A}^{\Gamma}_{f,G} \subseteq \mathcal{A}^{\Gamma}_{f_1,G} \cup \mathcal{A}^{\Gamma}_{f_2,G} \cup \bigcup_{G_1,G_2 \subset G \,:\, G_1 \cup G_2 = G} \mathcal{A}^{\Gamma}_{f_1,G_1} \bowtie \mathcal{A}^{\Gamma}_{f_2,G_2}.$$

6.3 Output of F

We now use the fact that F computes $\text{DISTCONN}(k,n)$. Our previous Claims 6.3 and 6.4 applied to arbitrary $\Gamma \in \{0,1\}^N$. We now shift perspective and consider *random* $\Gamma \in \{0,1\}^N_{1/n}$. That is, Γ is the random k-layered graph (i.e. subgraph of $P_{k,n}$) with edge probability $1/n$. Recall that $V_{k,n} = \{v^i : v \in V_k \text{ and } i \in [n]\}$ and s,t are the vertices v_0^1, v_k^1. Each $x \in [n]^{V_k}$ corresponds to a path of length k in $P_{k,n}$, where x is an st-path if and only if $x_0 = x_k = 1$ (writing x_i instead of x_{v_i} for the coordinates of x).

Observe that Γ almost surely contains no st-path. To see this, note that there are n^{k-1} potential st-paths, each of which is present in Γ with probability $(1/n)^k$. Therefore, the probability that Γ contains an st-path is at most $(1/n) = o(1)$.

For $x \in [n]^{V_k}$ with $x_0 = x_k = 1$, let us say that x is Γ-*independent* if Γ contains no path from $v_i^{(x_i)}$ to $v_i^{(x_j)}$ for any $0 \le i < j \le k$. By another simple union bound, the probability that a given x is Γ-independent is at most $\binom{k}{2}/n = o(1)$. It follows that almost surely 0.99 fraction of potential st-paths are Γ-independent.

Now suppose that Γ contains no st-path and that $x \in [n]^{V_k}$ with $x_0 = x_k = 1$ is a Γ-independent. Let $e_1,\dots,e_k$ be the k edges in x (i.e. $e_i = v_{i-1}^{(x_{i-1})} v_i^{(x_i)}$). We claim that $\Gamma \cup \{e_1,\dots,e_{i-1},e_{i+1},\dots,e_k\}$ contains no st-path for all $1 \le i \le k$. To see this, assume for the sake of contradiction that x' is an st-path in $\Gamma \cup \{e_1,\dots,e_{i-1},e_{i+1},\dots,e_k\}$. Let $e'_1,\dots,e'_k$ be the edges of x'. Since e_i is a non-edge of Γ, we have $e_i \ne e'_i$. Starting at the endpoint of e'_i, we can follow the path x' forwards until reaching a vertex in x; we can also follow x' backwards from the initial vertex of e'_i until reaching a vertex in x. This segment of x' is a path in Γ between two vertices of x, contradiction Γ-independence of x.

Since f_{out} computes $\text{DISTCONN}(k,n)$, it follows that

$$f_{\text{out}}(\Gamma \cup \{e_1,\dots,e_k\}) = 1 \quad \text{and} \quad f_{\text{out}}(\Gamma \cup \{e_1,\dots,e_{i-1},e_{i+1},\dots,e_k\}) = 0 \text{ for all } 1 \le i \le k.$$

This shows that the restricted function $f_{\text{out}}\lceil\rho^{\Gamma}_{P_k,x}$ depends on all k unrestricted variables (corresponding to the edges of x); in fact, $f_{\text{out}}\lceil\rho^{\Gamma}_{P_k,x}$ is the AND function. Therefore, $x \in \mathcal{A}^{\Gamma}_{f_{\text{out}},P_k}$ for every Γ-independent st-path x.

By this argument, we have proved:

Claim 6.5 (Output of F)**.** $\lim_{n\to\infty} \mathbb{P}_{\Gamma \in \{0,1\}^N_{1/n}} [\,\delta(\mathcal{A}^{\Gamma}_{f_{\text{out}},P_k}) \ge 0.99 n^{-2}\,] = 1.$

6.4 Reduction to Pathset Complexity

We now present the two main lemmas in the reduction from formula size to pathset complexity. Lemma 6.6, below, is the main technical lemma (the proof, which relies in part on the switching lemma, is given in §7). This lemma is the only place in the overall proof of Theorem 5.8 which depends on the assumption that F has bounded depth (though not on the fact that F is a formula as opposed to a circuit).

Lemma 6.6 (Pathsets $\mathcal{A}^{\Gamma}_{f,G}$ are Small)**.** *Suppose $f : \{0,1\}^N \to \{0,1\}$ is computable a circuit of size n^k and depth $\log n/(\log\log n)^6$. Then, for all $G \in \wp_k$,*

$$\mathbb{P}_{\Gamma \in \{0,1\}^N_{1/n}} [\,\mathcal{A}^{\Gamma}_{f,G} \textit{ is not } G\textit{-small}] \le O(n^{-2k}).$$

Lemma 6.7, below, is the nexus between formula size and pathset complexity. The proof involves a top-down argument, which is key to distinguishing formulas and circuits. (Though we will apply Lemma 6.7 to the formula F which we have been considering so far, Lemma 6.7 is stated in general terms for arbitrary Boolean functions with fan-in 2.)

Lemma 6.7. *Let F be any fan-in 2 formula and let $\Gamma \in \{0,1\}^N$. If $\mathcal{A}^{\Gamma}_{f,G} \in \mathscr{P}^{\text{small}}_{G}$ for all $f \in F$ and $G \in \wp_k$, then*

$$\chi_{P_k}(\mathcal{A}^{\Gamma}_{f_{\text{out}},P_k}) \leq 2^{O(k^2)} \cdot \text{depth}(F)^k \cdot \text{size}(F).$$

Proof. Assume $\mathcal{A}^{\Gamma}_{f,G} \in \mathscr{P}^{\text{small}}_{G}$ for all $f \in F$ and $G \in \wp_k$. Consider any $f \in F_{\text{gates}}$ with children f_1 and f_2. By Claim 6.4, together with the key properties (monotonicity), (sub-additivity) and (join rule) of pathset complexity, we have

$$\begin{aligned}
\chi_G(\mathcal{A}^{\Gamma}_{f,G}) &\leq \chi_G\Big(\mathcal{A}^{\Gamma}_{f_1,G} \cup \mathcal{A}^{\Gamma}_{f_2,G} \cup \bigcup_{G_1,G_2 \subset G\,:\, G_1 \cup G_2 = G} \mathcal{A}^{\Gamma}_{f_1,G_1} \bowtie \mathcal{A}^{\Gamma}_{f_2,G_2}\Big)\\
&\leq \chi_G(\mathcal{A}^{\Gamma}_{f_1,G}) + \chi_G(\mathcal{A}^{\Gamma}_{f_2,G}) + \sum_{G_1,G_2 \subset G\,:\, G_1 \cup G_2 = G} \Big(\chi_{G_1}(\mathcal{A}^{\Gamma}_{f_1,G_1}) + \chi_{G_2}(\mathcal{A}^{\Gamma}_{f_2,G_2})\Big)\\
&\leq \Big(\chi_G(\mathcal{A}^{\Gamma}_{f_1,G}) + 2^k \sum_{H \subset G} \chi_H(\mathcal{A}^{\Gamma}_{f_1,H})\Big) + \Big(\chi_G(\mathcal{A}^{\Gamma}_{f_2,G}) + 2^k \sum_{H \subset G} \chi_H(\mathcal{A}^{\Gamma}_{f_2,H})\Big).
\end{aligned}$$

If we start from $\chi_{P_k}(\mathcal{A}^{\Gamma}_{f_{\text{out}},P_k})$ and repeatedly apply the above inequality until reaching the inputs of F, we get a bound of the form

$$\chi_{P_k}(\mathcal{A}^{\Gamma}_{f_{\text{out}},P_k}) \leq \sum_{f \in F_{\text{in}},\, G \in \wp_k} c_{f,G} \cdot \chi_G(\mathcal{A}^{\Gamma}_{f,G})$$

for some $c_{f,G} \in \mathbb{Z}_{\geq 0}$. We claim that

$$c_{f,G} \leq \sum_{i,H_0,\dots,H_i\,:\, P_k = H_0 \supset \cdots \supset H_i = G} 2^{ik} \cdot \binom{\text{depth of } f \text{ in } F}{i} \leq 2^{O(k^2)} \cdot \text{depth}(F)^k.$$

To see this, consider any $f \in F_{\text{in}}$ and $G \in \wp_k$ and let $f_{\text{out}} = f_0, \dots, f_d = f$ be the branch in F from the output gate down to f. Then in the expansion of $\chi_{P_k}(\mathcal{A}^{\Gamma}_{f_{\text{out}},P_k})$, we get a contribution of 2^{ik} ($\leq 2^{k^2}$) from each sequence $(i, t_0, H_0, t_1, H_1, \dots, t_i, H_i)$ where $0 = t_0 < \cdots < t_i = d$ and $P_k = H_0 \supset \cdots \supset H_i = G$; here t_i is the location where the expansion of $\chi_{P_k}(\mathcal{A}^{\Gamma}_{f_{\text{out}},P_k})$ branches as we move from $\chi_{H_{i-1}}(\mathcal{A}^{\Gamma}_{f_{i-1},H_{i-1}})$ to $2^k \chi_{H_i}(\mathcal{A}^{\Gamma}_{f_i,H_i})$. Finally, we bound the number of $(t_0, \dots, t_i)$ by $\binom{d}{i}$ ($\leq \text{depth}(F)^k$) and the number of $(H_0, \dots, H_i)$ by 2^{ik} ($\leq 2^{k^2}$). Summing over i adds only a factor of k, so in total we get $c_{f,G} \leq 2^{O(k^2)} \cdot \text{depth}(F)^k$.

We now use the fact that $\sum_{G \in \wp_k} \chi_G(\mathcal{A}^{\Gamma}_{f,G}) = 1$ for all $f \in F_{\text{in}}$ (Claim 6.3) and $\text{size}(F) = |F_{\text{in}}|$ (since F is a formula!). Concluding the proof, we have

$$\chi_{P_k}(\mathcal{A}^{\Gamma}_{f_{\text{out}},P_k}) \leq 2^{O(k^2)} \cdot \text{depth}(F)^k \cdot \text{size}(F). \qquad \square$$

We conclude this section by giving the proof of Theorem 2.1 assuming our pathset complexity lower bound (Theorem 5.8) and main technical lemma (Lemma 6.6).

Reduction 6.8. Theorem 5.8 and Lemma 6.6 $\Longrightarrow$ Theorem 2.1.

Proof. Assuming Theorem 5.8 and Lemma 6.6, we must show that $\mathrm{size}(F) \geq n^{\Omega(\log k)}$. By Claim 6.5 and Lemma 6.6, there exists $\Gamma \in \{0,1\}^N$ such that $\delta(\mathcal{A}^{\Gamma}_{f_{\mathrm{out}},P_k}) \geq 0.99n^{-2}$ and $\mathcal{A}^{\Gamma}_{f,G} \in \mathscr{P}^{\mathrm{small}}_G$ for all $f \in F$ and $G \in \wp_k$. Fix any such Γ. We now have

$$\begin{aligned} \mathrm{size}(F) &\geq \frac{1}{2^{O(k^2)} \cdot \mathrm{depth}(f)^k} \cdot \chi_{P_k}(\mathcal{A}^{\Gamma}_{f_{\mathrm{out}},P_k}) && \text{(Lemma 6.7)} \\ &\geq \frac{1}{2^{O(k^2)} \cdot \mathrm{depth}(f)^k} \cdot \frac{n^{(1/4.41)\log k}}{2^{O(2^k)}} \cdot \delta(\mathcal{A}^{\Gamma}_{f_{\mathrm{out}},P_k}) && \text{(Theorem 5.8).} \end{aligned}$$

Using inequalities

$$\mathrm{depth}(F) \leq \log^2 n, \qquad \delta(\mathcal{A}^{\Gamma}_{f_{\mathrm{out}},P_k}) \geq 0.99n^{-2}, \qquad k \leq \log\log n,$$

we get the desired bound $\mathrm{size}(F) \geq n^{(1/4.41)\log k - O(1)}$. □

7 Small Pathsets from Random Restrictions

In this section, we prove Lemma 6.6 showing that, with high probability over random $\Gamma \in \{0,1\}^N_{1/n}$, pathsets $\mathcal{A}^{\Gamma}_{f,G}$ are small for all $f \in F$ and $G \in \wp_k$. The proof has the following scheme:

$$\begin{array}{|c|} \hline \underbrace{\begin{array}{ccc} \text{Janson's Inequality [9]} & \qquad & \text{Switching Lemma [7]} \\ \Downarrow & & \Downarrow \\ \text{Lemma 7.3} & & \text{Lemma 7.4} \end{array}} \\ \Downarrow \\ \text{Preliminary Lemma 7.5} \\ \Downarrow \\ \text{Main Technical Lemma 6.6} \\ \hline \end{array}$$

The central argument is contained in the proof of Preliminary Lemma 7.5 (from which Lemma 6.6 essentially follows as a corollary).

Remark 7.1. Lemma 6.6 is similar to the main technical lemma in the k-clique lower bound of [18]. One important difference is that here we require a upper tail bound where a upper bound on expectation was sufficient for the k-clique result. (Moreover, the tail bound that we require is stronger than what one gets by applying Markov's inequality to the upper bound on expectation.)

Recall that G-smallness consists of $2^{\Delta_G} - 1$ density constraints corresponding to the nonempty unions of components of G. We say that non-small pathset $\mathcal{A}$ is *G-critical* if it violates only the "top" constraint $\delta(\mathcal{A}) > \tilde{n}^{-\Delta_G}$. Formally:

Definition 7.2 (Critical Pathsets)**.** For a pattern graph G and pathset $\mathcal{A} \in \mathscr{P}_G$, we say that $\mathcal{A}$ is *G-critical* if $\delta(\mathcal{A}) > \tilde{n}^{-\Delta_G}$ and $\mu_{V_G \setminus S}(\mathcal{A}) \leq \tilde{n}^{s-\Delta_G}$ for all $1 \leq s < \Delta_G$ and every $S \subseteq V_G$ such that S is the union of s components of G.

Observe that if $\mathcal{A} \in \mathscr{P}_G$ is G-critical and $S' \subseteq V_G$ has nonempty intersection with exactly $s \geq 1$ components of G, then $\mu_{V_G \setminus S'}(\mathcal{A}) \leq \tilde{n}^{s-\Delta_G}$. This is seen by letting $S \subseteq V_G$ be the union of components which S' intersects and noting that $\mu_{V_G \setminus S'}(\mathcal{A}) \leq \mu_{V_G \setminus S}(\mathcal{A}) \leq \tilde{n}^{s-\Delta_G}$ by Lemma 4.4 and the definition of G-critical.

The following lemma, to be proved in §7.1, gives a concentration-of-measure inequality for critical pathsets. Recall that $\varepsilon \in [1/k, 1/2]$ is an arbitrary "smallness parameter" chosen in Def. 5.3(i).

Lemma 7.3. *Let G be a pattern graph and suppose $\mathcal{A}$ is a G-critical pathset. Let $q = (1/n)^{1+(\varepsilon/2k)}$. Then*

$$\mathop{\mathbb{P}}_{R \subseteq_q N} \Big[\#\big\{x \in \mathcal{A} : N_{G,x} \subseteq R\big\} \leq \frac{n^{\varepsilon/2}}{2} \Big] \leq \exp\Big(-\Omega\Big(\frac{n^{\varepsilon/2}}{2^k}\Big)\Big).$$

We now give the core argument in the proof of Lemma 6.6. We will use the following consequence of Håstad's Switching Lemma [7].

Lemma 7.4 (Consequence of the Switching Lemma)**.** *If a Boolean function f is computable by AC^0 circuits of depth d and size s, then for all $0 \leq q \leq p \leq 1/2$,*

$$\mathop{\mathbb{P}}_{\theta \in \mathcal{R}(p,q)} \big[\ |Live(f\lceil\theta)| \geq 2^r\ \big] \leq s \cdot O\big(r(q/p)^{1/d}\big)^r.$$

Proof. We first hit the circuit with a $\mathcal{R}(p_0, p)$ random restriction where $p_0 = \frac{1}{1-p}(\frac{p+q}{2} - pq)$. We when hit the circuit with a sequence of d unbiased $\mathcal{R}(1/2, (q/p)^{1/d})$ random restrictions, applying the switching lemma in the usual way. (The preliminary $\mathcal{R}(p_0, p)$ creates the correct bias of 1's and 0's, but does not simplify the circuit.) Letting θ be the composition of these random restrictions, we get a bound of $s \cdot O(r(q/p)^{1/d})^r$ on the probability that $f\lceil\theta$ has decision-tree depth $\geq r$. The lemma follows, since the number of live variables is at most 2 raised to the decision-tree depth. $\square$

We remark that Lemma 7.4 is not the most efficient bound that can obtained from the switching lemma, in particular for AC^0 formulas [20]. However, it suffices for our (preliminary) technical lemma.

Lemma 7.5. *Suppose $f : \{0,1\}^N \to \{0,1\}$ is computed by a circuit of size at most n^k and depth $\log n/(\log\log n)^6$. Let G, H be pattern graphs with $V_G \cap V_H = \emptyset$ and let $y \in [n]^{V_H}$. For $\Gamma \in \{0,1\}^N$, define G-pathset $\mathcal{A}^\Gamma$ by*

$$\mathcal{A}^\Gamma := \big\{x \in [n]^{V_G} : N_{G,x} \subseteq Live(f\lceil\rho^\Gamma_{G\cup H,xy})\big\}.$$

Then $\mathop{\mathbb{P}}_{\Gamma \in \{0,1\}^N_{1/n}} [\mathcal{A}^\Gamma$ *is G-critical*$] \leq O(n^{-3k})$.

We remark that the bound $O(n^{-3k})$ can be strengthened to $O(n^{-\Omega(k \log\log n)})$ or even better. We state the lemma with $O(n^{-3k})$ since this is all we require.

Proof. Define $\mathcal{I} \subseteq \{0,1\}^N$ by

$$\mathcal{I} := \big\{I \in \{0,1\}^N : I_\nu = 0 \text{ for all } \nu \in N \setminus N_{H,y}\big\}.$$

Note that $|\mathcal{I}| = 2^{|N_{H,y}|} = 2^{|E_H|} \leq 2^k$.

For $I \in \mathcal{I}$, define $f_I : \{0,1\}^N \to \{0,1\}$ by $f_I(\Gamma) = f(\Gamma \oplus I)$. For all $x \in [n]^{V_G}$ and $\Gamma \in \{0,1\}^N$, we have

$$N_{G,x} \cap \mathit{Live}(f\lceil\rho^{\Gamma}_{G\cup H,xy}) = \bigcup_{I\in\mathcal{I}} \mathit{Live}(f_I\lceil\rho^{\Gamma}_{G,x}). \tag{7}$$

For $R \subseteq N$, let $\theta^{\Gamma}_R : N \to \{0,1,*\}$ be the restriction taking value $*$ over R and equal to Γ over $N \setminus R$. Define pathsets $\mathcal{B}_R$ and $\mathcal{C}^{\Gamma}_R$ by

$$\begin{aligned}
\mathcal{B}_R &:= \big\{x \in [n]^{V_G} : N_{G,x} \subseteq R\big\},\\
\mathcal{C}^{\Gamma}_R &:= \big\{x \in [n]^{V_G} : N_{G,x} \subseteq \textstyle\bigcup_{I\in\mathcal{I}} \mathit{Live}(f_I\lceil\theta^{\Gamma}_R)\big\}.
\end{aligned}$$

It follows from (7) that $\mathcal{A}^{\Gamma} \cap \mathcal{B}_R \subseteq \mathcal{C}^{\Gamma}_R$. Also, since $|N_{G,x}| = |E_G| \le k$ and $|\mathcal{I}| \le 2^k$,

$$\big|\mathcal{C}^{\Gamma}_R\big|^{1/k} \le \Big|\bigcup_{I\in\mathcal{I}} \mathit{Live}(f_I\lceil\theta^{\Gamma}_R)\Big| \le 2^k \cdot \max_{I\in\mathcal{I}} \big|\mathit{Live}(f_I\lceil\theta^{\Gamma}_R)\big|. \tag{8}$$

We now consider independent random $\Gamma \in \{0,1\}^N_{1/n}$ and random $R \subseteq_q N$ where $q = (1/n)^{1+(\varepsilon/2k)}$. Note that θ^{Γ}_R has distribution $\mathcal{R}(1/n, q)$. Also, note that $\mathcal{A}^{\Gamma}$ and $\mathcal{B}_R$ are independent, as $\mathcal{A}^{\Gamma}$ depends only on Γ and $\mathcal{B}_R$ depends only on R.

We may assume that $k \le \log^{1/3} n$, since otherwise the lemma is trivial. In particular, $2^k = o(n^{\varepsilon/2})$ (recall that $\varepsilon = 1/\log k$) and hence $\exp(-\Omega(n^{\varepsilon/2}/2^k)) = o(1)$. We have

$$\begin{aligned}
\mathop{\mathbb{P}}_{\Gamma}\big[\,\mathcal{A}^{\Gamma} \text{ is } G\text{-critical}\big] &\le \mathop{\mathbb{P}}_{\Gamma}\Big[\mathop{\mathbb{P}}_{R}\Big[|\mathcal{A}^{\Gamma}\cap\mathcal{B}_R| \le \frac{n^{\varepsilon/2}}{2}\Big] \le \exp\Big(-\Omega\Big(\frac{n^{\varepsilon/2}}{2^k}\Big)\Big)\Big] && \text{(Lemma 7.3)}\\
&= \mathop{\mathbb{P}}_{\Gamma}\Big[\mathop{\mathbb{P}}_{R}\Big[|\mathcal{A}^{\Gamma}\cap\mathcal{B}_R| > \frac{n^{\varepsilon/2}}{2}\Big] \ge 1 - o(1)\Big]\\
&\le (1+o(1)) \mathop{\mathbb{P}}_{\Gamma,R}\Big[|\mathcal{A}^{\Gamma}\cap\mathcal{B}_R| > \frac{n^{\varepsilon/2}}{2}\Big] && \text{(Markov ineq.)}\\
&\le (1+o(1)) \mathop{\mathbb{P}}_{\Gamma,R}\Big[|\mathcal{C}^{\Gamma}_R| > \frac{n^{\varepsilon/2}}{2}\Big] && (\mathcal{A}^{\Gamma}\cap\mathcal{B}_R \subseteq \mathcal{C}^{\Gamma}_R)\\
&\le (1+o(1))2^k \max_{I\in\mathcal{I}} \mathop{\mathbb{P}}_{\Gamma,R}\Big[\big|\mathit{Live}(f_I\lceil\theta^{\Gamma}_R)\big| > \frac{n^{\varepsilon/2k}}{2^{k+1}}\Big] && \text{(by (8))}\\
&\le (\log n)^2 \max_{I\in\mathcal{I}} \mathop{\mathbb{P}}_{\Gamma,R}\Big[\big|\mathit{Live}(f_I\lceil\theta^{\Gamma}_R)\big| > n^{\varepsilon/4k}\Big]
\end{aligned}$$

where this last inequality uses $k \le \log\log n$ and $\varepsilon \ge 1/k$ (hence $2^{k+1} \le n^{\varepsilon/4k}$ for sufficiently large n).

For each $I \in \mathcal{I}$, f_I is computable by AC^0 formulas with the same size and depth as f, namely n^k and $\log n/(\log\log n)^6$. Since θ^{Γ}_R has distribution $\mathcal{R}(1/n, (1/n)^{1+(\varepsilon/2k)})$, Lemma 7.4 implies that

$$\begin{aligned}
\mathop{\mathbb{P}}_{\Gamma,R}\Big[\big|\mathit{Live}(f_I\lceil\theta^{\Gamma}_R)\big| > n^{\varepsilon/4k}\Big] &\le n^k \cdot O\Big(\big(n^{-\varepsilon/4k}\big)^{(\log\log n)^6/\log n}\Big)^{(\varepsilon/4k)\log n}\\
&\le O(n^{k-(\varepsilon/4k)^2(\log\log n)^6}) \le O(n^{-\Omega(k\log\log n)})
\end{aligned}$$

again using $\varepsilon \ge 1/k$ and $k \le \log\log n$. Finally, we get the bound

$$\mathop{\mathbb{P}}_{\Gamma}\big[\,\mathcal{A}^{\Gamma} \text{ is } G\text{-critical}\big] \le (\log n)^2 \cdot O(n^{-\Omega(k\log\log n)}) \le O(n^{-3k}). \qquad \square$$

Finally, we derive Lemma 6.6 from Lemma 7.5.

Proof of Lemma 6.6. Suppose $f : \{0,1\}^N \to \{0,1\}$ is computable a circuit of size n^k and depth $\log n/(\log\log n)^6$. Fix a pattern graph G. We must show

$$\mathop{\mathbb{P}}_{\Gamma\in\{0,1\}^N_{1/n}} [\,\mathcal{A}^\Gamma_{f,G} \text{ is not } G\text{-small}\,] \le O(n^{-2k}).$$

Suppose $\Gamma \in \{0,1\}^N$ is any layered graph such that $\mathcal{A}^\Gamma_{f,G}$ is not G-small. We claim that there exist $S \subseteq V_G$ and $z \in [n]^{V_G\setminus S}$ such that S is a nonempty union of components of G and the pathset $B^\Gamma_{S,z}$ is $G|_S$-critical where $G|_S$ is the induced subgraph of G on S and

$$\mathcal{B}^\Gamma_{S,z} := \big\{y \in [n]^S : N_{G|_S,y} \subseteq \mathit{Live}(f\lceil\rho^\Gamma_{G,yz})\big\}.$$

We first note that it suffices to prove this claim. Since there are $2^{\Delta_G} - 1$ ($\le 2^k$) choices for S and $\le n^k$ choices for z, assuming the claim we have

$$\begin{aligned}
\mathop{\mathbb{P}}_{\Gamma\in\{0,1\}^N_{1/n}} [\,\mathcal{A}^\Gamma_{f,G} \text{ is not } G\text{-small}\,] &\le \mathop{\mathbb{P}}_{\Gamma\in\{0,1\}^N_{1/n}} \Big[\bigvee_{S,z} \mathcal{B}^\Gamma_{S,z} \text{ is } G|_S\text{-critical}\Big]\\
&\le 2^k n^k O(n^{-3k}) \qquad \text{(by Lemma 7.5)}\\
&= O(n^{-2k}).
\end{aligned}$$

To see why the claim holds, assume that $\mathcal{A}^\Gamma_{f,G}$ is not G-small and consider the following procedure. Initially set $S \leftarrow V_G$ and $z \leftarrow ()$ (the empty tuple). If $\mathcal{B}^\Gamma_{S,z}$ is $G|_S$-critical, then we are done. Otherwise, since $\mathcal{B}^\Gamma_{S,z}$ is neither $G|_S$-small nor $G|_S$-critical, there is a proper subset $T \subset S$ such that T is a union of $t \ge 1$ components of $V_{G|_S}$ and $\mu_T(\mathcal{B}^\Gamma_{S,z}) > \tilde{n}^{-t}$. By definition of μ_T, there exists $y \in [n]^{S\setminus T}$ such that $\delta(\mathcal{B}^\Gamma_{S,z}|^y_T) > \tilde{n}^{-t}$. Note that T is a union of components of G and $yz \in [n]^{V_G\setminus T}$. Also, for all $u \in [n]^T$, we have $N_{G|_T,u} \subseteq N_{G|_S,uy}$ and hence

$$\begin{aligned}
u \in \mathcal{B}^\Gamma_{S,z}|^y_T &\implies uy \in \mathcal{B}^\Gamma_{S,z}\\
&\implies N_{G|_S,uy} \subseteq \mathit{Live}(f\lceil\rho^\Gamma_{G,uyz})\\
&\implies N_{G|_T,u} \subseteq \mathit{Live}(f\lceil\rho^\Gamma_{G,uyz})\\
&\implies u \in \mathcal{B}^\Gamma_{T,yz}.
\end{aligned}$$

Therefore, $\mathcal{B}^\Gamma_{S,z}|^y_T \subseteq \mathcal{B}^\Gamma_{T,yz}$. It follows that $\delta(\mathcal{B}^\Gamma_{T,yz}) > \tilde{n}^{-t}$ and, hence, $\mathcal{B}^\Gamma_{T,yz}$ is not $G|_T$-small. We now update $S \leftarrow T$ and $z \leftarrow yz$. Since $\mathcal{B}^\Gamma_{S,z}$ is not $G|_S$-small, we may repeat this process so long as $\mathcal{B}^\Gamma_{S,z}$ is not $G|_S$-critical. Since S shrinks with every step, eventually this process will terminate, at which point $\mathcal{B}^\Gamma_{S,z}$ is $G|_S$-critical (and S is nonempty by definition of $G|_S$-criticality). Thus, the claim holds and the lemma is proved. $\square$

7.1 Proof of Lemma 7.3

For the proof of Lemma 7.3 we use a concentration of measure inequality due to Janson [9].

Lemma 7.6 (Janson's Inequality [9]). *Let Ω be a finite universal set and let R be a random subset of Ω given by $\mathbb{P}[\, r \in R\,] = p_r$, these events mutually independent over $r \in \Omega$. Let $\{S_i\}_{i\in I}$ be an indexed family of subsets of Ω. Define λ and Υ by*

$$\lambda := \sum_{i\in I} \mathbb{P}\left[\, S_i \subseteq R \,\right], \qquad \Upsilon := \sum_{(i,j)\in I^2 \,:\, i\neq j,\, S_i\cap S_j\neq\emptyset} \mathbb{P}\left[\, S_i \cup S_j \subseteq R\,\right].$$

Then, for all $0 \le t \le \lambda$, $\mathbb{P}\Big[\#\big\{i \in I \,:\, S_i \subseteq R\big\} \le \lambda - t\Big] \le \exp\Big(-\dfrac{t^2}{2(\lambda+\Upsilon)}\Big)$.

Proof of Lemma 7.3. Let G be a nonempty pattern graph, let $\mathcal{A}$ be a G-critical pathset, and let $q = (1/n)^{1+(\varepsilon/2k)}$. We must show

$$\mathop{\mathbb{P}}_{R\subseteq_q N}\Big[\,\#\big\{x \in \mathcal{A} : N_{G,x} \subseteq R\big\} \le \frac{n^{\varepsilon/2}}{2}\,\Big] \le \exp\Big(-\Omega\Big(\frac{n^{\varepsilon/2}}{2^k}\Big)\Big).$$

As in Janson's Inequality, define λ and Υ by

$$\lambda := \sum_{x\in\mathcal{A}} \mathbb{P}\left[\, N_{G,x} \subseteq R\,\right], \qquad \Upsilon := \sum_{(x,y)\in\mathcal{A}^2 \,:\, x\neq y,\, N_{G,x}\cap N_{G,y}\neq\emptyset} \mathbb{P}\left[\, N_{G,x}\cup N_{G,y} \subseteq R\,\right].$$

Taking $t = \lambda/2$ in Lemma 7.6, we get

$$\mathbb{P}\Big[\,\#\big\{x \in \mathcal{A} \,:\, N_{G,x} \subseteq R\big\} \le \frac{\lambda}{2}\,\Big] \le \exp\Big(-\frac{1}{16}\min\Big\{\lambda,\, \frac{\lambda^2}{\Upsilon}\Big\}\Big). \tag{9}$$

Recall that $\Delta_G = |V_G| - |E_G|$ and $\tilde{n} = n^{1-\varepsilon}$. By G-criticality of $\mathcal{A}$,

$$|\mathcal{A}| = n^{|V_G|}\delta(\mathcal{A}) > n^{|V_G|}\tilde{n}^{-\Delta_G} = n^{|E_G|+\varepsilon\Delta_G}.$$

Note that $\mathbb{P}[\, N_{G,x} \subseteq R\,] = q^{|E_G|}$ for all $x \in \mathcal{A}$. Since $|E_G| \le k$ and $\Delta_G \ge 1$, it follows that

$$\lambda = |\mathcal{A}| \cdot q^{|E_G|} > n^{\varepsilon(\Delta_G - |E_G|/2k)} \ge n^{\varepsilon(\Delta_G-(1/2))} \ge n^{\varepsilon/2}. \tag{10}$$

To complete the proof, it suffices to show that $\dfrac{\lambda^2}{\Upsilon} \ge \dfrac{n^{\varepsilon/2}}{2^k}$.

For all $(x,y) \in \mathcal{A}^2$, let

$$T_{x,y} := \big\{vw \in E_G : x_v = y_v \text{ and } x_w = y_w\big\}.$$

Note that $x = y$ iff $T_{x,y} = E_G$, and $N_{G,x}\cap N_{G,y} \neq \emptyset$ iff $T_{x,y} \neq \emptyset$, and $|N_{G,x}\cup N_{G,y}| = 2|E_G| - |T_{x,y}|$. Next, note that $\Upsilon = \sum_{T\,:\,\emptyset\subset T\subset E_G} \Upsilon_T$ where

$$\Upsilon_T := \sum_{(x,y)\in\mathcal{A}^2 \,:\, T_{x,y}=T} \mathbb{P}\left[\, N_{G,x}\cup N_{G,y} \subseteq R\,\right] = \#\big\{(x,y) \in \mathcal{A}^2 : T_{x,y} = T\big\} \cdot q^{2|E_G|-|T|}.$$

Now consider any fixed $\emptyset \subset T \subset E_G$. Let $S = \bigcup_{vw \in T}\{v, w\}$ and let s be the number of components of G which S intersects. Note that $1 \le s \le |S| - |T|$, since $|S| - |T|$ equals the number of components in the induced subgraph $G|_S$. We have

$$\begin{aligned}
\#\big\{(x,y) \in \mathcal{A}^2 : T_{x,y} = T\big\} &\le \sum_{z \in [n]^S} \big|\mathcal{A}|^z_{V_G \setminus S}\big|^2 \\
&\le \Big(\sum_{z \in [n]^S} \big|\mathcal{A}|^z_{V_G \setminus S}\big|\Big)\Big(\max_{z \in [n]^S} \big|\mathcal{A}|^z_{V_G \setminus S}\big|\Big) \\
&= |\mathcal{A}| \cdot n^{|V_G| - |S|} \cdot \mu_{V_G \setminus S}(\mathcal{A}) \\
&\le |\mathcal{A}| \cdot n^{|V_G| - |S|} \cdot \tilde{n}^{s - \Delta_G} && \text{(by G-criticality of $\mathcal{A}$)} \\
&= |\mathcal{A}| \cdot n^{|E_G| - |S| + s + \varepsilon(\Delta_G - s)}.
\end{aligned}$$

It follows that

$$\begin{aligned}
\Upsilon_T &= \#\big\{(x,y) \in \mathcal{A}^2 : T_{x,y} = T\big\} \cdot q^{2|E_G| - |T|} \\
&\le |\mathcal{A}| \cdot n^{|E_G| - |S| + s + \varepsilon(\Delta_G - s)} \cdot q^{2|E_G| - |T|} \\
&\le \lambda \cdot n^{|T| - |S| + s + \varepsilon(\Delta_G - s)} && \text{(using $\lambda = |\mathcal{A}| \cdot q^{|E_G|}$ and $q \le n^{-1}$)} \\
&\le \lambda \cdot n^{\varepsilon(\Delta_G - 1)} && \text{(using $1 \le s \le |S| - |T|$).}
\end{aligned}$$

We now have

$$\begin{aligned}
\frac{\lambda^2}{\Upsilon} = \frac{\lambda^2}{\sum_{T : \emptyset \subset T \subset E_G} \Upsilon_T} &\ge \frac{\lambda \cdot n^{-\varepsilon(\Delta_G - 1)}}{2^k} && \text{(by the above)} \\
&\ge \frac{n^{\varepsilon/2}}{2^k} && \text{(since $\lambda \ge n^{\varepsilon(\Delta_G - (1/2))}$ by (10)).}
\end{aligned} \tag{11}$$

Plugging (10) and (11) into (9) completes the proof. □

8 Union Trees

At this point in the paper, it only remains to prove Theorem 5.8, our lower bound on pathset complexity χ. As the first step in the proof, we introduce the notion of *union trees* and *pathset complexity w.r.t. union trees*, denoted $\bar{\chi}$. Intuitively, a union tree is a blueprint for constructing a pattern graph via pairwise unions starting from individual edges. This leads to a more constrained notion of pathset complexity where the allowable joins are prescribed by a given union tree.

Fixing the union tree of allowable joins can only increase the cost of constructing a pathset, hence $\chi \le \bar{\chi}$. Counterintuitively, the lower bound on χ is derived from a lower bound on $\bar{\chi}$. This lower bound on $\bar{\chi}$ is the true combinatorial lower bound in this paper. (Unfortunately, in the shift from χ to $\bar{\chi}$ we lose a factor of $2^{O(2^k)}$, which is the reason that Theorem 5.8 only holds up to $k(n) \le \log\log n$.)

In this section, we present the definition of $\bar{\chi}$ and state our lower bound for $\bar{\chi}$ (Theorem 8.3). The reduction from Theorem 8.3 to Theorem 5.8 is given in §9. In §10 we prove some preliminary lemmas (on properties of $\bar{\chi}$ with respect to projections and restrictions). Finally we prove Theorem 8.3 in §11.

Definition 8.1 (Union Trees)**.** A *union tree* is a (rooted, unordered) binary tree whose leaves are labeled by edges of P_k (i.e. elements of $E_k = \{v_i v_{i+1} : 0 \le i < k\}$). Every union tree A is associated with a pattern graph denoted $G_A = (V_A, E_A)$ where E_A is the set of edges of P_k which label leaves in A.

The *empty union tree* (of size 0) is denoted $\emptyset$. Union trees of size 1 (corresponding to elements of E_k) are said to be *atomic*. Union trees of size ≥ 2 are *non-atomic*. Throughout, A and B represent non-empty union trees. Let $\{A, B\}$ $(= \{B, A\})$ denote the union tree with children A and B. Note that every non-atomic union tree has the form $\{A, B\}$ for some A and B; also, $G_{\{A,B\}} = G_A \cup G_B$.

For a union tree A, *sub-union trees* of A are sub-trees of A consisting a node in A and all nodes below that node with the inherited labeling of leaves. The sub-union tree and strict sub-union tree relations are denoted by $\preceq$ and $\prec$ respectively.

To simplify notation, for a union tree A we write $\mathscr{P}_A$ for $\mathscr{P}_{G_A}$ and proj_A for proj_{V_A} and ℓ_A for ℓ_{G_A}, etc. We consistently write $\mathcal{A}, \mathcal{B}, \mathcal{C}$ for pathsets with underlying union trees A, B, C respectively.

Definition 8.2 (Pathset Complexity w.r.t. Union Trees)**.** For every union tree A and pathset $\mathcal{A} \in \mathscr{P}_A$, the *pathset complexity* of $\mathcal{A}$ with respect to A, denoted $\bar{\chi}_A(\mathcal{A})$, is defined by the following induction:

(i) $\bar{\chi}_\emptyset(\{()\}) := 0$, that is, the pathset complexity of $\{()\}$ w.r.t. the empty union tree $\emptyset$ is 0.

(ii) If A is atomic and $|\mathcal{A}| = 1$, then $\bar{\chi}_A(\mathcal{A}) := 1$.

(iii) For non-atomic $A = \{B, C\}$,

$$\bar{\chi}_A(\mathcal{A}) := \min_{(\mathcal{B}_i, \mathcal{C}_i)_i} \sum_i \max\{\bar{\chi}_B(\mathcal{B}_i),\ \bar{\chi}_C(\mathcal{C}_i)\}$$

where $(\mathcal{B}_i, \mathcal{C}_i)_i$ ranges over sequences such that $\mathcal{B}_i \in \mathscr{P}_B^{\text{small}}$, $\mathcal{C}_i \in \mathscr{P}_C^{\text{small}}$ and $\mathcal{A} \subseteq \bigcup_i \mathcal{B}_i \bowtie \mathcal{C}_i$.

In Appendices A–C we present some key examples of union trees and prove upper and lower bounds for $\bar{\chi}$ with respect to some special classes of union trees. The material in these appendices is not directly needed for our main results. However, these appendices serve as a warm-up and motivation for the lower bound that follows.

The following inequalities (analogous to the inequalities following Definition 5.6 of χ) are essentially built into Definition 8.2 of $\bar{\chi}$:

(base case)	$\bar{\chi}_\emptyset(\{()\}) \le 0$ and $\bar{\chi}_A(\mathcal{A}) \le 1$	if A is atomic and $\lvert\mathcal{A}\rvert = 1$,
(monotonicity)	$\bar{\chi}_A(\mathcal{A}') \le \bar{\chi}_A(\mathcal{A})$	if $\mathcal{A}' \subseteq \mathcal{A}$,
(sub-additivity)	$\bar{\chi}_A(\mathcal{A}_1 \cup \mathcal{A}_2) \le \bar{\chi}_A(\mathcal{A}_1) + \bar{\chi}_A(\mathcal{A}_2)$	for all $\mathcal{A}_1, \mathcal{A}_2$,
(join rule)	$\bar{\chi}_{\{A,B\}}(\mathcal{A} \bowtie \mathcal{B}) \le \max\{\bar{\chi}_A(\mathcal{A}),\ \bar{\chi}_B(\mathcal{B})\}$	if $\mathcal{A} \in \mathscr{P}_A^{\text{small}}$, $\mathcal{B} \in \mathscr{P}_B^{\text{small}}$.

The essential difference between χ and $\bar{\chi}$ is that χ allows arbitrary joins, while $\bar{\chi}$ only allows joins as prescribed by the given union tree. Viewed as a minimum construction cost (see Remark 5.9), this means that $\bar{\chi}$ has more highly constrained rules of construction compared with χ. Consequently, $\chi_{G_A}(\mathcal{A}) \le \bar{\chi}_A(\mathcal{A})$ for every union tree A and $\mathcal{A} \in \mathscr{P}_G$. Note that this inequality goes in the wrong direction for the purpose of proving a lower bound on χ. In §9 we give a different inequality between χ and $\bar{\chi}$ in the right direction.

Theorem 8.3 (Lower Bound for $\bar{\chi}$)**.** *For every union tree A and pathset $\mathcal{A} \in \mathscr{P}_A$,*

$$\bar{\chi}_A(\mathcal{A}) \geq \tilde{n}^{(1/4.41)\log(\ell_A)+\Delta_A} \cdot \delta(\mathcal{A}).$$

The game plan for the rest of the paper is as follows: in §9 we derive our lower bound for χ (Theorem 5.8) from Thereom 8.3. In §10 we establish some important properties of $\bar{\chi}$. Finally, in §11 we give the proof of Theorem 8.3.

Remark 8.4 (Dual Characterization of $\bar{\chi}$)**.** Similar to the dual characterization of χ mentioned in Remark 5.7, $\bar{\chi}$ has a *dual characterization* as the unique pointwise maximal function from $\{(A, \mathcal{A}) : A$ is a union tree and $\mathcal{A} \in \mathscr{P}_A\}$ to $\mathbb{R}$ which satisfies inequalities (base case), (monotonicity), (sub-additivity) and (join rule). This fact is established by a straightforward induction on union trees (omitted here since we don't actually use this dual characterization in our lower bound).

This dual characterization suggests an obvious "direct method" for proving a lower bound on $\bar{\chi}$: find an explicit function from pairs $(A, \mathcal{A})$ to $\mathbb{R}$ and show that this function satisfies inequalities (base case), (monotonicity), (sub-additivity) and (join rule). This is analogous to the "direct method" of proving a formula size lower bound via a *complexity measure*, defined as a function M from {Boolean functions on n variables} to $\mathbb{R}$ satisfying inequalities $M(f \wedge g) \leq M(f) + M(g)$ and $M(f \vee g) \leq M(f) + M(g)$ in addition to base case inequalities $M(f) \leq 0$ if f is constant and $M(f) \leq 1$ if f is a coordinate function (see [10]).

Using the direct method, we were only able to prove lower bounds on $\bar{\chi}$ for a few restricted classes union trees (see Appendix B). For general union trees, we could not prove a lower bound along the lines of Theorem 8.3 using the direct method. We still do not know of any explicit function which satisfies (base case), (monotonicity), (sub-additivity) and (join rule) and maps $(A, [n]^{P_k})$ to $n^{\Omega(\log k)}$ for all union trees A with graph P_k. A priori, it is not even clear whether any such nice explicit function exists.[6]

The proof of Theorem 8.3 which we present in §11 does not proceed via the direct method. In particular, neither the function $\tilde{n}^{(1/4.41)\log(\ell_A)+\Delta_A} \cdot \delta(\mathcal{A})$ nor $\tilde{n}^{\Phi_A} \cdot \delta(\mathcal{A})$ (defined in §11.1) satisfies inequality (join rule). Rather, our proof involves a more subtle induction on union trees.

9 From χ to $\bar{\chi}$

In this section, we prove:

Reduction 9.1. Theorem 8.3 (lower bound on $\bar{\chi}$) $\Longrightarrow$ Theorem 5.8 (lower bound on χ).

The following definition of *strict union tree* is only needed in this section. Rather than A, B, C, we write α, β, γ for this special class of union trees.

Definition 9.2. A union tree α is *strict* if $G_{\alpha''} \subset G_{\alpha'}$ for all $\alpha'' \prec \alpha' \preceq \alpha$. For a pattern graph G, let *Strict*(G) denote the set of strict union trees α with graph G.

It is important that the number of strict union trees with a given pattern graph is bounded (though doubly exponential in $|E_G|$).

[6] A natural approach is to consider functions of the form $n^{c_A} \cdot \nu(\mathcal{A})$ where c_A is a constant depending only on A and $\nu : \mathscr{P}_A \to \mathbb{R}$ is a monotone sub-additive function, such as δ or μ_S or π_S or any norm on $\mathbb{R}^{[n]^{V_A}}$ (viewing $\mathscr{P}_A \cong \{0,1\}^{[n]^{V_A}}$ as a subset of $\mathbb{R}^{[n]^{V_A}}$). For such functions, one only needs to show (join rule). The (base case) can be handled by appropriate scaling.

Lemma 9.3. *For every pattern graph G with r edges, there are only $2^{O(2^r)}$ strict union trees with graph G.*

Proof. Denote by $s(r)$ the number of strict union trees supported on any fixed set of r edges. Note that $|\mathit{Strict}(G)|$ depends only on $|E_G|$ and that $s(r)$ is an increasing function of r. We have $s(1) = 1$ and, for $r \geq 2$,

$$s(r) = \sum_{I,J \subset [r]\,:\,I \cup J = [r]} s(|I|)s(|J|) \leq 3^r s(r-1)^2$$

for all $r \geq 2$. Therefore,

$$s(r) \leq \prod_{i=1}^{r} 3^{i2^{(r-i)}} = 3^{2^r \sum_{i=1}^{r} i2^{-i}} = 2^{O(2^r)}. \qquad \square$$

We now give the main lemma needed for Reduction 9.1.

Lemma 9.4. *For every pattern graph G and pathset $\mathcal{A} \in \mathscr{P}_G$, there is an indexed family $\{\mathcal{A}^{(\alpha)}\}_{\alpha \in \mathit{Strict}(G)}$ of sub-pathsets $\mathcal{A}^{(\alpha)} \subseteq \mathcal{A}$ such that*

$$\mathcal{A} = \bigcup_{\alpha \in \mathit{Strict}(G)} \mathcal{A}^{(\alpha)} \quad \text{and} \quad \forall \alpha \in \mathit{Strict}(G),\ \bar{\chi}_\alpha(\mathcal{A}^{(\alpha)}) \leq \chi_G(\mathcal{A}).$$

Proof. By induction on $|E_G|$. The lemma is trivial if $|E_G| \leq 1$ (since in this case $|\mathit{Strict}(G)| = 1$). For the induction step, suppose G is a pattern graph with ≥ 2 edges. By Definition 5.6 of χ, there exists a sequence $(H_i, K_i, \mathcal{B}_i, \mathcal{C}_i)_i$ with

$$H_i, K_i \subset G, \quad H_i \cup K_i = G, \quad \mathcal{B}_i \in \mathscr{P}^{\text{small}}_{H_i}, \quad \mathcal{C}_i \in \mathscr{P}^{\text{small}}_{K_i}$$

such that

$$\mathcal{A} \subseteq \bigcup_i \mathcal{B}_i \bowtie \mathcal{C}_i \quad \text{and} \quad \chi_G(\mathcal{A}) = \sum_i \max\{\chi_{H_i}(\mathcal{B}_i),\, \chi_{K_i}(\mathcal{C}_i)\}.$$

For each $\alpha = \{\beta, \gamma\} \in \mathit{Strict}(G)$, define $\mathcal{A}^{(\alpha)}$ inductively by

$$\mathcal{A}^{(\alpha)} := \mathcal{A} \cap \bigcup_{i\,:\,H_i = G_\beta,\, K_i = G_\gamma} \mathcal{B}_i^{(\beta)} \bowtie \mathcal{C}_i^{(\gamma)}.$$

First, we show that $\mathcal{A} = \bigcup_{\alpha \in \mathit{Strict}(G)} \mathcal{A}^{(\alpha)}$. The inclusion $\supseteq$ is obvious. For the inclusion $\subseteq$, consider any $x \in \mathcal{A}$. Then x belongs to $\mathcal{B}_i \bowtie \mathcal{C}_i$ for some i. This means that $x_{V_{H_i}} \in \mathcal{B}_i$ and $x_{V_{K_i}} \in \mathcal{C}_i$. By the induction hypothesis, there exist $\beta \in \mathit{Strict}(H_i)$ and $\gamma \in \mathit{Strict}(K_i)$ such that $x_{V_{H_i}} \in \mathcal{B}_i^{(\beta)}$ and $x_{V_{K_i}} \in \mathcal{C}_i^{(\gamma)}$. Let $\alpha = \{\beta, \gamma\}$ and note that $\alpha \in \mathit{Strict}(G)$. Since $x \in \mathcal{B}_i^{(\beta)} \bowtie \mathcal{C}_i^{(\gamma)}$, it follows that $x \in \mathcal{A}^{(\alpha)}$, proving the inclusion $\subseteq$.

Finally, for all $\alpha \in \mathit{Strict}(G)$, we show $\bar{\chi}_\alpha(\mathcal{A}^{(\alpha)}) \leq \chi_G(\mathcal{A})$ as follows:

$$\begin{aligned}
\bar{\chi}_\alpha(\mathcal{A}^{(\alpha)}) &\leq \bar{\chi}_\alpha\Big(\bigcup_{i\,:\,H_i = G_\beta,\, K_i = G_\gamma} \mathcal{B}_i^{(\beta)} \bowtie \mathcal{C}_i^{(\gamma)}\Big) && \text{(monotonicity)}\\
&\leq \sum_{i\,:\,H_i = G_\beta,\, K_i = G_\gamma} \bar{\chi}_\alpha(\mathcal{B}_i^{(\beta)} \bowtie \mathcal{C}_i^{(\gamma)}) && \text{(sub-additivity)}
\end{aligned}$$

Noting that $\mathcal{B}_i^{(\beta)}$ and $\mathcal{C}_i^{(\gamma)}$ are small (since $\mathcal{B}_i^{(\beta)} \subseteq \mathcal{B}_i \in \mathscr{P}_{H_i}^{\text{small}}$ and $\mathcal{C}_i^{(\gamma)} \subseteq \mathcal{C}_i \in \mathscr{P}_{K_i}^{\text{small}}$), we continue:

$$\begin{aligned}
&\leq \sum_{i\,:\,H_i=G_\beta,\,K_i=G_\gamma} \max\{\bar{\chi}_\beta(\mathcal{B}_i^{(\beta)}),\, \bar{\chi}_\gamma(\mathcal{C}_i^{(\gamma)})\} && \text{(join rule)}\\
&\leq \sum_{i\,:\,H_i=G_\beta,\,K_i=G_\gamma} \max\{\chi_{H_i}(\mathcal{B}_i),\, \chi_{K_i}(\mathcal{C}_i)\} && \text{(ind. hyp.)}\\
&\leq \chi_G(\mathcal{A}).
\end{aligned}$$

□

The next corollary follows directly from Lemma 9.4.

Corollary 9.5. *For every pattern graph G and pathset $\mathcal{A} \in \mathscr{P}_G$, there is a strict union tree $\alpha \in$ Strict(G) and a sub-pathset $\mathcal{A}' \subseteq \mathcal{A}$ such that $\bar{\chi}_\alpha(\mathcal{A}') \leq \chi_G(\mathcal{A})$ and $\delta(\mathcal{A}) \leq |\mathit{Strict}(G)| \cdot \delta(\mathcal{A}')$.* □

We conclude this section with the proof of Reduction 9.1.

Proof of Reduction 9.1. Assume Theorem 8.3 and consider arbitrary $\mathcal{A} \in \mathscr{P}_{P_k}$. By Corollary 9.5, there exist $\alpha \in \mathit{Strict}(P_k)$ and $\mathcal{A}' \subseteq \mathcal{A}$ such that $\chi_\alpha(\mathcal{A}') \leq \chi_{P_k}(\mathcal{A})$ and $\delta(\mathcal{A}) \leq |\mathit{Strict}(P_k)| \cdot \delta(\mathcal{A}') \leq 2^{O(2^k)} \cdot \delta(\mathcal{A}')$ (Lemma 9.3). We now have

$$\begin{aligned}
\chi_{P_k}(\mathcal{A}) \geq \chi_\alpha(\mathcal{A}') &\geq \tilde{n}^{(1/4.41)\log(\ell_\alpha)+\Delta_\alpha} \cdot \delta(\mathcal{A}') && \text{(Theorem 8.3)}\\
&\geq \frac{\tilde{n}^{(1/4.41)\log(k)+1}}{2^{O(2^k)}} \cdot \delta(\mathcal{A})\\
&\geq \frac{n^{(1/4.41)\log k}}{2^{O(2^k)}} \cdot \delta(\mathcal{A}) && (\text{as } \tilde{n} = n^{1-\varepsilon} = n^{1-(1/\log k)}).
\end{aligned}$$

This shows that Theorem 5.8 holds, which completes the proof of the reduction. □

10 Projection and Restriction

In this section we establish two key properties of $\bar{\chi}$: it is monotone decreasing with respect to projections to sub-union trees (Lemma 10.2) and restriction to unions of components (Lemma 10.6). We also introduce an operation on union trees $A \ominus B$ (Definition 10.7), read as "A restricted away from B". This notation will be extremely convenient in §11.

10.1 $\bar{\chi}$ Decreases Under Projection

Claim 10.1. *For every non-atomic union tree $\{A, B\}$ and pathset $\mathcal{C} \in \mathscr{P}_{\{A,B\}}$, we have $\bar{\chi}_A(\mathrm{proj}_A(\mathcal{C})) \leq \bar{\chi}_{\{A,B\}}(\mathcal{C})$.*

Proof. By Definition 8.2(iii) of $\bar{\chi}_{\{A,B\}}(\mathcal{C})$, there is a sequence $(\mathcal{A}_i, \mathcal{B}_i)_i$ such that

$$\mathcal{A}_i \in \mathscr{P}_A^{\text{small}}, \quad \mathcal{B}_i \in \mathscr{P}_B^{\text{small}}, \quad \mathcal{C} \subseteq \bigcup_i \mathcal{A}_i \bowtie \mathcal{B}_i \quad \text{and} \quad \bar{\chi}_{\{A,B\}}(\mathcal{C}) = \sum_i \max\{\bar{\chi}_A(\mathcal{A}_i),\, \bar{\chi}_B(\mathcal{B}_i)\}.$$

Note that $\mathrm{proj}_A(\mathcal{C}) \subseteq \mathrm{proj}_A(\bigcup_i \mathcal{A}_i \bowtie \mathcal{B}_i) \subseteq \bigcup_i \mathcal{A}_i$. By monotonicity and sub-additivity of $\bar{\chi}_A$, it follows that

$$\bar{\chi}_A(\mathrm{proj}_A(\mathcal{C})) \leq \bar{\chi}_A(\bigcup_i \mathcal{A}_i) \leq \sum_i \bar{\chi}_A(\mathcal{A}_i) \leq \bar{\chi}_{\{A,B\}}(\mathcal{C}).$$

□

Lemma 10.2 ($\bar{\chi}$ decreases under projections)**.** *For every union tree A and pathset $\mathcal{A} \in \mathscr{P}_A$ and sub-union tree $A' \preceq A$, $\bar{\chi}_{A'}(\mathrm{proj}_{A'}(\mathcal{A})) \leq \bar{\chi}_A(\mathcal{A})$.*

Proof. Induction using Claim 10.1 and the observation that $\mathrm{proj}_{S'}(\mathcal{A}) = \mathrm{proj}_{S'}(\mathrm{proj}_S(\mathcal{A}))$ for all $S' \subseteq S \subseteq V_A$. □

10.2 $\bar{\chi}$ Decreases Under Restriction

For a union tree A and a pathset $\mathcal{A} \in \mathscr{P}_A$, Lemma 10.2 concerns projections of $\mathcal{A}$ of the form $\mathrm{proj}_{A'}(\mathcal{A})$ where A' is a sub-union tree of A. The restrictions of $\mathcal{A}$ that we consider next are not restrictions of the form $\mathcal{A}|^z_{A'}$ where $z \in [n]^{V_A \setminus V_{A'}}$. Note that $\mathcal{A}|^z_{A'} \subseteq \mathrm{proj}_{A'}(\mathcal{A})$, so we already have $\bar{\chi}_{A'}(\mathcal{A}|^z_{A'}) \leq \bar{\chi}_A(\mathcal{A})$ by Lemma 10.2 and monotonicity of $\bar{\chi}_{A'}$.

Rather than restrictions over sub-union trees, we instead consider restrictions of the form $\mathcal{A}|^z_S$ where $z \in [n]^{V_A \setminus S}$ and $S \subseteq V_A$ is a union of components of G_A. We define an operation of restriction on union trees; the restriction $A {\restriction} S$ is a union tree with $V_{A\restriction S} = S$. Even though $A {\restriction} S$ is not necessarily a sub-union tree of A, we will show that $\bar{\chi}_{A\restriction S}(\mathcal{A}|^z_S) \leq \bar{\chi}_A(\mathcal{A})$.

Definition 10.3 (Restriction of Union Trees)**.**

(i) For all $S \subseteq V_k$, let $\overline{S}$ denote the complement $V_k \setminus S$ of S in V_k.

(ii) For a union tree A, we say that S is *A-respecting* if $V_A \cap S$ is a union of components of G_A.

Note that S is A-respecting $\Longleftrightarrow$ $\overline{S}$ is A-respecting $\Longleftrightarrow$ every leaf in A is labeled by an edge $v_iv_{i+1} \in E_k$ such that $\{v_i, v_{i+1}\} \subseteq S$ or $\{v_i, v_{i+1}\} \subseteq \overline{S}$. Also note that if S is $\{A, B\}$-respecting, then it is both A-respecting and B-respecting and $\{A, B\}{\restriction}S = \{A{\restriction}S, B{\restriction}S\}$.

(iii) If S is A-respecting, we denote by $A{\restriction}S$ the union tree obtained from A by pruning all leaves labeled by edges whose endpoints are not contained in S. This pruning operation does not simplify the pattern above the leaves (i.e. there is no propagation or change in the tree structure of the pattern.)

For example, if A is the union tree $\{\{v_1v_2, v_5v_6\}, \{v_2v_3, v_6v_7\}\}$ and S is the A-respecting set $\{v_1, v_2, v_3\}$, then $A{\restriction}S = \{v_1v_2, v_2v_3\}$. Note that $A{\restriction}S = \{v_1v_2, v_2v_3\}$ also when S is the A-respecting set $\{v_1, v_2, v_3, v_4\}$; in general, $A{\restriction}S = A{\restriction}(V_A \cap S)$. Also note that $A{\restriction}S$ is not a sub-union tree of A in this example.

Recall our convention concerning notation $\mathcal{A}|^z_S$ (see Definition 4.3): for every union tree A and pathset $\mathcal{A} \in \mathscr{P}_A$ and $S \subseteq V_k$ and $z \in [n]^{\overline{S}}$, the pathset $\mathcal{A}|^z_S \in \mathscr{P}_{A\restriction S}$ is defined by $\mathcal{A}|^z_S := \mathcal{A}|^{z'}_{V_A \cap S} = \{y \in [n]^{V_A \cap S} : yz' \in \mathcal{A}\}$ where $z' = z_{V_A \setminus S}$.

Lemma 10.4 (Smallness is preserved under restriction)**.** *For every union tree A and small pathset $\mathcal{A} \in \mathscr{P}^{\mathrm{small}}_A$ and A-respecting $S \subseteq V_k$ and $z \in [n]^{\overline{S}}$, we have $\mathcal{A}|^z_S \in \mathscr{P}^{\mathrm{small}}_{A\restriction S}$.*

Proof. Immediate from Definition 5.3 of small pathsets. □

Remark 10.5. Smallness is preserved under joins (Lemma 5.5) and restrictions to union of components (Lemma 10.4). However, smallness is not preserved under projection to unions of components. A counterexample is the union tree $A = \{v_1v_2, v_3v_4\}$ and pathset $\mathcal{A} = \{x \in [n]^{V_A} : x_1 = x_3 \text{ and } x_2 = x_4\} \in \mathscr{P}^{\mathrm{small}}_A$. Letting A' be the atomic sub-union tree v_1v_2 of A, we have $\pi_{A'}(\mathcal{A}) = 1$, hence $\mathrm{proj}_{A'}(\mathcal{A}) \notin \mathscr{P}^{\mathrm{small}}_{A'}$.

Lemma 10.6 ($\bar{\chi}$ decreases under restrictions)**.** *For every union tree A and pathset $\mathcal{A} \in \mathscr{P}_A$ and A-respecting $S \subseteq V_k$ and $z \in [n]^{\overline{S}}$, we have $\bar{\chi}_{A\restriction S}(\mathcal{A}|^z_S) \le \bar{\chi}_A(\mathcal{A})$.*

Proof. By induction on union trees. The lemma is trivial for empty and atomic union trees. For the induction step, consider a non-atomic union tree $\{A, B\}$ and assume the lemma holds for A and B. Let $\mathcal{C} \in \mathscr{P}_{\{A,B\}}$, let S be a $\{A, B\}$-respecting subset of V_k, and let $z \in [n]^{\overline{S}}$. By Def. 8.2(iii) of $\bar{\chi}_{\{A,B\}}(\mathcal{C})$, there is a sequence $(\mathcal{A}_i, \mathcal{B}_i)_i$ such that

$$\mathcal{A}_i \in \mathscr{P}^{\text{small}}_A, \quad \mathcal{B}_i \in \mathscr{P}^{\text{small}}_B, \quad \mathcal{C} \subseteq \bigcup_i \mathcal{A}_i \bowtie \mathcal{B}_i \quad \text{and} \quad \bar{\chi}_{\{A,B\}}(\mathcal{C}) = \sum_i \max\{\bar{\chi}_A(\mathcal{A}_i),\, \bar{\chi}_B(\mathcal{B}_i)\}.$$

By Lemma 10.4, $\mathcal{A}_i|^z_S \in \mathscr{P}^{\text{small}}_{A\restriction S}$ and $\mathcal{B}_i|^z_S \in \mathscr{P}^{\text{small}}_{B\restriction S}$. We now have

$$\begin{aligned}
\bar{\chi}_{\{A,B\}\restriction S}(\mathcal{C}|^z_S) &\le \bar{\chi}_{\{A,B\}\restriction S}\Big(\bigcup_i (\mathcal{A}_i \bowtie \mathcal{B}_i)|^z_S\Big) && \text{(monotonicity)}\\
&\le \sum_i \bar{\chi}_{\{A,B\}\restriction S}((\mathcal{A}_i \bowtie \mathcal{B}_i)|^z_S) && \text{(sub-additivity)}\\
&= \sum_i \bar{\chi}_{\{A\restriction S, B\restriction S\}}(\mathcal{A}_i|^z_S \bowtie \mathcal{B}_i|^z_S) && \\
&\le \sum_i \max\{\bar{\chi}_{A\restriction S}(\mathcal{A}_i|^z_S),\, \bar{\chi}_{B\restriction S}(\mathcal{B}_i|^z_S)\} && \text{(join rule)}\\
&\le \sum_i \max\{\bar{\chi}_A(\mathcal{A}_i),\, \bar{\chi}_B(\mathcal{B}_i)\} && \text{(ind. hyp.)}\\
&= \bar{\chi}_{\{A,B\}}(\mathcal{C}).
\end{aligned}$$

□

10.3 The Operation $A \ominus B$

We introduce an operation $A \ominus B$ on union trees ("A restricted away from B").

Definition 10.7. For union trees A and B, we write $A \ominus B$ for the union tree $A\restriction S$ where $S \subseteq V_A$ consists of the components of G_A which do not intersect V_B.

For example, if $A = \{\{v_1v_2, v_4v_5\}, \{v_2v_3, v_5v_6\}\}$ (so G_A is the union of paths $v_1v_2v_3$ and $v_4v_5v_6$) and $B = \{v_6v_7\}$, then $A \ominus B = \{v_1v_2, v_2v_3\}$.

Lemma 10.8. *For all union trees $C = \{A, B\}$ and $A' \preceq A$ and $B' \preceq B$,*

$$\Delta_C \le \Delta_{A'} + \Delta_{B' \ominus A'} + \Delta_{C \ominus \{A',B'\}}.$$

Proof. This follows from the observation that each component of G_C contains at least one vertex in $G_{A'}$, $G_{B'\ominus A'}$ or $G_{C\ominus\{A',B'\}}$, and each component in any of these three graphs is contained in a component of G_C. □

Lemma 10.9. *For all union trees $C = \{A, B\}$ and $A' \preceq A$ and $B' \preceq B$ and pathsets $\mathcal{A} \in \mathscr{P}_A$ and $\mathcal{B} \in \mathscr{P}_B$,*

$$\delta(\mathcal{A} \bowtie \mathcal{B}) \le \pi_{A'}(\mathcal{A}) \cdot \mu_{B'\ominus A'}(\text{proj}_{B'}(\mathcal{B})) \cdot \mu_{C\ominus\{A',B'\}}(\mathcal{A} \bowtie \mathcal{B}).$$

Proof. By Lemma 4.6,

$$\delta(\mathcal{A} \bowtie \mathcal{B}) \leq \pi_{A'}(\mathcal{A}) \cdot \mu_{V_{B'} \setminus V_{A'}}(\mathrm{proj}_{B'}(\mathcal{B})) \cdot \mu_{V_C \setminus (V_{A'} \cup V_{B'})}(\mathcal{A} \bowtie \mathcal{B}).$$

Since $V_{B' \ominus A'} \subseteq V_{B'} \setminus V_{A'}$ and $V_{C \ominus \{A', B'\}} \subseteq V_C \setminus (V_{A'} \cup V_{B'})$, by Lemma 4.4,

$$\mu_{V_{B'} \setminus V_{A'}}(\mathrm{proj}_{B'}(\mathcal{B})) \leq \mu_{B' \ominus A'}(\mathrm{proj}_{B'}(\mathcal{B})),$$
$$\mu_{V_C \setminus (V_{A'} \cup V_{B'})}(\mathcal{A} \bowtie \mathcal{B}) \leq \mu_{C \ominus \{A', B'\}}(\mathcal{A} \bowtie \mathcal{B}).$$

Combining these inequalities finishes the proof. □

11 Lower Bound for $\bar{\chi}$

In this section we prove Theorem 8.3, our lower bound for $\bar{\chi}$. Recall that ℓ_A denote the length of the longest path in G_A, i.e., the number of edges in the largest component of G_A.

Theorem 8.3. (restated) *For every union tree A and pathset $\mathcal{A} \in \mathscr{P}_A$,*

$$\bar{\chi}_A(\mathcal{A}) \geq \tilde{n}^{(1/4.41)\log(\ell_A) + \Delta_A} \cdot \delta(\mathcal{A}).$$

To prove Theorem 8.3, first we define an auxiliary function Φ : {union trees} $\to \mathbb{R}$. We then prove two lemmas: $\bar{\chi}_A(\mathcal{A}) \geq \tilde{n}^{\Phi_A}\delta(\mathcal{A})$ (Lemma 11.2) and $\Phi_A \geq \frac{1}{4.41}\log(\ell_A) + \Delta_A$ (Lemma 11.4).

11.1 Definition of Φ_A

Definition 11.1. Let Φ : {union trees} $\to \mathbb{R}$ be the unique minimal function such that the following hold:

- $\Phi_A = 0$ if A is empty, and $\Phi_A = 2$ if A is atomic,
- for every non-atomic union tree $C = \{A, B\}$ and sub-union trees $A' \preceq A$ and $B' \preceq B$,

$$(\dagger)^C_{A',B} \qquad \Phi_C \geq \Phi_{A'} + \Delta_{B \ominus A'} + \Delta_{C \ominus \{A', B\}},$$
$$(\ddagger)^C_{A',B'} \qquad \Phi_C \geq \frac{1}{2}\Big(\Phi_{A'} + \Phi_{B' \ominus A'} + \Delta_C + \Delta_{C \ominus \{A', B'\}}\Big).$$

We refer to ($\dagger$) and ($\ddagger$) as the "one-sided" and "balanced" inequalities. Note that since $\{A, B\}$ and $\{B, A\}$ are considered to be the same union tree, we also have the reverse inequalities $(\dagger)^C_{B',A}$ and $(\ddagger)^C_{B',A'}$. For better readability, we sometimes write Φ_A instead of $\Phi(A)$.

Some remarks on this definition:

- Minimality of Φ among functions satisfying these inequalities means that for every non-atomic union tree $C = \{A, B\}$, at least one of the four inequalities $(\dagger)^C_{A',B}$, $(\dagger)^C_{B',A}$, $(\ddagger)^C_{A',B'}$, $(\ddagger)^C_{B',A'}$ is tight (i.e. holds with equality) for some $A' \preceq A$ and $B' \preceq B$.
- Note that Φ is monotone decreasing with respect to sub-union trees, that is, $\Phi_{A'} \leq \Phi_A$ for all $A' \preceq A$ (by inequalities ($\dagger$)).

- Φ increases by means of the contribution of Δ's: if we remove the Δ's from $(\dagger)^C_{A',B}$ and $(\ddagger)^C_{A',B'}$ (replacing these inequalities by $\Phi_C \geq \Phi_{A'}$ and $\Phi_C \geq \frac{1}{2}(\Phi_{A'} + \Phi_{B'})$ respectively), then we would have $\Phi_A = 2$ for every nonempty union tree A. Intuitively, in the attempt to lower bound Φ_A, the objective of the game is to pick up as many Δ's as possible.

- For the union trees A_k and B_k defined in Appendix A, we have $\Phi_{A_k} \geq \Phi_{A_{\lceil k/4 \rceil}} + 1$ by $(\dagger)$ and $\Phi_{B_k} \geq \Phi_{B_{\lceil (k-1)/2 \rceil}} + \frac{1}{2}$ by $(\ddagger)$ for all $k \geq 4$. It follows that $\Phi_{A_k} \geq \frac{1}{2}\log k - O(1)$ and $\Phi_{B_k} \geq \frac{1}{2}\log k - O(1)$.

11.2 Showing $\bar{\chi}_A(\mathcal{A}) \geq \tilde{n}^{\Phi_A}\delta(\mathcal{A})$

We now prove the most important lemma in the overall proof of Theorem 8.3. Lemma 11.2 accounts for the definition of Φ_A (essentially Φ_A is the maximum function for which the argument of Lemma 11.2 is valid). The two cases $(\dagger)$ and $(\ddagger)$ in the proof are inspired by the special cases proved in Appendix B.

Lemma 11.2. *For every union tree A and pathset $\mathcal{A} \in \mathscr{P}_A$, $\bar{\chi}_A(\mathcal{A}) \geq \tilde{n}^{\Phi(A)}\delta(\mathcal{A})$.*

Proof. We argue by induction on union trees. The base case where A is empty or atomic is trivial. For the induction step, consider a non-atomic union tree $C = \{A, B\}$ and assume the lemma holds for all smaller union trees.

We claim that it suffices to show that

$$\tilde{n}^{\Phi(C)}\delta(\mathcal{A} \bowtie \mathcal{B}) \leq \max\{\bar{\chi}_A(\mathcal{A}),\, \bar{\chi}_B(\mathcal{B})\} \tag{12}$$

for all $\mathcal{A} \in \mathscr{P}^{\text{small}}_A$, $\mathcal{B} \in \mathscr{P}^{\text{small}}_B$. To see that this suffices, consider any $\mathcal{C} \in \mathscr{P}_C$. By Definition 8.2 of pathset complexity, there exists a covering $\mathcal{C} \subseteq \bigcup_i \mathcal{A}_i \bowtie \mathcal{B}_i$ by joins of small pathsets $\mathcal{A}_i$ and $\mathcal{B}_i$ such that $\bar{\chi}_C(\mathcal{C}) = \sum_i \bar{\chi}_A(\mathcal{A}_i) + \bar{\chi}_B(\mathcal{B}_i)$. Note that

$$\delta(\mathcal{C}) \leq \delta(\bigcup_i \mathcal{A}_i \bowtie \mathcal{B}_i) \leq \sum_i \delta(\mathcal{A}_i \bowtie \mathcal{B}_i).$$

Assuming (12) holds for all $\mathcal{A}_i$ and $\mathcal{B}_i$, we have

$$\tilde{n}^{\Phi(C)}\delta(\mathcal{C}) \leq \sum_i \tilde{n}^{\Phi(C)}\delta(\mathcal{A}_i \bowtie \mathcal{B}_i) \leq \sum_i \max\{\bar{\chi}_A(\mathcal{A}_i),\, \bar{\chi}_B(\mathcal{B}_i)\} = \bar{\chi}_C(\mathcal{C}).$$

We now turn to proving inequality (12). Fix small pathsets $\mathcal{A} \in \mathscr{P}^{\text{small}}_A$ and $\mathcal{B} \in \mathscr{P}^{\text{small}}_B$. Note that $\mathcal{A} \bowtie \mathcal{B} \in \mathscr{P}^{\text{small}}_C$ by Lemma 5.5. Recall that at least one of the four inequalities $(\dagger)^C_{A',B}$, $(\ddagger)^C_{A',B'}$, $(\dagger)^C_{B',A}$, $(\ddagger)^C_{B',A'}$ is tight for some $A' \preceq A$ and $B' \preceq B$. By symmetry of the argument, we consider only the first two possibilities without loss of generality .

Case $(\dagger)$ (one-sided induction case): Assume that there exists $A' \preceq A$ such that $(\dagger)^C_{A',B}$ is tight, that is,

$$\Phi_C = \Phi_{A'} + \Delta_{B \ominus A'} + \Delta_{C \ominus \{A',B\}}. \tag{13}$$

By Lemma 10.9, we have

$$\delta(\mathcal{A} \bowtie \mathcal{B}) \leq \pi_{A'}(\mathcal{A}) \cdot \mu_{B \ominus A'}(\mathcal{B}) \cdot \mu_{C \ominus \{A',B\}}(\mathcal{A} \bowtie \mathcal{B}).$$

Since $\mathcal{B}$ is B-small and $\mathcal{A} \bowtie \mathcal{B}$ is C-small, we have

$$\mu_{B\ominus A'}(\mathcal{B}) \leq \tilde{n}^{-\Delta(B\ominus A')} \quad \text{and} \quad \mu_{C\ominus\{A',B\}}(\mathcal{A} \bowtie \mathcal{B}) \leq \tilde{n}^{-\Delta(C\ominus\{A',B\})}.$$

Combining these inequalities (and substituting $\delta(\mathrm{proj}_{A'}(\mathcal{A}))$ for $\pi_{A'}(\mathcal{A})$), we have

$$\delta(\mathcal{A} \bowtie \mathcal{B}) \leq \tilde{n}^{-\Delta(B\ominus A')-\Delta(C\ominus\{A',B\})}\delta(\mathrm{proj}_{A'}(\mathcal{A})). \tag{14}$$

Using the fact that $\bar{\chi}$ decreases under projections, together with the induction hypothesis, we have

$$\begin{aligned}
\tilde{n}^{\Phi(C)}\delta(\mathcal{A} \bowtie \mathcal{B}) &= \tilde{n}^{\Phi(A')+\Delta(B\ominus A')+\Delta(C\ominus\{A',B\})}\delta(\mathcal{A} \bowtie \mathcal{B}) && \text{(by (13))}\\
&\leq \tilde{n}^{\Phi(A')}\delta(\mathrm{proj}_{A'}(\mathcal{A})) && \text{(by (14))}\\
&\leq \bar{\chi}_{A'}(\mathrm{proj}_{A'}(\mathcal{A})) && \text{(ind. hyp.)}\\
&\leq \bar{\chi}_{A}(\mathcal{A}) && \text{(Lemma 10.2)}\\
&\leq \max\{\bar{\chi}_A(\mathcal{A}),\ \bar{\chi}_B(\mathcal{B})\}.
\end{aligned}$$

Therefore, (12) holds in this case.

Case (‡) (balanced induction case): Assume that there exist $A' \preceq A$ and $B' \preceq B$ such that $(\ddagger)^{C}_{A',B'}$ is tight, that is,

$$\Phi_C = \frac{1}{2}\Big(\Phi_{A'} + \Phi_{B'\ominus A'} + \Delta_C + \Delta_{C\ominus\{A',B'\}}\Big). \tag{15}$$

By Lemma 10.9, we have

$$\delta(\mathcal{A} \bowtie \mathcal{B}) \leq \pi_{A'}(\mathcal{A}) \cdot \mu_{B'\ominus A'}(\mathrm{proj}_{B'}(\mathcal{B})) \cdot \mu_{C\ominus\{A',B'\}}(\mathcal{A} \bowtie \mathcal{B}).$$

By definition of $\mu_{B'\ominus A'}$, there exists $z \in [n]^{V_{B'}\setminus V_{B'\ominus A'}}$ such that

$$\mu_{B'\ominus A'}(\mathrm{proj}_{B'}(\mathcal{B})) = \delta(\mathrm{proj}_{B'}(\mathcal{B})|^{z}_{B'\ominus A'}).$$

C-smallness of $\mathcal{A} \bowtie \mathcal{B}$ implies both

$$\delta(\mathcal{A} \bowtie \mathcal{B}) \leq \tilde{n}^{-\Delta(C)} \quad \text{and} \quad \mu_{C\ominus\{A',B'\}}(\mathcal{A} \bowtie \mathcal{B}) \leq \tilde{n}^{-\Delta(C\ominus\{A',B'\})}.$$

Taking square roots and combining these inequalities, we have

$$\delta(\mathcal{A} \bowtie \mathcal{B}) \leq \sqrt{\tilde{n}^{-\Delta(C)-\Delta(C\ominus\{A',B'\})} \cdot \pi_{A'}(\mathcal{A}) \cdot \delta(\mathrm{proj}_{B'}(\mathcal{B})|^{z}_{B'\ominus A'})}. \tag{16}$$

Using the fact that $\bar{\chi}$ decreases under projections and restrictions (Lemmas 10.2 and 10.6), together with the induction hypothesis, we have

$$\begin{aligned}
\tilde{n}^{\Phi(A')}\pi_{A'}(\mathcal{A}) = \tilde{n}^{\Phi(A')}\delta(\mathrm{proj}_{A'}(\mathcal{A})) &\leq \bar{\chi}_{A'}(\mathrm{proj}_{A'}(\mathcal{A})) && \text{(ind. hyp.)}\\
&\leq \bar{\chi}_{A}(\mathcal{A}) && \text{(Lemma 10.2)}
\end{aligned} \tag{17}$$

and also

$$\begin{aligned}
\tilde{n}^{\Phi(B'\ominus A')}\delta(\mathrm{proj}_{B'}(\mathcal{B})|^{z}_{B'\ominus A'}) &\leq \bar{\chi}_{B'\ominus A'}(\mathrm{proj}_{B'}(\mathcal{B})|^{z}_{B'\ominus A'}) && \text{(ind. hyp.)}\\
&\leq \bar{\chi}_{B'}(\mathrm{proj}_{B'}(\mathcal{B})) && \text{(Lemma 10.6)}\\
&\leq \bar{\chi}_{B}(\mathcal{B}) && \text{(Lemma 10.2).}
\end{aligned} \tag{18}$$

We now finish the proof using the inequality of arithmetic and geometric means (AM-GM inequality):

$$
\begin{aligned}
\tilde{n}^{\Phi(C)}\delta(\mathcal{A} \bowtie \mathcal{B}) &= \sqrt{\tilde{n}^{\Phi(A')+\Phi(B'\ominus A')+\Delta(C)+\Delta(C\ominus\{A',B'\})}} \cdot \delta(\mathcal{A} \bowtie \mathcal{B}) && \text{(by (15))} \\
&\le \sqrt{\tilde{n}^{\Phi(A')+\Phi(B'\ominus A')} \cdot \pi_{A'}(\mathcal{A}) \cdot \delta(\mathrm{proj}_{B'}(\mathcal{B})|^{z}_{B'\ominus A'})} && \text{(by (16))} \\
&\le \frac{1}{2}\Big(\tilde{n}^{\Phi(A')}\pi_{A'}(\mathcal{A}) + \tilde{n}^{\Phi(B'\ominus A')}\delta(\mathrm{proj}_{B'}(\mathcal{B})|^{z}_{B'\ominus A'})\Big) && \text{(AM-GM ineq.)} \\
&\le \frac{1}{2}\Big(\bar{\chi}_A(\mathcal{A}) + \bar{\chi}_B(\mathcal{B})\Big) && \text{(by (17), (18))} \\
&\le \max\{\bar{\chi}_A(\mathcal{A}),\, \bar{\chi}_B(\mathcal{B})\}.
\end{aligned}
$$

Therefore, (12) holds in this case also, which concludes the proof. □

11.3 Showing $\Phi_A \geq \frac{1}{4.41}\log(\ell_A) + \Delta_A$

We now complete the proof of Theorem 8.3 by proving Lemma 11.4 ($\Phi_A \geq \frac{1}{4.41}\log(\ell_A) + \Delta_A$ for all union trees A). We require one preliminary lemma.

Lemma 11.3. *For every union tree A and A-respecting $S \subseteq V_k$, we have $\Phi_A \geq \Phi_{A\restriction S} + \Delta_{A\restriction\overline{S}}$.*

Proof. We argue by induction on union trees. The lemma is trivial when A is empty or atomic. For the induction step, consider any non-atomic union tree $C = \{A, B\}$ and assume the lemma holds for all smaller union trees.

Let S be any C-respecting subset of V_k. Note that S is C'-respecting for any $C' \preceq C$. Also note that $C\restriction S = \{A\restriction S, B\restriction S\}$ and that every sub-union tree of $A\restriction S$ has the form $A'\restriction S$ where $A' \preceq A$ (and similarly for $B\restriction S$). We will also use the fact that $(A' \ominus B')\restriction S = (A'\restriction S) \ominus (B'\restriction S)$ for all $A' \preceq A$ and $B' \preceq B$.

From the definition of $\Phi_{C\restriction S}$, it follows that at least one the four inequalities

$$(\dagger)^{C\restriction S}_{A'\restriction S, B\restriction S}, \quad (\ddagger)^{C\restriction S}_{A'\restriction S, B'\restriction S}, \quad (\dagger)^{C\restriction S}_{B'\restriction S, A\restriction S}, \quad (\ddagger)^{C\restriction S}_{B'\restriction S, A'\restriction S}$$

is tight for some $A' \preceq A$ and $B' \preceq B$. Once again, without loss of generality, we consider just the first two possibilities.

First, consider the case that there exists $A' \preceq A$ for which $(\dagger)^{C\restriction S}_{A'\restriction S, B\restriction S}$ is tight, that is,

$$\Phi_{C\restriction S} = \Phi_{A'\restriction S} + \Delta_{(B\ominus A')\restriction S} + \Delta_{(C\ominus\{A',B\})\restriction S}. \tag{19}$$

In this case, we have

$$
\begin{aligned}
\Phi_C &\geq \Phi_{A'} + \Delta_{B\ominus A'} + \Delta_{C\ominus\{A',B\}} && \text{(by } (\dagger)^{C}_{A',B}) \\
&\geq \Phi_{A'} + \Delta_{B\ominus A'} + \Delta_{C\ominus\{A',B\}} \\
&\quad + \Delta_{C\restriction\overline{S}} - \Delta_{A'\restriction\overline{S}} - \Delta_{(B\ominus A')\restriction\overline{S}} - \Delta_{(C\ominus\{A',B\})\restriction\overline{S}} && \text{(Lemma 10.8)} \\
&= \Phi_{A'} - \Delta_{A'\restriction\overline{S}} + \Delta_{(B\ominus A')\restriction S} + \Delta_{(C\ominus\{A',B\})\restriction S} + \Delta_{C\restriction\overline{S}} \\
&\geq \Phi_{A'\restriction S} + \Delta_{(B\ominus A')\restriction S} + \Delta_{(C\ominus\{A',B\})\restriction S} + \Delta_{C\restriction\overline{S}} && \text{(ind. hyp.)} \\
&= \Phi_{C\restriction S} + \Delta_{C\restriction\overline{S}} && \text{(by (19)).}
\end{aligned}
$$

Finally, consider the alternative that there exist $A' \preceq A$ and $B' \preceq B$ for which $(\ddagger)^{C\restriction S}_{A'\restriction S, B'\restriction S}$ is tight, that is,

$$\Phi_{C\restriction S} = \frac{1}{2}\Big(\Phi_{A'\restriction S} + \Phi_{(B'\ominus A')\restriction S} + \Delta_{C\restriction S} + \Delta_{(C\ominus\{A',B'\})\restriction S}\Big). \tag{20}$$

In this case, we have

$$\begin{aligned}
\Phi_C &\geq \tfrac{1}{2}\big(\Phi_{A'} + \Phi_{B'\ominus A'} + \Delta_C + \Delta_{C\ominus\{A',B'\}}\big) && \text{(by } (\ddagger)^{C}_{A',B'}) \\
&\geq \tfrac{1}{2}\big(\Phi_{A'} + \Phi_{B'\ominus A'} + (\Delta_{C\restriction S} + \Delta_{C\restriction \overline{S}}) + \Delta_{C\ominus\{A',B'\}}\big) + \\
&\quad \tfrac{1}{2}\big(\Delta_{C\restriction \overline{S}} - \Delta_{A'\restriction \overline{S}} - \Delta_{(B'\ominus A')\restriction \overline{S}} - \Delta_{(C\ominus\{A',B'\})\restriction \overline{S}}\big) && \text{(Lemma 10.8)} \\
&= \tfrac{1}{2}\big(\Phi_{A'} - \Delta_{A'\restriction \overline{S}} + \Phi_{B'\ominus A'} - \Delta_{(B'\ominus A')\restriction \overline{S}} + \Delta_{C\restriction S} + \Delta_{(C\ominus\{A',B'\})\restriction S}\big) + \Delta_{C\restriction \overline{S}} \\
&\geq \tfrac{1}{2}\big(\Phi_{A'\restriction S} + \Phi_{(B'\ominus A')\restriction S} + \Delta_{C\restriction S} + \Delta_{(C\ominus\{A',B'\})\restriction S}\big) + \Delta_{C\restriction \overline{S}} && \text{(ind. hyp.)} \\
&= \Phi_{C\restriction S} + \Delta_{C\restriction \overline{S}} && \text{(by (20))}.
\end{aligned}$$

Having shown $\Phi_C \geq \Phi_{C\restriction S} + \Delta_{C\restriction \overline{S}}$ in both cases, we are done. □

Lemma 11.4. *For every union tree A, $\Phi_A \geq \frac{1}{c}\log(\ell_A) + \Delta_A$ where $c = 2\log(\sqrt{13}+1) \leq 4.41$.*

An earlier version of this article had the constant $c = 6$ in Lemma 11.4. An anonymous referee pointed out an optimization in the proof which gives the better constant c. (We remark that the best possible constant c in this lemma is ≥ 2 by Proposition 5.11 and Corollary A.3.)

Proof. Here c is chosen such that $\frac{1}{2} - \frac{1}{2^{c/2}} - \frac{1}{2^{c-1}} = \frac{1}{2^{c-2}}$.

We argue by induction on union trees. The base case where A is empty or atomic is trivial. For the induction step, let A be a non-atomic union tree and assume the lemma holds for all smaller union trees. We will consider a sequence of cases. In each case, after showing that $\Phi_A \geq \frac{1}{c}\log(\ell_A) + \Delta_A$ under a given hypothesis, we will proceed assuming the negation of that hypothesis. The sequences of cases is summarized at the end of the proof.

First, consider the case that G_A is disconnected (i.e. $\Delta_A \geq 2$). Let S be the largest component of G_A. We have

$$\begin{aligned}
\Phi_A &\geq \Phi_{A\restriction S} + \Delta_{A\restriction \overline{S}} && \text{(Lemma 11.3)} \\
&\geq \tfrac{1}{c}\log(\ell_{A\restriction S}) + \Delta_{A\restriction S} + \Delta_{A\restriction \overline{S}} && \text{(ind. hyp.)} \\
&= \tfrac{1}{c}\log(\ell_A) + \Delta_A.
\end{aligned}$$

This proves the lemma in the case where G_A is disconnected.

Therefore, we proceed under the assumption that G_A is connected (i.e. $\Delta_A = 1$). Without loss of generality, we assume that $G_A = P_k$ (i.e. $\ell_A = k$). Our goal is to show that

$$\Phi_A \geq \tfrac{1}{c}\log(k) + 1.$$

Consider the case that there exists a sub-union tree $A' \preceq A$ such that $|E_{A'}| \geq \frac{1}{2^{c-1}}k$ and $\Delta_{A'} \geq 2$. Note that $\ell_{A'} \geq |E_{A'}|/\Delta_{A'}$ (i.e. the number of edges in the largest component of $G_{A'}$ is at least the number of edges in $G_{A'}$ divided by the number of components in $G_{A'}$). We have

$$\begin{aligned}
\Phi_A \geq \Phi_{A'} &\geq \tfrac{1}{c}\log(\ell_{A'}) + \Delta_{A'} && \text{(ind. hyp.)} \\
&\geq \tfrac{1}{c}\log(k) - \tfrac{c-1}{c} - \tfrac{1}{c}\log(\Delta_{A'}) + \Delta_{A'} && (\ell_{A'} \geq |E_{A'}|/\Delta_{A'} \geq \tfrac{1}{2^{c-1}}k\Delta_{A'}) \\
&\geq \tfrac{1}{c}\log(k) - \tfrac{c-1}{c} - \tfrac{1}{c}\log(2) + 2 && (\Delta_{A'} \geq 2 \text{ and } x - \tfrac{1}{c}\log x \text{ increasing for } x \geq 2) \\
&= \tfrac{1}{c}\log(k) + 1.
\end{aligned}$$

This proves the lemma in this case.

Therefore, we proceed under the following assumption:

$$(\circledast) \qquad \text{for all } A' \preceq A, \text{ if } |E_{A'}| \geq \tfrac{1}{2^{c-1}}k \text{ then } \Delta_{A'} = 1.$$

Going forward, the following notation will be convenient: for a proper sub-union tree $B \prec A$, let $B^{\uparrow}$ denote the parent of B in A, and let $B^{\sim}$ denote the sibling of B in A. Note that $B^{\uparrow} = \{B, B^{\sim}\} \preceq A$.

By walking down the union tree A, we can proper sub-union trees $B, Z \prec A$ such that

$$v_0 \in V_B, \qquad v_k \in V_Z, \qquad |E_B|, |E_Z| < \tfrac{1}{2^{c/2}}k, \qquad |E_{B^{\uparrow}}|, |E_{Z^{\uparrow}}| \geq \tfrac{1}{2^{c/2}}k.$$

Fix any choice of such B and Z. Note that $G_{B^{\uparrow}}$ and $G_{Z^{\uparrow}}$ are connected by ($\circledast$). In particular, $G_{B^{\uparrow}}$ is a path of length $|E_{B^{\uparrow}}|$ with initial endpoint v_0, and $G_{Z^{\uparrow}}$ is a path of length $|E_{Z^{\uparrow}}|$ with final endpoint v_k.

Consider the case that $B^{\uparrow}$ and $Z^{\uparrow}$ are vertex-disjoint. Note that $\frac{1}{2^{c/2}}k \geq \frac{1}{2^{c-1}}k$, so the assumption ($\circledast$) implies that $B^{\uparrow}$ and $Z^{\uparrow}$ are connected and $\ell_{B^{\uparrow}}, \ell_{Z^{\uparrow}} \geq \frac{1}{2^{c/2}}k$. Let Y denote the least common ancestor of $B^{\uparrow}$ and $Z^{\uparrow}$ in A. We have

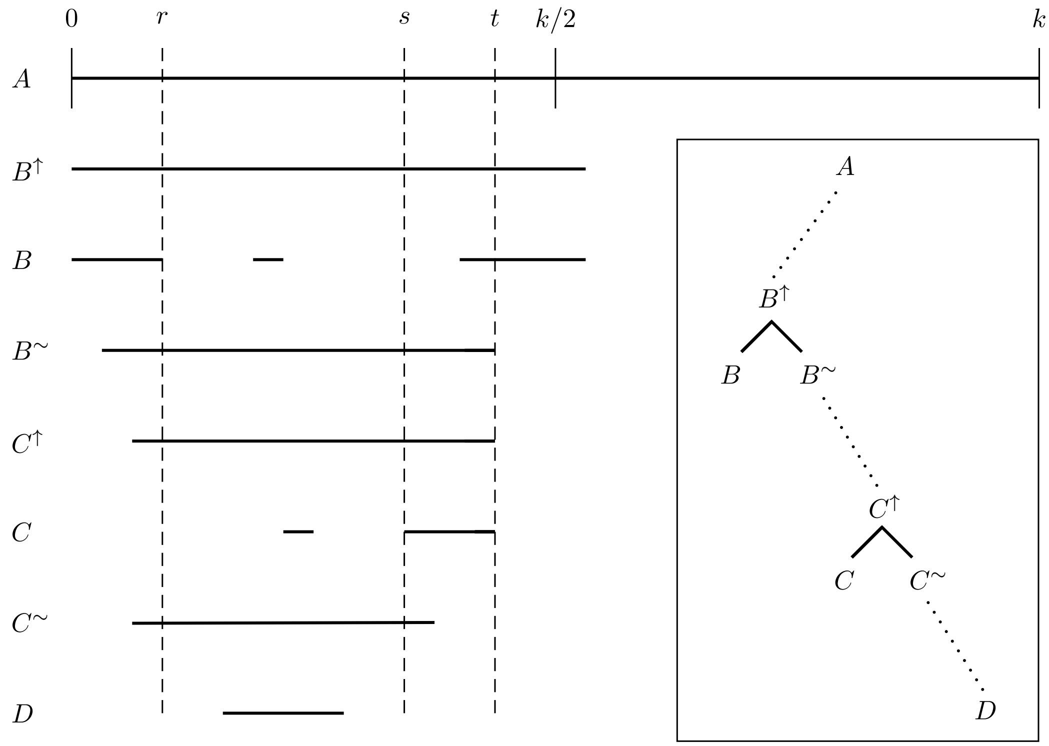
$$\begin{aligned}
\Phi_A \geq \Phi_Y &\geq \tfrac{1}{2}\big(\Phi_{B^{\uparrow}} + \Phi_{Z^{\uparrow} \ominus B^{\uparrow}} + \Delta_Y + \Delta_{Y \ominus \{B^{\uparrow}, Z^{\uparrow}\}}\big) && (\text{by } (\ddagger)^{Y}_{B^{\uparrow}, Z^{\uparrow}})\\
&\geq \tfrac{1}{2}\big(\Phi_{B^{\uparrow}} + \Phi_{Z^{\uparrow}}\big) + \tfrac{1}{2} && (\Delta_Y \geq 1 \text{ and } Z^{\uparrow} \ominus B^{\uparrow} = Z^{\uparrow})\\
&\geq \tfrac{1}{2}\big(\tfrac{1}{c}\log(\ell_{B^{\uparrow}}) + \Delta_{B^{\uparrow}} + \tfrac{1}{c}\log(\ell_{Z^{\uparrow}}) + \Delta_{Z^{\uparrow}}\big) + \tfrac{1}{2} && (\text{ind. hyp.})\\
&\geq \tfrac{1}{2}\big(\tfrac{1}{c}\log(\tfrac{1}{2^{c/2}}k) + 1 + \tfrac{1}{c}\log(\tfrac{1}{2^{c/2}}k) + 1\big) + \tfrac{1}{2}\\
&= \tfrac{1}{c}\log(k) + 1.
\end{aligned}$$

Therefore, we proceed under the assumption that $B^{\uparrow}$ and $Z^{\uparrow}$ are not vertex-disjoint. It follows that $\ell_{B^{\uparrow}} \geq k/2$ or $\ell_{Z^{\uparrow}} \geq k/2$. Without loss of generality, we assume that $\ell_{B^{\uparrow}} \geq k/2$. (We now forget about Z and $Z^{\uparrow}$.)

Before continuing, let's take stock of the assumptions we have made so far:

$$G_A = P_k, \qquad (\circledast), \qquad B \preceq A, \qquad v_0 \in V_B, \qquad |E_B| < \tfrac{1}{2^{c-1}}k, \qquad |E_{B^{\uparrow}}| = \ell_{B^{\uparrow}} \geq k/2.$$

Going forward, we will define vertices v_r, v_s, v_t where $0 < r < s < t \leq k$. The following illustration might be helpful for what follows:



We first define $v_r \in B$ and $v_t \in B^\sim$ as follows: Let $\{v_0, \ldots, v_r\}$ be the component of G_B containing v_0. (That is, the component of v_0 in G_B is a path whose initial vertex is v_0; let v_r be the final vertex in this path.) Let v_t be the vertex in $V_{B^\sim}$ with maximal index t (i.e. farthest away from v_0).

Note that E_B contains edges $v_i v_{i+1}$ for all $i \in \{0, \ldots, r-1\} \cup \{t, \ldots, \lceil k/2 \rceil - 1\}$. (In the event that $t < k/2$, since $G_{B^\uparrow} = G_B \cup G_{B^\sim}$ is a path of length $\geq k/2$ and $G_{B^\sim}$ does not contain vertices $v_{t+1}, \ldots, v_{\lceil k/2 \rceil}$, it follows that G_B contains all edges between v_t and $v_{\lceil k/2 \rceil}$.) Therefore, $r + (k/2) - t \leq |E_B| < \frac{1}{2^{c/2}} k$. It follows that

$$t - r > (\tfrac{1}{2} - \tfrac{1}{2^{c/2}})k.$$

Next, note that $|E_{B^\sim}| \geq |E_{B^\uparrow}| - |E_B| \geq (\frac{1}{2} - \frac{1}{2^{c/2}})k > \frac{1}{2^{c-1}} k$. We now walk down $B^\sim$ to find a proper sub-union tree $C \prec B^\sim$ such that

$$v_t \in V_C, \qquad |E_C| < \tfrac{1}{2^{c-1}} k, \qquad |E_{C^\uparrow}| \geq \tfrac{1}{2^{c-1}} k.$$

Fix any choice of such C. Note that $G_{C^\uparrow}$ is connected by ($\circledast$).

Consider the case that $|E_{C^\uparrow}| < (\frac{1}{2} - \frac{1}{2^{c/2}})k$. Since $G_{C^\uparrow}$ is connected and $v_t \in V_{C^\uparrow}$ and $t - r > (\frac{1}{2} - \frac{1}{2^{c/2}})k$, it follows that $V_{C^\uparrow} \cap \{v_0, \ldots, v_r\} = \emptyset$ and hence $\Delta_{B \ominus C^\uparrow} \geq 1$. We have

$$\begin{aligned}
\Phi_A \geq \Phi_{B^\uparrow} &\geq \Phi_{C^\uparrow} + \Delta_{B \ominus C^\uparrow} + \Delta_{B^\uparrow \ominus \{B, C^\uparrow\}} && \text{(by } (\dagger)^{B^\uparrow}_{C^\uparrow, B}) \\
&\geq \Phi_{C^\uparrow} + 1 \\
&\geq \tfrac{1}{c} \log(\ell_{C^\uparrow}) + \Delta_{C^\uparrow} + 1 && \text{(ind. hyp.)} \\
&\geq \tfrac{1}{c} \log(\tfrac{1}{2^{c-1}} k) + 2 \\
&> \tfrac{1}{c} \log(k) + 1.
\end{aligned}$$

Therefore, we proceed under the assumption that $|E_{C^\uparrow}| \geq (\frac{1}{2} - \frac{1}{2^{c/2}})k$. Since $E_{C^\uparrow} = E_C \cup E_{C^\sim}$, we have

$$|E_{C^\sim}| \geq |E_{C^\uparrow}| - |E_C| > (\tfrac{1}{2} - \tfrac{1}{2^{c/2}} - \tfrac{1}{2^{c-1}})k = \tfrac{1}{2^{c-2}}k.$$

We now define vertex $v_s \in V_C$. Since v_t is the vertex of $G_{B^\sim}$ with maximal index, it follows that $v_t v_{t+1} \notin E_{B^\sim}$ and hence $v_t v_{t+1} \notin E_C$ (since $C \prec B^\sim$). Therefore, the component of G_C containing v_t is a path with final vertex v_t; let v_s be the initial vertex in this path. That is, $\{v_s, \ldots, v_t\}$ is the component of G_C which contains v_t. Recall that $t - r > (\frac{1}{2} - \frac{1}{2^{c/2}})k$ and note that $t - s \leq |E_C| < \frac{1}{2^{c-1}}k$. Therefore,

$$s - r = (t - r) - (t - s) > (\tfrac{1}{2} - \tfrac{1}{2^{c/2}} - \tfrac{1}{2^{c-1}})k = \tfrac{1}{2^{c-2}}k.$$

We now claim that there exists a proper sub-union tree $D \prec C^\sim$ such that

$$\tfrac{1}{2^{c-1}}k \leq |E_D| < \tfrac{1}{2^{c-2}}k.$$

To see this, note that there exists a chain of sub-union trees $C^\sim = D_0 \succ D_1 \succ \cdots \succ D_j$ such that D_j is atomic and $D_{i-1} = D_i^\uparrow$ and $|E_{D_i}| \geq |E_{D_i^\sim}|$ for all $i \in \{1, \ldots, j\}$. Since $|E_{D_0}| > \frac{1}{2^{c-2}}k$ and $|E_{D_j}| = 1$ and $|E_{D_{i-1}}| \leq |E_{D_i}| + |E_{D_i^\sim}| \leq 2|E_{D_i}|$, it must be the case that there exists $i \in \{1, \ldots, j\}$ such that $\frac{1}{2^{c-1}}k \leq |E_{D_i}| < \frac{1}{2^{c-2}}k$.

Since $|E_D| \geq \frac{1}{2^{c-1}}k$, (⊛) implies that G_D is connected. Since $|E_D| < \frac{1}{2^{c-2}}k$ and $s - r > \frac{1}{2^{c-2}}k$, it follows that V_D cannot contain both v_r and v_s. We are now down to our final two cases: either $v_r \notin V_D$ or $v_s \notin V_D$.

First, suppose that $v_r \notin V_D$. We have $\Delta_{B \ominus D} \geq 1$ and hence

$$\begin{aligned}
\Phi_A \geq \Phi_{B^\uparrow} &\geq \Phi_D + \Delta_{B \ominus D} + \Delta_{B^\uparrow \ominus \{B,D\}} && \text{(by } (\dagger)^{B^\uparrow}_{D,B}) \\
&\geq \Phi_D + 1 \\
&\geq \tfrac{1}{c}\log(\ell_D) + \Delta_D + 1 && \text{(ind. hyp.)} \\
&\geq \tfrac{1}{c}\log(\tfrac{1}{2^{c-1}}k) + 2 \\
&> \tfrac{1}{c}\log(k) + 1.
\end{aligned}$$

Finally, we are left with the alternative that $v_s \notin V_D$. In this case $\Delta_{C \ominus D} \geq 1$ and hence (substituting C for B in the above), we have

$$\Phi_A \geq \Phi_{C^\uparrow} \geq \Phi_D + \Delta_{C \ominus D} + \Delta_{C^\uparrow \ominus \{C,D\}} \geq \Phi_D + 1 > \tfrac{1}{c}\log(k) + 1.$$

We have now covered all cases. In summary, we considered cases in the following sequence:

1. $\Delta_A \geq 2$ — else assume wlog $G_A = P_k$,
2. $\exists A' \prec A$ with $\Delta_{A'} \geq 2$ and $\ell_{A'} \geq \frac{1}{2^{c-1}}k$ — else assume (⊛),
3. $B^\uparrow$ and $Z^\uparrow$ are vertex-disjoint — else assume wlog $|E_{B^\uparrow}| \geq k/2$,
4. $|E_{C^\uparrow}| < (\frac{1}{2} - \frac{1}{2^{c/2}})k$ — else assume $|E_{C^\uparrow}| \geq (\frac{1}{2} - \frac{1}{2^{c/2}})k$,
5. $v_r \notin V_D$ or $v_s \notin V_D$.

Since $\Phi_A \geq \frac{1}{c}\log(\ell_A) + \Delta_A$ in each case, the proof is complete. □

As we have now proved Lemmas 11.2 and 11.4, this completes the proof of Theorem 8.3 and hence also of Theorem 2.1.

12 Conclusion

We proved the first super-polynomial separation in the power of bounded-depth Boolean formulas vs. circuits via technique based on the notion of pathset complexity. A natural question for future research is whether the pathset complexity technique can be used to derive lower bounds for distance $k(n)$ connectivity in other models of computation. In a subsequent work of the author [19], pathset complexity was used to prove the first average-case lower bounds under product distributions against the class $\mathsf{monotone\text{-}NC}^1$ of polynomial-size monotone formulas.

We remark that the results in this paper also extend to the *average-case* setting. Let $p(n) = \Theta(n^{-(k+1)/k})$ be the exact threshold function such that

$$\mathop{\mathbb{P}}_{G=G(n,p)}[\, G \in \textsc{stconn}(k(n)) \,] = 1/2$$

where $G(n,p)$ is the Erdős-Rényi random graph with edge probability $p(n)$. Our proof of Theorem 2.1 is easily adapted to give the same $n^{(1/4.41)\log k - O(1)}$ lower bound for bounded-depth formulas F which satisfy

$$\mathop{\mathbb{P}}_{G=G(n,p)}[\, F(G) = 1 \iff G \in \textsc{stconn}(k(n)) \,] \geq 1/2 + \varepsilon$$

for any constant $\varepsilon > 0$. Using the idea behind Proposition 5.11, we can construct formulas F of size $n^{(1/2)\log k + O(1)}$ and depth $O(\log k)$ which solve $\textsc{stconn}(k(n))$ in a strong average-case sense:

$$\mathop{\mathbb{P}}_{G=G(n,p)}[\, F(G) = 1 \iff G \in \textsc{stconn}(k(n)) \,] \geq 1 - \exp(-n^{\Omega(1)}).$$

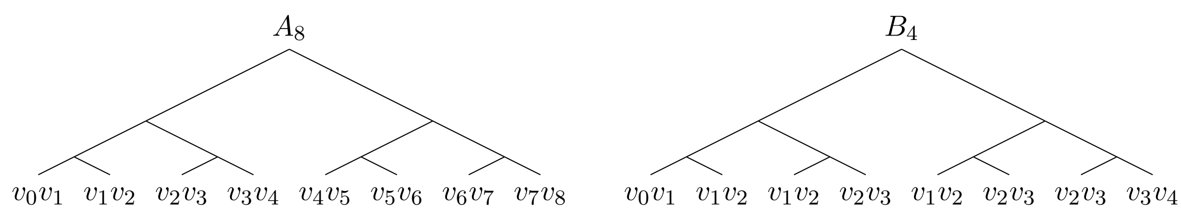
It would be interesting to close the gap between $\frac{1}{4.41}\log k$ and $\frac{1}{2}\log k$ in these upper and lower bounds.

A Key Examples

We introduce two key examples of union trees, denoted A_k and B_k, and present upper bounds for $\bar{\chi}$ with respect to these union trees. In the next section, we prove lower bounds for two classes of union trees which generalize A_k and B_k. The arguments in these special cases show up in the two cases (†) and (‡) of our main lower bound (Theorem 8.3).

Notation A.1. Recall Notation 5.12 for *s-shifted* pattern graphs $G^{\triangleright s}$ and pathsets $\mathcal{A}^{\triangleright s}$. For a union tree A and integer s, we define the *s-shifted* union tree $A^{\triangleright s}$ analogously by replacing each label $v_i v_{i+1}$ with the label $v_{i+s}v_{i+s+1}$.

Definition A.2 (Union Trees A_k and B_k)**.** We define union trees A_k and B_k for all $k \geq 1$ by the following induction. Let $A_1 = B_1 :=$ the atomic union tree labeled by v_0v_1. For $k \geq 2$, let $A_k := \{A_j, A_{k-j}^{\triangleright j}\}$ where $j = \lceil k/2 \rceil$, and let $B_k := \{B_{k-1}, B_{k-1}^{\triangleright 1}\}$. For example, the explicit pictures of A_8 and B_4 are:



Intuitively, the union tree A_k corresponds to the recursive doubling algorithm for $\textsc{distconn}(k,n)$. Note that we have essentially already encountered this union tree in the proof of Proposition 5.11 (our upper bound for χ_{P_k}). In fact, this proof shows:

Corollary A.3. *For all $\mathcal{A} \in \mathscr{P}_{A_k}$, $\bar{\chi}_{A_k}(\mathcal{A}) \le O(n^{(1/2)\lceil \log k \rceil + 2})$.* □

The union tree B_k has a different nature than A_k. Whereas sub-union trees A_j and $A^{\triangleright j}_{k-j}$ of A_k overlap at only a single vertex v_j, sub-union trees B_{k-1} and $B^{\triangleright 1}_{k-1}$ of B_k overlap to the maximum possible extent. Despite this difference, it turns out that there is also a reasonable upper bound for $\bar{\chi}_{B_k}$.

Proposition A.4. *For all $\mathcal{B} \in \mathscr{P}_{B_k}$, $\bar{\chi}_{B_k}(\mathcal{B}) \le n^{\ln k+1}$.*

Proof. We present a similar argument to the proof of Proposition 5.11. For all $k \ge 1$, define $\mathcal{B}_k \in \mathscr{P}_{B_k}$ by

$$\mathcal{B}_k := \{x \in [n]^{V_k} : x_0, \dots, x_k \le n^{1-1/(k+1)}\}.$$

We have $\mathcal{B}_{k-1} \bowtie \mathcal{B}^{\triangleright 1}_{k-1} = \{x \in [n]^{V_k} : x_0, \dots, x_k \le n^{1-1/k}\}$. For all $1 \le t_0, \dots, t_k \le n^{1/k(k+1)}$, let

$$\mathrm{Copy}_{t_0,\dots,t_k}(\mathcal{B}_{k-1} \bowtie \mathcal{B}^{\triangleright 1}_{k-1}) := \Big\{x \in [n]^{V_k} : t_i - 1 < \frac{x_i}{n^{1-1/k}} \le t_i \text{ for all } 0 \le i \le k\Big\}.$$

Note that

$$\mathcal{B}_k = \bigcup_{1 \le t_0,\dots,t_k \le n^{1/k(k+1)}} \mathrm{Copy}_{t_0,\dots,t_k}(\mathcal{B}_{k-1} \bowtie \mathcal{B}^{\triangleright 1}_{k-1}).$$

Using (sub-additivity) and (join rule), together with the invariance of $\bar{\chi}$ under coordinate-wise permutations of $[n]$ and under shifts, we have

$$\bar{\chi}_{B_k}(\mathcal{B}_k) \le n^{1/k} \bar{\chi}_{B_{k-1}}(\mathcal{B}_{k-1}).$$

This recurrence, together with the base case $\bar{\chi}_{B_1}(\mathcal{B}_1) = n$, implies

$$\bar{\chi}_{B_k}(\mathcal{B}_k) \le n^{1+(1/2)+\dots+(1/k)} \le n^{\ln k+1}.$$

Noting that $[n]^{V_k}$ is covered by n copies of $\mathcal{B}_k$, we have $\bar{\chi}_{B_k}([n]^{V_k}) \le n^{\ln k+1}$. The proposition then follows using (monotonicity). □

In Appendix B we prove matching lower bounds for $\bar{\chi}_{A_k}$ and $\bar{\chi}_{B_k}$. In fact, these lower bounds apply to two classes of union trees which include A_k and B_k. While the upper bounds for $\bar{\chi}_{A_k}$ and $\bar{\chi}_{B_k}$ are quite similar, our lower bound arguments are significantly different. The arguments in these two special cases—a "one-sided" induction for $\bar{\chi}_{A_k}$ and a "balanced" induction using the AM-GM inequality for $\bar{\chi}_{B_k}$—show up in the two cases (†) and (‡) of our general lower bound (Theorem 8.3). For this reason, the reader might find the results in Appendix B to be a helpful warm-up.

Remark A.5. The pathsets $\mathcal{A}_k$ and $\mathcal{B}_k$ which show up in the proofs of our upper bounds are of a particularly simple form: they are *rectangular* subsets of $[n]^{V_k}$. In Appendix C we discuss a notion of *rectangular pathset complexity* $\bar{\chi}^{\mathrm{rect}}$. Proving lower bounds for $\bar{\chi}^{\mathrm{rect}}$ turns our to be much easier than for $\bar{\chi}$. We present an example (the "palindrome pathset") which illustrates the difficulty in attempting to generalize this easier lower bound to the non-rectangular setting.

B Lower Bound for $\bar{\chi}$: Special Cases

We prove easier special cases of our lower bound for $\bar{\chi}$ with respect to two classes of union trees which include the key examples A_k and B_k introduced in §A. Although the results of this appendix are not used in the main body of the paper, the arguments in the proof show up in the two cases (†) and (‡) of our general lower bound.

Definition B.1.

(i) For a union tree A,

— let $V_{\text{ends}}(A) \subseteq V_A$ denote the set of *endpoints* in G_A (i.e. vertices of in-degree or out-degree zero), and let $V_{\text{interior}}(A) := V_A \setminus V_{\text{ends}}(A)$ denote the set of *interior vertices* in G_A,

— let $\mathcal{I}(A)$ denote the set of *intervals* in G_A (i.e. nonempty subsets of V_A which are connected in G_A).

Note that $\ell_A = \max_{I \in \mathcal{I}(A)} |I| - 1$ and $\Delta_A = |V_{\text{ends}}(A)| \,/\, 2$.

(ii) The classes of *end-joining* and *fully connected* union trees are defined as follows:

— A is *end-joining* if no edge of P_k labels more than one leaf of A (equivalently, $E_{A_1} \cap E_{A_2} = \emptyset$ for all non-atomic sub-union trees $\{A_1, A_2\} \preceq A$),

— A is *fully connected* if $G_{A'}$ is connected (i.e. $\Delta_{A'} = 1$) for all sub-union trees $A' \preceq A$.

Note that union trees A_k and B_k are both fully connected, while only A_k is end-joining (for $k \geq 3$).

(iii) Functions $\psi_A, \xi_A : \mathscr{P}_A \to \mathbb{R}$ are defined as follows:

— for end-joining union trees A,

$$\psi_A(\mathcal{A}) := \tilde{n}^{\frac{1}{2}\left(\log(\ell_A) + \Delta_A\right)} \sqrt{\mathbb{E}_{z \in [n]^{V_{\text{ends}}(A)}} \left[\, \delta(\mathcal{A}|^{z}_{V_{\text{interior}}(A)})^2 \,\right]},$$

— for fully connected union trees A,

$$\xi_A(\mathcal{A}) := \max_{I \in \mathcal{I}(A)} \tilde{n}^{\frac{1}{4}\left(\log(|I|+1) + |I \cap V_{\text{ends}}(A)|\right)} \cdot \pi_I(\mathcal{A}).$$

For non-end-joining union trees A, we set $\psi_A(\mathcal{A}) := 0$, and for non-fully connected union trees A, we set $\xi_A(\mathcal{A}) := 0$.

Proposition B.2. *Both ψ and ξ are lower bounds on pathset complexity. That is, for every union tree A and pathset $\mathcal{A} \in \mathscr{P}_A$, we have $\psi_A(\mathcal{A}) \leq \bar{\chi}_A(\mathcal{A})$ and $\xi_A(\mathcal{A}) \leq \bar{\chi}_A(\mathcal{A})$. In particular,*

$$\bar{\chi}_{A_k}([n]^{V_k}) \geq \psi_{A_k}([n]^{V_k}) \geq \tilde{n}^{\frac{1}{2}(\log(k)+1)} \geq n^{\frac{1}{2}\log k},$$

$$\bar{\chi}_{B_k}([n]^{V_k}) \geq \xi_{B_k}([n]^{V_k}) \geq \tilde{n}^{\frac{1}{4}(\log(k+1)+2)} \geq n^{\frac{1}{4}\log k}.$$

Recall from Remark 8.4 the dual characterization of $\bar{\chi}$ as the unique coordinate-wise maximal function from pairs $(A, \mathcal{A})$ to $\mathbb{R}$ which satisfies inequalities (base case), (monotone), (sub-additive) and (join rule). It is easy to see that ψ and ξ satisfy the first three of these inequalities. To prove Proposition B.2, it suffices to show that ψ and ξ also satisfy inequality (join rule). We begin with ψ.

Lemma B.3. *For every non-atomic end-joining union tree $C = \{A, B\}$ and small pathsets $\mathcal{A} \in \mathscr{P}_A^{\text{small}}$ and $\mathcal{B} \in \mathscr{P}_B^{\text{small}}$,*

$$\psi_C(\mathcal{A} \bowtie \mathcal{B}) \leq \max\{\psi_A(\mathcal{A}), \psi_B(\mathcal{B})\}.$$

Proof. Without loss of generality, assume that $\ell_A \geq \ell_B$. After making three observations, will show that $\psi_C(\mathcal{A} \bowtie \mathcal{B}) \leq \psi_A(\mathcal{A})$.

First, note that each connected component of G_C $(= G_A \cup G_B)$ is the union of at most $\Delta_A + \Delta_B - \Delta_C + 1$ components of G_A and G_B. It follows that $\ell_C \leq (\Delta_A + \Delta_B - \Delta_C + 1) \cdot \ell_A$.

Since C is end-joining, $V_{\text{ends}}(C)$ is the symmetric difference of $V_{\text{ends}}(A)$ and $V_{\text{ends}}(B)$. By the Cauchy-Schwarz inequality,

$$\begin{aligned}
&\mathop{\mathbb{E}}_{c \in [n]^{V_{\text{ends}}(C)}} \left[\, \delta((\mathcal{A} \bowtie \mathcal{B})|^{c}_{V_{\text{interior}}(C)})^2 \,\right] \\
&\qquad = \mathop{\mathbb{E}}_{\substack{x \in [n]^{V_{\text{ends}}(A) \setminus V_{\text{ends}}(B)} \\ y \in [n]^{V_{\text{ends}}(B) \setminus V_{\text{ends}}(A)}}} \left[\left(\mathop{\mathbb{E}}_{z \in [n]^{V_{\text{ends}}(A) \cap V_{\text{ends}}(B)}} \left[\, \delta(\mathcal{A}|^{xz}_{V_{\text{interior}}(A)}) \cdot \delta(\mathcal{B}|^{yz}_{V_{\text{interior}}(B)}) \,\right] \right)^2 \right] \\
&\qquad \leq \mathop{\mathbb{E}}_{\substack{x \in [n]^{V_{\text{ends}}(A) \setminus V_{\text{ends}}(B)} \\ z \in [n]^{V_{\text{ends}}(A) \cap V_{\text{ends}}(B)}}} \left[\, \delta(\mathcal{A}|^{xz}_{V_{\text{interior}}(A)})^2 \,\right] \mathop{\mathbb{E}}_{\substack{y \in [n]^{V_{\text{ends}}(B) \setminus V_{\text{ends}}(A)} \\ z \in [n]^{V_{\text{ends}}(A) \cap V_{\text{ends}}(B)}}} \left[\, \delta(\mathcal{B}|^{yz}_{V_{\text{interior}}(B)})^2 \,\right] \\
&\qquad = \mathop{\mathbb{E}}_{a \in [n]^{V_{\text{ends}}(A)}} \left[\, \delta(\mathcal{A}|^{a}_{V_{\text{interior}}(A)})^2 \,\right] \mathop{\mathbb{E}}_{b \in [n]^{V_{\text{ends}}(B)}} \left[\, \delta(\mathcal{B}|^{b}_{V_{\text{interior}}(B)})^2 \,\right].
\end{aligned}$$

We next note that B-smallness of $\mathcal{B}$ implies

$$\mathop{\mathbb{E}}_{b \in [n]^{V_{\text{ends}}(B)}} \left[\, \delta(\mathcal{B}|^{b}_{V_{\text{interior}}(B)})^2 \,\right] \leq \mathop{\mathbb{E}}_{b \in [n]^{V_{\text{ends}}(B)}} \left[\, \delta(\mathcal{B}|^{b}_{V_{\text{interior}}(B)}) \,\right] = \delta(\mathcal{B}) \leq \tilde{n}^{-\Delta_B}.$$

Putting these inequalities together, we have

$$\begin{aligned}
\psi_C(\mathcal{A} \bowtie \mathcal{B}) &= \tilde{n}^{\frac{1}{2}\left(\log(\ell_C) + \Delta_C\right)} \sqrt{\mathbb{E}_{c \in [n]^{V_{\text{ends}}(C)}} \left[\, \delta(\mathcal{A} \bowtie \mathcal{B}|^{c}_{V_{\text{interior}}(C)})^2 \,\right]} \\
&\leq \tilde{n}^{\frac{1}{2}\left(\log(\ell_A) + \log(\Delta_A + \Delta_B - \Delta_C + 1) + \Delta_C - \Delta_B\right)} \sqrt{\mathbb{E}_{a \in [n]^{V_{\text{ends}}(A)}} \left[\, \delta(\mathcal{A}|^{a}_{V_{\text{interior}}(A)})^2 \,\right]} \\
&= \tilde{n}^{\frac{1}{2}\left(\log(\Delta_A + \Delta_B - \Delta_C + 1) + \Delta_C - \Delta_B - \Delta_A\right)} \cdot \psi_A(\mathcal{A}) \\
&\leq \psi_A(\mathcal{A})
\end{aligned}$$

using the fact that $\log(s+1) \leq s$ for every integer $s \geq 0$. We get $\psi_C(\mathcal{A} \bowtie \mathcal{B}) \leq \max\{\psi_A(\mathcal{A}), \psi_B(\mathcal{B})\}$ as required. □

We next show that ξ satisfies inequality (join rule).

Lemma B.4. *For every non-atomic fully connected union tree $C = \{A, B\}$ and small pathsets $\mathcal{A} \in \mathscr{P}_A^{\text{small}}$ and $\mathcal{B} \in \mathscr{P}_B^{\text{small}}$,*

$$\xi_C(\mathcal{A} \bowtie \mathcal{B}) \leq \frac{\xi_A(\mathcal{A}) + \xi_B(\mathcal{B})}{2}.$$

Proof. Fix $I \in \mathcal{I}(C)$ such that

$$\xi_C(\mathcal{A} \bowtie \mathcal{B}) = \tilde{n}^{\frac{1}{4}\big(\log(|I|+1)+|I\cap V_{\text{ends}}(C)|\big)}\pi_I(\mathcal{A} \bowtie \mathcal{B}).$$

We consider various cases depending on $|I \cap V_{\text{ends}}(C)| \in \{0, 1, 2\}$. The most important case is where $|I \cap V_{\text{ends}}(C)| = 2$ (i.e. I contains both endpoints of G_C). Because G_C is connected, this means that $I = V_C$ $(= V_A \cup V_B)$ and hence $\pi_I(\mathcal{A} \bowtie \mathcal{B}) = \delta(\mathcal{A} \bowtie \mathcal{B})$.

Within this case, the most important sub-case is where $|E_A|, |E_B| \geq \frac{1}{2}|E_C|$. In this sub-case, we argue as follows. Without loss of generality, $V_C = \{v_0, \ldots, v_k\}$ (i.e. G_C is the path P_k) and $v_0 \in V_A$ and $v_k \in V_B$. Let $j = \lfloor \frac{k-1}{2} \rfloor$ and $J = \{v_0, \ldots, v_j\}$ and $K = \{v_{k-j}, \ldots, v_k\}$ and note that $J \in \mathcal{I}(A)$ and $K \in \mathcal{I}(B)$. Since $v_0 \in J \cap V_{\text{ends}}(A)$ and $v_k \in K \cap V_{\text{ends}}(B)$, we have

$$\log(k+2) \leq \log(|J|+1) + |J \cap V_{\text{ends}}(A)| \quad \text{and} \quad \log(k+2) \leq \log(|K|+1) + |K \cap V_{\text{ends}}(B)|. \tag{21}$$

Next, observe that $\delta(\mathcal{A} \bowtie \mathcal{B}) \leq \tilde{n}^{-1}$ by C-smallness of $\mathcal{A} \bowtie \mathcal{B}$. We also have the bound $\delta(\mathcal{A} \bowtie \mathcal{B}) \leq \pi_J(\mathcal{A}) \cdot \pi_K(\mathcal{B})$ by Lemmas 4.4 and 4.5 (since $J \cap K = \emptyset$). Taking the geometric mean of these two inequalities, we have

$$\delta(\mathcal{A} \bowtie \mathcal{B}) \leq \tilde{n}^{-1/2}\sqrt{\pi_J(\mathcal{A}) \cdot \pi_K(\mathcal{B})}.$$

Putting these pieces together, we have

$$\begin{aligned}
\xi_C(\mathcal{A} \bowtie \mathcal{B}) &= \tilde{n}^{\frac{1}{4}\big(\log(k+2)+2\big)}\delta(\mathcal{A} \bowtie \mathcal{B}) \\
&\leq \tilde{n}^{\frac{1}{4}\log(k+2)}\sqrt{\pi_J(\mathcal{A}) \cdot \pi_K(\mathcal{B})} \\
&\leq \frac{1}{2}\Big(\tilde{n}^{\frac{1}{4}\log(k+2)}\pi_J(\mathcal{A}) + \tilde{n}^{\frac{1}{4}\log(k+2)}\pi_K(\mathcal{B})\Big) && \text{(AM-GM ineq.)} \\
&\leq \frac{1}{2}\Big(\tilde{n}^{\frac{1}{4}\big(\log(|J|+1)+|J\cap V_{\text{ends}}(A)|\big)}\pi_J(\mathcal{A}) + \tilde{n}^{\frac{1}{4}\big(\log(|K|+1)+|K\cap V_{\text{ends}}(B)|\big)}\pi_K(\mathcal{B})\Big) && \text{(by (21))} \\
&\leq \frac{1}{2}\Big(\xi_A(\mathcal{A}) + \xi_B(\mathcal{B})\Big) && \text{(ind. hyp.)}
\end{aligned}$$

In all other cases (i.e. when $|I \cap V_{\text{ends}}(C)| < 2$ or $\min\{|E_A|, |E_B|\} < |E_C|/2$), the inequality is proved by finding $J \in \mathcal{I}(A)$ or $K \in \mathcal{I}(B)$ such that $|I \cap V_{\text{ends}}(C)| < |J \cap V_{\text{ends}}(A)|$ or $|K \in V_{\text{ends}}(B)|$. We omit the analysis of these cases, since the arguments are not relevant to our main pathset complexity lower bound. $\square$

Having shown that ψ and ξ both satisfying (join rule), the proof of Proposition B.2 is complete. Combining our upper and lower bounds for $\bar{\chi}_{A_k}$ and $\bar{\chi}_{B_k}$ (Corollary A.3 and Propositions A.4 and B.2), we have

Corollary B.5. *With respect to union trees A_k and B_k, the pathset complexity of the complete P_k-pathset $[n]^{V_k}$ has the following bounds:*

$$\begin{aligned}
n^{\frac{1}{2}\log k - O(1)} &\leq \bar{\chi}_{A_k}([n]^{V_k}) \leq k n^{\frac{1}{2}\log k + O(1)}, \\
n^{\frac{1}{4}\log k - O(1)} &\leq \bar{\chi}_{B_k}([n]^{V_k}) \leq 2^k n^{\ln k + O(1)}.
\end{aligned}$$

Since $\Phi_{B_k} = \frac{1}{2}\log k - O(1)$ (as noted in §11.1), Theorem 8.3 gives the stronger lower bound $\bar{\chi}_{B_k}([n]^{V_k}) \geq \tilde{n}^{(1/2)\log k - O(1)} = n^{(1/2)\log k - O(1)}$. Even after extensively studying this special case, we were unable to narrow the gap between $\frac{1}{2}\log k$ and $\ln k$ ($\approx 0.69 \log k$) in the exponent of n in $\bar{\chi}_{B_k}([n]^{V_k})$.

C Rectangular Pathsets

A set $X \subseteq [n]^V$ is *rectangular* if there exist sets $S_i \subseteq [n]$, $i \in V$, such that $X = \{x \in [n]^V : x_i \in S_i$ for all $i \in V\}$. For a pattern graph G, let $\mathscr{R}_G = \{\mathcal{A} \in \mathscr{P}_G : \mathcal{A}$ is rectangular$\}$ and $\mathscr{R}_G^{\text{small}} = \mathscr{R}_G \cap \mathscr{P}_G^{\text{small}}$. For $\mathcal{A} \in \mathscr{R}_G$, we define *rectangular pathset complexity* $\bar{\chi}_G^{\text{rect}}(\mathcal{A})$ exactly like pathset complexity $\bar{\chi}_G(\mathcal{A})$ (Definition 5.6) except with $\mathscr{R}_G$ and $\mathscr{R}_G^{\text{small}}$ replacing $\mathscr{P}_G$ and $\mathscr{P}_G^{\text{small}}$. Analogously, we define $\bar{\chi}_A^{\text{rect}}(\mathcal{A})$ for union trees A. Note that $\bar{\chi}_A(\mathcal{A}) \le \bar{\chi}_A^{\text{rect}}(\mathcal{A})$ for all $\mathcal{A} \in \mathscr{R}_A$.

Remark C.1. I venture to guess that $\bar{\chi}_A(\mathcal{A}) = \bar{\chi}_A^{\text{rect}}(\mathcal{A})$ for all $\mathcal{A} \in \mathscr{R}_A$, but have no idea how to prove this.

We have remarked that our upper bounds on $\bar{\chi}_{A_k}$ and $\bar{\chi}_{B_k}$ (Corollary A.3 and Proposition A.4) involved only rectangular pathsets. It follows that the same upper bounds apply to $\bar{\chi}_{A_k}^{\text{rect}}$ and $\bar{\chi}_{B_k}^{\text{rect}}$.

As for lower bounds on $\bar{\chi}^{\text{rect}}$, this turns out to be significantly easier than our lower bound for $\bar{\chi}$. Similar to our lower bound for fully connected union trees in Appendix B, we can lower bound $\bar{\chi}_G^{\text{rect}}(\mathcal{A})$ for all $\mathcal{A} \in \mathscr{R}_G$ in terms of the projection densities $\pi_S(\mathcal{A})$ where $S \in \mathcal{I}(G)$ via a function similar to $\xi_A(\mathcal{A})$. A key difference when it comes rectangular pathsets is that $\pi_S = \mu_S$ (projection density = maximum restriction density) and hence smallness of rectangular pathsets is preserved under projections to a union of components (cp. Remark 10.5). This fact turns out to greatly simplify the task of proving a lower bound for $\bar{\chi}^{\text{rect}}$.

The next example shows that projections of non-rectangular pathsets can be tricky. This illustrates the difficulty in generalizing the lower bound for $\bar{\chi}^{\text{rect}}$ to the non-rectangular setting.

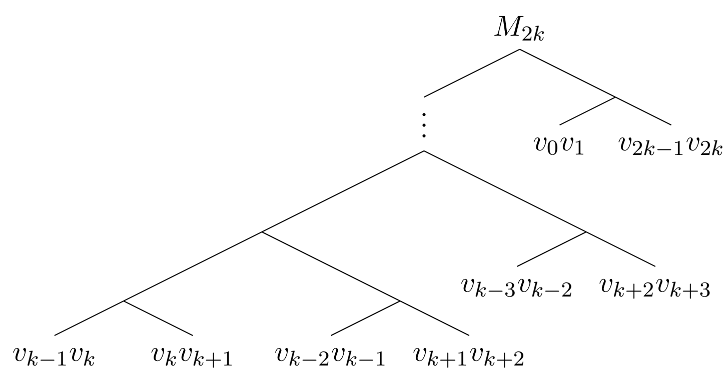
Example C.2. For $k \ge 1$, let $\mathcal{P}al_{2k} \in \mathscr{P}_{P_{2k}}$ be the "palindrome pathset"

$$\mathcal{P}al_{2k} = \big\{x \in [n]^{\{v_0,\dots,v_{2k}\}} : x_{k-i} = x_{k+i} \text{ for all } 0 \le i \le k\big\}.$$

The palindrome pathset $\mathcal{P}al_{2k}$ has low density, while having the maximum projection over vertices $v_0, \dots, v_k$:

$$\delta(\mathcal{P}al_{2k}) = n^{-k} \quad \text{and} \quad \pi_{\{v_0,\dots,v_k\}}(\mathcal{P}al_{2k}) = 1.$$

It turns out that $\mathcal{P}al_{2k}$ is inexpensive to construct, given the right union tree. Let M_{2k} be the union tree



It is easy to show that $\bar{\chi}_{M_{2k}}(\mathcal{P}al_{2k}) \le O(n^2)$. On the other hand, for any fully connected union tree C with graph P_{2k} (such as A_{2k} or B_{2k}), the lower bound of Appendix B implies

$$\bar{\chi}_C(\mathcal{P}al_{2k}) \ge \xi_C(\mathcal{P}al_{2k}) \ge \tilde{n}^{\frac{1}{4}\left(\log(|\{v_0,\dots,v_k\}|+1)+|\{v_0,\dots,v_k\}\cap V_{\text{ends}}(A)|\right)} \cdot \pi_{\{v_0,\dots,v_k\}}(\mathcal{A}) = n^{\Omega(\log k)}.$$

Acknowledgements

This work was carried out while the author was at the National Institute of Informatics in Tokyo, supported by the JST ERATO Kawarabayashi Large Graph Project. I am grateful to Osamu Watanabe and Rahul Santhanam for helpful discussions, Stasys Jukna for stimulating discussions at a time when I was stuck on the proof. Thanks also to Igor Carboni Oliveira for valuable feedback on an early draft of this paper. Finally and not least, I thank the anonymous referees for their close reading and many suggested improvements, including a simplification of Claim 6.5 and optimization of the bound in Lemma 11.4.

References

[1] Miklós Ajtai. First-order definability on finite structures. *Annals of Pure and Applied Logic*, 45(3):211–225, 1989.

[2] Paul Beame, Russell Impagliazzo, and Toniann Pitassi. Improved depth lower bounds for small distance connectivity. *Computational Complexity*, 7(4):325–345, 1998.

[3] Stephen Bellantoni, Toniann Pitassi, and Alasdair Urquhart. Approximation and small-depth frege proofs. *SIAM Journal on Computing*, 21(6):1161–1179, 1992.

[4] Xi Chen, Igor Oliveira, Rocco A. Servedio, and Li-Yang Tan. Near-optimal small-depth lower bounds for small distance connectivity. In *Proc. 48th ACM Symposium on Theory of Computing*, pages 612–625, 2016.

[5] Jeff Edmonds, Chung Keung Poon, and Dimitris Achlioptas. Tight lower bounds for st-connectivity on the NNJAG model. *SIAM Journal on Computing*, 28(6):2257–2284, 1999.

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

[7] Johan Håstad. *Computational limitations of small-depth circuits*. MIT press, 1987.

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

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

[10] Stasys Jukna. *Boolean Function Complexity: Advances and Frontiers*, volume 27. Springer-Verlag Berlin Heidelberg, 2012.

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

[12] Mrinal Kumar and Shubhangi Saraf. On the power of homogeneous depth 4 arithmetic circuits. *SIAM Journal on Computing*, 46(1):336–387, 2017.

[13] Noam Nisan and Avi Wigderson. Lower bounds on arithmetic circuits via partial derivatives. *Computational Complexity*, 6(3):217–234, 1996.

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

[15] Ran Raz and Avi Wigderson. Probabilistic communication complexity of boolean relations. In *Foundations of Computer Science, 1989., 30th Annual Symposium on*, pages 562–567. IEEE, 1989.

[16] Omer Reingold. Undirected connectivity in log-space. *Journal of the ACM (JACM)*, 55(4):17, 2008.

[17] John Riordan and Claude E Shannon. The number of two-terminal series-parallel networks. *J. Math. Phys*, 21(2):83–93, 1942.

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

[19] Benjamin Rossman. Correlation bounds against monotone NC^1. In *Proc. 30th Annual Computational Complexity Conference*, pages 392–411, 2015.

[20] Benjamin Rossman. The average sensitivity of bounded-depth formulas. *Computational Complexity*, 27(2):209–223, 2018.

[21] Benjamin Rossman and Srikanth Srinivasan. Separation of $\mathrm{AC}^0[\oplus]$ formulas and circuits. In *LIPIcs-Leibniz International Proceedings in Informatics*, volume 80. Schloss Dagstuhl-Leibniz-Zentrum fuer Informatik, 2017.

[22] Petr Savickỳ and Alan R Woods. The number of boolean functions computed by formulas of a given size. *Random Structures & Algorithms*, 13(3-4):349–382, 1998.

[23] Walter J Savitch. Relationships between nondeterministic and deterministic tape complexities. *Journal of computer and system sciences*, 4(2):177–192, 1970.

[24] Eli Shamir and Marc Snir. On the depth complexity of formulas. *Mathematical Systems Theory*, 13(1):301–322, 1979.

[25] PM Spira. On time-hardware complexity tradeoffs for boolean functions. In *Proceedings of the 4th Hawaii Symposium on System Sciences*, pages 525–527, 1971.

[26] Prasoon Tiwari and Martin Tompa. A direct version of Shamir and Snir's lower bounds on monotone circuit depth. *Information Processing Letters*, 49(5):243–248, 1994.

[27] Avi Wigderson. The complexity of graph connectivity. In *Proceedings of the 17th International Symposium on Mathematical Foundations of Computer Science*, pages 112–132. Springer-Verlag, 1992.