U.U.D.M. Project Report 2026:9
Degree project 30 credits
Master`s Programme in Mathematics
May 2026

Caterpillar-counting to detect subgraphs

Erik Sundström

Department of Mathematics, Uppsala University

Supervisor: Fiona Skerman
Examiner: Rolf Larsson

Caterpillar-counting to detect subgraphs

Erik Sundström

May 2026

Abstract

This paper was written as a master thesis in mathematics. Detecting graphs with planted dense subgraphs is a problem that consists of distinguishing between Erdős-Rényi random graphs and random graphs with a high-density subgraph inserted on a randomly chosen vertex subset. Bhashkara et al. [1] introduced a class of trees named r/s-caterpillars to solve this by a means of subgraph counting. We use these graphs to take on a different approach of the proof, utilizing the concept of signed subgraph counting among others. We also closer investigate these r/s-caterpillars to formally conclude some properties about them.

1 Introduction

The so-called *planted clique problem* consists of distinguishing random graphs from graphs with a planted clique, i.e., a complete induced subgraph. It is an algorithmic problem lying within the areas of computational complexity theory and random graph theory, and a solution requires the proportion of correct guesses to tend to 1 as the graph size increases. We think of a graph as a tuple consisting of a finite set of vertices and a set of edges, where each edge represents a relationship between two distinct vertices.

This paper examines a related problem where the aim again is to distinguish between two different graph distributions. This time, however, the planted clique is replaced by a planted dense subgraph. This way, the problem is generalized in a way where we have a parameter that adjusts how prominent the planted subgraph is. Naturally, there will be regimes where the problem is unsolvable, for example if the density of the planted subgraph lies too close to the overall graph density.

We constrain ourselves to working within the framework of so-called low-degree polynomials, defined in Section 1.2. Briefly explained, these algorithms are polynomials which degrees are small relative to the number of vertices, allowing them to be more computationally manageable by computers. In this sense, the problem of solving for detection is not merely a theoretical one, we are also interested in a solution that can be used in practice.

Approaching this problem can be done in many different ways. One way is to count the number of a certain subgraph and use that number decide the guess. This paper aims to utilize a subgraph-counting technique to solve the detection problem. Namely, the subgraphs in use will be a certain type of tree structure which is called r/s-caterpillars and is introduced in [1]. However, we examine a different approach of proving detection, one which constitutes the central result in this paper.

In the subgraph-counting proof of [1], a flaw was discovered where a version of the Chernoff bound was used despite not having independent random variables (which the bound requires). In particular, in the proof of Claim 3.2 in [1], the Chernoff bound is applied to the cardinality $|S(t-1)|$, even though $|S(t-1)|$ is a sum of random variables that are dependent. This led to our interest in reproving the detection result, based on the same caterpillar-counting method used in [1]. This stands as one of the main contributions of this paper, along with bounds for when the caterpillar-counting method is successful.

One aspect that sets our approach apart from the proof in [1] is that we make a so-called signed subgraph count. This technique was introduced in [2] and is useful to limit the variance of the count.

Studying the subclass of caterpillars introduced in [1] turned out to be an interesting area in itself, and a number of results regarding them are formally established and proved in Section 1.1.1.

1.1 r/s-caterpillars

In general, a *caterpillar graph*, or a *caterpillar*, is a tree where every vertex either is part of a central backbone (consisting of a path graph) or directly connected to a vertex in the backbone by a single edge. A vertex in a caterpillar can thus be categorized as either a backbone vertex or a leaf vertex. A specific class of caterpillars are defined below, following [1], which is dependent on two parameters r and s.

Definition 1.1. For relatively prime integers $0 < r < s$, an *r/s-caterpillar* is a tree which is constructed inductively in the following way:

Start with a single vertex which is assigned to be the left-most backbone vertex in the caterpillar. The construction will consist of s steps. At step $i \in \{1, \ldots, s\}$, do the following:

If the interval $[(i-1)r/s, ir/s]$ contains an integer, add a leaf to the rightmost backbone vertex in the caterpillar. Otherwise, add a new rightmost backbone vertex.

To give a better sense for these caterpillars, we proceed by presenting a detailed construction of the 2/5-caterpillar as an example, together with a labelled graph in Figure 1 to show in which order the vertices are added in the algorithm. To more easily distinguish between backbone and leaf vertices we colour them black and red, respectively.

Initially, we have only one backbone vertex (labelled 0 in Figure 1) and for the first step we consider the interval $[0, 2/5]$. This contains the integer 0, hence we add a leaf to the lone backbone vertex.

The second interval to consider is $[2/5, 4/5]$. As this does not contain any integer, we now add a new backbone vertex to obtain the graph in Figure 1.

The process goes on by noting that the integer 1 is contained in the interval $[4/5, 6/5]$ and we thus connect leaf to the right-most (newest) backbone vertex.

The integer-free interval $[6/5, 8/5]$ leads us to add a new backbone vertex.

Finally, as the interval $[8/5, 10/5]$ does contain the integer 2, a final leaf is added.

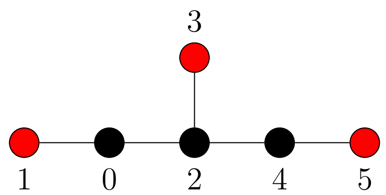


Figure 1: A labelled 2/5-caterpillar with black backbone vertices and red leaves

For more concrete examples of caterpillars, a complete list of caterpillars with $s \leq 12$ is found in Section 4. A useful aspect of this construction is that we have two parameters r and s that we can tweak depending on the situation we are in. This will turn out to be very useful when proving detection results later on.

1.1.1 General caterpillar results

This section is dedicated to cover some results about r/s-caterpillars. The most straight-forward one follows below, a collection of facts that is used in [1] without explicit proofs.

Lemma 1.2. *r/s-caterpillars have a total of s edges, $s - r$ backbone vertices and $r + 1$ leaves.*

Proof. The construction of an r/s-caterpillar consists of s steps, and exactly one edge is added in each step. The graph will thus contain s number of edges.

The number of leaves in an r/s-caterpillar corresponds to the count of integers in the intervals $[(i-1)r/s, i/s]$, $i = 1, \ldots, s$. As r and s are set to be coprime, no such given interval will contain an integer in either endpoint except for when i is equal to 1 or s. Therefore the total number of leaves will be equal to the total number of integers from 0 to r, which is $r+1$.

As the caterpillar is a tree with s edges, the total number of vertices in the graph is $s+1$. We can thereby conclude that the number of backbone vertices is the total number of vertices minus the number of leaves, namely $(s+1)-(r+1)=s-r$. □

r/s-caterpillars are symmetric, this fact is taken for granted in [1] without proof. The result is carefully stated and proven in Lemma 1.3 below. By leaf-degree we mean the number of adjacent leaves.

Lemma 1.3. *r/s-caterpillars are symmetric, i.e., the i-th backbone vertex and the $s-r+1-i$-th backbone vertex have the same leaf-degree for all $i=1,\dots,s-r$.*

Proof. Consider the interval $[(t-1)r/s, tr/s]$ for some arbitrary $t \in \{1,\dots,s\}$. We need to show that this interval contains an integer if and only if the interval $[(s-t+1-1)r/s,(s-t+1)r/s]$ does.

The latter can be rewritten as

$$[(s-t)r/s,(s-t+1)r/s] = [r-tr/s, r-(t-1)r/s].$$

Shifting this interval by r steps does not change the number of integers in it. This leaves us with the interval

$$[-tr/s, -(t-1)r/s],$$

which contains the same numbers as the first interval, only with a negation sign in front. It therefore also contains the same number of integers as the first one. □

Lemma 1.4 and Corollary 1.5 are new contributions to the theory of r/s-caterpillars.

Lemma 1.4. *The backbone vertices in an r/s-caterpillar has a leaf-degree that is bounded by the degree of the leftmost backbone vertex, which is equal to $\lfloor s/(s-r)\rfloor$.*

Proof. Let d be the leaf-degree of the leftmost backbone vertex in the caterpillar. That implies that both of the intervals $[0, dr/s]$ and $[0,(d+1)r/s]$ contains exactly d integers each. Suppose for a contradiction that there is some backbone vertex which leaf-degree is higher than d. Then there is some integer a such that the interval $[ar/s,(a+d+1)r/s]$ contains $d+1$ integers.

We consider now the same interval shifted by $ar/s-\lfloor ar/s\rfloor$, namely $[\lfloor ar/s\rfloor, \lfloor ar/s\rfloor + (d+1)r/s]$. Making this shift cannot have decreased the number of integers since the integer $\lfloor ar/s\rfloor$ was added at one end and a maximum of one integer at the other end can possibly have been removed since $ar/s - \lfloor ar/s \rfloor < 1$. Now, shifting an interval by some integer number of steps does not change the number of integers in the interval. Hence the interval $[0, (d+1)r/s]$ must also contain at least $d+1$ integers. This is a contradiction, however, since we asserted that this interval contained exactly d integers.

To get an expression for the leaf-degree of the left-most backbone vertex, we can think of the caterpillar-construction process in the following way: We start from 0 and add r/s at each step, leaving a distance $1 - r/s$ to the next integer. We add a new backbone vertex when these distances add up to 1 since we at that point will not reach the next integer. This will happen after $1/(1-r/s) = s/(s-r)$ steps. Note that as r and s are coprime, the only way $s/(s-r)$ can be an integer is when r and s have a difference of 1. Thus, the leaf-degree will be the biggest integer smaller than or equal to $s/(s-r)$, namely $\lfloor s/(s-r) \rfloor$. □

Corollary 1.5 is a direct consequence of Lemma 1.4.

Corollary 1.5. *$s/(s-r)$ is an upper bound for the leaf-degree of any backbone vertex in an r/s-caterpillar.*

1.2 Low-degree polynomial testing

In this paper we will be restricting ourselves to the so-called low-degree polynomial models of computation. What these models have in common is that the output is computable via a polynomial whose degree is "low". In particular, the degree should not exceed the logarithm of the dimension, which in our case is the number of vertices in the graph [7].

In this setting, our input is the adjacency matrix $A = A(G)$ of the given graph G. A is a square matrix with $A_{ij} = 1$ if there is an edge between vertices i and j, otherwise $A_{ij} = 0$. Given a graph G with $V(G) = [n] = \{1, \ldots, n\}$, we could for example want to count the number of subgraphs of a certain type, say the triangle K_3. We are then able to take the sum over all triplets of vertices,

$$f(A) = \sum_{i<j<k} A_{ij}A_{jk}A_{ik}.$$

The above sum equals the total number of triangles in G. As we can see, $f(A)$ is a polynomial of degree 3 and thus a "valid" low degree polynomial for large values of n. In general, a polynomial that counts the number of copies of some subgraph with d edges has degree d. These polynomials should be viewed as algorithms, where the degree of the polynomial is a way of measuring the complexity of the algorithm.

The example above illustrates that subgraph-counting algorithms fall into the category of low-degree polynomials, and these specific algorithms are what we are interested in. However, instead of counting the number of triangles, we will be counting appearances of r/s-caterpillars. These were introduced in full detail in subsection 1.1.

1.2.1 Detection success

To be able to tell when a certain detection method works, we need a formal definition of what "success" means. The definition is gathered from a survey by Alex Wein [8]. In section 3.3 in [8], a distinction is made between strong and weak detection. This paper will only cover the strong version, hence any mention of detection refers to strong detection. Before being able to present the definition, we require some setup, however.

In general, we consider two sequences of distributions, $\mathcal{P} = (\mathcal{P}_n)_{n\geq 1}$ and $\mathcal{Q} = (\mathcal{Q}_n)_{n\geq 1}$, where $\mathcal{P}_n$ and $\mathcal{Q}_n$ are distributions over some subset of $\mathbb{R}^N$ for some $N = N_n$. We will consider the case where $n \to \infty$.

Definition 1.6. A *test* is a function $t : \mathbb{R}^N \to \{0, 1\}$ that takes a sample from $\mathcal{P}$ or $\mathcal{Q}$ as input and outputs a "guess": 1 for $\mathcal{P}$ or 0 for $\mathcal{Q}$.

Strong detection of a test t is said to be achieved when the probability of incorrect guesses tends to zero, i.e.,

$$\mathbb{P}(t(Y) = 0|Y \sim \mathcal{P}) + \mathbb{P}(t(Y) = 1|Y \sim \mathcal{Q}) = o(1) \qquad \text{as } n \to \infty.$$

To be fully formal, one should specify that a test is a sequence of functions, one for each value of n. Likewise, when speaking of polynomials $f : \mathbb{R}^N \to \mathbb{R}$, we really mean a sequence f_n of polynomials. This is usually omitted for simplicity, however.

We can now define the notion of success for a polynomial. It is phrased in terms of the polynomial strongly or weakly separating the two given distributions, where the former implies the latter.

Notation. For a given distribution $\mathcal{P}$, we denote by $\mathbb{E}_\mathcal{P}$ and $\mathrm{Var}_\mathcal{P}$ the expectation and variance under $\mathcal{P}$, respectively.

Definition 1.7. A polynomial $f : \mathbb{R}^N \to \mathbb{R}$ is said to

- *strongly separate* distributions $\mathcal{P}$ and $\mathcal{Q}$ if

$$\sqrt{\max\left\{\mathrm{Var}_\mathcal{P}(f(Y)), \mathrm{Var}_\mathcal{Q}(f(Y))\right\}} = o\left(|\mathbb{E}_\mathcal{P}[f(Y)] - \mathbb{E}_\mathcal{Q}[f(Y)]|\right),$$

- *weakly separate* distributions $\mathcal{P}$ and $\mathcal{Q}$ if

$$\sqrt{\max\left\{\mathrm{Var}_\mathcal{P}(f(Y)), \mathrm{Var}_\mathcal{Q}(f(Y))\right\}} = O\left(|\mathbb{E}_\mathcal{P}[f(Y)] - \mathbb{E}_\mathcal{Q}[f(Y)]|\right),$$

as $n \to \infty$.

We will not go into detail why these conditions are natural to determine if a polynomial is successful or not. We may note, however, that strong separation implies that there is some threshold function of $f(Y)$ that separates $\mathcal{P}$ and $\mathcal{Q}$ with probability approaching 1. This is a consequence of Chebyshev's inequality.

A useful way of proving strong separation is through the next theorem.

Theorem 1.8. *Suppose that $\mathcal{P}$ and $\mathcal{Q}$ are distributions and $f : \mathbb{R}^N \to \mathbb{R}$ is some polynomial that satisfies one of the following equalities.*

$$\lim_{n\to\infty} \frac{\mathbb{E}_{\mathcal{P}}[f] - \mathbb{E}_{\mathcal{Q}}[f]}{\sqrt{Var_{\mathcal{P}}(f)}} = \infty$$

$$\lim_{n\to\infty} \frac{\mathbb{E}_{\mathcal{P}}[f] - \mathbb{E}_{\mathcal{Q}}[f]}{\sqrt{Var_{\mathcal{Q}}(f)}} = \infty$$

$$\lim_{n\to\infty} \frac{\mathbb{E}_{\mathcal{Q}}[f] - \mathbb{E}_{\mathcal{P}}[f]}{\sqrt{Var_{\mathcal{P}}(f)}} = \infty$$

$$\lim_{n\to\infty} \frac{\mathbb{E}_{\mathcal{Q}}[f] - \mathbb{E}_{\mathcal{P}}[f]}{\sqrt{Var_{\mathcal{Q}}(f)}} = \infty$$

Then f strongly separates $\mathcal{P}$ and $\mathcal{Q}$.

Proof. It holds that any of the four limits above can be bounded from above by the same limit, namely

$$\lim_{n\to\infty} \frac{|\mathbb{E}_{\mathcal{P}}[f] - \mathbb{E}_{\mathcal{Q}}[f]|}{\sqrt{\max\{\mathrm{Var}_{\mathcal{P}}(f), \mathrm{Var}_{\mathcal{Q}}(f)\}}}.$$

Assuming one of the four limits in the theorem tends to ∞, the above limit then does as well. By definition, this is equivalent to

$$\sqrt{\max\{\mathrm{Var}_{\mathcal{P}}(f(Y)), \mathrm{Var}_{\mathcal{Q}}(f(Y))\}} = o\left(|\mathbb{E}_{\mathcal{P}}[f(Y)] - \mathbb{E}_{\mathcal{Q}}[f(Y)]|\right),$$

and hence we are done. □

1.3 Setup of the model

There are different variants of the random distinguishing problem. We will present a standard model, called The Random Planted model, gathered from section 3 in [1].

We aim in this problem to distinguish between two random distributions, $\mathcal{D}_0$ and $\mathcal{D}_1$. These are introduced in definitions 1.9 and 1.10, respectively.

Notation. By $G(n,p)$ we denote here the so-called Erdős-Rényi random graph model on n vertices, where each edge is present with probability p. If G is a sample graph from the model $G(n,p)$ one may write $G \sim G(n,p)$.

Definition 1.9. In the null model $\mathcal{D}_0 = \mathcal{D}_0(\beta)$, we set $G \sim G(n,p)$ with $p = n^{-\beta}$, $0 < \beta < 1$.

Definition 1.10. In the planted model $\mathcal{D}_1 = \mathcal{D}_1(\alpha,\beta,\gamma)$, $G \sim G(n,n^{-\beta})$ as before. A set S is drawn from $V(G)$ by including each vertex independently with probability k/n. The subgraph on S is then replaced with a random graph $H \sim G(|V(S)|, n^{-\alpha})$. The parameter k is related to n by $k = n^{\gamma}$.

We will consider the case where $\alpha < \beta$, since that implies that the planted subgraph is denser than the rest of the graph. A sketch of the $\mathcal{D}_1$ distribution is provided in Figure 2.

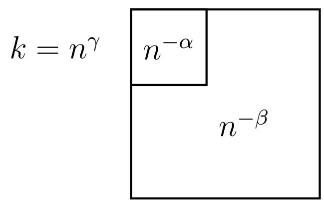


Figure 2: Sketch of the planted model

1.4 Earlier work

Wein et al. [3] covers the question of subgraph-detecting in detail, with focus on low-degree polynomial testing. There, they establish sharp necessary and sufficient conditions for success of low-degree polynomial tests.

Their main result is encapsulated in Theorem 1.11, giving concrete conditions on α, β and γ for when there are successful, computationally efficient tests.

Theorem 1.11 (cf. [3, Thm 2.1])**.** *Suppose that we observe a random graph drawn from either $\mathcal{D}_0$ or $\mathcal{D}_1$ (defined as in Definitions 1.9 and 1.10).*

- *Suppose that either*

  (1) $\gamma \geq 1/2$ and $\alpha > \beta/2 + 2\gamma - 1$, or

  (2) $\gamma < 1/2$ and $\alpha > \beta\gamma$.

  Then there is no low-degree polynomial that weakly separates $\mathcal{D}_0$ and $\mathcal{D}_1$.

- *Suppose that either*

  (1) $\gamma \geq 1/2$ *and* $\alpha < \beta/2 + 2\gamma - 1$*, or*

  (2) $\gamma < 1/2$ *and* $\alpha < \beta\gamma$*.*

  Then there exists some low-degree polynomial that strongly separates $\mathcal{D}_0$ *and* $\mathcal{D}_1$*.*

Wein et al. [3] takes a more general approach and considers the hypergraph version of the detection problem. As we are merely interested in the graph case, we narrowed down the previous theorem to regular graphs to more precisely fit our purposes.

2 Preliminary lemmas

This section will cover some results that will prove to be useful in the upcoming parts of the paper. Firstly, the first and second moment methods are stated.

Lemma 2.1 (The first moment method, cf. [5, eq 3.1])**.** *If* X *is a non-negative, integer valued random variable, then the inequality*

$$\mathbb{P}(X > 0) \leq \mathbb{E}[X]$$

holds.

Proving Lemma 2.1 consists of a direct application of Markov's inequality. The lemma is often used to show that some sequence of random variables $\{X_n\}_n$ converges to 0 if the expected values of X_n converges to 0.

Lemma 2.2 (The second moment method, cf. [5, eq 3.2])**.** *If* X *is a non-negative random variable (that is not identically zero), then the inequality*

$$\mathbb{P}(X > 0) \geq \frac{(\mathbb{E}[X])^2}{\mathbb{E}[X^2]}$$

holds.

The proof of this result is quite straight-forward with the use of Chebyshev's inequality. If one can show that the right-hand side of the inequality tends to zero for some random variables X_n, one can conclude that $X_n > 0$ a.a.s. The following two well-known binomial bounds will be useful, for which we provide a proof for completeness.

Lemma 2.3. *For $1 \leq k \leq n$, it holds that*

$$\left(\frac{n}{k}\right)^k \leq \binom{n}{k} \leq \left(\frac{en}{k}\right)^k. \tag{1}$$

Proof. To show the first inequality, we can rewrite $\binom{n}{k}$ to get

$$\begin{aligned}\binom{n}{k} &= \frac{n!}{k!(n-k)!}\\ &= \frac{n}{k} \cdot \frac{n-1}{k-1} \cdots \frac{n-(k-1)}{1}\\ &\geq \frac{n}{k} \cdot \frac{n}{k} \cdots \frac{n}{k}\\ &= \left(\frac{n}{k}\right)^k.\end{aligned}$$

The inequality holds simply due to the fact that each factor of the product in the second row is bigger than n/k.

For the second inequality, we make use of the Taylor expansion of the function e^k,

$$e^k = 1 + k + \frac{k^2}{2!} + \frac{k^3}{3!} + \cdots.$$

For $k > 0$, all terms are positive and it hence holds that $e^k \geq k^k/k!$. Utilizing this inequality, the binomial coefficient $\binom{n}{k}$ can be rewritten in the following way:

$$\begin{aligned}\binom{n}{k} &= \frac{n!}{k!(n-k)!}\\ &= \frac{1}{k^k} \cdot \frac{k^k}{k!} \cdot n(n-1)\cdots(n-(k-1))\\ &= \left(\frac{n}{k}\right)^k \cdot \frac{k^k}{k!} \cdot (1-1/n)\cdots(1-(k-1)/n)\\ &\leq \left(\frac{n}{k}\right)^k \cdot e^k\\ &= \left(\frac{en}{k}\right)^k,\end{aligned}$$

thus completing the proof. □

What follows now is a variant of the so-called Chernoff bounds which can be found in e.g. [6].

Lemma 2.4 (cf. [6, Lemma 4.1]). *Let $b > 0$, and let random variables $X_1, X_2, \dots, X_k$ be independent, with $0 \le X_j \le b$ for each j. Let $S = \sum_{j=1}^{k} X_j$ and $\mu = \mathbb{E}[S]$. Then*

$$\mathbb{P}(|S-\mu| \ge \epsilon\mu) \le 2e^{-\frac{1}{3}\epsilon^2\mu/b} \text{ for } 0 < \epsilon \le 1;$$

or equivalently

$$\mathbb{P}(|S-\mu| \ge x\sqrt{\mu}) \le 2e^{-\frac{1}{3}x^2/b} \text{ for } 0 < x \le \sqrt{\mu}.$$

A special case of Lemma 2.4 is acquired if we set $b = 1$, a version that will be utilized later on.

2.1 Stochastic dominance

We here introduce the notion of stochastic dominance. The information in this subsection is gathered from [4]. As a reminder for the reader, the cumulative distribution function F_X of a random variable X is defined as $F_X(x) = \mathbb{P}(X \le x)$.

Definition 2.5. Let X and Y be two real-valued random variables with common sample space Ω and with cumulative distribution functions F_X and F_Y, respectively. We say that X has *first-order stochastic dominance over* Y if $F_X(x) \le F_Y(x)$ for all $x \in \Omega$.

A basic example would be the following. Imagine that there are two coins, A and B. Coin A is marked with 2 on one side and 5 on the other side. Coin B is marked with 1 on one side and 4 on the other side. Let us now say that two people are flipping one coin each and noting the resulting numbers each time, we may call the random variables C_A and C_B. In some cases, coin B will yield a higher result than A, but C_A still has first-order stochastic dominance over C_B as coin A is at least as likely as B to exceed any given value.

In a more general sense, first-order stochastic dominance of X over Y means that X is at least as likely as Y to reach or exceed any value x for all $x \in \mathbb{R}$. As not all random variables may be compared in this sense, stochastic dominance yields a partial order between random variables. One may also consider stochastic dominance of higher orders, but whenever stochastic dominance is mentioned in this paper we refer to first-order stochastic dominance.

Theorem 2.6 serves as an example of how this notion can be used in practice.

Theorem 2.6. *Let $(X_n)_{n\in\mathbb{N}}$ and $(Y_n)_{n\in\mathbb{N}}$ be two sets of non-negative random variables and let X_n have first-order stochastic dominance over Y_n for each $n \in \mathbb{N}$. If X_n converges to 0 in distribution as $n \to \infty$, then Y_n converges to 0 in distribution as $n \to \infty$ as well.*

Proof. As X_n converges in distribution to 0, we have that $\lim_{n\to\infty} F_{X_n}(x) = 1$ for all $x \ge 0$. Thus, $\lim_{n\to\infty} F_{Y_n}(x) \ge 1$ for all $x \ge 0$, implying that Y_n indeed converges to 0 in distribution. $\square$

3 Proof for detection

This section is dedicated to showing that the caterpillar count method can be used for detection in the random planted model.

For the setup of the proof, some definitions are first needed.

Definition 3.1. For a given a r/s-caterpillar graph H, we define $L = L(H)$ as the set of leaf vertices of H.

L is partitioned in a way which corresponds to the number of backbone vertices in H. We set $L = L_1 \sqcup L_2 \sqcup \ldots \sqcup L_{s-r}$, where each L_i, $i = 1, \ldots, s-r$, contains all leaf vertices in H that are adjacent to the i-th backbone vertex.

Definition 3.2. For a given graph G and a caterpillar graph H with leaf vertex set L, we say that a ϕ is an *L-embedding of H in G* if

$$\phi : L \to V(G)$$

is an injective map.

Definition 3.3. Given a graph G, an r/s-caterpillar H with leaf vertex set L and an L-embedding ϕ of H in G, the sets $\mathcal{U}_1, \mathcal{U}_2, \ldots, \mathcal{U}_{s-r}$ are defined as

$$\mathcal{U}_i = \mathcal{U}_i(G, H, \phi) = \{u \in V(G) \mid u \sim v \quad \forall v \in \phi(L_i)\}$$

for all $i = 1, \ldots, s-r$.

One may think of the set $\mathcal{U}_i$ in Definition 3.3 as the candidates for the i-th backbone vertex in a leaf-rooted caterpillar in G. We proceed with a lemma that covers a partial result in the upcoming proof.

Lemma 3.4. *Let $G \sim G(n,p)$, H be an r/s-caterpillar with leaf set L and ϕ an L-embedding of H in G. Then for each $i = 1, \ldots, s-r$,*

$$|\mathcal{U}_i| \sim Bin(n - |L_i|, p^{|L_i|}).$$

Moreover, each random variable $|\mathcal{U}_i|$ is stochastically dominated by X_i, where

$$X_i \sim Bin(n, p^{|L_i|})$$

for all $i = 1, \ldots, s-r$.

Proof. For a fixed vertex $u \in V(G) \setminus \phi(L_i)$ for some given $1 \leq i \leq s-r$, the probability that $u \in \mathcal{U}_i$ equals the probability of all edges uv, $v \in \phi(L_i)$, being present. As each edge exists independently with probability p, we have that $\mathbb{P}(u \in \mathcal{U}_i) = p^{|L_i|}$. Now, since there are $n - |L_i|$ vertices to choose from outside of $\phi(L_i)$,

$|\mathcal{U}_i|$ will be a sum of i.i.d. random variables which are all Bernoulli distributed with parameter $p^{|L_i|}$. Thus, $|\mathcal{U}_i|$ will be binomially distributed as

$$|\mathcal{U}_i| \sim \mathrm{Bin}(n - |L_i|, p^{L_i})$$

The second part follows almost immediately, as X_i can be seen as a sum of strictly more equally distributed random variables than $|\mathcal{U}_i|$ and hence will always have a greater probability of exceeding any given value. □

What follows is the first part of the general statement that we want to prove.

Theorem 3.5. *Fix $\epsilon > 0$, relatively prime integers $0 < r < s$ and $L' \subseteq [n]$. Furthermore, let $G \sim G(n,p)$ for $p \leq n^{r/s-1-\epsilon}$, H be an r/s-caterpillar with leaf set L and $\phi : L \to L'$ be an L-embedding of H in G. Then, the number of leaf-rooted r/s-caterpillars on $\phi(L)$ in G will tend to 0 as $n \to \infty$.*

Proof. We start by defining the random variable Y by

$$\begin{aligned} Y = {} & \text{Number of tuples } (u_1, \ldots, u_{s-r}) \text{ of distinct vertices in } V(G) \setminus \phi(L) \\ & \text{such that } u_1 \in \mathcal{U}_1, \ldots, u_{s-r} \in \mathcal{U}_{s-r},\ u_k \sim u_{k+1} \text{ for all } k = 1, \ldots, s-r-1. \end{aligned}$$

Y thus counts the number of leaf-rooted r/s-caterpillars on $\phi(L)$. We define the random variable W by

$$\begin{aligned} W = {} & \text{Number of tuples } (u_1, \ldots, u_{s-r}) \text{ of (not necessarily distinct) vertices in } V(G) \\ & \text{such that } u_1 \in \mathcal{U}_1, \ldots, u_{s-r} \in \mathcal{U}_{s-r},\ u_k \sim u_{k+1} \text{ for all } k = 1, \ldots, s-r-1. \end{aligned}$$

Note that we allow vertices from $\phi(L)$ in W but not in Y. The random variable W will stochastically dominate Y, and we can express W as

$$W = \sum_{u_1 \in \mathcal{U}_1, \ldots, u_{s-r} \in \mathcal{U}_{s-r}} \prod_{k=1}^{s-r-1} \mathbf{1}[u_k \sim u_{k+1}].$$

Set the random variables X_i as in Lemma 3.4, namely

$$X_i \sim \mathrm{Bin}(n, p^{|L_i|}),$$

for all $i = 1, \ldots, s-r$. We now consider $\mathcal{E}$, the event that $X_i > 2np^{|L_i|}$ for some $i = 1, \ldots, s-r$. Using Lemma 2.4, the probability of $\mathcal{E}^{\mathsf{c}}$ is bounded from below in the following way:

$$\mathbb{P}(\mathcal{E}^{\mathsf{c}}) \geq \prod_{i=1}^{s-r} \left(1 - \exp\left(-\frac{1}{3} np^{|L_i|}\right)\right).$$

The above product can be rewritten as an inclusion-exclusion type of expansion, which yields

$$\begin{aligned}\mathbb{P}(\mathcal{E}^{\mathsf{c}}) &\geq \sum_{k=0}^{s-r}(-1)^k \sum_{1\leq i_1\leq\ldots\leq i_k\leq s-r} \exp\left(-\frac{1}{3}n\left(p^{|L_{i_1}|}+\ldots+p^{|L_{i_k}|}\right)\right)\\ &= 1+\sum_{k=1}^{s-r}(-1)^k \sum_{1\leq i_1\leq\ldots\leq i_k\leq s-r} \exp\left(-\frac{1}{3}n\left(p^{|L_{i_1}|}+\ldots+p^{|L_{i_k}|}\right)\right)\end{aligned}$$

We see that each one of the terms in the last sum approaches 0 as n tends to infinity. Hence, this probability will tend to 1, i.e.,

$$\mathbb{P}(\mathcal{E}^{\mathsf{c}}) = 1 + o(1). \tag{2}$$

We proceed by introducing the random variable W', defined by

$$W' = \sum_{u_1\in\mathcal{U}'_1,\ldots,u_{s-r}\in\mathcal{U}'_{s-r}} \prod_{k=1}^{s-r-1} \mathbf{1}[u_k \sim u_{k+1}].$$

Here, $|\mathcal{U}'_i| = 2np^{|L_i|}$ for each $i = 1,\ldots,s-r$. Now, provided that $\mathcal{E}^{\mathsf{c}}$ holds, W will be stochastically dominated by W'. To see this, we first note that the statement that X_i is stochastically dominating $|\mathcal{U}_i|$ for all $i = 1,\ldots,s-r$ is equivalent to

$$\mathbb{P}(X_i \leq x) \leq \mathbb{P}(|\mathcal{U}_i| \leq x)$$

for all x. Assuming that $\mathcal{E}^{\mathsf{c}}$ holds, we have for all $i = 1,\ldots,s-r$

$$1 = \mathbb{P}(X_i \leq 2np^{|L_i|}) \leq \mathbb{P}(|\mathcal{U}_i| \leq 2np^{|L_i|}).$$

The above implies that

$$|\mathcal{U}_i| \leq 2np^{|L_i|}$$

for all $i = 1,\ldots,s-r$. Using this fact, we obtain for each w

$$
\begin{aligned}
\mathbb{P}(W \leq w \mid \mathcal{E}^{\mathsf{c}}) &= \mathbb{P}\left(\sum_{u_1 \in \mathcal{U}_1, \dots, u_{s-r} \in \mathcal{U}_{s-r}} \prod_{k=1}^{s-r-1} \mathbf{1}[u_k \sim u_{k+1}] \leq w \mid X_i \leq 2np^{|L_i|} \ \forall i = 1, \dots, s-r \right) \\
&\geq \mathbb{P}\left(\sum_{u_1 \in \mathcal{U}_1, \dots, u_{s-r} \in \mathcal{U}_{s-r}} \prod_{k=1}^{s-r-1} \mathbf{1}[u_k \sim u_{k+1}] \leq w \mid |\mathcal{U}_i| \leq 2np^{|L_i|} \ \forall i = 1, \dots, s-r \right) \\
&\geq \mathbb{P}\left(\sum_{u_1 \in \mathcal{U}_1, \dots, u_{s-r} \in \mathcal{U}_{s-r}} \prod_{k=1}^{s-r-1} \mathbf{1}[u_k \sim u_{k+1}] \leq w \mid |\mathcal{U}_i| = 2np^{|L_i|} \ \forall i = 1, \dots, s-r \right) \\
&= \mathbb{P}(W' \leq w)\,.
\end{aligned}
$$

This concludes that W' does indeed stochastically dominate W provided that $\mathcal{E}^{\mathsf{c}}$ holds. Expressing the inequality in terms of the complements yields

$$
\mathbb{P}(W > w \mid \mathcal{E}^{\mathsf{c}}) \leq \mathbb{P}(W' > w)\,. \tag{3}
$$

We now take the expected value of W':

$$
\begin{aligned}
\mathbb{E}[W'] &= \prod_{i=1}^{s-r} 2np^{|L_i|} p^{s-r-1} \\
&= 2^{s-r} n^{s-r} p^{r+1} p^{s-r-1} \\
&= 2^{s-r} n^{s-r} p^{s} \\
&\leq 2^{s-r} n^{s-r} (n^{r/s-1-\epsilon})^{s} \\
&= 2^{s-r} n^{s-r} n^{r-s-s\epsilon} \\
&= 2^{s-r} n^{-s\epsilon} \\
&= o(1)
\end{aligned} \tag{4}
$$

We here used the fact that the number of leaf vertices in $\phi(L)$ is $r+1$. Now, we can relate $\mathbb{P}(W > 0)$ and $\mathbb{P}(W > 0 \mid \mathcal{E}^{\mathsf{c}})$ by

$$
\begin{aligned}
\mathbb{P}(W > 0) &= \mathbb{P}(W > 0 \mid \mathcal{E}^{\mathsf{c}}) \cdot \mathbb{P}(\mathcal{E}^{\mathsf{c}}) + \mathbb{P}(W > 0 \mid \mathcal{E}) \cdot \mathbb{P}(\mathcal{E}) \\
&\leq \mathbb{P}(W > 0 \mid \mathcal{E}^{\mathsf{c}}) \cdot (1 + o(1)) \\
&\leq \mathbb{P}(W > 0 \mid \mathcal{E}^{\mathsf{c}})\,,
\end{aligned} \tag{5}
$$

where we made use of equation 2 in the second step.
All the components needed to bound $\mathbb{P}(Y > 0)$ are now in place.

$$
\begin{aligned}
\mathbb{P}(Y > 0) &\leq \mathbb{P}(W > 0) \\
&\leq \mathbb{P}(W > 0 \mid \mathcal{E}^{\mathrm{c}}) \\
&\leq \mathbb{P}(W' > 0) \\
&\leq \mathbb{E}[W'] \\
&\leq o(1).
\end{aligned}
$$

The above inequalities relies in order on the fact that Y stochastically dominates W, equations (5) and (3), Theorem 2.1 as well as equation (4).

We have now shown that Y is 0 with high probability. Since L' was chosen arbitrarily, we can thus conclude that the number of rooted caterpillars on any given set of leaf vertices will tend to 0. □

We now move on to the second part of the proof for detection, covering the planted dense subgraph model this time. Figure 3 serves as a reminder of the model in use. The edge probability in the dense subgraph is $n^{-\alpha}$ while the edge probability elsewhere is $n^{-\beta}$. Moreover, the expected size of the dense subgraph is n^{γ}.

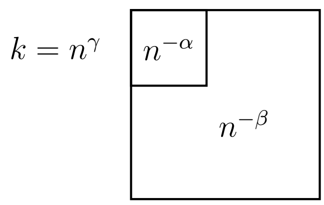


Figure 3: Sketch of the planted model

We fix the leaf vertex set $L = \{v_0, v_1, \ldots, v_r\}$ and set $f = f_{L,r,s,\beta}$ as

$$
f(G) = \sum_{\underline{u}} T_{L,\underline{u}}, \tag{6}
$$

where $\underline{u} = (u_1, \ldots, u_{s-r})$ is a tuple of vertices chosen from $V(G) \setminus L$. We define $T_{L,\underline{u}}$ as

$$
\begin{aligned}
T_{L,\underline{u}} &= \prod_{i=1}^{s-r} \prod_{v \in L_i} R_{v,u_i} \\
&= \prod_{i=1}^{s-r} \prod_{v \in L_i} (A_{vu_i} - n^{-\beta})
\end{aligned}
$$

Each L_i are defined as in Definition 3.1 here.

Let us now reflect a bit on what f actually does. This time it is not fully as simple as f counting the number of r/s-caterpillars in G. By subtracting $n^{-\beta}$ from each factor, we are utilizing a trick introduced in [2] and is referred to as *signed counting*. As we will see, the expectation of f in the homogeneous setting is 0, and the variance is significantly reduced compared to the regular caterpillar count due to the cancellations introduced. As such, counting the signed r/s-caterpillars will be our method for distinguishing between our two models $\mathcal{D}_0$ and $\mathcal{D}_1$.

Henceforth, index 1 corresponds to the planted subgraph model $\mathcal{D}_1$ and index 0 corresponds to the Erdős-Rényi model $\mathcal{D}_0$. Now, to see that $E_0[f] = 0$ we start by noting that for fixed vertices u and v, $\mathbb{E}_0[R_{v,u}] = 0$:

$$\begin{aligned}
\mathbb{E}_0[R_{v,u}] &= \mathbb{E}_0[A_{vu} - n^{-\beta}] \\
&= \mathbb{E}_0[A_{vu}] - E_0[n^{-\beta}] \\
&= (\mathbb{P}(v \sim u) \cdot 1 + \mathbb{P}(v \nsim u) \cdot 0) - n^{-\beta} \\
&= n^{-\beta} - n^{-\beta} \\
&= 0.
\end{aligned}$$

We here use the definition of the expected value as well as the fact that the second term is constant.

Making use of the above equality, the expected value of f in $\mathcal{D}_0$ becomes 0.

$$\begin{aligned}
\mathbb{E}_0[f] &= \sum_{\underline{u} \in V(G) \setminus L} \prod_{i=1}^{s-r} \prod_{v \in L_i} \mathbb{E}[R_{v,u_i}] \\
&= \sum_{\underline{u} \in V(G) \setminus L} \prod_{i=1}^{s-r} \prod_{v \in L_i} 0 \\
&= 0
\end{aligned} \tag{7}$$

We are now ready to formulate the second large result of the paper.

Theorem 3.6. *Fix the positive parameters* $0 < \alpha < \beta < 1$ *and* $0 < \gamma < 1/2$ *such that*

$$\alpha < \beta\gamma\,. \tag{8}$$

Then there exists relatively prime integers $0 < r < s$ *such that the low-degree polynomial* f*, defined as in* (6)*, strongly separates the distributions* $\mathcal{D}_0(\beta)$ *and* $\mathcal{D}_1(\alpha, \beta, \gamma)$*. Equivalently,* f *satisfies*

$$\frac{\mathbb{E}_1[f] - \mathbb{E}_0[f]}{\sqrt{Var_0(f)}} \to \infty$$

as $n \to \infty$.

Proof of Theorem 3.6. Firstly, since α, β and γ are all constants, and $\gamma < 1$, we may pick r and s such that

$$\alpha < \beta\gamma - \frac{r+1}{s}(\gamma - 1) \,. \tag{9}$$

As we already saw in (7) that $\mathbb{E}_0[f] = 0$, we have that

$$\mathrm{Var}_0(f) = \mathbb{E}_0[f^2] - \mathbb{E}_0[f]^2 = \mathbb{E}_0[f^2].$$

For this expression, we need to consider different overlaps of vertices between two r/s-caterpillars. However, for a copy of two caterpillars that do not align at every edge, the expected value of that term will be zero as it contains some linear factor corresponding to a non-overlapping edge. To see this, note that if edge uv is present in one caterpillar and not the other, then the term corresponding to this pair of caterpillars will contain the factor R_{uv}. Thus, the expected value of this term will be 0 as seen below. C_1 here denotes the caterpillar where the edge uv is present, and C_2 denotes the one without it.

$$\begin{aligned}
\mathbb{E}_0\left[R_{uv} \prod_{ij \in E(C_1)\setminus\{uv\}} R_{ij} \prod_{kl \in E(C_2)} R_{kl}\right] &= \mathbb{E}_0[R_{uv}] \cdot \mathbb{E}_0\left[\prod_{ij \in E(C_1)\setminus\{uv\}} R_{ij} \prod_{kl \in E(C_2)} R_{kl}\right] \\
&= 0 \cdot \mathbb{E}_0\left[\prod_{ij \in E(C_1)\setminus\{uv\}} R_{ij} \prod_{kl \in E(C_2)} R_{kl}\right] \\
&= 0
\end{aligned}$$

The calculations make use of the independence of the existence of each edge, allowing the split of factors in the first equality.

In view of the above fact, we only have to consider cases where we have complete overlap of edges, and thus $\mathbb{E}_0[T_{L,\underline{u}} \cdot T_{L,\underline{u}'}] = 0$ for $\underline{u} \neq \underline{u}'$.

For a given tuple of backbone vertices $\underline{u}$, we now aim to arrive at an upper bound for $\mathbb{E}_0[T^2_{L,\underline{u}}]$:

$$
\begin{aligned}
\mathbb{E}_0[T_{L,\underline{u}}^2] &= \mathbb{E}_0\left[\prod_{i=1}^{s-r}\prod_{v\in L_i}(A_{vu_i}-n^{-\beta})^2\right] \\
&= \prod_{i=1}^{s-r}\prod_{v\in L_i}\mathbb{E}_0\left[(A_{vu_i}-n^{-\beta})^2\right] \\
&= \prod_{i=1}^{s-r}\prod_{v\in L_i}\mathbb{E}_0\left[A_{vu_i}^2-2n^{-\beta}A_{vu_i}+n^{-2\beta}\right] \\
&= \prod_{i=1}^{s-r}\prod_{v\in L_i}\left(\mathbb{E}_0[A_{vu_i}]-2n^{-\beta}\mathbb{E}_0[A_{vu_i}]+n^{-2\beta}\right) \\
&= \prod_{i=1}^{s-r}\prod_{v\in L_i}\left(n^{-\beta}-2n^{-\beta}\cdot n^{-\beta}+n^{-2\beta}\right) \\
&= \prod_{i=1}^{s-r}\prod_{v\in L_i}n^{-\beta}(1-n^{-\beta}) \\
&\le \prod_{i=1}^{s-r}\prod_{v\in L_i}n^{-\beta} \\
&= n^{-\beta s}.
\end{aligned} \tag{10}
$$

To verify the above equations, note that $A_{vu_i}^2 = A_{vu_i}$ as A_{vu_i} only takes values 0 and 1. Also, in the last step we simply need to realize that there are exactly s factors in our expression, and that none of the factors no longer depend on v.

The number of ways we can choose these $s-r$ backbone vertices among the $n-(r+1)$ available vertices in the graph is $\binom{n-(r+1)}{s-r}$. By utilizing Lemma 2.3 we get

$$
\begin{aligned}
\binom{n-(r+1)}{s-r} &\le \left(\frac{e(n-(r+1))}{s-r}\right)^{s-r} \\
&\le \left(\frac{en}{s-r}\right)^{s-r} \\
&\le (3n)^{s-r}.
\end{aligned} \tag{11}
$$

With the help of (10) and (11) we now get an upper bound for $\mathrm{Var}_0(f) = E_0[f^2]$,

$$
\begin{aligned}
\mathrm{Var}_0(f) &= \sum_{\overline{u}} \mathbb{E}_0[T_{L,\underline{u}}^2] \\
&\leq \sum_{\underline{u}} n^{-\beta s} \\
&\leq (3n)^{s-r} n^{-\beta s} \\
&= 3^{s-r} n^{s-r-\beta s}.
\end{aligned} \tag{12}
$$

Recall that $S = S(G)$ is the vertex set of the planted dense subgraph. Regarding $\mathbb{E}_1[f]$, one can get a lower bound by only considering the cases where we pick the leaf set in S, i.e., making use of the inequality

$$
\mathbb{E}_1[f] \geq \mathbb{E}_1[f|L \subseteq S]\mathbb{P}(L \subseteq S).
$$

As for $\mathbb{E}_1[f|L \subseteq S]$, we make an important observation. For any $\underline{u}$ with some vertex not belonging to S, it will not contribute to the sum that $\mathbb{E}_1[f|L \subseteq S]$ consists of. To clearly illustrate this, let $\underline{u}' = (u_1', \ldots, u_{s-r}')$ be a tuple containing some vertex $u' \in V(G) \setminus S$. We obtain

$$
\begin{aligned}
\mathbb{E}_1[T_{L,\underline{u}}|L \subseteq S] &= \mathbb{E}_1\left[\prod_{i=1}^{s-r} \prod_{v \in L_i} (A_{vu_i'} - n^{-\beta})|L \subseteq S\right] \\
&= \prod_{i=1}^{s-r} \prod_{v \in L_i} \mathbb{E}_1[A_{vu_i'} - n^{-\beta}|L \subseteq S].
\end{aligned}
$$

The above expression will contain the factor $(A_{vu'} - n^{-\beta}) = n^{-\beta} - n^{-\beta} = 0$, making the entire product 0. With this in mind, we only have to consider cases where each vertex in the tuples $\underline{u}$ belong to S. Thus

$$
\begin{aligned}
\mathbb{E}_1[T_{L,\underline{u}}|L \subseteq S] &= \mathbb{P}(V(\underline{u}) \subseteq S) \cdot \prod_{i=1}^{s-r} \prod_{v \in L_i} (n^{-\alpha} - n^{-\beta}) \\
&= (n^{\gamma-1})^{|V(\underline{u})|} \cdot (n^{-\alpha} - n^{-\beta})^s \\
&= n^{(\gamma-1)(s-r)} n^{-\alpha s}(1 + o(1)).
\end{aligned} \tag{13}
$$

We here used the fact that each vertex in $V(\underline{u})$ belongs to S with probability $n^{\gamma-1}$. It remains to also bound $\binom{n-(r+1)}{s-r}$ from below. This is done using Lemma 2.3, as seen below.

$$
\begin{aligned}
\binom{n-(r+1)}{s-r} &\geq \binom{0.99n}{s-r} \\
&\geq \left(\frac{0.99n}{s-r}\right)^{s-r} \\
&= \left(\frac{0.99}{s-r}\right)^{s-r} n^{s-r} \\
&= Cn^{s-r}
\end{aligned} \tag{14}
$$

We note here that the first inequality holds true for large enough n and that C is a constant. Now we finally get that

$$
\begin{aligned}
\mathbb{E}_1[f|L \subseteq S] &= \sum_{\underline{u}} \mathbb{E}_1[T_{L,\underline{u}}|L \subseteq S] \\
&\geq Cn^{s-r} n^{(\gamma-1)(s-r)} n^{-\alpha s}(1+o(1)) \\
&= Cn^{\gamma(s-r)-\alpha s}(1+o(1)),
\end{aligned} \tag{15}
$$

making use of the equality (13) and the bound in (14).

Each vertex of L belongs to S with independent probability $k/n = n^{\gamma-1}$, hence the probability of all vertices in L belonging to S is

$$
\mathbb{P}(L \subseteq S) = (n^{\gamma-1})^{|L|} = n^{(\gamma-1)(r+1)}.
$$

Recalling the inequality 9, we may now choose the parameter ϵ to be sufficiently small, satisfying the inequality

$$
\alpha < \beta\gamma - \epsilon\left(\gamma - \frac{1}{2}\right) + \frac{r+1}{s}(\gamma - 1) \ .
$$

This can be reorganized in the following way;

$$
\begin{aligned}
\alpha &< \beta\gamma - \epsilon\left(\gamma - \frac{1}{2}\right) + \frac{r+1}{s}(\gamma-1) && (16) \\
&= \frac{\beta}{2} + \beta\gamma - \frac{\beta}{2} - \epsilon\left(\gamma - \frac{1}{2}\right) + \frac{r+1}{s}(\gamma-1) && (17) \\
&= \frac{\beta}{2} + (\beta-\epsilon)\left(\gamma - \frac{1}{2}\right) + \frac{r+1}{s}(\gamma-1)\,. && (18) \\
& && (19)
\end{aligned}
$$

Using the fact that $p = n^{-\beta} \leq n^{r/s-1-\epsilon}$, we have that

$$\begin{aligned}
-\beta &\leq r/s - 1 - \epsilon \\
\beta &\geq 1 - r/s + \epsilon \\
\beta - \epsilon &\geq \frac{s-r}{s}\,.
\end{aligned}$$

Inserting this in the inequality in (18) while recalling that the factor $\gamma - \frac{1}{2}$ is negative, we obtain

$$\begin{aligned}
\alpha &< \frac{\beta}{2} + \frac{s-r}{s}\left(\gamma - \frac{1}{2}\right) + \frac{r+1}{s}\left(\gamma - \frac{1}{2} - \frac{1}{2}\right) \\
&= \frac{\beta}{2} + \frac{s-r}{s}\left(\gamma - \frac{1}{2}\right) + \frac{r+1}{s}\left(\gamma - \frac{1}{2}\right) - \frac{r+1}{2s} \\
&= \frac{\beta}{2} + \left(1 + \frac{1}{s}\right)\left(\gamma - \frac{1}{2}\right) - \frac{r+1}{2s} \\
&= \frac{\beta}{2} + \left(\gamma - \frac{1}{2}\right) + \frac{1}{s}\left(\gamma - \frac{1}{2}\right) - \frac{1}{s}\left(\frac{r}{2} + \frac{1}{2}\right) \\
&= \frac{\beta}{2} + \left(\gamma - \frac{1}{2}\right) + \frac{1}{s}\left(\gamma - 1 - \frac{r}{2}\right) \\
&= \frac{\beta}{2} - \frac{r}{2s} - \frac{1}{2} + \gamma + \frac{\gamma - 1}{s}
\end{aligned}$$

Reorganizing the terms in the inequality now yields

$$\gamma - \alpha + \frac{\gamma - 1}{s} > \frac{1}{2} + \frac{r}{2s} - \frac{\beta}{2}\,.$$

By rearranging of terms the above becomes

$$\gamma s - \alpha s + \gamma - 1 > \frac{s}{2} + \frac{r}{2} - \frac{\beta s}{2}\,,$$

which in turn is equivalent to

$$\gamma s - \gamma r - \alpha s + \gamma r + \gamma - r - 1 > \frac{s}{r} - \frac{r}{2} - \frac{\beta s}{2}\,.$$

Finally, the inequality can be rewritten to obtain

$$\gamma(s-r) - \alpha s + (\gamma - 1)(r+1) > \frac{1}{2}(s - r - \beta s)\,.$$

This now implies that

$$n^{\gamma(s-r)-\alpha s+(\gamma-1)(r+1)} \gg n^{\frac{1}{2}(s-r-\beta s)},$$

and thus we can conclude that $\mathbb{E}_1[f|L \subseteq S]\mathbb{P}(L \subseteq S) \gg \sqrt{\mathrm{Var}_0(f)}$. By 1.8, our polynomial f strongly separates the two distributions. □

References

[1] Aditya Bhaskara, Moses Charikar, Eden Chlamtac, Uriel Feige, and Aravindan Vijayaraghavan. Detecting high log-densities: an $\mathrm{o(n^{\frac{1}{4}})}$ approximation for densest k-subgraph. In *Proceedings of the Forty-Second ACM Symposium on Theory of Computing*, STOC '10, page 201–210, New York, NY, USA, 2010. Association for Computing Machinery.

[2] Sébastien Bubeck, Jian Ding, Ronen Eldan, and Miklós Rácz. Testing for high-dimensional geometry in random graphs, 2015.

[3] Abhishek Dhawan, Cheng Mao, and Alexander S. Wein. Detection of dense subhypergraphs by low-degree polynomials, 2023.

[4] Josef Hadar and William R. Russell. Rules for ordering uncertain prospects. *The American Economic Review*, 59(01):25–34, 1969.

[5] Svante Janson, Tomasz Łuczak, and Andrzej Ruciński. *Random graphs*. Wiley-Interscience series in discrete mathematics. Wiley, New York, 2000.

[6] Colin McDiarmid and Fiona Skerman. Modularity and partially observed graphs, 2024.

[7] Tselil Schramm and Alexander S. Wein. Computational barriers to estimation from low-degree polynomials. *The Annals of Statistics*, 50(3), June 2022.

[8] Alexander S. Wein. Computational complexity of statistics: New insights from low-degree polynomials, 2025.

4 Appendix

4.1 Examples of caterpillars

Provided here are examples of r/s-caterpillars for different values of r and s.

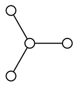
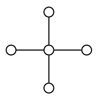
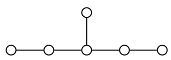
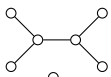
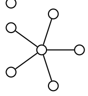
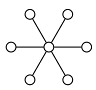
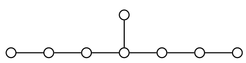
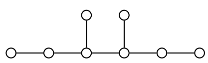
r	s	Caterpillar graph
1	2	
1	3	
2	3	
1	4	
3	4	
1	5	
2	5	
3	5	
4	5	
1	6	
5	6	
1	7	
2	7	
3	7	

Table 1: r/s-caterpillars for regimes $r < s \leq 12$

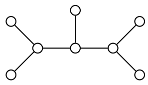
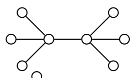
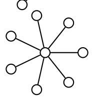
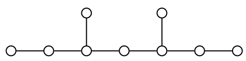
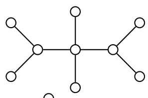
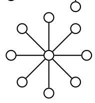
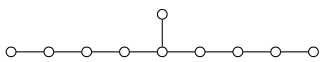
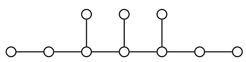
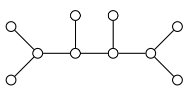
r	s	Caterpillar graph
4	7	
5	7	
6	7	
1	8	
3	8	
5	8	
7	8	
1	9	
2	9	
4	9	
5	9	

Table 2: r/s-caterpillars for regimes $r < s \leq 12$

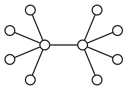
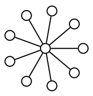
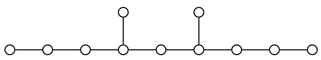
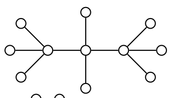
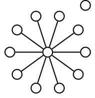
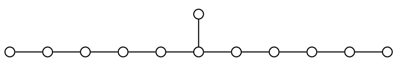
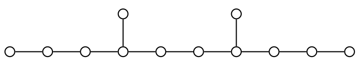
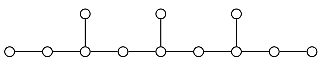
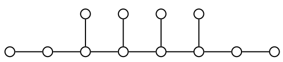
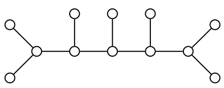
r	s	Caterpillar graph
7	9	
8	9	
1	10	
3	10	
7	10	
9	10	
1	11	
2	11	
3	11	
4	11	
5	11	
6	11	

Table 3: r/s-caterpillars for regimes $r < s \leq 12$

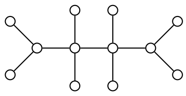
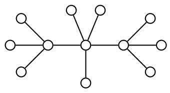
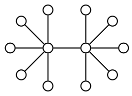
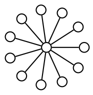
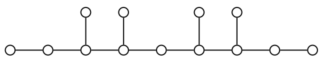
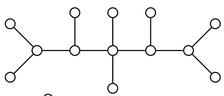
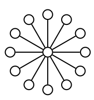
r	s	Caterpillar graph
7	11	
8	11	
9	11	
10	11	
1	12	
5	12	
7	12	
11	12	

Table 4: r/s-caterpillars for regimes $r < s \leq 12$