Math. Stat. Learn. 8 (2025), 105–164
DOI 10.4171/MSL/51

Counting stars is constant-degree optimal for detecting any planted subgraph

Xifan Yu, Ilias Zadik, and Peiyuan Zhang

Abstract. We study the computational limits of the following general hypothesis testing problem. Let $H = H_n$ be an *arbitrary* undirected graph. We study the detection task between a "null" Erdős–Rényi random graph $G(n, p)$ and a "planted" random graph which is the union of $G(n, p)$ together with a random copy of $H = H_n$. Our notion of planted model is a generalization of a plethora of recently studied models initiated with the study of the planted clique model (Jerrum, 1992), which corresponds to the special case where H is a k-clique and $p = 1/2$.

Over the last decade, several papers have studied the power of low-degree polynomials for limited choices of H's in the above task. In this work, we adopt a unifying perspective and characterize the power of *constant degree* polynomials for the detection task, when $H = H_n$ *is any arbitrary graph* and for *any* $p = \Omega(1)$. Perhaps surprisingly, we prove that an optimal constant degree polynomial is always given by simply *counting stars* in the input random graph. As a direct corollary, we conclude that the class of constant-degree polynomials is only able to "sense" the degree distribution of the planted graph H, and no other graph theoretic property of it.

Contents

Mathematics Subject Classification 2020: 62F03 (primary); 05C85, 60C05, 68R10 (secondary).

Keywords: planted subgraph, low-degree polynomial, computational hardness.

1. Introduction

During the last decade, researchers have revealed the existence of an intriguing phenomenon in several hypothesis testing tasks called *a computational-statistical trade-off*, which is a parameter regime where some test statistic is known to succeed, but, conjecturally, no computationally efficient[1] test statistics can work. This phenomenon interestingly appears in multiple different contexts across high-dimensional statistics, including community detection settings [2, 32], principal component analysis frameworks [5], regression models [12], and more.

To understand the nature of these phenomena, researchers have focused on simple statistical models that exhibit them. A simple, yet rich and canonical, family of such settings appears in community detection, and specifically are the so-called "planted" subgraph detection (or hypothesis testing) tasks where the goal is to detect the presence of a subgraph planted in an otherwise Erdős–Rényi random graph $G(n, p)$ [2, 5, 31, 32].[2] The motivation to study these planted subgraph tasks is to be able to identify the presence of an unusually large community in an otherwise homogeneous graph. A notable and very well-studied example of such a task is the planted clique problem [16], where one seeks to detect between a "null" model which is the Erdős–Rényi $G(n, 1/2)$, and a "planted subgraph" model which is the *union* of $G(n, 1/2)$ with a randomly chosen k-clique. Albeit a natural first choice, the assumption that the planted subgraph is a clique may stand as too restricted for many applications; the hidden community may be subject to some other structure that could significantly alter its detection limits. As an example, the desire to extract information from mobile objects in physics led to the study of the planted *matching* problem [7, 10, 19, 22]. For other reasons, researchers have studied multiple other planted subgraph models as well such that: (a) the planted dense subgraph problem [13], where one plants in $G(n, p)$ an instance of $G(n, q)$ for $q > p$, (b) the planted tree model [21], where one plants in $G(n, p)$ a D-ary tree, (c) the planted Hamiltonian cycle problem [3] where one plants a Hamiltonian cycle.

[1] For us, we say a test statistic is computational efficient if it is polynomial-time computable.

[2] For $n \in \mathbb{N}$, $p = p_n$, an instance of the Erdős–Rényi $G(n, p)$ is an n-vertex undirected random graph where each edge appears independently with probability p.

Notably, while at a high-level the methods to understand each planted subgraph model share similarities, the actual technical statistical analysis is often quite intricate and tailored to the specifics of the planted graph structure. For this reason, it appears hard to conclude general statistical principles for community detection based on this line of work. It is natural to wonder if one could simultaneously study all planted subgraph detection tasks by focusing on the properties of a general framework. Motivated by exactly this desire, the authors of [25] studied the information-theoretic transitions of a general planted subgraph model, which includes all the above mentioned planted models as special cases. While [25] focused on the recovery task of estimating the hidden subgraph, we focus here on the detection variant of it.

Definition 1.1 (Planted subgraph detection task). Let $n \in \mathbb{N}$, $p = p_n \in (0, 1)$ and $H = H_n$ be an *arbitrary* undirected graph. We consider the following detection task:

(1) (Null distribution $\mathbb{Q}$) In this case, the statistician observes an instance from the Erdős–Rényi random graph distribution $\mathbb{Q} = G(n, p)$.

(2) (Planted-H distribution $\mathbb{P}$) In this case, the statistician observes the union of an Erdős–Rényi random graph $G(n, p)$ with a random copy of H. The random copy of H is chosen uniformly at random from all the labelled copies of H in the complete graph.

It is rather straightforward to see how the general Definition 1.1 contains the mentioned detection tasks as special cases; e.g., when H is a k-clique and $p = 1/2$, we recover the planted clique task.

Searching for universal structure: Null and planted models. It is worth mentioning that this desire for generality shares roots with a fascinating line of work on the Kahn–Kalai conjecture from probabilistic combinatorics (see [18] for the conjecture, [28] for a recent breakthrough proof, as well as [26] for connection to Bayesian inference). The context of the conjecture has similarities with our setting. It studies our null distribution $\mathbb{Q} = G(n, p)$, and it is about characterizing the thresholds p for which an instance of a $G(n, p)$ contains a specific subgraph $H = H_n$ of interest. Similar again to the literature of planted models, a plethora of works have studied the thresholds for specific choices of subgraphs (e.g., see the classic work on Hamiltonian cycles [29] and the very technical recent work on spanning trees of bounded degree [24]). The Kahn–Kalai conjecture offers a formula for the threshold for *any subgraph* H. It is remarkable how the, now proven, Kahn–Kalai conjecture directly implies multiple previous notable results in random graph theory as direct corollaries (including the mentioned examples). It is also remarkable that the proof of the conjecture was only a few pages long. This line of work offers at least an argument that seeking a general and unifying structure in the analysis of such random graph models can be very fruitful.

Returning now to planted subgraph models, similar to [25], the question of finding a general (now *statistical*) structure underlying all these models, similar to the line of work on the Kahn–Kalai conjecture, is the primary motivation of our work. While [25] studied the information-theoretic limits of planted subgraph recovery tasks, in this work we investigate a common structure on their computational limits, i.e., in their *computational-statistical trade-offs*. Unfortunately, given that the $\mathcal{P} \neq \mathcal{NP}$ question remains unsolved, identifying the "true" computational limit of any detection task, characterizing when some polynomial-time test succeeds or not, appears to be well beyond the current mathematical abilities. For this reason, researchers on computational-statistical trade-offs have turned to studying multiple powerful restricted class of test statistics, often containing the best known polynomial-time test, and offering their proven failure point as evidence that the existing computational limits are fundamental.

Low-degree polynomials. Motivated by connections with a celebrated family of semidefinite programs, called the Sum-of-Squares hierarchy, the study of the powerful class of *low-degree polynomials* to construct test statistics has played a key role in this direction. First, it can be verified in a plethora of cases that the best known polynomial-time test statistics (e.g., based on spectral methods, or message passing methods) can be well approximated by low-degree polynomials (commonly $O(1)$ or $O(\log n)$ degree suffices). On top of that, the class of low-degree polynomials is believed to be very powerful, and a now well-known "low-degree conjecture" [14], [20] states that for a general class of detection problems when all degree-$O(\log n)$ polynomials fail to strongly separate the two distributions (see Definition 2.1 below), then no polynomial-time test will be able to detect between the two. The performance of the class of degree-$O(1)$ polynomials has also been used as (less strong but still quite interesting) evidence of hardness. For example, a recent work [23] established that Approximate Message Passing is optimal among degree-$O(1)$ polynomial in a spiked matrix estimation setting.

For these reasons, multiple papers have studied so far the power of low-degree polynomials to achieve strong separation for a number of different planted subgraph detection tasks. For example, in the planted clique model it is known that if $k = \Omega(\sqrt{n})$, some degree-$O(\log n)$ polynomial succeeds, while if $k = o(\sqrt{n})$ all degree-$O(\log n)$ polynomials fail to strongly separate the two distributions (see e.g., [20] and references therein). It should be noted though that for every new planted subgraph detection task that has been analyzed a new careful analysis is usually needed, which often brings its own challenges (similar to the literature on the Kahn–Kalai conjecture). The main focus of this work is to explore the *simultaneous* study of the class of low-degree polynomials for all planted subgraph detection tasks i.e., for any planted subgraph H.

Absence of structural positive results. Moreover, prior work on low-degree polynomials has built a powerful technique, based on what is called the low-degree advantage (or low-degree likelihood ratio) [20] (see Definition 2.2), to prove the *failure*[3] of the class of low-degree polynomials for detection tasks. To be more precise, as long as the degree-D advantage remains bounded, we know that no degree-D polynomial can strongly separate $\mathbb{Q}$ and $\mathbb{P}$. Yet, our understanding of how to argue about *positive results* for the class of low-degree polynomials is significantly more limited and much less automated. One natural candidate would be to consider the low-degree advantage again and use it as a criterion if it is unbounded. Unfortunately, this suggestion is not generally true. For example, for a regime of the so-called planted dense hypergraph problem, the low-degree advantage explodes (due to rare events) but in fact no low-degree polynomial succeeds [9]. Understanding whether the low-degree advantage exploding is a sufficient criterion for the success of low-degree polynomials when we focus on planted subgraph detection tasks is also a partial motivation for the present work.

A related struggle is that even if a low-degree polynomial is "predicted" to work, there is no known general tool to understand the structure of this "optimal" low-degree polynomial. Yet, the best known algorithms (and therefore their polynomial approximations) appear to be significantly different among different settings. For example, the best known polynomial-time algorithm for detecting or recovering the planted clique when $k = c\sqrt{n}$ for $c > 0$ small, is a spectral method combined with a postprocessing step [1], while for the recovery task in the planted Hamiltonian cycle problem it is a linear program relaxation of a TSP problem [3]. This is a significant issue, as for any new detection or recovery task that statisticians are facing, it appears as they need to design the "correct" polynomial-time test statistic mostly based on their intuition and the unique properties of each subgraph H.

Summarizing the above, this work is motivated by the following three key questions on planted subgraph detection tasks:

(Q1) For any $H = H_n$, can we automate when a degree-D polynomial works?

(Q2) Can we characterize the structure of an optimal degree-D polynomial?

(Q3) Which features of $H = H_n$ should a degree-D polynomial be exploiting?

Main contributions (informal). In this work, our main contribution is to offer an answer to the above questions (Q1), (Q2), (Q3) for any planted subgraph detection task with arbitrary $H = H_n$ and for any $p = \Omega(1)$, when we focus on the class of *degree-D $= O(1)$ polynomials*. Informally, under these assumptions, a summary of our contributions is as follows:

[3]From this point on in the introduction, by "success" (or "failure") of a polynomial in a detection task we strictly refer to whether it strongly separates $\mathbb{P}$ and $\mathbb{Q}$ (or not), per Definition 2.1.

(i) We start with our main result (Theorem 3.1). We prove that for all choices of $H = H_n$, an optimal degree-$D = O(1)$ polynomial is always given by the *signed count of a t-star graph* in the input graph for some $t \in \{1, D\}$. In other words, some degree-D polynomial succeeds in detecting if and only if the *signed edge count* or the *signed D-star count* works. This reveals an interesting new statistical principle shared by all planted subgraph detection tasks, and an easy-to-check criterion for the success of constant degree polynomials.

(ii) A moment analysis of the star counting polynomials, together with our main result, implies that the success of constant degree polynomials is a function solely of the *degree distribution* of the planted H (see Theorem 4.2). In other words, for any two H_1 and H_2 with the same degree distribution, either some degree-D polynomial succeeds in both detection tasks corresponding to H_1 and H_2, or all degree-D polynomials fail in both detection tasks corresponding to H_1 and H_2. We find this a surprising conclusion of our work, given all the potential other features of H a constant degree polynomial could be exploiting (e.g., the local-neighborhood structure of each vertex).

(iii) We describe how our results implies a series of old and new results on low-degree polynomials for planted detection tasks (see Section 4.1).

(iv) We prove that our main result is tight. We provide counterexamples for the optimality of star counts when either $p = o(1)$ or $D = \omega(1)$ (see Section 5).

Further comparison with previous work. We would like to expand here briefly on our literature review. We are not aware of any other work attempting to understand the above questions for planted subgraph detection tasks per Definition 1.1 in that level of generality. Yet, we should mention a relevant work by [15].

First, while the author of [15] also defined the union model, they focus their results on a similar, yet incomparable, general planted subgraph setting where one seeks to detect between an Erdős–Rényi $G(n, p)$ and a planted distribution where one plants a copy of a subgraph H as an *induced subgraph* in $G(n, p)$.

We first note this is not the same setting as Definition 1.1, as in our planted model we observe the *union* of a copy of H with $G(n, p)$. In particular, in our model the planted H is not assumed to be an induced subgraph of the input graph. Interestingly notice that the "induced" and "union" models are two distinct generalizations of the planted clique setting.

The power of a low-degree polynomial in the induced setting for all $n^{-o(1)} \leq p \leq 1 - n^{-o(1)}$ and for all H that have edge density bounded away from p is analyzed in [15]. In that case, the author proved that the computational trade-off is similar to the case of the planted clique model. For any such H, by simply counting edges one can detect whenever $|V(H)| = \omega(\sqrt{n \log n})$, and the authors proved that a spectral

method can improve this to $|V(H)| = \Omega(\sqrt{n})$. Moreover, if $|V(H)| = o(\sqrt{n})$ the author established that no degree-$O(\log n)$ polynomial works for the H's of interest. We remark again that we analyze the incomparable union model, and on top of this, we note that our results do not restrict H at all.

2. Preliminaries

To describe our main contribution in more detail, we need first to give a few definitions. Recall that a graph on n vertices can be represented as a list of Boolean variables

$$G = (G_{\{i,j\}})_{\{i,j\}\in\binom{[n]}{2}},$$

where each variable is the indicator variable of one edge in the complete graph K_n. We start with the notion of the strong separation which will be our focus of "success" for a polynomial in a planted subgraph detection task.

Definition 2.1 (Strong separation). For two distributions $\mathbb{P}, \mathbb{Q}$ supported on graphs $\{0,1\}^{\binom{n}{2}}$, we say that an $\binom{n}{2}$-variate polynomial $f(G)$ strongly separates the distributions $\mathbb{P}$ and $\mathbb{Q}$ if it holds that

$$\max\left\{\sqrt{\mathrm{Var}_{\mathbb{P}}(f)}, \sqrt{\mathrm{Var}_{\mathbb{Q}}(f)}\right\} = o\big(|\mathbb{E}_{\mathbb{P}}(f) - \mathbb{E}_{\mathbb{Q}}(f)|\big).$$

One should think of strong separation as a stronger condition for detection. A simple application of Chebychev's inequality implies that if f strongly separates $\mathbb{P}, \mathbb{Q}$, then thresholding f suffices to distinguish between the two distributions with vanishing Type I and II errors.

Relevant to strong separation is the concept of the advantage of a test function.

Definition 2.2 ((Low-degree) advantage). For two distributions $\mathbb{P}, \mathbb{Q}$ supported on graphs $\{0,1\}^{\binom{n}{2}}$, the advantage of a real-valued test function $f\colon \{0,1\}^{\binom{n}{2}} \to \mathbb{R}$ for testing distribution $\mathbb{P}$ against distribution $\mathbb{Q}$ is given by

$$\mathrm{Adv}(f) := \frac{\mathbb{E}_{\mathbb{P}}[f]}{\sqrt{\mathbb{E}_{\mathbb{Q}}[f^2]}},$$

and the degree-D advantage for testing distribution $\mathbb{P}$ against distribution $\mathbb{Q}$ (also called, the low-degree likelihood ratio) is

$$\mathrm{Adv}^{\leq D} := \max_{f\in\mathbb{R}[X]_{\leq D}} \mathrm{Adv}(f) = \max_{f\in\mathbb{R}[X]_{\leq D}} \frac{\mathbb{E}_{\mathbb{P}}[f]}{\sqrt{\mathbb{E}_{\mathbb{Q}}[f^2]}}.$$

It is known (see, e.g., [8, Lemma 7.3]) that if $\mathrm{Adv}^{\leq D} = O(1)$, then no degree-$D$ polynomial can strongly separate $\mathbb{P}$ and $\mathbb{Q}$.

Walsh–Fourier basis. We will use the Walsh–Fourier basis $(\chi_S)_{S\subseteq\binom{[n]}{2}}$ (resp. the degree-D Walsh–Fourier basis $(\chi_S)_{S\subseteq\binom{[n]}{2}:|S|\le D}$) with respect to the Erdős–Rényi distribution $G(n,p)$, which is defined by

$$\chi_S(G) := \prod_{\{i,j\}\in S} \frac{G_{\{i,j\}}-p}{\sqrt{p(1-p)}}.$$

Signed subgraph counts. Of special role in this work are the degree-D polynomials called (signed) subgraph counts. That is, for any shape[4] $\mathcal{S}$ with at most D edges, the signed count of $\mathcal{S}$ is the degree-D polynomial given by

$$f_{\mathcal{S}}(G) = \sum_{S\subseteq\binom{[n]}{2}:S\cong\mathcal{S}} \chi_S(G),$$

where $S\cong\mathcal{S}$ denotes the graph isomorphism relation.

Finally, we remind the reader of the definition of a star graph.

Definition 2.3. A star graph with t edges is a tree with $t+1$ vertices consisting of t leaves and 1 internal "central" vertex, as shown in Figure 1. In this paper, this graph is denoted as $K_{1,t}$, as it can be viewed as the complete bipartite graph with 1 vertex in one part and t vertices in the other. We will call $K_{1,t}$ a t-star graph.

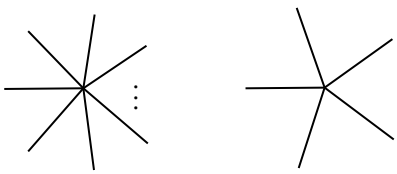

Figure 1. A generic star graph on the left, and the star graph $K_{1,5}$ on the right.

3. Main result: Optimality of star counts

Our main result is a generic result that holds for any $p=\Omega(1)$ and any $D=O(1)$. We prove that for all planted subgraph detection tasks (per Definition 1.1), i.e., for any $H=H_n$, there exists $t\le D$ for which the t-star signed subgraph count is optimal among all degree-D polynomials to strongly separate $\mathbb{P}$ and $\mathbb{Q}$. Formally, our main result is the following.

[4] A shape is an edge-induced graph.

Theorem 3.1. *Suppose $H = H_n$ is an arbitrary subgraph, $D = O(1)$ and $p = \Omega(1)$. Then, the following hold for testing $\mathbb{P}$ and $\mathbb{Q}$ in the planted subgraph detection task corresponding to planting a copy of H per Definition 1.1:*

- *If $\limsup_{n\to\infty} \mathrm{Adv}^{\le D} < \infty$, then no degree-$D$ polynomial $f \in \mathbb{R}[X]_{\le D}$ achieves strong separation between $\mathbb{P}$ and $\mathbb{Q}$ (see [8, Lemma 7.3]).*
- *If $\lim_{n\to\infty} \mathrm{Adv}^{\le D} = \infty$, then either the signed edge count $f_{\mathcal{K}_{1,1}}$ or the signed D-star count $f_{\mathcal{K}_{1,D}}$ has unbounded advantage, i.e.,*

$$\lim_{n\to\infty} \max\{\mathrm{Adv}(f_{\mathcal{K}_{1,1}}), \mathrm{Adv}(f_{\mathcal{K}_{1,D}})\} = \infty.$$

Moreover, suppose $\mathcal{S} \in \{\mathcal{K}_{1,1}, \mathcal{K}_{1,D}\}$ satisfies

$$\frac{\max\{\mathrm{Adv}(f_{\mathcal{K}_{1,1}}), \mathrm{Adv}(f_{\mathcal{K}_{1,D}})\}}{\mathrm{Adv}(f_{\mathcal{S}})} = O(1),$$

then $f_{\mathcal{S}}$ achieves strong separation between $\mathbb{P}$ and $\mathbb{Q}$.

A few remarks are in order.

Remark 3.2. As we mentioned in the introduction, it is not generally true that a growing advantage $\mathrm{Adv}^{\le D} = \omega(1)$ implies that some degree-D polynomial achieves strong separation. One such example has recently been studied in [9], where one plants an $H \sim G(n^\gamma, n^{-\alpha})$ in a $G(n, n^{-\beta})$ for constants $\alpha, \beta, \gamma \in (0, 1)$. The authors prove that whenever $\alpha > \beta\gamma$ and $0 < \gamma < 1/2$, no degree-$n^{o(1)}$ polynomial can strongly separate $\mathbb{P}$ from $\mathbb{Q}$. Yet, it is easy to check that, e.g., when $\alpha = 5/16$, $\gamma = 1/4$, $\beta = 1$, the constant degree advantage diverges to infinity with n. See Appendix C for more details on this.

We also note that the above is in interesting contrast to the well-known fact (the first bullet point in Theorem 3.1) that bounded degree-D advantage rules out the existence of degree-D polynomials that achieve strong separation [8, Lemma 7.3].

Our Theorem 3.1 shows that, interestingly, for all planted subgraph detection tasks with $p = \Omega(1)$, the "converse" does indeed hold for $D = O(1)$: whenever the degree-D advantage tends to infinity, there indeed exists a degree-D polynomial that achieves strong separation. In particular our result offers, to the best of our knowledge, the first complete characterization of the power of constant-degree polynomials in such settings.

Remark 3.3. Theorem 3.1 implies that for all planted H's, the simple choice of counting signed star graphs is always an optimal choice for strong separation between $\mathbb{P}$ and $\mathbb{Q}$, among all constant-degree polynomials. To the best of our knowledge, this is the first result revealing this universal optimality of counting stars for all planted subgraph detection tasks in our regime.

It is natural to wonder what is the reason counting stars enjoy such general optimality. We point the reader to Section 7, and specifically Proposition 7.1 for details and a proof sketch. We only would like to note here that the star structure appears naturally as a subgraph $\mathcal{S}$ whose signed count has (almost[5]) maximum advantage $\max_{\mathcal{S}} \mathrm{Adv}(f_{\mathcal{S}})$ among all constant-sized graphs.

Remark 3.4. Our theorem needs two assumptions for the optimality of signed t-star counts. First, that the class is constant degree polynomials, i.e., $D = O(1)$, and second that $p = \Omega(1)$. It turns out that both assumptions are necessary: if either is not satisfied, then the signed count of stars may fail to be optimal among constant-degree polynomials (see Section 5).

3.1. A simple criterion using only the degree profile of H

The fact that for any $D = O(1)$, the signed count of star graphs is optimal among degree-D polynomials for strong separation implies a very simple criterion for the success of the class of polynomial test functions. To present this, we fix an arbitrary subgraph H and consider the planted subgraph detection task for H with any noise level $p = \Omega(1)$. Suppose a statistician wishes to understand the power of degree-D polynomial test functions for this setting. Based on the literature, the currently natural approach would be as follows. First, the statistician would try to show that in one regime the low-degree advantage

$$\mathrm{Adv}^{\leq D} = \max_{f \in \mathbb{R}^{\leq D}[X]} \mathrm{Adv}(f)$$

remains bounded and then, in the remaining regime, to design (from scratch!) a constant-degree polynomial that works.

Based on Theorem 3.1, we arrive at a much simpler and automated approach for how to understand both directions when $D = O(1)$. It is in fact sufficient for the statistician to only calculate the advantage of the signed edge counts and the advantage of the signed D-star count. Indeed, by our Theorem 3.1 there exists a degree-D polynomial that can strongly separate $\mathbb{P}$ and $\mathbb{Q}$, if and only if

$$\max\{\mathrm{Adv}(f_{\mathcal{K}_{1,1}}), \mathrm{Adv}(f_{\mathcal{K}_{1,D}})\} = \omega(1), \tag{3.1}$$

and by comparing the advantages of the two polynomials $f_{\mathcal{K}_{1,1}}$ and $f_{\mathcal{K}_{1,D}}$, the statistician would also arrive at an optimal test function.

[5]By "almost" we mean that whenever the low-degree advantage tends to infinity, the signed count of some star graph also has its advantage tending to infinity.

On top of that, one can give a more explicit condition than (3.1) using Proposition 7.9 and Lemma 7.7. Together with the above discussion, this leads to the following simple and general condition for all planted subgraph detection tasks.

Corollary 3.5 (A simple condition on the degree profile of H). *For any $H = H_n$, $p = \Omega(1)$, $D = O(1)$, and the corresponding planted subgraph detection task, the following holds. There exists a degree-D polynomial that achieves strong separation if and only if*

$$\lim_{n\to\infty} \max\left\{ \frac{\sum_{v\in V(H)} d_v}{n(p/(1-p))^{1/2}}, \frac{\sum_{v\in V(H)} d_v^D}{n^{(1+D)/2}(p/(1-p))^{D/2}} \right\} = \infty, \tag{3.2}$$

where $d_v := \deg_H(v)$ denotes the degree of v in H. Moreover, if (3.2) holds, by choosing $t^* \in \{1, D\}$ that satisfies

$$\frac{\max\left\{ \frac{\sum_{v\in V(H)} d_v}{n(p/(1-p))^{1/2}}, \frac{\sum_{v\in V(H)} d_v^D}{n^{(1+D)/2}(p/(1-p))^{D/2}} \right\}}{\frac{\sum_{v\in V(H)} d_v^{t^*}}{n^{(1+t^*)/2}(p/(1-p))^{t^*/2}}} = O(1),$$

we can strongly separate $\mathbb{P}$ and $\mathbb{Q}$ using $f_{\mathcal{K}_{1,t^*}}$.

A potential striking aspect of Corollary 3.5 is that to judge whether some constant-degree polynomial works *one needs to only know the degree profile of H*, and no other graph property of it. For example, for a d-regular H with v vertices, constant degree polynomials can strongly separate $\mathbb{P}$ and $\mathbb{Q}$ either *for all* such graphs H or *for no* such H at all. Other more specific properties of H (e.g., the H's clique number, or its girth, or even spectral properties like H being an expander or not), which could naturally motivate the study of several other candidate degree-D polynomials, make *no difference* in whether some constant-degree polynomial can strongly separate or not.

4. Characterization of the optimal signed star count

Definition 4.1. We say a test T is an optimal test among a class of tests $\mathcal{T}$ for testing $\mathbb{P}$ against $\mathbb{Q}$ if whenever there exists a test $\widetilde{T} \in \mathcal{T}$ that strongly separates $\mathbb{P}$ and $\mathbb{Q}$, the test T also does so.

Based on Corollary 3.5, for constant D, an optimal degree-D polynomial test for detecting the planted subgraph is given by the signed count of t-stars for the maximizer of the condition (3.2) among $t \in \{1, D\}$. We will present a characterization of an optimal test that is (almost[6]) entirely based on the maximum degree of the planted

[6]Excluding a small "gray" area, see Figure 2.

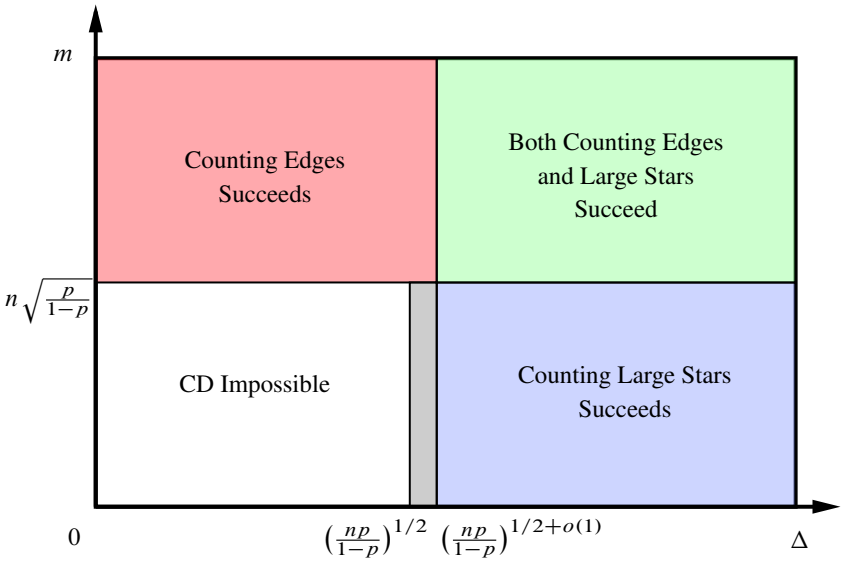

Figure 2. Phase-transition diagram characterized by Theorem 4.2. The x-axis is the maximum degree Δ and the y-axis is the number of edges m of the planted subgraph H. CD is an abbreviation of constant degree.

subgraph H. If the maximum degree is below the threshold $\approx (np/(1-p))^{1/2}$, one should simply count edges. On the other hand, if the maximum degree is much bigger than $\approx (np/(1-p))^{1/2}$, an optimal test is to count signed large stars (see Figure 2).

Theorem 4.2. *Denote $\Delta = \max_{i \in V(H)} d_i$, where $d_i = \deg_H(i)$ and $m = |E(H)|$. Then, for $p = \Omega(1)$, we have the following characterization for optimal constant degree polynomial tests for the planted subgraph detection problem of testing $\mathbb{P}$ against $\mathbb{Q}$ as defined in Definition 1.1:*

- *If $\Delta = O((n(p/(1-p)))^{1/2})$, then an optimal test is to count signed edges, i.e.,*

$$f = \sum_{\{i,j\} \in \binom{V}{2}} \chi_{\{i,j\}}.$$

  In particular, some constant-degree polynomial achieves strong separation if and only if $m = \omega((n(p/(1-p)))^{1/2})$.

- *If $\Delta \geq (n(p/(1-p)))^{1/2+\varepsilon}$ for some constant $\varepsilon > 0$, then an optimal test is to count signed "large" stars , i.e.,*

$$f = \sum_{S \subseteq \binom{V}{2} : S \cong s} \chi_S,$$

where $\mathcal{S} = \mathcal{K}_{1,t}$ for a large enough constant t. Moreover, it is always possible to set $t = \lceil 3/2\varepsilon \rceil$ so that the above f achieves strong separation.

The result of Theorem 4.2 is best visualized in Figure 2.

Remark 4.3. Although the characterization in Theorem 4.2 does not capture a middle region of

$$\omega\left(\left(n\frac{p}{1-p}\right)^{1/2}\right) \le \Delta \le \left(n\frac{p}{1-p}\right)^{1/2+o(1)},$$

an optimal test for this region can still be found by comparing the advantage achieved by the signed count of edges and that achieved by the signed count of D-stars as stated in Theorem 3.1.

4.1. Applications

To show the applicability of Theorem 4.2, we apply it to a number of examples.

Planted dense subgraph. Let $n, k = k_n \in \mathbb{N}$ and $p = p_n, q = q_n \in (0, 1)$. The first example we consider is the planted dense subgraph (PDS) setting, denoted by $\mathrm{PDS}(k, p, q)$. To describe it, we first draw H from $G(k, q)$. Then we consider the planted detection task per Definition 1.1 for planted H and p. Notice that this is called a PDS setting, as in the planted model $\mathbb{P}$, the induced subgraph on the k vertices corresponding to H is "denser" compared to the rest edges of the graph. Indeed, in the planted instance, every edge using only vertices of H appears marginally with probability $p + (1-p)q > p$ while the rest of the edges with probability p.

Notice that, as long as $kq = \omega(\log k)$, H has maximum degree $(1 + o(1))(k-1)q$ and edges $(1 + o(1))k^2q/2$ with high probability as n grows to infinity (see Lemma B.1 in the appendix). Using Theorem 4.2 we directly conclude the following.

Corollary 4.4. *If $p = \Omega(1)$, a constant-degree polynomial can achieve strong separation in* $\mathrm{PDS}(k, p, q)$ *if and only if*

$$k = \omega\left(\frac{\sqrt{n}}{\sqrt{q}(1-p)^{1/4}}\right).$$

Moreover, if $k = \omega(\sqrt{n}/\sqrt{q}(1-p)^{1/4})$, counting edges achieves strong separation.

For instance, if $k = n^\beta$, $p = 1 - n^{-\gamma}$ and $q = n^{-\alpha}$ for constants $\alpha, \beta, \gamma \in (0, 1)$ we immediately conclude the phase diagram for this $\mathrm{PDS}(k, p, q)$ for constant-degree polynomials: they work if and only if $\beta > 1/2 + \alpha/2 + \gamma/4$ (see Figure 3).

It is worth noting that PDS has been commonly studied under the slightly different following definition we call $\mathrm{PDS}'(k, p, p')$ for $0 < p = p_n < p' = p'_n < 1$ and $k = k_n$ (see, e.g., [6]). $\mathrm{PDS}'(k, p, p')$ is the detection task between the null $\mathbb{Q} = G(n, p)$

and the planted model $\mathbb{P}$ where now the observed graph is sampled like a $G(n, p)$ except that the edges between k vertices, chosen uniformly at random, are sampled now with probability $p' > p$. Notice this definition, differently from the one using Definition 1.1, is not assuming that the planted instance is the union of $G(n, p)$ with a randomly placed $H \sim G(n,q)$. Yet, for $q = (p' - p)/(1 - p)$ the two planted models of $\mathrm{PDS}(k, p, q)$ and $\mathrm{PDS}'(k, p, p')$ have marginally (per edge) the same law and it is natural to expect to have a similar computational transition.

Indeed, for $\mathrm{PDS}'(k, p, p')$ in the specialized regime

$$\Omega(1) = p = 1 - \Omega(1), \quad p' - p = \Theta(n^{-\alpha}), \quad \text{and} \quad k = \Theta(n^{\beta})$$

it is known, by using an average-case reduction from planted clique [6], that there is conjectured hard phase if and only if $\beta < 1/2 + \alpha/2$. By choosing the matching parameter $q = (p' - p)/(1 - p) = \Theta(n^{-\alpha})$ and $\gamma = 0$, we arrive at the same conclusion for $\mathrm{PDS}(k, p, q)$: constant degree polynomials fail if and only if $\beta < 1/2 + \alpha/2$ and counting edges works in the opposite regime.

We are not aware of any comparable low-degree lower bound result for either $\mathrm{PDS}'(k, p, p')$ and $\mathrm{PDS}(k, p, q)$.

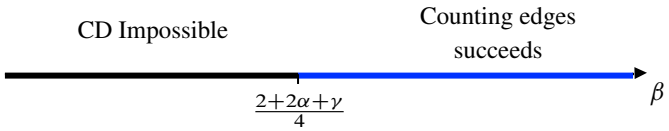

Figure 3. Phase-transition diagram for $\mathrm{PDS}(k, p, q)$, with $k = n^{\beta}$, $p = 1 - n^{-\gamma}$ and $q = n^{-\alpha}$.

Planted clique and planted independent set. When we choose $q = 1$ and $p = 1/2$ for $\mathrm{PDS}(k, p, q)$, the setting simply corresponds to the detection task of the planted clique problem. Corollary 4.4 then directly yields the well-known $k = \Theta(\sqrt{n})$ computational phase transition for constant-degree polynomials for planted clique [1, 4].

Interestingly, for $p = 1 - d/n$, $d = o(n)$ and $q = 1$, $\mathrm{PDS}(k, p, q)$ now maps to the detection task of the planted independent set problem. Indeed, $\mathrm{PDS}(k, 1 - d/n, 1)$ is the detection setting between $G(n, 1 - d/n)$ and $G(n, 1 - d/n)$ union a random k-clique. By equivalently considering the complements of the observation graphs, we need to detect between $G(n, d/n)$ and $G(n, d/n)$ where k vertices are constrained to induce an independent set, known as the planted independent set model. This and several related settings has been well studied in the literature (see, e.g., when $d = O(1)$, [33] for low-degree lower bounds for the search problem in the null model, and [17] for certification lower bounds in the null model). Corollary 4.4 implies that constant degree polynomials can detect if and only if $k = \omega(n^{3/4}/d^{1/4})$,

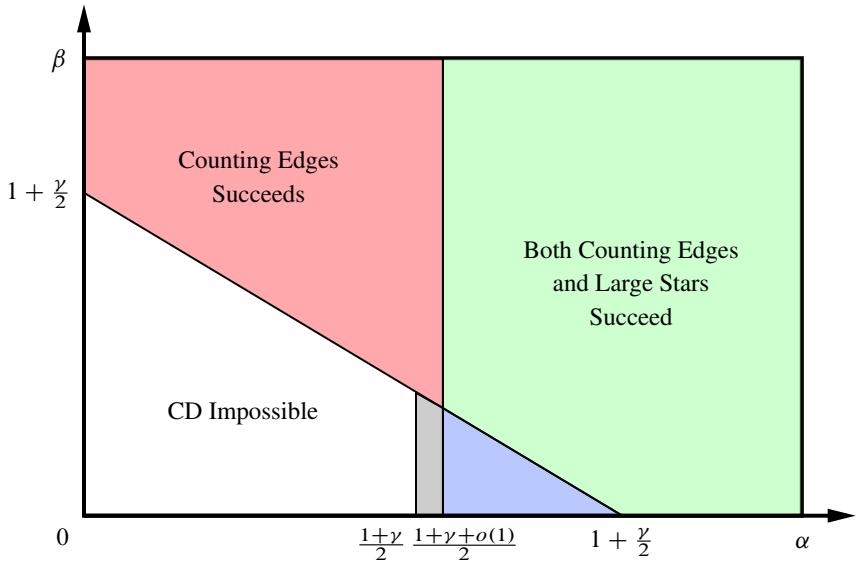

Figure 4. Phase-transition diagram of planted bipartite clique, where $H = K_{a,b}$ with $a \geq b$. The small blue triangle represents the region that only counting large stars succeeds. We use parameter configuration $p = 1 - n^{-\gamma}$, $a = n^{\alpha}$ and $b = n^{\beta}$.

which recovers the reduction-based hardness of planted independent set conditional on planted clique conjecture shown in [6].

Planted bipartite clique. Given the previous examples where counting edges is always constant-degree optimal, one may wonder if the count of large stars is an optimal constant-degree polynomial in some natural setting. For this reason, we now consider a planted bipartite clique setting $\mathrm{PBC}(a, b, p)$ for $a = a_n$, $b = b_n \in \mathbb{N}$ and $p = p_n \in (0, 1)$ with $a \geq b$, where we simply choose the planted H to be the bipartite clique $K_{a,b}$. The maximum degree of H is then clearly $\Delta = a$. Using Theorem 4.2 we arrive at the following richer computational diagram (see Figure 4).

Corollary 4.5. *For $\alpha, \beta, \gamma \in (0, 1)$, if $p = 1 - n^{-\gamma}$, $a = n^{\alpha}$, $b = n^{\beta}$, we have for* $\mathrm{PBC}(a, b, p)$*:*

- *if $2\alpha \leq 1 + \gamma$ and $\alpha + \beta \leq 1 + \gamma/2$, constant-degree polynomial test fails;*
- *$\alpha + \beta > 1 + \gamma/2$ if and only if counting edges achieves strong separation;*
- *if $2\alpha > 1 + \gamma + \varepsilon$ for some constant $\varepsilon > 0$, counting large stars achieves strong separation.*

We note that the planted bipartite clique has been studied in [30] in the case when p is bounded away from 1, where they showed a similar statistical computa-

tional gap as in the planted clique model. On the other hand, the example we describe here applies for all $p = \Omega(1)$.

5. Tightness of Theorem 3.1

In this section, we prove the tightness of our main result (Theorem 3.1). Recall, that according to Theorem 3.1, under the assumptions $p = \Omega(1)$ and $D = O(1)$, the degree-D polynomials achieve strong separation for detecting a planted subgraph in $G(n, p)$ if and only if the signed count polynomial of a $t \leq D$-star does so too. We provide counterexamples showing that if either $p = \Omega(1)$ or $D = O(1)$ is not satisfied, counting stars could fail to achieve strong separation even when some other degree-D polynomial does so.

Failure of counting stars under vanishing p. Assume $p = n^{-\gamma}$ for a constant $\gamma \in (0, 1)$. Then we show that for some appropriate constant size k, the planted detection model with H being a k-clique has the following behavior.

Lemma 5.1. *For any $p = n^{-\gamma}$ where $\gamma \in (0, 1)$ is a constant, there exists a constant k such that, for the planted subgraph detection task per Definition 1.1 with H being a k-clique, counting constant-sized stars fails to achieve strong separation, whereas some constant-degree polynomial test achieves strong separation.*

Failure of counting stars under non-constant degree D. Here, we fix $p = 1/2$ and $k = C\sqrt{n}$, where $C > 0$ is a large enough constant. By focusing on the performance of degree-$D = O(\log n)$ polynomials for the planted subgraph detection task where H is a k-clique the following holds.

Lemma 5.2. *Let H be a clique of size $C\sqrt{n}$ and $p = 1/2$, where $C > 0$ is any constant. Consider the planted subgraph detection task where H is planted in $G(n, p)$. Then, for any $t = t_n \geq 1$ (potentially growing with n), counting t-stars fails to achieve strong separation. However, for large enough $C > 0$, a degree-$O(\log n)$ polynomial achieves strong separation.*

6. Proof preliminaries

We now move to the proof sections of our results. We start by introducing some necessary definitions and notations.

Let $n, k \in \mathbb{N}$ be natural numbers. We will denote $[n] := \{1, 2, \ldots, n\}$. We will use $\binom{n}{k}$ to denote the number of ways to choose k elements from n elements, and $n_{(k)} := n(n-1)\ldots(n-k+1)$ to denote the k-th falling factorial of n. By conven-

tion, the 0-th falling factorial of any number is equal to 1 and $n_{(k)} = 0$ for any $k > n$. Let X be a set. We will denote $\binom{X}{k} := \{A \subseteq X : |A| = k\}$.

Asymptotic notations. We will use standard asymptotic notations $O, o, \Omega, \omega, \Theta$. Let $(a_n)_{n\in\mathbb{N}}$, $(b_n)_{n\in\mathbb{N}}$ be two sequences of positive real numbers. We sometimes write $a_n \lesssim b_n$ when $a_n = O(b_n)$, $a_n \ll b_n$ when $a_n = o(b_n)$, $a_n \gtrsim b_n$ when $a_n = \Omega(b_n)$, $a_n \gg b_n$ when $a_n = \omega(b_n)$, and $a_n \sim b_n$ when $\lim_{n\to\infty}(a_n/b_n) = 1$.

Graph theory basics. A graph is a pair of (V, E) where V is the vertex set and $E \subseteq \binom{V}{2}$ is the edge set. Sometimes we will identify the vertex set V with $[n]$. K_n will denote the complete graph on n vertices, and $K_{s,t}$ will be the complete bipartite graph with one part having s vertices and the other having t vertices.

We say that H is a subgraph of G, denoted as $H \subseteq G$, if

$$V(H) \subseteq V(G) \quad \text{and} \quad E(H) \subseteq E(G).$$

We say that H is an induced subgraph of G, if

$$V(H) \subseteq V(G) \quad \text{and} \quad E(H) = E(G) \cap \binom{V(H)}{2}.$$

For $S \subseteq V(G)$, we denote the unique induced subgraph of G with vertex set S as $G[S]$. We say H is a *spanning* subgraph of G, if it is a subgraph of G and $V(H) = V(G)$. We say a graph G is *edge-induced* if there are no isolated vertices in G.

A homomorphism from graph G_1 to graph G_2 is a mapping $f: V(G_1) \to V(G_2)$ that preserves adjacency relation, i.e., $\{f(u), f(v)\} \in E(G_2)$ if $\{u, v\} \in E(G_1)$. An isomorphism is a bijective homomorphism whose inverse is also a homomorphism, and we use $G_1 \cong G_2$ to denote that G_1 and G_2 are isomorphic. An automorphism of a graph G is an isomorphism from G to itself. The set of automorphisms of a graph G equipped with composition forms a group called the automorphism group of G, denoted as $\mathrm{Aut}(G)$.

Subgraph copies: Notation. We will use throughout a notion of shape and labelled/unlabelled copies of a shape.

Definition 6.1. A shape $\mathcal{S}$ is an edge-induced graph, i.e., a graph without isolated vertices. By a slight abuse of notation, we sometimes use $\mathcal{S}$ to refer to the edge set $E(\mathcal{S})$ as an edge-induced graph is determined by its edge set. For $D \in \mathbb{N}$, we use $\mathbb{G}_{\le D}$ to denote the collections of shapes with at most D edges up to isomorphism.

Let $\mathcal{S}$ be a shape, and G be a graph. An unlabelled copy of $\mathcal{S}$ in G is a subgraph $S \subseteq G$ such that $\mathcal{S}$ is isomorphic to S. A labelled copy of $\mathcal{S}$ in G is a pair (S, γ) of a subgraph $S \subseteq G$ together with a labelling $\gamma: V(S) \to V(\mathcal{S})$, such that γ is an

isomorphism from S to $\mathcal{S}$. Note that for a labelled copy (S, γ) of $\mathcal{S}$, γ^{-1} is an injective homomorphism from $\mathcal{S}$ to G, and conversely every injective homomorphism defines a labelled copy.

Unless stated otherwise, in this paper we will use copies to refer to labelled copies.

Following the notation in [27] we define for a shape $\mathcal{S}$ and a graph G,

$$M_{\mathcal{S},G} := \#\{\text{copies of } \mathcal{S} \text{ inside } G\}, \tag{6.1}$$

and when n is clear from context, $M_{\mathcal{S}} := M_{\mathcal{S},K_n} = n_{(|V(\mathcal{S})|)}$.

Example 6.2. In the complete graph K_n, the number of (labelled) copies of a triangle K_3 is $n_{(3)} = n(n-1)(n-2)$, whereas the number of triangles is $\binom{n}{3} = \frac{n(n-1)(n-2)}{6}$.

Definition 6.3 (Isomorphic pairs of copies). Let $\mathcal{S}_1$ and $\mathcal{S}_2$ be two shapes, and G and $\widehat{G}$ be graphs. Let $((S_1, \gamma_1), (S_2, \gamma_2))$ be a pair of copy of $\mathcal{S}_1$ and copy of $\mathcal{S}_2$ in G, and similarly $((\widehat{S}_1, \widehat{\gamma}_1), (\widehat{S}_2, \widehat{\gamma}_2))$ be a pair of copy of $\mathcal{S}_1$ and copy of $\mathcal{S}_2$ in $\widehat{G}$. We say that $((S_1, \gamma_1), (S_2, \gamma_2))$ and $((\widehat{S}_1, \widehat{\gamma}_1), (\widehat{S}_2, \widehat{\gamma}_2))$ are isomorphic if

- for every $u \in V(S_1) \cap V(S_2)$, $\widehat{\gamma}_1^{-1}(\gamma_1(u)) = \widehat{\gamma}_2^{-1}(\gamma_2(u))$;
- for every $v \in V(\widehat{S}_1) \cap V(\widehat{S}_2)$, $\gamma_1^{-1}(\widehat{\gamma}_1(u)) = \gamma_2^{-1}(\widehat{\gamma}_2(u))$.

Example 6.4. To illustrate Definition 6.3, consider the two shapes $\mathcal{S}_1$ and $\mathcal{S}_2$, and the 4 pairs of copies of them in Figure 5. (1) and (2) are isomorphic, as the pairs of labellings on all the vertices in the intersection are the same. (1) and (3) are not isomorphic, as in (1) a vertex in the intersection has a pair of labels (b, g) but in (3) a vertex in the intersection has a pair of labels (b, f). (1) and (4) are not isomorphic, as in (1) the two vertex sets intersect whereas in (4) the two vertex sets are disjoint.

Remark 6.5. It is not hard to see that the isomorphism of pairs of copies in Definition 6.3 forms an equivalence relation of the pairs of copies of two shapes in all graphs. Therefore, the set of pairs of copies $((S_1, \gamma_1), (S_2, \gamma_2))$ of $\mathcal{S}_1$ and $\mathcal{S}_2$ in all graphs can be partitioned into isomorphism classes, i.e., equivalence classes defined by the isomorphism relation.

Definition 6.6 (Intersecting pattern). Let $\mathcal{S}_1$ and $\mathcal{S}_2$ be two shapes. Then, an intersecting pattern of $\mathcal{S}_1$ and $\mathcal{S}_2$ is an isomorphism class of the pairs of copies of $\mathcal{S}_1$ and $\mathcal{S}_2$ in all graphs.

The set of all intersecting patterns of $\mathcal{S}_1$ and $\mathcal{S}_2$ is denoted as $\text{Inter}(\mathcal{S}_1, \mathcal{S}_2)$, and an element in it is denoted as $i(S_1, S_2)$. We use non-calligraphic $i(S_1, S_2)$ in order to distinguish a specific intersecting pattern from the given shapes $\mathcal{S}_1$ and $\mathcal{S}_2$.

Remark 6.7. Recall that an intersecting pattern is characterized by the exact way the vertex labels of two labelled copies of two shapes intersect. Let us now consider two

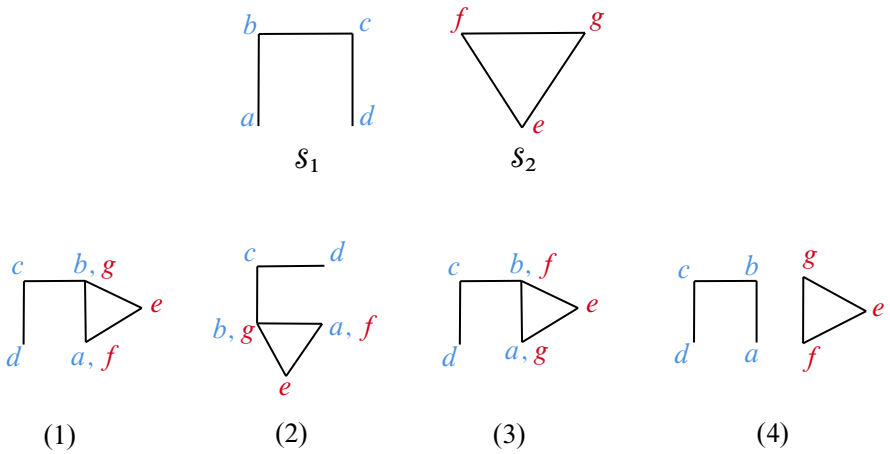

Figure 5. An example of 4 pairs of copies of $\mathcal{S}_1$ and $\mathcal{S}_2$ numbered by (1), (2), (3), (4), where the labellings of copies of $\mathcal{S}_1$ are marked with blue letters, and the labellings of copies of $\mathcal{S}_2$ are marked with red letters.

arbitrary shapes $\mathcal{S}_1$ and $\mathcal{S}_2$, with $s_1 = |V(\mathcal{S}_1)|$ and $s_2 = |V(\mathcal{S}_2)|$. Notice that the total number of possible intersecting patterns of $\mathcal{S}_1$ and $\mathcal{S}_2$ is $\sum_k \binom{s_1}{k}\binom{s_2}{k}k!$, as two pairs of copies of $\mathcal{S}_1$ and $\mathcal{S}_2$ are isomorphic if and only if the labels in their intersection are in to one to one correspondence (see Definition 6.3 and Example 6.4). Indeed, given the above, we just need to count the total number of ways to choose a subset of $V(\mathcal{S}_1)$ of size k, a subset of $V(\mathcal{S}_2)$ of size k, and a bijective function (that specifies how the vertices are "glued" in the intersection) between these two subsets, for all values of k. Importantly, if $\mathcal{S}$ is a shape with a constant number of edges (thus a constant number of vertices), then the number of intersecting patterns of pairs of copies of $\mathcal{S}$ with $\mathcal{S}$ is also bounded by a constant.

Definition 6.8. Let $\mathcal{S}_1$ and $\mathcal{S}_2$ be shapes, and $i(S_1, S_2)$ be an intersecting pattern of pairs of copies of $\mathcal{S}_1$ and $\mathcal{S}_2$. For $i(S_1, S_2)$, we define the symmetric difference shape $S_1 \triangle S_2$ as the shape obtained by first taking the symmetric difference of the edge sets of a pair of copies that have the intersecting pattern $i(S_1, S_2)$ and deleting isolated vertices after the symmetric difference operation. Similarly, we define the union shape $S_1 \cup S_2$ as the shape obtained by taking the union of the edge sets of a pair of copies that have the intersecting pattern $i(S_1, S_2)$.

Remark 6.9. Let $\mathcal{S}_1$ and $\mathcal{S}_2$ be shapes, $i(S_1, S_2)$ be an intersecting pattern, and G be a graph. The number of pairs of copies of $\mathcal{S}_1$ and $\mathcal{S}_2$ with the intersecting pattern $i(S_1, S_2)$ in G is equal to $M_{S_1 \cup S_2, G}$.

7. Proof of the main theorem

7.1. Proof strategy

We briefly describe our proof strategy for Theorem 3.1. The first bullet point is a standard result in the literature of low-degree polynomials, cf. [8, Lemma 7.3], and our focus will be on the second bullet point of Theorem 3.1.

The first key step of the proof is proving that if the low-degree advantage $\mathrm{Adv}^{\leq D}$ explodes, then the advantage of the count of a t-star $\mathrm{Adv}(f_{\mathcal{K}_{1,t}})$ explodes for some $t \leq D$. It is easy to see that if $\mathrm{Adv}^{\leq D} = \omega(1)$ for some $D = O(1)$, then the signed count of some shape $\mathcal{S}$ with at most D edges satisfies $\mathrm{Adv}(f_{\mathcal{S}}) = \omega(1)$. This follows by expanding the low-degree advantage (Proposition 7.5 and Proposition 7.9) to get

$$(\mathrm{Adv}^{\leq D})^2 = \sum_{\mathcal{S} \in \mathbb{G}_{\leq D}} (\mathrm{Adv}(f_{\mathcal{S}}))^2,$$

and using that there are a constant number of shapes with at most D edges. Our main idea is to use a careful recursive argument, stated in Proposition 7.1, that proves that as long as

$$\mathrm{Adv}(f_{\mathcal{S}}) = \omega(1) \tag{7.1}$$

for some $\mathcal{S}$, then it also holds

$$\mathrm{Adv}(f_{\mathcal{K}_{1,t}}) = \omega(1)$$

for some star shape $\mathcal{K}_{1,t}$. To prove this we use two "advantage-preserving" reductions that allow us to start with any shape S satisfying (7.1) and recursively switch to (1) any vertex-spanning subgraph of S which still satisfies (7.1) and (2) any connected component of S which still satisfies (7.1), as shown in Corollary 7.14 and Corollary 7.15, respectively. *This is the key step where the star shapes arise, as it is easy to prove that the minimal connected and spanning sub-shapes of any subgraph S are trees of diameter at most* 2*, which are exactly the star shapes.* Using a convexity argument, we further show in Proposition 7.1 that it suffices to consider the signed count of star graphs for two extreme cases: either the edge graph, or the D-star graph.

The second step is to prove that either the signed edge count or the signed D-star count, $f_{\mathcal{K}_{1,t}}$ for $t \in \{1, D\}$, achieves strong separation. Note that having an advantage that tends to infinity already implies by definition part of the strong separation condition, i.e.,

$$\sqrt{\mathrm{Var}_{\mathbb{Q}}[f_{\mathcal{K}_{1,t}}]} = o\big(|\mathbb{E}_{\mathbb{P}}[f_{\mathcal{K}_{1,t}}] - \mathbb{E}_{\mathbb{Q}}[f_{\mathcal{K}_{1,t}}]|\big).$$

To show the other side of the strong separation

$$\sqrt{\mathrm{Var}_{\mathbb{P}}[f_{\mathcal{K}_{1,t}}]} = o\big(|\mathbb{E}_{\mathbb{P}}[f_{\mathcal{K}_{1,t}}] - \mathbb{E}_{\mathbb{Q}}[f_{\mathcal{K}_{1,t}}]|\big),$$

we need to show equivalently

$$\frac{\mathbb{E}_{\mathbb{P}}[f^2_{\mathcal{K}_{1,t}}]}{\mathbb{E}_{\mathbb{P}}[f_{\mathcal{K}_{1,t}}]^2} \le 1 + o(1).$$

We begin by expanding $\mathbb{E}_{\mathbb{P}}[f^2_{\mathcal{K}_{1,t}}]$ as the sum of the expectations $\mathbb{E}[\chi_S \chi'_S]$ for all pairs of $S, S' \cong \mathcal{K}_{1,t}$. We show that the pairs with empty (edge) intersection contributes at most $1 + o(1)$ to the ratio above. For the pairs with non-empty intersection, we show that they contribute $o(1)$ to the ratio above using Proposition 7.2, under the assumption that

$$\lim_{n\to\infty} \max\{\mathrm{Adv}(f_{\mathcal{K}_{1,1}}), \mathrm{Adv}(f_{\mathcal{K}_{1,D}})\} = \infty,$$

and an approximate maximization condition

$$\frac{\max\{\mathrm{Adv}(f_{\mathcal{K}_{1,1}}), \mathrm{Adv}(f_{\mathcal{K}_{1,D}})\}}{\max_{1\le t\le D} \mathrm{Adv}(f_{\mathcal{K}_{1,t}})} = \Omega(1).$$

We highlight that this is a key technical part of the proof and the simplicity of the star shape appears essential. For example, one way this manifests itself is that there are only a few different cases that two star graphs can be intersecting (see, e.g., Figures 6, 7 and 8 below) which greatly simplifies the second moment expansion (see Proposition 7.2).

7.2. Key lemmas

Recall the notation $M_{\mathcal{S},G}$ in (6.1). The first key proposition proves that an exploding degree-D advantage implies that for $\mathcal{S}$ either being the edge graph $\mathcal{K}_{1,1}$ or the D-star $\mathcal{K}_{1,D}$, the quantity $(M^2_{\mathcal{S},H}/M_{\mathcal{S}})((1-p)/p)^{|\mathcal{S}|}$ is exploding. This quantity is simply the rescaled squared advantage of star count $f_{\mathcal{S}}$, since by Proposition 7.9, we have

$$\mathrm{Adv}(f_{\mathcal{S}})^2 = \frac{\mathbb{E}_{\mathbb{P}}[f_{\mathcal{S}}]^2}{\mathbb{E}_{\mathbb{Q}}[f^2_{\mathcal{S}}]} = \frac{M^2_{\mathcal{S},H}}{M_{\mathcal{S}} \cdot |\mathrm{Aut}(\mathcal{S})|}\Big(\frac{1-p}{p}\Big)^{|\mathcal{S}|}.$$

Proposition 7.1. *Suppose $D = O(1)$ and $p = \Omega(1)$. If $\mathrm{Adv}^{\le D} \to \infty$, then among the edge graph $\mathcal{K}_{1,1}$ and the D-star $\mathcal{K}_{1,D}$, we have*

$$\max\{\mathrm{Adv}(f_{\mathcal{K}_{1,1}}), \mathrm{Adv}(f_{\mathcal{K}_{1,D}})\} \to \infty.$$

Moreover, if $\mathrm{Adv}^{\le D} \to \infty$, there exists a constant $C = C(D)$ such that for all large enough n, we have

$$\max_{1\le t\le D} \mathrm{Adv}(f_{\mathcal{K}_{1,t}}) \le C \cdot \max\{\mathrm{Adv}(f_{\mathcal{K}_{1,1}}), \mathrm{Adv}(f_{\mathcal{K}_{1,D}})\}.$$

The next is an important and quite technical proposition that reveals some structure between overlapping copies of the star shape $\mathcal{S}$ with exploding advantage from Proposition 7.1. This proposition is the key step in a second moment calculation used in the proof of Theorem 3.1.

Proposition 7.2. *Suppose $D = O(1)$, $C = O(1)$, and $p = \Omega(1)$. Suppose $\mathcal{S} \cong \mathcal{K}_{1,t}$ with $1 \le t \le D$ is a star shape that satisfies $\mathrm{Adv}(f_{\mathcal{S}'}) \le C \cdot \mathrm{Adv}(f_{\mathcal{S}})$ for all star shapes $\mathcal{S}' \cong \mathcal{K}_{1,t'}$ with at most $|\mathcal{S}| = t$ edges, and suppose $\mathrm{Adv}(f_{\mathcal{S}}) \to \infty$. Let $i(S_1, S_2) \in \mathrm{Inter}(\mathcal{S}, \mathcal{S})$ be an intersecting pattern such that S_1 and S_2 have non-empty intersection. Then,*

$$\frac{n^{|V(S_1 \cup S_2)| - |V(S_1 \triangle S_2)|} M_{S_1 \triangle S_2, H}((1-p)/p)^{|S_1 \triangle S_2|/2}}{M_{\mathcal{S},H}^2((1-p)/p)^{|\mathcal{S}|}} = o(1).$$

7.3. Auxiliary lemmas

In this section, we have a series of auxiliary lemmas needed for the proof of Theorem 3.1. The proofs for some of the lemmas are deferred to Section A in the appendix.

We start with a standard result, which we state as a fact.

Fact 7.3. *The Walsh–Fourier basis $(\chi_S)_{S \subseteq \binom{[n]}{2}}$ (resp. the degree-D Walsh–Fourier basis $(\chi_S)_{S \subseteq \binom{[n]}{2} : |S| \le D}$) with respect to the Erdős–Rényi distribution $\mathbb{Q} := G(n,p)$ form an orthonormal basis for $\mathbb{R}[X]$ (resp. $\mathbb{R}[X]_{\le D}$) with respect to the inner product $\langle \cdot, \cdot \rangle_{\mathbb{Q}}$ defined by $\langle f, g \rangle_{\mathbb{Q}} = \mathbb{E}_{\mathbb{Q}}[fg]$.*

The following lemma is taken from [27], and will be useful throughout the calculations in this paper. Recall the definition of $M_{\mathcal{S},H}$ in (6.1).

Lemma 7.4. *For a fixed $S \subseteq \binom{[n]}{2}$ isomorphic to some shape $\mathcal{S}$, and $(G, \mathbf{H})$ drawn from $\mathbb{P}$, the probability that S is a subgraph of the planted $\mathbf{H}$ is*

$$\mathbb{P}(S \subseteq \mathbf{H}) = \frac{M_{\mathcal{S},H}}{M_{\mathcal{S}}}.$$

Proof of Lemma 7.4. We have

$$M_{\mathcal{S},H} = \mathbb{E}[\#\{\text{copies of } \mathcal{S} \text{ inside } \mathbf{H}\}] = M_{\mathcal{S}} \cdot \mathbb{P}(S \subseteq \mathbf{H}). \qquad \blacksquare$$

The following proposition calculates the low-degree advantage for all planted subgraph detection tasks.

Proposition 7.5. *For the planted subgraph detection task, the square of the degree-D advantage for testing distribution $\mathbb{P}$ against distribution $\mathbb{Q}$ is*

$$(\mathrm{Adv}^{\le D})^2 = \sum_{\mathcal{S} \in \mathbb{G}_{\le D}} \frac{M_{\mathcal{S},H}^2}{M_{\mathcal{S}} \cdot |\mathrm{Aut}(\mathcal{S})|} \left(\frac{1-p}{p}\right)^{|\mathcal{S}|}.$$

The following proposition gives a way of double counting the number of pairs of copies of two shapes in a graph in terms of their intersecting patterns.

Proposition 7.6. *Let $\mathcal{S}_1$ and $\mathcal{S}_2$ be shapes, and G be a graph. Then,*

$$M_{\mathcal{S}_1,G} M_{\mathcal{S}_2,G} = \sum_{i(S_1,S_2)\in \mathrm{Inter}(\mathcal{S}_1,\mathcal{S}_2)} M_{S_1\cup S_2,G}.$$

Proof of Proposition 7.6. The left-hand side is the number of pairs of copies of $\mathcal{S}_1$ and $\mathcal{S}_2$. The right-hand side counts the same number by enumerating over intersecting patterns of $\mathcal{S}_1$ and $\mathcal{S}_2$, and then counting the number of pairs isomorphic to a specific intersecting pattern. ∎

Next, we state a lemma that expresses the number of copies of a star shape in any graph in terms of its degree sequence.

Lemma 7.7. *Let H be a graph, and $\mathcal{S} \cong \mathcal{K}_{1,t}$ be a star shape. Then,*

$$M_{\mathcal{S},H} = \sum_{i\in V(H)} (d_i)_{(t)},$$

where $d_i = \deg_H(i)$.

The following lemmas will be useful when dealing with sums of falling factorials that arise from Lemma 7.7.

Lemma 7.8. *Let $d_1, \dots, d_k$ be a sequence of natural numbers taking values at most k for some function $k = k(n)$, and $t \in \mathbb{N}$ be a constant. If $\sum_{i\in[k]} d_i^t = \omega(k)$, then*

$$\sum_{i\in[k]} (d_i)_{(t)} = (1-o(1))\Big(\sum_{i\in[k]} d_i^t\Big).$$

7.4. Proof of Theorem 3.1: Putting it all together

For the signed count polynomial $f_{\mathcal{S}}$ of some shape $\mathcal{S}$, the following proposition expresses its first, second moments under $\mathbb{Q}$ and the first moment under $\mathbb{P}$ in simple formulas, whose proof is deferred to Section A in the appendix.

Proposition 7.9. *Let $\mathcal{S}$ be a shape. Then the following identities hold:*

$$\mathbb{E}_{\mathbb{Q}}[f_{\mathcal{S}}] = 0, \quad \mathbb{E}_{\mathbb{Q}}[f_{\mathcal{S}}^2] = \frac{M_{\mathcal{S}}}{|\operatorname{Aut}(\mathcal{S})|},$$
$$\mathbb{E}_{\mathbb{P}}[f_{\mathcal{S}}] = \frac{M_{\mathcal{S},H}}{|\operatorname{Aut}(\mathcal{S})|}\Big(\frac{1-p}{p}\Big)^{|\mathcal{S}|/2},$$
$$\operatorname{Adv}(f_{\mathcal{S}}) = \frac{M_{\mathcal{S},H}}{M_{\mathcal{S}}^{1/2}\cdot|\operatorname{Aut}(\mathcal{S})|^{1/2}}\Big(\frac{1-p}{p}\Big)^{|\mathcal{S}|/2}.$$

Now we are ready to present the full proof of our main theorem.

Proof of Theorem 3.1. If $\limsup_{n\to\infty} \mathrm{Adv}^{\le D} < \infty$, by [8, Lemma 7.3] we conclude no degree-D polynomial f achieves strong separation between $\mathbb{P}$ and $\mathbb{Q}$.

From now on we assume that $\lim_{n\to\infty} \mathrm{Adv}^{\le D} = \infty$. Suppose $\mathcal{S} \in \{\mathcal{K}_{1,1}, \mathcal{K}_{1,D}\}$ satisfies

$$\frac{\max\{\mathrm{Adv}(f_{\mathcal{K}_{1,1}}), \mathrm{Adv}(f_{\mathcal{K}_{1,D}})\}}{\mathrm{Adv}(f_{\mathcal{S}})} = O(1).$$

By Proposition 7.1,

$$\max\{\mathrm{Adv}(f_{\mathcal{K}_{1,1}}), \mathrm{Adv}(f_{\mathcal{K}_{1,D}})\} \to \infty, \tag{7.2}$$

and for some constant C_1, for all large enough n,

$$\max_{1\le t\le D} \mathrm{Adv}(f_{\mathcal{K}_{1,t}}) \le C_1 \cdot \max\{\mathrm{Adv}(f_{\mathcal{K}_{1,1}}), \mathrm{Adv}(f_{\mathcal{K}_{1,D}})\}. \tag{7.3}$$

We now aim to show that $f_{\mathcal{S}}$, the signed count of $\mathcal{S}$, achieves strong separation between $\mathbb{P}$ and $\mathbb{Q}$. From Proposition 7.9, we have

$$\mathbb{E}_{\mathbb{Q}}[f_{\mathcal{S}}] = 0, \quad \mathbb{E}_{\mathbb{Q}}[f_{\mathcal{S}}^2] = \frac{M_{\mathcal{S}}}{|\mathrm{Aut}(\mathcal{S})|},$$
$$\mathbb{E}_{\mathbb{P}}[f_{\mathcal{S}}] = \frac{M_{\mathcal{S},H}}{|\mathrm{Aut}(\mathcal{S})|}\Big(\frac{1-p}{p}\Big)^{|\mathcal{S}|/2}.$$

Note that

$$\mathrm{Var}_{\mathbb{Q}}[f_{\mathcal{S}}] = \mathbb{E}_{\mathbb{Q}}[f_{\mathcal{S}}^2] - \mathbb{E}_{\mathbb{Q}}[f_{\mathcal{S}}]^2 = \frac{M_{\mathcal{S}}}{|\mathrm{Aut}(\mathcal{S})|},$$
$$|\mathbb{E}_{\mathbb{P}}[f_{\mathcal{S}}] - \mathbb{E}_{\mathbb{Q}}[f_{\mathcal{S}}]| = \frac{M_{\mathcal{S},H}}{|\mathrm{Aut}(\mathcal{S})|}\Big(\frac{1-p}{p}\Big)^{|\mathcal{S}|/2}.$$

The condition in (7.2) implies one side of the strong separation

$$\sqrt{\mathrm{Var}_{\mathbb{Q}}[f_{\mathcal{S}}]} = o\big(|\mathbb{E}_{\mathbb{P}}[f_{\mathcal{S}}] - \mathbb{E}_{\mathbb{Q}}[f_{\mathcal{S}}]|\big), \tag{7.4}$$

where we use that $|\mathrm{Aut}(\mathcal{S})| = O(1)$ as $\mathcal{S}$ has a constant number of edges.

It remains to show that

$$\sqrt{\mathrm{Var}_{\mathbb{P}}[f_{\mathcal{S}}]} = o\big(|\mathbb{E}_{\mathbb{P}}[f_{\mathcal{S}}] - \mathbb{E}_{\mathbb{Q}}[f_{\mathcal{S}}]|\big).$$

Since $\mathbb{E}_{\mathbb{P}}[f_{\mathcal{S}}] - \mathbb{E}_{\mathbb{Q}}[f_{\mathcal{S}}] = \mathbb{E}_{\mathbb{P}}[f_{\mathcal{S}}]$ and $\mathrm{Var}_{\mathbb{P}}[f_{\mathcal{S}}] = \mathbb{E}_{\mathbb{P}}[f_{\mathcal{S}}^2] - \mathbb{E}_{\mathbb{P}}[f_{\mathcal{S}}]^2$, to show $\mathrm{Var}_{\mathbb{P}}[f_{\mathcal{S}}] = o(|\mathbb{E}_{\mathbb{P}}[f_{\mathcal{S}}] - \mathbb{E}_{\mathbb{Q}}[f_{\mathcal{S}}]|^2)$ is equivalent to proving

$$\frac{\mathbb{E}_{\mathbb{P}}[f_{\mathcal{S}}^2]}{\mathbb{E}_{\mathbb{P}}[f_{\mathcal{S}}]^2} \le 1 + o(1).$$

Recall $f_{\mathcal{S}} = \sum_{S \subseteq \binom{V}{2} : S \cong \mathcal{S}} \chi_S$. We now examine this ratio

$$\frac{\mathbb{E}_{\mathbb{P}}[f_{\mathcal{S}}^2]}{\mathbb{E}_{\mathbb{P}}[f_{\mathcal{S}}]^2} = \frac{\sum_{S,S' \subseteq \binom{V}{2} : S,S' \cong \mathcal{S}} \mathbb{E}_{\mathbb{P}}[\chi_S \chi_{S'}]}{(M^2_{\mathcal{S},H}/|\operatorname{Aut}(\mathcal{S})|^2)((1-p)/p)^{|\mathcal{S}|}}. \tag{7.5}$$

Let us compute

$$\begin{aligned}
\mathbb{E}_{\mathbb{P}}[\chi_S \chi_{S'}] &= \mathbb{E}_{\mathbb{P}}\left[\prod_{\{i,j\} \in S} \frac{G_{i,j} - p}{\sqrt{p(1-p)}} \prod_{\{i',j'\} \in S'} \frac{G_{i',j'} - p}{\sqrt{p(1-p)}} \right] \\
&= \mathbb{E}_{\mathbb{P}}\left[\prod_{\{i,j\} \in S \triangle S'} \frac{G_{i,j} - p}{\sqrt{p(1-p)}} \prod_{\{i',j'\} \in S \cap S'} \left(\frac{G_{i',j'} - p}{\sqrt{p(1-p)}} \right)^2 \right] \\
&= \mathbb{E}_{\mathbb{P}}\left[\chi_{S \triangle S'} \prod_{\{i,j\} \in S \cap S'} \left(\frac{G_{i,j} - p}{\sqrt{p(1-p)}} \right)^2 \right] \\
&= \mathbb{E}_{\mathbf{H}} \mathbb{E}_{\mathbb{P}}\left[\chi_{S \triangle S'} \prod_{\{i,j\} \in S \cap S'} \left(\frac{G_{i,j} - p}{\sqrt{p(1-p)}} \right)^2 \,\middle|\, \mathbf{H} \right].
\end{aligned} \tag{7.6}$$

Observe that the conditional expectation above evaluates to 0 whenever $S \triangle S'$ is not fully contained inside $\mathbf{H}$. For a fixed embedding $\mathbf{H}$ of H, if $S \triangle S' \subseteq \mathbf{H}$, then

$$\begin{aligned}
&\mathbb{E}_{\mathbb{P}}\left[\chi_{S \triangle S'} \prod_{\{i,j\} \in S \cap S'} \left(\frac{G_{i,j} - p}{\sqrt{p(1-p)}} \right)^2 \,\middle|\, \mathbf{H} \right] \\
&\quad = \left(\frac{1-p}{p} \right)^{|S \triangle S'|/2} \mathbb{E}_{\mathbb{P}}\left[\prod_{\{i,j\} \in S \cap S'} \left(\frac{G_{i,j} - p}{\sqrt{p(1-p)}} \right)^2 \,\middle|\, \mathbf{H} \right] \\
&\quad = \left(\frac{1-p}{p} \right)^{|S \triangle S'|/2} \left(\frac{1-p}{p} \right)^{|(S \cap S') \cap E(\mathbf{H})|},
\end{aligned} \tag{7.7}$$

where the last equality follows from

$$\mathbb{E}\left[\frac{(G_{i,j} - p)^2}{p(1-p)} \,\middle|\, \mathbf{H} \right] = \begin{cases} (1-p)/p & \text{if } \{i,j\} \in E(\mathbf{H}), \\ 1 & \text{otherwise.} \end{cases}$$

Moreover, since $D = O(1)$ and $p = \Omega(1)$, there exists some constant $C_2 = C_2(D,p)$ such that

$$\left(\frac{1-p}{p} \right)^{|(S \cap S') \cap E(\mathbf{H})|} \le \max\left\{ 1, \left(\frac{1-p}{p} \right)^D \right\} \le C_2, \tag{7.8}$$

and for S, S' with $S \cap S' = \varnothing$, we have

$$\left(\frac{1-p}{p} \right)^{|(S \cap S') \cap E(\mathbf{H})|} = 1. \tag{7.9}$$

Inserting (7.7) and (7.8) back to (7.6), we get

$$\begin{aligned}
\mathbb{E}_{\mathbb{P}}[\chi_S\chi_{S'}] &= \mathbb{E}_{\mathbf{H}}\mathbb{E}_{\mathbb{P}}\left[\chi_{S\triangle S'}\prod_{\{i,j\}\in E(S\cap S')}\left(\frac{G_{i,j}-p}{\sqrt{p(1-p)}}\right)^2 \,\middle|\, \mathbf{H}\right] \\
&\le C_2\cdot\mathbb{E}_{\mathbf{H}}\left[\mathbf{1}\{S\triangle S'\subseteq\mathbf{H}\}\left(\frac{1-p}{p}\right)^{|S\triangle S'|/2}\right] \\
&= C_2\cdot\left(\frac{1-p}{p}\right)^{|S\triangle S'|/2}\mathbb{P}(S\triangle S'\subseteq\mathbf{H}) \\
&= C_2\cdot\left(\frac{1-p}{p}\right)^{|S\triangle S'|/2}\frac{M_{S\triangle S',H}}{M_{S\triangle S'}}, \qquad (7.10)
\end{aligned}$$

where in the last line we use Lemma 7.4. Specifically, in the case that $S\cap S'=\varnothing$, by (7.9), we have

$$\mathbb{E}_{\mathbb{P}}[\chi_S\chi_{S'}] = \left(\frac{1-p}{p}\right)^{|S\triangle S'|/2}\frac{M_{S\triangle S',H}}{M_{S\triangle S'}}. \qquad (7.11)$$

Substituting the bound (7.10) and (7.11) back to our ratio (7.5), we get

$$\begin{aligned}
\frac{\mathbb{E}_{\mathbb{P}}[f_{\mathcal{S}}^2]}{\mathbb{E}_{\mathbb{P}}[f_{\mathcal{S}}]^2} &= \frac{\sum_{S,S'\subseteq\binom{V}{2}:S,S'\cong\mathcal{S}}\mathbb{E}_{\mathbb{P}}[\chi_S\chi_{S'}]}{(M_{\mathcal{S},H}^2/|\operatorname{Aut}(\mathcal{S})|^2)((1-p)/p)^{|\mathcal{S}|}} \\
&= \frac{1}{M_{\mathcal{S},H}^2((1-p)/p)^{|\mathcal{S}|}}\cdot\sum_{\substack{((S,\gamma),(S',\gamma')):\\ S,S'\cong\mathcal{S}}}\mathbb{E}_{\mathbb{P}}[\chi_S\chi_{S'}] \\
&= \frac{1}{M_{\mathcal{S},H}^2((1-p)/p)^{|\mathcal{S}|}}\cdot\sum_{\substack{i(S_1,S_2)\\ \in\mathrm{Inter}(\mathcal{S},\mathcal{S})}}\sum_{\substack{(S,S'):\\ (S,S')\\ \cong i(S_1,S_2)}}\mathbb{E}_{\mathbb{P}}[\chi_S\chi_{S'}] \\
&\le \frac{1}{M_{\mathcal{S},H}^2((1-p)/p)^{|\mathcal{S}|}}\cdot\left(\sum_{\substack{i(S_1,S_2)\\ \in\mathrm{Inter}(\mathcal{S},\mathcal{S}):\\ S_1\cap S_2=\varnothing}}\sum_{\substack{(S,S'):\\ (S,S')\\ \cong i(S_1,S_2)}}\frac{M_{S_1\triangle S_2,H}}{M_{S_1\triangle S_2}}\left(\frac{1-p}{p}\right)^{|S_1\triangle S_2|/2}\right. \\
&\qquad \left. + C_2\cdot\sum_{\substack{i(S_1,S_2)\\ \in\mathrm{Inter}(\mathcal{S},\mathcal{S}):\\ S_1\cap S_2\neq\varnothing}}\sum_{\substack{(S,S'):\\ (S,S')\\ \cong i(S_1,S_2)}}\frac{M_{S_1\triangle S_2,H}}{M_{S_1\triangle S_2}}\left(\frac{1-p}{p}\right)^{|S_1\triangle S_2|/2}\right),
\end{aligned}$$

where in the second line the summation is over pairs of copies of $\mathcal{S}$, rather than sets (unlabelled copies) isomorphic to $\mathcal{S}$, which cancels out the $|\operatorname{Aut}(\mathcal{S})|^2$ from the first line.

Let us examine the first term, corresponding to pairs with empty intersection:

$$\begin{aligned}
&\sum_{\substack{i(S_1,S_2)\\ \in \mathrm{Inter}(\mathcal{S},\mathcal{S}):\\ S_1\cap S_2=\varnothing}} \sum_{\substack{(S,S'):\\ (S,S')\\ \cong i(S_1,S_2)}} \frac{M_{S_1\triangle S_2,H}}{M_{S_1\triangle S_2}} \\
&\quad= \sum_{\substack{i(S_1,S_2)\\ \in \mathrm{Inter}(\mathcal{S},\mathcal{S}):\\ S_1\cap S_2=\varnothing}} \sum_{\substack{(S,S'):\\ (S,S')\\ \cong i(S_1,S_2)}} \frac{M_{S_1\cup S_2,H}}{M_{S_1\cup S_2}} \le \sum_{\substack{i(S_1,S_2)\\ \in \mathrm{Inter}(\mathcal{S},\mathcal{S})}} \sum_{\substack{(S,S'):\\ (S,S')\\ \cong i(S_1,S_2)}} \frac{M_{S_1\cup S_2,H}}{M_{S_1\cup S_2}} \\
&\quad= \sum_{\substack{i(S_1,S_2)\\ \in \mathrm{Inter}(\mathcal{S},\mathcal{S})}} M_{S_1\cup S_2} \frac{M_{S_1\cup S_2,H}}{M_{S_1\cup S_2}} = \sum_{\substack{i(S_1,S_2)\\ \in \mathrm{Inter}(\mathcal{S},\mathcal{S})}} M_{S_1\cup S_2,H} \\
&\quad= M_{S_1,H} M_{S_2,H} = M_{\mathcal{S},H}^2,
\end{aligned}$$

where in the second-to-last equality we use Proposition 7.6. As a result, the first term can be bounded by

$$\begin{aligned}
&\frac{1}{M_{\mathcal{S},H}^2((1-p)/p)^{|\mathcal{S}|}} \cdot \sum_{\substack{i(S_1,S_2)\\ \in \mathrm{Inter}(\mathcal{S},\mathcal{S}):\\ S_1\cap S_2=\varnothing}} \sum_{\substack{(S,S'):\\ (S,S')\\ \cong i(S_1,S_2)}} \frac{M_{S_1\triangle S_2,H}}{M_{S_1\triangle S_2}} \Big(\frac{1-p}{p}\Big)^{|S_1\triangle S_2|/2} \\
&\quad\le \frac{1}{M_{\mathcal{S},H}^2((1-p)/p)^{|\mathcal{S}|}} \cdot M_{\mathcal{S},H}^2 \Big(\frac{1-p}{p}\Big)^{|\mathcal{S}|} = 1.
\end{aligned}$$

Thus, we obtain a bound

$$\begin{aligned}
\frac{\mathbb{E}_{\mathbb{P}}[f_{\mathcal{S}}^2]}{\mathbb{E}_{\mathbb{P}}[f_{\mathcal{S}}]^2} \le 1 + \frac{C_2}{M_{\mathcal{S},H}^2((1-p)/p)^{|\mathcal{S}|}} \\
\cdot \sum_{\substack{i(S_1,S_2)\\ \in \mathrm{Inter}(\mathcal{S},\mathcal{S}):\\ S_1\cap S_2\neq\varnothing}} \sum_{\substack{(S,S'):\\ (S,S')\\ \cong i(S_1,S_2)}} \frac{M_{S_1\triangle S_2,H}}{M_{S_1\triangle S_2}} \Big(\frac{1-p}{p}\Big)^{|S_1\triangle S_2|/2}.
\end{aligned} \tag{7.12}$$

Next let us examine the second term, corresponding to pairs with non-empty intersection:

$$\begin{aligned}
&\sum_{\substack{i(S_1,S_2)\\ \in \mathrm{Inter}(\mathcal{S},\mathcal{S}):\\ S_1\cap S_2\neq\varnothing}} \sum_{\substack{(S,S'):\\ (S,S')\\ \cong i(S_1,S_2)}} \frac{M_{S_1\triangle S_2,H}}{M_{S_1\triangle S_2}} \Big(\frac{1-p}{p}\Big)^{|S_1\triangle S_2|/2} \\
&\quad= \sum_{\substack{i(S_1,S_2)\\ \in \mathrm{Inter}(\mathcal{S},\mathcal{S}):\\ S_1\cap S_2\neq\varnothing}} M_{S_1\cup S_2} \cdot \frac{M_{S_1\triangle S_2,H}}{M_{S_1\triangle S_2}} \Big(\frac{1-p}{p}\Big)^{|S_1\triangle S_2|/2}
\end{aligned}$$

$$\leq \sum_{\substack{i(S_1,S_2)\\ \in \mathrm{Inter}(\mathcal{S},\mathcal{S}):\\ S_1\cap S_2\neq\varnothing}} n^{|V(S_1\cup S_2)|-|V(S_1\triangle S_2)|} M_{S_1\triangle S_2,H}\Big(\frac{1-p}{p}\Big)^{|S_1\triangle S_2|/2},$$

which is a sum over a constant number (depending on the constant D) of terms. Moreover, under the condition (7.2) and (7.3), each term is $o(M^2_{\mathcal{S},H}((1-p)/p)^{|\mathcal{S}|})$ by Proposition 7.2. Thus, the second term is bounded by

$$\begin{aligned}
&\frac{C_2}{M^2_{\mathcal{S},H}((1-p)/p)^{|\mathcal{S}|}} \cdot \sum_{\substack{i(S_1,S_2)\\ \in \mathrm{Inter}(\mathcal{S},\mathcal{S}):\\ S_1\cap S_2\neq\varnothing}} \sum_{\substack{(S,S'):\\ (S,S')\\ \cong i(S_1,S_2)}} \frac{M_{S_1\triangle S_2,H}}{M_{S_1\triangle S_2}}\Big(\frac{1-p}{p}\Big)^{|S_1\triangle S_2|/2}\\
&\leq \frac{C_2}{M^2_{\mathcal{S},H}((1-p)/p)^{|\mathcal{S}|}}\\
&\quad\cdot \sum_{\substack{i(S_1,S_2)\\ \in \mathrm{Inter}(\mathcal{S},\mathcal{S}):\\ S_1\cap S_2\neq\varnothing}} n^{|V(S_1\cup S_2)|-|V(S_1\triangle S_2)|} M_{S_1\triangle S_2,H}\Big(\frac{1-p}{p}\Big)^{|S_1\triangle S_2|/2}\\
&= o(1).
\end{aligned}$$

Plugging it back to (7.12), we get the desired bound

$$\frac{\mathbb{E}_{\mathbb{P}}[f_{\mathcal{S}}^2]}{\mathbb{E}_{\mathbb{P}}[f_{\mathcal{S}}]^2} \leq 1 + o(1),$$

which, as discussed at the beginning of the proof, implies the other side of the strong separation

$$\sqrt{\mathrm{Var}_{\mathbb{P}}[f_{\mathcal{S}}]} = o\big(|\mathbb{E}_{\mathbb{P}}[f_{\mathcal{S}}] - \mathbb{E}_{\mathbb{Q}}[f_{\mathcal{S}}]|\big). \tag{7.13}$$

Since both conditions (7.4) and (7.13) hold, we conclude that $f_{\mathcal{S}}$, the signed count of the star shape $\mathcal{S}$, where $\mathcal{S}$ is either the edge graph $\mathcal{K}_{1,1}$ or the D-star $\mathcal{K}_{1,D}$, achieves strong separation between $\mathbb{P}$ and $\mathbb{Q}$. ∎

7.5. Proof of the key lemmas

We will need the following two claims about some combinatorial properties of the quantities $M_{\mathcal{S},H}$. This will be directly useful for proving the intuition that focusing on stars is all we need, leading to Proposition 7.1.

Claim 7.10. *Let $\mathcal{S}$ be a shape, and $\mathcal{S}'$ be a spanning sub-shape of $\mathcal{S}$. Then, $M_{\mathcal{S}',H} \geq M_{\mathcal{S},H}$ for any graph H.*

Claim 7.11. *Let $\mathcal{S}$ be a shape that is the disjoint union of two shapes $\mathcal{S}_1$ and $\mathcal{S}_2$ that are vertex-disjoint inside $\mathcal{S}$. Then, $M_{\mathcal{S}_1,H} M_{\mathcal{S}_2,H} \geq M_{\mathcal{S},H}$ for any graph H. If moreover $\mathcal{S}$ has a constant number of edges, then $M_{\mathcal{S}_1} M_{\mathcal{S}_2} \leq (1+o(1))M_{\mathcal{S}}$ as $n\to\infty$.*

Proof of Claim 7.10. Let $S \subseteq H$ together with a labelling $\gamma\colon V(S) \to V(\mathcal{S})$ be a copy of $\mathcal{S}$ in H. Since $\mathcal{S}'$ is a spanning sub-shape of $\mathcal{S}$, there is a spanning subgraph $S' \subseteq S$ that is isomorphic to $\mathcal{S}'$ and moreover inherits γ as the isomorphism mapping. It is not hard to see from the argument above for every copy of $\mathcal{S}$, we find a distinct copy of $\mathcal{S}'$, and $M_{\mathcal{S}',H} \geq M_{\mathcal{S},H}$. ∎

Proof of Claim 7.11. Let $S \subseteq H$ with a labelling $\gamma\colon V(S) \to V(\mathcal{S})$ be a copy of $\mathcal{S}$ inside H. As $\mathcal{S}$ is the vertex-disjoint union of two shapes $\mathcal{S}_1$ and $\mathcal{S}_2$, S consists of two vertex-disjoint subgraphs $\gamma^{-1}(\mathcal{S}_1)$ and $\gamma^{-1}(\mathcal{S}_2)$, and γ induces two labellings $\gamma_1\colon \gamma^{-1}(\mathcal{S}_1) \to \mathcal{S}_1$ and $\gamma_2\colon \gamma^{-1}(\mathcal{S}_2) \to \mathcal{S}_2$. Thus, $(\gamma^{-1}(\mathcal{S}_i), \gamma_i)$ is a copy of $\mathcal{S}_i$ in H for $i \in \{1,2\}$. It is not hard to see that for every copy of $\mathcal{S}$, we find a distinct pair of copies of $\mathcal{S}_1$ and $\mathcal{S}_2$, and $M_{\mathcal{S}_1,H} M_{\mathcal{S}_2,H} \geq M_{\mathcal{S},H}$.

On the other hand, if $\mathcal{S}$ has only a constant number of edges, then in the complete graph K_n, we have

$$\begin{aligned} M_{\mathcal{S}_1} M_{\mathcal{S}_2} &= (n)_{(|V(\mathcal{S}_1)|)}(n)_{(|V(\mathcal{S}_2)|)} \leq n^{|V(\mathcal{S}_1)|+|V(\mathcal{S}_2)|} \\ &\leq \big(1+o(1)\big)(n)_{|V(\mathcal{S})|} = \big(1+o(1)\big)M_{\mathcal{S}}. \end{aligned}$$ ∎

Immediately, the two claims above yield the following two simple corollaries, which will be used to prove Proposition 7.1.

Corollary 7.12. *Let $\mathcal{S}$ be a shape, and $\mathcal{S}'$ be a spanning sub-shape of $\mathcal{S}$. Then,*

$$\frac{M^2_{\mathcal{S}',H}}{M_{\mathcal{S}'}} \geq \frac{M^2_{\mathcal{S},H}}{M_{\mathcal{S}}}.$$

Corollary 7.13. *Let $\mathcal{S}$ be a shape with a constant number of edges that is the disjoint union of two shapes $\mathcal{S}_1$ and $\mathcal{S}_2$ that are vertex-disjoint inside $\mathcal{S}$. Then,*

$$\left(\frac{M^2_{\mathcal{S}_1,H}}{M_{\mathcal{S}_1}}\left(\frac{1-p}{p}\right)^{|\mathcal{S}_1|}\right)\cdot\left(\frac{M^2_{\mathcal{S}_2,H}}{M_{\mathcal{S}_2}}\left(\frac{1-p}{p}\right)^{|\mathcal{S}_2|}\right) \geq \big(1-o(1)\big)\frac{M^2_{\mathcal{S},H}}{M_{\mathcal{S}}}\left(\frac{1-p}{p}\right)^{|\mathcal{S}|}.$$

Combining the advantage of signed subgraph count computed in Proposition 7.9 and the fact that shapes with a constant number of edges have automorphism groups of constant size, we state the following direct consequences of Corollary 7.12 and Corollary 7.13, which may be of independent interest.

Corollary 7.14. *Let $\mathcal{S}$ be a shape, and $\mathcal{S}'$ be a spanning sub-shape of $\mathcal{S}$. Suppose $p = \Omega(1)$. If $|\mathcal{S}| = O(1)$, then* $\mathrm{Adv}(f_{\mathcal{S}'}) = \Omega(\mathrm{Adv}(f_{\mathcal{S}}))$.

Corollary 7.15. *Let $\mathcal{S}$ be a shape that is the disjoint union of two shapes $\mathcal{S}_1$ and $\mathcal{S}_2$ that are vertex-disjoint inside $\mathcal{S}$. If $|\mathcal{S}| = O(1)$, then*

$$\mathrm{Adv}(f_{\mathcal{S}_1})\cdot\mathrm{Adv}(f_{\mathcal{S}_2}) = \Omega\big(\mathrm{Adv}(f_{\mathcal{S}})\big).$$

Proof of Corollary 7.14. By Proposition 7.9, we have

$$\mathrm{Adv}(f_{\mathcal{S}'})^2 = \frac{M^2_{\mathcal{S}',H}}{M_{\mathcal{S}'\cdot|\operatorname{Aut}(\mathcal{S}')|}}\Big(\frac{1-p}{p}\Big)^{|\mathcal{S}'|}, \quad \mathrm{Adv}(f_{\mathcal{S}})^2 = \frac{M^2_{\mathcal{S},H}}{M_{\mathcal{S}}\cdot|\operatorname{Aut}(\mathcal{S})|}\Big(\frac{1-p}{p}\Big)^{|\mathcal{S}|}.$$

Notice that the automorphism group of any constant sized shape has size bounded by a constant, and that

$$\Big(\frac{1-p}{p}\Big)^{|\mathcal{S}|} \le \Big(\frac{1-p}{p}\Big)^{|\mathcal{S}'|}$$

since $p = \Omega(1)$ and $|\mathcal{S}'| \le |\mathcal{S}|$. We may thus use Corollary 7.12 to conclude

$$\begin{aligned}\mathrm{Adv}(f_{\mathcal{S}'})^2 &= \frac{M^2_{\mathcal{S}',H}}{M_{\mathcal{S}'\cdot|\operatorname{Aut}(\mathcal{S}')|}}\Big(\frac{1-p}{p}\Big)^{|\mathcal{S}'|}\\ &\gtrsim \frac{M^2_{\mathcal{S},H}}{M_{\mathcal{S}\cdot|\operatorname{Aut}(\mathcal{S})|}}\Big(\frac{1-p}{p}\Big)^{|\mathcal{S}|} = \mathrm{Adv}(f_{\mathcal{S}})^2.\end{aligned}$$

■

Proof of Corollary 7.15. Again the proof follows straightforwardly by using Proposition 7.9 and noticing that the automorphism groups involved have constant sizes. ■

Proof of Proposition 7.1. By Proposition 7.5 and Proposition 7.9, we have

$$\begin{aligned}(\mathrm{Adv}^{\le D})^2 &= \sum_{\mathcal{S}\in\mathbb{G}_{\le D}} \frac{M^2_{\mathcal{S},H}}{M_{\mathcal{S}}\cdot|\operatorname{Aut}(\mathcal{S})|}\Big(\frac{1-p}{p}\Big)^{|\mathcal{S}|}\\ &= \sum_{\mathcal{S}\in\mathbb{G}_{\le D}} \mathrm{Adv}(f_{\mathcal{S}})^2.\end{aligned}$$

Since there are at most a constant number of shapes with at most D edges for constant D, if $\mathrm{Adv}^{\le D} \to \infty$, then

$$\max_{\mathcal{S}\in\mathbb{G}_{\le D}} \mathrm{Adv}(f_{\mathcal{S}}) \to \infty.$$

Now let $\mathcal{S}$ be a shape with at most D edges that maximizes $\mathrm{Adv}(f_{\mathcal{S}})$ in the above. With Corollary 7.14 and Corollary 7.15 in hand, we will show a properly chosen sub-shape of $\mathcal{S}$ satisfies the condition of the corollary.

If $\mathcal{S}$ is already star graph, then we are done. If $\mathcal{S}$ is not a connected shape, then we may recurse on one of the connected components $\mathcal{S}'$ of $\mathcal{S}$ while ensuring that $\mathrm{Adv}(f_{\mathcal{S}}) \to \infty$ using Corollary 7.15. So now let us assume $\mathcal{S}$ is connected. Let $\mathcal{T}$ be a spanning tree of the shape $\mathcal{S}$. By Corollary 7.14,

$$\mathrm{Adv}(f_{\mathcal{T}}) \gtrsim \mathrm{Adv}(f_{\mathcal{S}}) \to \infty.$$

If the diameter of $\mathcal{T}$ is at least 3, let us consider a path $a-b-c-d$ of length 3 in $\mathcal{T}$. Note that after the edge $\{b,c\}$ is deleted from $\mathcal{T}$, what remains is still a spanning sub-shape $\mathcal{T}-\{b,c\}$, and we may recurse on it by Corollary 7.14.

Repeating the process above, we will end up with a shape $\mathcal{S}$ that satisfies

$$\mathrm{Adv}(f_{\mathcal{S}}) \to \infty,$$

and moreover is either a star graph or a tree of diameter at most 2. Note that star graphs are precisely trees with diameters at most 2. This shows that

$$\max_{\mathcal{S}\cong\mathcal{K}_{1,t}:t\le D} \mathrm{Adv}(f_{\mathcal{S}}) \to \infty. \tag{7.14}$$

Next, we will show that (7.14) implies that

$$\max\{\mathrm{Adv}(f_{\mathcal{K}_{1,1}}), \mathrm{Adv}(f_{\mathcal{K}_{1,D}})\} \to \infty,$$

and

$$\max_{1\le t\le D} \mathrm{Adv}(f_{\mathcal{K}_{1,t}}) \le C\cdot\max\{\mathrm{Adv}(f_{\mathcal{K}_{1,1}}), \mathrm{Adv}(f_{\mathcal{K}_{1,D}})\}$$

for some constant C that depends on D. For $\mathcal{S}\cong\mathcal{K}_{1,t}$, let us consider the following function

$$g(t) := \frac{M_{\mathcal{K}_{1,t},H}}{M_{\mathcal{K}_{1,t}}^{1/2}}\Big(\frac{1-p}{p}\Big)^{|\mathcal{K}_{1,t}|/2}.$$

By Proposition 7.9, $g(t) = |\mathrm{Aut}(\mathcal{K}_{1,t})|^{1/2}\cdot\mathrm{Adv}(f_{\mathcal{K}_{1,t}})$. Thus, for $t\le D = O(1)$, since the automorphism group of $\mathcal{K}_{1,t}$ has bounded size, there exists $C_0>0$ such that

$$C_0 \le \mathrm{Adv}(f_{\mathcal{K}_{1,t}})/g(t) \le 1. \tag{7.15}$$

We will show that $g(t)$ is sandwiched between two convex functions. Using this strategy, we will show that whenever (7.14) holds, we have $\max\{g(1), g(D)\}\to\infty$. By Lemma 7.7, we have

$$g(t) = \big(1+o(1)\big)\cdot\frac{\sum_{i\in V(H)}(d_i)_{(t)}}{n^{(1+t)/2}}\Big(\frac{1-p}{p}\Big)^{t/2}.$$

For the ease of notation, let us denote

$$\widehat{g}(t) := \frac{\sum_{i\in V(H)}(d_i)_{(t)}}{n^{(1+t)/2}}\Big(\frac{1-p}{p}\Big)^{t/2},$$

so that $g(t)\sim\widehat{g}(t)$. We will prove that, for fixed constant D, there exists constants $C_1, C_2>0$ such that

$$C_1\cdot h(t) - C_2 \le \widehat{g}(t) \le h(t), \quad \forall t\in[D], \tag{7.16}$$

where we denote

$$h(t) := \frac{\sum_{i\in V(H)} d_i^t}{n^{(1+t)/2}}\Big(\frac{1-p}{p}\Big)^{t/2}.$$

The second inequality of (7.16) is trivial using the definition of falling factorial. We then focus on the first inequality. Let us group all the vertices $i \in V(H)$ according to their degrees d_i:

$$\sum_{i\in V(H)} (d_i)_{(t)} = \sum_{i:d_i\geq t} (d_i)_{(t)} + \sum_{i:d_i<t} (d_i)_{(t)}.$$

For any $i \in V(H)$ with $d_i \geq t$, by Lemma 9.1,

$$(d_i)_{(t)} \geq d_i^t \cdot e^{-t^2/2(d_i-t+1)} \geq d_i^t \cdot e^{-t^2/2} \geq d_i^t \cdot e^{-D^2/2},$$

since $t \leq D$ and $t \leq d_i$. For the vertices with degrees less than t, we compute

$$\begin{aligned}\frac{\sum_{i\in V(H):d_i<t} d_i^t}{n^{(1+t)/2}}\Big(\frac{1-p}{p}\Big)^{t/2} &\leq \frac{|V(H)|\cdot t^t}{n^{(1+t)/2}}\Big(\frac{1-p}{p}\Big)^{t/2}\\ &\leq \frac{|V(H)|t}{n}\cdot\Big(\frac{t^2}{np}\Big)^{(t-1)/2} = O(1),\end{aligned}$$

since $p = \Omega(1)$, $|V(H)| \leq n$, and $d_i < t \leq D = O(1)$. Combining the results, we get

$$\begin{aligned}h(t) &= \frac{\sum_{i\in V(H)} d_i^t}{n^{(1+t)/2}}\Big(\frac{1-p}{p}\Big)^{t/2}\\ &= \frac{\sum_{i\in V(H):d_i<t} d_i^t}{n^{(1+t)/2}}\Big(\frac{1-p}{p}\Big)^{t/2} + \frac{\sum_{i\in V(H):d_i\geq t} d_i^t}{n^{(1+t)/2}}\Big(\frac{1-p}{p}\Big)^{t/2}\\ &\leq O(1) + e^{D^2/2}\cdot\frac{\sum_{i\in V(H):d_i\geq t}(d_i)_{(t)}}{n^{(1+t)/2}}\Big(\frac{1-p}{p}\Big)^{t/2}\\ &= O(1) + e^{D^2/2}\cdot\frac{\sum_{i\in V(H)}(d_i)_{(t)}}{n^{(1+t)/2}}\Big(\frac{1-p}{p}\Big)^{t/2}\\ &= O(1) + e^{D^2/2}\cdot\hat{g}(t),\end{aligned} \tag{7.17}$$

where the second last equality holds because $(d_i)_{(t)} = 0$ for $d_i < t$. Note that $e^{D^2/2}$ is a constant when D is a constant. Therefore, by (7.17), given fixed D, there exists constants $C_1, C_2 > 0$ such that

$$C_1 \cdot h(t) - C_2 \leq \hat{g}(t) \leq h(t), \quad \forall t \in [D],$$

which finishes the proof of (7.16).

Finally, we notice that the following is convex:

$$h(t) = \frac{\sum_{i\in V(H)} d_i^t}{n^{(1+t)/2}}\Big(\frac{1-p}{p}\Big)^{t/2}.$$

We will make use of the convexity of $h(t)$ and the sandwiching bound (7.16) to conclude the proof. When (7.14) holds, we have

$$\begin{aligned}
&\max_{1\le t\le D} \mathrm{Adv}(f_{\mathcal{K}_{1,t}}) \to \infty \\
&\quad\Rightarrow \max_{1\le t\le D} g(t) \to \infty && \text{by the bound (7.15),} \\
&\quad\Rightarrow \max_{1\le t\le D} \widehat{g}(t) \to \infty && \text{since } g(t) \sim \widehat{g}(t), \\
&\quad\Rightarrow \max_{1\le t\le D} h(t) \to \infty && \text{by the bound (7.16),} \\
&\quad\Rightarrow \max\{h(1), h(D)\} \to \infty && \text{by the convexity of } h(t), \\
&\quad\Rightarrow \max\{g(1), g(D)\} \to \infty && \text{again by (7.16) and } g(t) \sim \widehat{g}(t), \\
&\quad\Rightarrow \max\{\mathrm{Adv}(f_{\mathcal{K}_{1,1}}), \mathrm{Adv}(f_{\mathcal{K}_{1,D}})\} \to \infty && \text{again by (7.15).} \qquad (7.18)
\end{aligned}$$

We could similarly check that

$$\frac{\max_{1\le t\le D} \mathrm{Adv}(f_{\mathcal{K}_{1,t}})}{\max\{\mathrm{Adv}(f_{\mathcal{K}_{1,1}}), \mathrm{Adv}(f_{\mathcal{K}_{1,D}})\}} \le C_0 \cdot \frac{\max_{1\le t\le D} g(t)}{\max\{g(1), g(D)\}}$$

since $C_0 \le \mathrm{Adv}(f_{\mathcal{K}_{1,t}})/g(t) \le 1$,

$$\le \big(C_0 + o(1)\big) \frac{\max_{1\le t\le D} \widehat{g}(t)}{\max\{\widehat{g}(1), \widehat{g}(D)\}}$$

since $g(t) \sim \widehat{g}(t)$,

$$\le \big(C_0 + o(1)\big) \frac{\max_{1\le t\le D} h(t)}{\max\{(C_1 \cdot h(1) - C_2), (C_1 \cdot h(D) - C_2)\}}$$

by (7.16) and that $\max\{h(1), h(D)\} \to \infty$,

$$\le \frac{2C_0}{C_1} \cdot \frac{\max_{1\le t\le D} h(t)}{\max\{h(1), h(D)\}} \le \frac{2C_0}{C_1}$$

since $h(t)$ is convex. In other words, there exists a constant C_3 such that, if

$$\max_{1\le t\le D} \mathrm{Adv}(f_{\mathcal{K}_{1,t}}) \to \infty,$$

then for large enough n,

$$\max_{1\le t\le D} \mathrm{Adv}(f_{\mathcal{K}_{1,t}}) \le C_3 \cdot \max\{\mathrm{Adv}(f_{\mathcal{K}_{1,1}}), \mathrm{Adv}(f_{\mathcal{K}_{1,D}})\}. \qquad (7.19)$$

Moreover, notice that C_0 and C_1 depend on D only, and thus C_3 depends on D only. In conclusion, we showed that (7.14) implies (7.18) and (7.19). ∎

Proof of Proposition 7.2. Let $d_i = \deg_H(i)$ for $i \in V(H)$. By Lemma 7.7,

$$M_{\mathcal{S},H} = \sum_{i \in V(H)} (d_i)_{(t)}.$$

First, let us notice that if $S_1 \triangle S_2 = \varnothing$, then $S_1 = S_2 \cong \mathcal{S}$, and the desired bound easily follows, as in this situation the ratio of interest is bounded by

$$\begin{aligned}\frac{n^{|V(S_1 \cup S_2)|} M_{\varnothing,H}}{M_{\mathcal{S},H}^2 ((1-p)/p)^{|\mathcal{S}|}} &\le \big(1+o(1)\big) \frac{M_{\mathcal{S}}}{M_{\mathcal{S},H}^2 ((1-p)/p)^{|\mathcal{S}|}} \\ &= \big(1+o(1)\big) \frac{1}{|\operatorname{Aut}(\mathcal{S})| \cdot \operatorname{Adv}(f_{\mathcal{S}})^2} = o(1),\end{aligned}$$

where we use Proposition 7.9 in the second to last line and $\operatorname{Adv}(f_{\mathcal{S}}) \to \infty$ in the last line. So from now on let us assume $S_1 \triangle S_2 \neq \varnothing$. If $S_1, S_2 \cong \mathcal{K}_{1,t}$ have non-empty intersection and $S_1 \triangle S_2 \neq \varnothing$, there are two cases.

Case 1. In this case, S_1 and S_2 do not share the same root. Note that

$$|V(S_1 \cup S_2)| = |V(S_1 \triangle S_2)| \quad \text{and} \quad |S_1 \triangle S_2| = 2|\mathcal{S}| - 2$$

in this case. Consider the shapes $S_1' = S_1 \setminus S_2$ and $S_2' = S_2[V(S_2) \setminus V(S_1')]$, as illustrated in Figure 7.

Notice that S_1' and S_2' are vertex disjoint, and $S_1' \sqcup S_2'$ is a spanning sub-shape of $S_1 \triangle S_2$. By Claim 7.10 and Claim 7.11, we have

$$M_{S_1 \triangle S_2, H} \le M_{S_1' \sqcup S_2', H} \le M_{S_1', H} M_{S_2', H}.$$

Then, using $|V(S_1 \cup S_2)| = |V(S_1 \triangle S_2)|$, $|S_1 \triangle S_2| = 2|\mathcal{S}| - 2$, and that $M_{S_1 \triangle S_2, H} \le M_{S_1', H} M_{S_2', H}$, we get

$$\begin{aligned}&\frac{n^{|V(S_1 \cup S_2)| - |V(S_1 \triangle S_2)|} M_{S_1 \triangle S_2, H} ((1-p)/p)^{|S_1 \triangle S_2|/2}}{M_{\mathcal{S},H}^2 ((1-p)/p)^{|\mathcal{S}|}} \\ &\quad \le \frac{M_{S_1', H} M_{S_2', H}}{M_{\mathcal{S},H}^2 ((1-p)/p)} \\ &\quad \le \big(1+o(1)\big) \frac{\sqrt{(M_{S_1',H}^2 / M_{S_1'})((1-p)/p)^{|S_1'|} \cdot (M_{S_2',H}^2 / M_{S_2'})((1-p)/p)^{|S_2'|}}}{(M_{\mathcal{S},H}^2 / M_{\mathcal{S}})((1-p)/p)^{|\mathcal{S}|}} \\ &\qquad \cdot \frac{((1-p)/p)^{|\mathcal{S}| - (|S_1'| + |S_2'|)/2 - 1}}{n^{|V(\mathcal{S})| - (|V(S_1')| + |V(S_2')|)/2}},\end{aligned} \tag{7.20}$$

where the last line uses $M_S = (1 - o(1)) n^{|V(S)|}$ for constant sized shapes S.

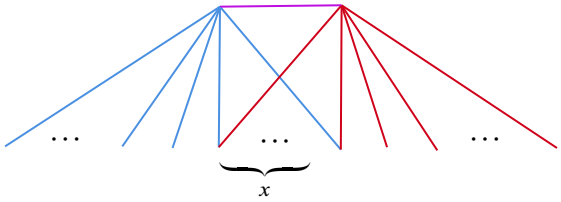

Figure 6. Case 1 for the intersecting pattern $i(S_1, S_2)$ of two t-stars with non-empty intersection. The edges contained solely in S_1, solely in S_2, and in the intersection of S_1 and S_2, are marked by blue, red, and purple respectively.

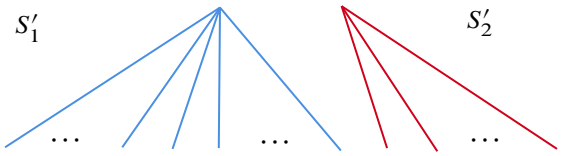

Figure 7. Vertex disjoint S_1' and S_2' whose union is a spanning sub-shape of $S_1 \triangle S_2$ in Figure 6.

Since for some constant C, $\mathcal{S}$ satisfies $\mathrm{Adv}(f_{\mathcal{S}'}) \le C \cdot \mathrm{Adv}(f_{\mathcal{S}})$ for all star shapes $\mathcal{S}' \cong \mathcal{K}_{1,t'}$ with at most $|\mathcal{S}|$ edges, by Proposition 7.9 and that the automorphism groups involved have bounded size for $|\mathcal{S}| \le D = O(1)$, we have that for some constant $C' > 0$,

$$\frac{(M^2_{\mathcal{S}',H}/M_{\mathcal{S}'})((1-p)/p)^{|\mathcal{S}'|}}{(M^2_{\mathcal{S},H}/M_{\mathcal{S}})((1-p)/p)^{|\mathcal{S}|}} \le C' \tag{7.21}$$

for all star shapes $\mathcal{S}' \cong \mathcal{K}_{1,t'}$ with at most $|\mathcal{S}|$ edges. As S_1', S_2' are also star shapes with at most $|\mathcal{S}|$ number of edges, we have

$$\frac{\sqrt{(M^2_{S_1',H}/M_{S_1'})((1-p)/p)^{|S_1'|} \cdot (M^2_{S_2',H}/M_{S_2'})((1-p)/p)^{|S_2'|}}}{(M^2_{\mathcal{S},H}/M_{\mathcal{S}})((1-p)/p)^{|\mathcal{S}|}} \le C'.$$

Also observe that as

$$|\mathcal{S}| - \frac{|S_1'| + |S_2'|}{2} - 1 \ge 0 \quad \text{and} \quad |V(\mathcal{S})| - \frac{|V(S_1')| + |V(S_2')|}{2} \ge 1,$$

we have

$$\frac{((1-p)/p)^{|\mathcal{S}|-(|S_1'|+|S_2'|)/2-1}}{n^{|V(\mathcal{S})|-(|V(S_1')|+|V(S_2')|)/2}} \lesssim \frac{1}{n},$$

where we use the fact that $(1-p)/p = O(1)$ as $p = \Omega(1)$, and $|\mathcal{S}| \le D = O(1)$.

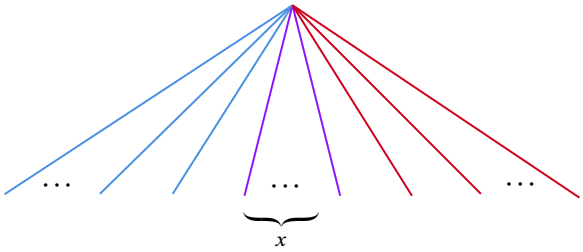

Figure 8. Case 2 for the intersecting pattern $i(S_1, S_2)$ of two t-stars with non-empty intersection. The edges contained solely in S_1, solely in S_2, and in the intersection of S_1 and S_2, are marked by blue, red, and purple, respectively.

Plugging these bounds back into (7.20), we get

$$\begin{aligned}
&\frac{n^{|V(S_1\cup S_2)|-|V(S_1\triangle S_2)|}M_{S_1\triangle S_2,H}((1-p)/p)^{|S_1\triangle S_2|/2}}{M_{\mathcal{S},H}^2((1-p)/p)^{|\mathcal{S}|}}\\
&\le \big(1+o(1)\big)\frac{\sqrt{(M_{S_1',H}^2/M_{S_1'})((1-p)/p)^{|S_1'|}\cdot(M_{S_2',H}^2/M_{S_2'})((1-p)/p)^{|S_2'|}}}{(M_{\mathcal{S},H}^2/M_{\mathcal{S}})((1-p)/p)^{|\mathcal{S}|}}\\
&\quad\cdot\frac{((1-p)/p)^{|\mathcal{S}|-(|S_1'|+|S_2'|)/2-1}}{n^{|V(\mathcal{S})|-(|V(S_1')|+|V(S_2')|)/2}}\\
&\lesssim \frac{C'}{n}=o(1).
\end{aligned}$$

Case 2. Under this case, star-shaped copies S_1 and S_2 share the same root vertex, as shown in the Figure 8. Let $x := |V(S_1 \cup S_2)| - |V(S_1 \triangle S_2)|$. Note that $0 < x < t$. Then, $S_1 \triangle S_2 \cong K_{1,2t-2x}$, and by Lemma 7.7, we have

$$M_{S_1\triangle S_2,H} = M_{K_{1,2t-2x},H} = \sum_{i\in V(H)}(d_i)_{(2t-2x)}.$$

If $x \ge t/2$, then $S_1 \triangle S_2 \cong K_{1,2t-2x}$ is a star graph with at most t edges, and therefore by the inequality (7.21), for some constant C', $\mathcal{S}$ satisfies

$$\frac{(M_{\mathcal{S}',H}^2/M_{\mathcal{S}'})((1-p)/p)^{|\mathcal{S}'|}}{(M_{\mathcal{S},H}^2/M_{\mathcal{S}})((1-p)/p)^{|\mathcal{S}|}}\le C'$$

for all star shapes $\mathcal{S}' \cong \mathcal{K}_{1,t'}$ with at most $|\mathcal{S}| = t$ edges, we have

$$\frac{M_{S_1\triangle S_2,H}^2}{M_{S_1\triangle S_2}}\Big(\frac{1-p}{p}\Big)^{|S_1\triangle S_2|}\le C'\cdot\frac{M_{\mathcal{S},H}^2}{M_{\mathcal{S}}}\Big(\frac{1-p}{p}\Big)^{|\mathcal{S}|}.$$

Rearranging the inequality, we get

$$\begin{aligned} M_{S_1 \triangle S_2, H} &\le \sqrt{C'} \cdot M_{\mathcal{S},H} \sqrt{\frac{M_{S_1 \triangle S_2}}{M_{\mathcal{S}}} \Big(\frac{1-p}{p}\Big)^{|\mathcal{S}|-|S_1 \triangle S_2|}} \\ &\le \big(\sqrt{C'} + o(1)\big) \cdot n^{(|V(S_1 \triangle S_2)|-|V(\mathcal{S})|)/2} \Big(\frac{1-p}{p}\Big)^{(|\mathcal{S}|-|S_1 \triangle S_2|)/2} M_{\mathcal{S},H} \\ &= \big(\sqrt{C'} + o(1)\big) \cdot n^{(t-2x)/2} \Big(\frac{1-p}{p}\Big)^{(2x-t)/2} M_{\mathcal{S},H}, \end{aligned}$$

and

$$\begin{aligned} &\frac{n^{|V(S_1 \cup S_2)|-|V(S_1 \triangle S_2)|} M_{S_1 \triangle S_2, H} ((1-p)/p)^{|S_1 \triangle S_2|/2}}{M_{\mathcal{S},H}^2 ((1-p)/p)^{|\mathcal{S}|}} \\ &\quad \le \big(\sqrt{C'} + o(1)\big) \cdot \frac{(n(p/(1-p)))^x (n(p/(1-p)))^{t/2-x}}{M_{\mathcal{S},H}} \\ &\quad \le \big(\sqrt{C'} + o(1)\big) \cdot \frac{(n(p/(1-p)))^{t/2}}{M_{\mathcal{S},H}} \le o(1), \end{aligned}$$

as

$$\begin{aligned} \frac{M_{\mathcal{S},H}^2}{(n(p/(1-p)))^t} &\ge \big(1 - o(1)\big) \cdot \frac{M_{\mathcal{S},H}^2}{M_{\mathcal{S}}} \Big(\frac{1-p}{p}\Big)^{|\mathcal{S}|} \\ &\ge \big(1 - o(1)\big) \cdot \mathrm{Adv}(f_{\mathcal{S}})^2 \cdot |\operatorname{Aut}(\mathcal{S})| \to \infty \end{aligned}$$

by our assumption.

Therefore, we now assume $x < t/2$. Since

$$\begin{aligned} \sum_{i \in V(H)} (d_i)_{(2t-2x)} &\ge \sum_{i \in V(H)} (d_i)_{(t)} \\ &= M_{\mathcal{S},H} \ge \omega\Big(n^{(1+t)/2} \Big(\frac{p}{1-p}\Big)^{t/2}\Big) = \omega(n), \end{aligned}$$

we may apply Lemma 7.8 and calculate

$$\begin{aligned} &\frac{n^{|V(S_1 \cup S_2)|-|V(S_1 \triangle S_2)|} M_{S_1 \triangle S_2, H} ((1-p)/p)^{|S_1 \triangle S_2|/2}}{M_{\mathcal{S},H}^2 ((1-p)/p)^{|\mathcal{S}|}} \\ &\quad = \frac{(n(p/(1-p)))^x \sum_{i \in V(H)} (d_i)_{(2t-2x)}}{(\sum_{i \in V(H)} (d_i)_{(t)})^2} \\ &\quad \le \big(1 + o(1)\big) \cdot \frac{(n(p/(1-p)))^x \sum_{i \in V(H)} d_i^{2t-2x}}{(\sum_{i \in V(H)} d_i^t)^2}. \end{aligned}$$

Let $\Delta := \max_{i \in V(H)} d_i$. We have

$$\frac{(n(p/(1-p)))^x \sum_{i \in V(H)} d_i^{2t-2x}}{(\sum_{i \in V(H)} d_i^t)^2} \le \frac{(n(p/(1-p)))^x (\sum_{i \in V(H)} d_i^t)\Delta^{t-2x}}{(\sum_{i \in V(H)} d_i^t)^2} = \frac{(n(p/(1-p)))^x \Delta^{t-2x}}{\sum_{i \in V(H)} d_i^t}. \tag{7.22}$$

Recalling the assumption that the star-shape $\mathcal{S} \cong \mathcal{K}_{1,t}$ satisfies

$$\frac{M_{\mathcal{S},H}^2}{M_{\mathcal{S}}} \Big(\frac{1-p}{p}\Big)^{|\mathcal{S}|} = \mathrm{Adv}(f_{\mathcal{S}})^2 \cdot |\operatorname{Aut}(\mathcal{S})| \to \infty,$$

and using Lemma 7.8, we have

$$\omega(1) \le \frac{M_{\mathcal{S},H}^2}{M_{\mathcal{S}}} \Big(\frac{1-p}{p}\Big)^{|\mathcal{S}|} = \frac{(\sum_{i \in V(H)} (d_i)_{(t)})^2}{(n)_{(t+1)} (p(1-p))^t} \le \big(1 + o(1)\big) \cdot \frac{(\sum_{i \in V(H)} d_i^t)^2}{n^{t+1}(p/(1-p))^t}.$$

Thus, we obtain

$$\frac{1}{\sum_{i \in V(H)} d_i^t} \le o\bigg(\frac{1}{n^{(t+1)/2}(p/(1-p))^{t/2}}\bigg),$$

and plugging it back to (7.22), we get

$$\frac{(n(p/(1-p)))^x \sum_{i \in V(H)} d_i^{2t-2x}}{(\sum_{i \in V(H)} d_i^t)^2} \le \frac{(n(p/(1-p)))^x \Delta^{t-2x}}{\sum_{i \in V(H)} d_i^t} \le o\bigg(\frac{\Delta^{t-2x}}{n^{(t+1)/2-x}(p/(1-p))^{t/2-x}}\bigg). \tag{7.23}$$

On the other hand, notice that $\sum_{i \in V(H)} d_i^t \ge \Delta^t$, and therefore

$$\frac{(n(p/(1-p)))^x \sum_{i \in V(H)} d_i^{2t-2x}}{(\sum_{i \in V(H)} d_i^t)^2} \le \frac{(n(p/(1-p)))^x \Delta^{t-2x}}{\sum_{i \in V(H)} d_i^t} \le \frac{(n(p/(1-p)))^x}{\Delta^{2x}}. \tag{7.24}$$

Combining (7.23) and (7.24),

$$\frac{(n(p/(1-p)))^x \sum_{i \in V(H)} d_i^{2t-2x}}{(\sum_{i \in V(H)} d_i^t)^2} \le \min\bigg\{o\bigg(\frac{\Delta^{t-2x}}{n^{(t+1)/2-x}(p/(1-p))^{t/2-x}}\bigg), \frac{(n(p/(1-p)))^x}{\Delta^{2x}}\bigg\}. \tag{7.25}$$

If $\Delta \geq n^{1/2+1/(2t)}(p/(1-p))^{1/2}$, then

$$\frac{(n(p/(1-p)))^x}{\Delta^{2x}} \leq \frac{(n(p/(1-p))^x}{n^{x+x/t}(p/(1-p))^x} = \frac{1}{n^{x/t}} = o(1),$$

since $x > 0$ and $t \leq D = O(1)$. Otherwise, if $\Delta \leq n^{1/2+1/(2t)}(p/(1-p))^{1/2}$, we can bound

$$\begin{aligned} o\left(\frac{\Delta^{t-2x}}{n^{(t+1)/2-x}(p(1-p))^{t/2-x}}\right) &\leq o\left(\frac{n^{(t+1)/2-x-x/t}(p/(1-p))^{t/2-x}}{n^{(t+1)/2-x}(p/(1-p))^{t/2-x}}\right) \\ &= o\left(\frac{1}{n^{x/t}}\right) = o(1). \end{aligned}$$

Since in both regimes of Δ the expression in (7.25) is $o(1)$, we conclude that

$$\begin{aligned} &\frac{n^{|V(S_1\cup S_2)|-|V(S_1\triangle S_2)|}M_{S_1\triangle S_2,H}((1-p)/p)^{|S_1\triangle S_2|/2}}{M_{\mathcal{S},H}^2((1-p)/p)^{|\mathcal{S}|}} \\ &\quad\leq \big(1+o(1)\big)\cdot\frac{(n(p/(1-p)))^x\sum_{i\in V(H)}d_i^{2t-2x}}{(\sum_{i\in V(H)}d_i^t)^2} \leq o(1), \end{aligned}$$

which finishes the proof for Case 2.

Combining the case discussions, we obtain the result. ∎

8. Proof of the main corollary

Proof of Corollary 3.5. We prove both directions of the stated condition.

If $\limsup_{n\to\infty}\frac{\sum_{v\in V(H)}d_v^t}{n^{(1+t)/2}(p/(1-p))^{t/2}} < \infty$ for $t \in \{1, D\}$, then by Lemma 7.7, we have

$$\begin{aligned} &\limsup_{n\to\infty}\max\left\{\frac{M_{\mathcal{K}_{1,1},H}^2}{M_{\mathcal{K}_{1,1}}}\left(\frac{1-p}{p}\right)^1, \frac{M_{\mathcal{K}_{1,D},H}^2}{M_{\mathcal{K}_{1,D}}}\left(\frac{1-p}{p}\right)^D\right\} \\ &\quad= \limsup_{n\to\infty}\max_{t\in\{1,D\}}\left\{\big(1+o(1)\big)\left(\frac{\sum_{v\in V(H)}(d_v)_{(t)}}{n^{(1+t)/2}(p/(1-p))^{t/2}}\right)^2\right\} \\ &\quad\leq \limsup_{n\to\infty}\max_{t\in\{1,D\}}\left\{\big(1+o(1)\big)\left(\frac{\sum_{v\in V(H)}d_v^t}{n^{(1+t)/2}(p/(1-p))^{t/2}}\right)^2\right\} < \infty. \end{aligned}$$

Thus, by Proposition 7.9, we have $\max\{\mathrm{Adv}(f_{\mathcal{K}_{1,1}}), \mathrm{Adv}(f_{\mathcal{K}_{1,D}})\} = O(1)$ since the automorphism groups of $\mathcal{K}_{1,1}$ and $\mathcal{K}_{1,D}$ have bounded sizes. By the contrapositive of Proposition 7.1, the degree-D advantage in this case must be bounded away from infinity. By Theorem 3.1, degree-D polynomial tests do not achieve strong separation between $\mathbb{P}$ and $\mathbb{Q}$.

If $\lim_{n\to\infty} \max_{t\in\{1,D\}}\{\frac{\sum_{v\in V(H)} d_v^t}{n^{(1+t)/2}(p/(1-p))^{t/2}}\} = \infty$, by Lemma 7.7 and Lemma 7.8,

$$\begin{aligned}
&\lim_{n\to\infty} \max\left\{ \frac{M^2_{\mathcal{K}_{1,1},H}}{M_{\mathcal{K}_{1,1}}}\left(\frac{1-p}{p}\right)^1, \frac{M^2_{\mathcal{K}_{1,D},H}}{M_{\mathcal{K}_{1,D}}}\left(\frac{1-p}{p}\right)^D \right\} \\
&= \lim_{n\to\infty} \max_{t\in\{1,D\}} \left\{ \big(1+o(1)\big)\left(\frac{\sum_{v\in V(H)} (d_v)_{(t)}}{n^{(1+t)/2}(p/(1-p))^{t/2}} \right)^2 \right\} \\
&= \lim_{n\to\infty} \max_{t\in\{1,D\}} \left\{ \big(1-o(1)\big)\left(\frac{\sum_{v\in V(H)} d_v^t}{n^{(1+t)/2}(p/(1-p))^{t/2}} \right)^2 \right\} = \infty.
\end{aligned}$$

By Proposition 7.5, the square of the degree-D advantage can be written as

$$(\mathrm{Adv}^{\le D})^2 = \sum_{\mathcal{S}\in\mathbb{G}_{\le D}} \frac{M^2_{\mathcal{S},H}}{M_{\mathcal{S}}\cdot|\operatorname{Aut}(\mathcal{S})|}\left(\frac{1-p}{p}\right)^{|\mathcal{S}|}, \tag{8.1}$$

which clearly tends to infinity as $n\to\infty$, since for $t\in\{1,D\}$, $\mathcal{K}_{1,t}$ is one shape that contributes to the summation (8.1), its automorphism group has constant size, and all the terms are non-negative. If $t^*\in\{1,D\}$ satisfies

$$\frac{\max\left\{ \frac{\sum_{v\in V(H)} d_v}{n(p/(1-p))^{1/2}}, \frac{\sum_{v\in V(H)} d_v^D}{n^{(1+D)/2}(p/(1-p))^{D/2}} \right\}}{\frac{\sum_{v\in V(H)} d_v^{t^*}}{n^{(1+t^*)/2}(p/(1-p))^{t^*/2}}} = O(1), \tag{8.2}$$

then by Proposition 7.9 and Lemma 7.7, we also have

$$\begin{aligned}
&\frac{\max\{\mathrm{Adv}(f_{\mathcal{K}_{1,1}}), \mathrm{Adv}(f_{\mathcal{K}_{1,D}})\}}{\mathrm{Adv}(f_{\mathcal{K}_{1,t^*}})} \\
&\le \big(1+o(1)\big)\cdot \frac{\max\left\{ \frac{\sum_{v\in V(H)} d_v}{|\operatorname{Aut}(\mathcal{K}_{1,1})|^{1/2}\cdot n(p/(1-p))^{1/2}}, \frac{\sum_{v\in V(H)} (d_v)_{(D)}}{|\operatorname{Aut}(\mathcal{K}_{1,D})|^{1/2}\cdot n^{(1+D)/2}(p/(1-p))^{D/2}} \right\}}{\frac{\sum_{v\in V(H)} (d_v)_{(t^*)}}{n^{(1+t^*)/2}(p/(1-p))^{t^*/2}}} \\
&\lesssim \frac{\max\left\{ \frac{\sum_{v\in V(H)} d_v}{n(p/(1-p))^{1/2}}, \frac{\sum_{v\in V(H)} d_v^D}{n^{(1+D)/2}(p/(1-p))^{D/2}} \right\}}{C_1\cdot \frac{\sum_{v\in V(H)} d_v^{t^*}}{n^{(1+t^*)/2}(p/(1-p))^{t^*/2}} - C_2}
\end{aligned}$$

for some constants $C_1, C_2 > 0$ by the bound (7.16),

$$= O(1),$$

where in the last line we use that (8.2) together with the assumption

$$\max_{t\in\{1,D\}}\left\{ \frac{\sum_{v\in V(H)} d_v^t}{n^{(1+t)/2}(p/(1-p))^{t/2}} \right\} \to \infty$$

implies that $\frac{\sum_{v\in V(H)} d_v^{t^*}}{n^{(1+t^*)/2}(p/(1-p))^{t^*/2}} \to \infty$, and thus

$$C_1 \cdot \frac{\sum_{v\in V(H)} d_v^{t^*}}{n^{(1+t^*)/2}(p/(1-p))^{t^*/2}} - C_2 = \Theta\left(\frac{\sum_{v\in V(H)} d_v^{t^*}}{n^{(1+t^*)/2}(p/(1-p))^{t^*/2}}\right).$$

Finally, since $\mathrm{Adv}^{\le D} \to \infty$ and

$$\frac{\max\{\mathrm{Adv}(f_{\mathcal{K}_{1,1}}), \mathrm{Adv}(f_{\mathcal{K}_{1,D}})\}}{\mathrm{Adv}(f_{\mathcal{K}_{1,t^*}})} = O(1),$$

by Theorem 3.1, the polynomial $f_{\mathcal{K}_{1,t^*}}$ strongly separates $\mathbb{P}$ and $\mathbb{Q}$. ∎

9. Proof of the characterization theorem

Here we prove our characterization theorem of optimal tests based on the maximum degree.

Proof of Theorem 4.2. By Corollary 3.5, whenever strong separation can be achieved by degree-D polynomials for some constant D, an optimal test is

$$f_{\mathcal{K}_{1,t^*}} = \sum_{S\subseteq\binom{V}{2}: S\cong\mathcal{K}_{1,t^*}} \chi_S,$$

where $t^* \in \{1, D\}$ and satisfies

$$\frac{\max\left\{\frac{\sum_{v\in V(H)} d_v}{n(p/(1-p))^{1/2}}, \frac{\sum_{v\in V(H)} d_v^D}{n^{(1+D)/2}(p/(1-p))^{D/2}}\right\}}{\frac{\sum_{v\in V(H)} d_v^{t^*}}{n^{(1+t^*)/2}(p/(1-p))^{t^*/2}}} = O(1). \tag{9.1}$$

Now let us consider for which $t^* \in \{1, D\}$ the condition (9.1) is achieved.

If $\Delta \lesssim (n(p/(1-p)))^{1/2}$, then

$$\begin{aligned}\frac{\sum_{i\in V(H)} d_i^D}{n^{(1+D)/2}(p/(1-p))^{D/2}} &= \frac{1}{\sqrt{n}} \sum_{i\in V(H)} \left(\frac{d_i}{(n(p/(1-p)))^{1/2}}\right)^D \\ &\lesssim \frac{1}{\sqrt{n}} \sum_{i\in V(H)} \left(\frac{d_i}{(n(p/(1-p)))^{1/2}}\right) \\ &= \frac{\sum_{i\in V(H)} d_i}{n(p/(1-p))^{1/2}}.\end{aligned}$$

Therefore, $t^* = 1$ achieves the condition (9.1), and the signed count of edges is an optimal test in this case. Let us now address when the signed count of edges achieves

strong separation given that it is an optimal test. By Corollary 3.5, $f_{\mathcal{K}_{1,1}}$ achieves strong separation if and only if

$$\frac{\sum_{i\in V(H)} d_i}{n(p/(1-p))^{1/2}} \to \infty.$$

Since $\sum_{i\in V(H)} d_i = 2m$ where $m = |E(H)|$, we thus conclude that the signed count of edges achieves strong separation if and only if $m = \omega(n(p(1-p))^{1/2})$.

On the other hand, if $\Delta \geq (n(p/(1-p)))^{1/2+\varepsilon}$ for some constant $\varepsilon > 0$,

$$\begin{aligned}\frac{\sum_{i\in V(H)} d_i^D}{n^{(1+D)/2}(p/(1-p))^{D/2}} &\geq \frac{\Delta^D}{n^{(1+D)/2}(p/(1-p))^{D/2}} \\ &= \frac{1}{\sqrt{n}}\left(\frac{\Delta}{(n(p/(1-p)))^{1/2}}\right)^D \\ &\geq \frac{1}{\sqrt{n}}(n(p/(1-p)))^{\varepsilon D}.\end{aligned}$$

In particular, setting $D = \lceil 3/2\varepsilon \rceil$, we have

$$\begin{aligned}\frac{\sum_{i\in V(H)} d_i^D}{n^{(1+D)/2}(p/(1-p))^{D/2}} &\geq \frac{1}{\sqrt{n}}\left(n\frac{p}{1-p}\right)^{\varepsilon D} \geq \frac{1}{\sqrt{n}}\left(n\frac{p}{1-p}\right)^{3/2} \\ &= n\left(\frac{p}{1-p}\right)^{3/2} \gtrsim \frac{\sum_{i\in V(H)} d_i}{n(p/(1-p))^{1/2}},\end{aligned}$$

where in the last line we use that $\sum_{i\in V(H)} d_i \leq n(n-1)$ and $p = \Omega(1)$. Therefore, $t^* = D = \lceil 3/2\varepsilon \rceil$ achieves the condition (9.1), and the signed count of D-stars is an optimal test in this case. Moreover, from the second to last line above, we know

$$\frac{\sum_{i\in V(H)} d_i^D}{n^{(1+D)/2}(p/(1-p))^{D/2}} \to \infty.$$

By Corollary 3.5, we conclude that the signed count of D-stars, $f_{\mathcal{K}_{1,D}}$, achieves strong separation. ∎

Lemma 9.1. *Let a, b be non-negative integers with $a \geq b$. Then it holds that*

$$(a)_{(b)} = \frac{a!}{(a-b)!} \geq a^b \cdot \exp\big(-b^2/2(a-b+1)\big).$$

Proof. We compute

$$\log\left(\frac{a!}{(a-b)!}\right) = b\log a + \sum_{i=0}^{b-1} \log(1-i/a)$$

using the fact that $\log(1-x) \geq -x/(1-x)$ holds for any $x \in (0,1)$,

$$\begin{aligned} &\geq b \log a - \sum_{i=0}^{b-1} \frac{i}{a-i} \\ &\geq b \log a - \sum_{i=0}^{b-1} \frac{i}{a-b+1} \\ &\geq b \log a - \frac{b^2}{2(a-b+1)}. \end{aligned}$$

We conclude by taking exponential of both sides. ∎

10. Proof of tightness of the main theorem

In this section, we prove the tightness of our main theorem: if either $p = \Omega(1)$ or $D = O(1)$ is not satisfied, counting stars could fail to strong separate $\mathbb{P}$ and $\mathbb{Q}$ in the planted subgraph detection task while some other degree-D polynomial does so. Our proofs consider two natural planted subgraph detection settings:

- a constant-sized clique planted in a sparse $G(n,p)$,
- a clique of size $k = \Theta(\sqrt{n})$ planted in $G(n,1/2)$,

and show in each case that counting stars does not capture the strong separability of the problem with respect to degree-D polynomials, with $D = O(1)$ in the first setting and $D = O(\log n)$ in the second setting. The main effort in these proofs lies in careful second moment analysis which confirms strong separation is achieved by some natural degree-D polynomial (which of course is not achieved by counting stars) in each setting.

Proof of Lemma 5.1. For $p = n^{-\gamma}$ where $\gamma \in (0,1)$ is a constant, let $k = 4/\gamma$. Consider the planted subgraph detection task with a clique of size k planted in $G(n,p)$. We want to show under this setting, it holds that (1) counting stars fails to achieve strong separation and (2) there exists a constant degree polynomial test that strongly separates the two hypotheses.

For (1), we calculate the advantage of signed count of a star shape $\mathcal{K}_{1,t}$ using Proposition 7.9 and Lemma 7.7:

$$\begin{aligned} \left(\mathrm{Adv}(f_{\mathcal{K}_{1,t}})\right)^2 &= \frac{(\sum_{i\in V(H)} (d_i)_{(t)})^2}{|\operatorname{Aut}(\mathcal{K}_{1,t})| \cdot n^{1+t}} \left(\frac{1-p}{p}\right)^t \\ &\leq \frac{(\sum_{i\in V(H)} d_i^t)^2}{n^{1+t}} \left(\frac{1-p}{p}\right)^t \leq \frac{k^2}{n} \cdot \left(\frac{k^2}{n^{1-\gamma}}\right)^t. \end{aligned} \tag{10.1}$$

In particular, we plug in $k = 4/\gamma$ and can easily verify that $k^2 \le n^{1-\gamma} \le n$ holds. Therefore, the advantage (10.1) of counting a star shape $\mathcal{K}_{1,t}$ is $O(1)$ for any t, which implies that counting star fails to achieve strong separation.

For (2), consider the simple polynomial test f which counts the *unsigned* number of *unlabelled* copies of k-cliques in G, which can be expressed as

$$f(G) = \sum_{U \subseteq V : |U| = k} \mathbf{1}\{G[U] \text{ is a clique}\},$$

and corresponds to a degree-$\binom{k}{2}$ polynomial. We will show that f achieves strong separation for detecting a planted clique of size k in $G(n, p)$. To this end, we need to compute the first and the second moments of f under $\mathbb{P}$ and $\mathbb{Q}$. We observe that we always have $f \ge 1$ under $G \sim \mathbb{P}$, so $\mathbb{E}_{\mathbb{P}}[f] \ge 1$. Recall that $p = n^{-\gamma} = o(n^{-2/(k-1)})$ for our choice of k. Under $\mathbb{Q}$, we have

$$\begin{aligned}
\mathbb{E}_{\mathbb{Q}}[f] &= \sum_{U \subseteq V : |U| = k} \mathbb{E}_{\mathbb{Q}}[\mathbf{1}\{G[U] \text{ is a clique}\}] \\
&= \sum_{U \subseteq V : |U| = k} p^{\binom{k}{2}} \le n^k p^{\binom{k}{2}} = o(1), \\
\mathbb{E}_{\mathbb{Q}}[f^2] &= \sum_{U, U' \subseteq V : |U| = |U'| = k} \mathbb{E}_{\mathbb{Q}}[\mathbf{1}\{G[U] \text{ is clique}\}\mathbf{1}\{G[U'] \text{ is clique}\}] \\
&= \sum_{U, U' \subseteq V : |U| = |U'| = k} p^{2\binom{k}{2} - \binom{|U \cap U'|}{2}} \\
&= \sum_{i=0}^{k} \sum_{\substack{U, U' \subseteq V : \\ |U| = |U'| = k, \\ |U \cap U'| = i}} p^{2\binom{k}{2} - \binom{i}{2}} \\
&\le \sum_{i=0}^{k} \binom{n}{i} \binom{n}{k-i} \binom{n}{k-i} p^{k(k-1) - i(i-1)/2} \\
&\le \sum_{i=0}^{k} n^{2k-i} p^{k(k-1) - i(i-1)/2} \\
&= \sum_{i=0}^{k} \left(n^k p^{\binom{k}{2}}\right)^{2 - i/k} p^{i(k-i)/2} = o(1),
\end{aligned}$$

and we get $\mathrm{Var}_{\mathbb{Q}}[f] = o(1)$. Now we turn to the second moment of f under $\mathbb{P}$. We first notice that for any realization of the planted $\mathbf{H}$, we have

$$\mathbb{E}_{\mathbb{P}}[f^2] = \mathbb{E}_{\mathbb{P}}[f^2 | \mathbf{H}].$$

Thus, we may equivalently consider another distribution $\mathbb{P}'$ where we fix the planted k-clique to be on the first k vertices $\{1,\dots,k\} \subseteq V := [n]$. Let us denote this *fixed* set of vertices where the k-clique under $\mathbb{P}'$ is planted on as W. We may now compute

$$
\begin{aligned}
&\mathbb{E}_{\mathbb{P}}[f^2] = \mathbb{E}_{\mathbb{P}'}[f^2] \\
&= \sum_{\substack{U,U' \subseteq V: \\ |U|=|U'|=k}} \mathbb{E}_{\mathbb{P}'}[\mathbf{1}\{G[U] \text{ is clique}\}\mathbf{1}\{G[U'] \text{ is clique}\}] \\
&= \sum_{x=0}^{k} \sum_{y=0}^{k} \sum_{\substack{i \le \min\{k-x,k-y\}, \\ j \le \min\{x,y\}}} \\
&\qquad \sum_{\substack{U,U' \subseteq V: \\ |U|=|U'|=k, \\ |U\cap W|=x, |U'\cap W|=y, \\ |U\cap U'\setminus W|=i, |U\cap U'\cap W|=j}} \mathbb{E}_{\mathbb{P}'}[\mathbf{1}\{G[U] \text{ is clique}\}\mathbf{1}\{G[U'] \text{ is clique}\}] \\
&= 1 + 2 \sum_{0 \le x < k} \sum_{\substack{U \subseteq V: \\ |U|=k, \\ |U\cap W|=x}} \mathbb{E}_{\mathbb{P}'}[\mathbf{1}\{G[U] \text{ is clique}\}] \\
&\quad + \sum_{0 \le x < k} \sum_{0 \le y < k} \sum_{\substack{i \le \min\{k-x,k-y\}, \\ j \le \min\{x,y\}}} \\
&\qquad \sum_{\substack{U,U' \subseteq V: \\ |U|=|U'|=k, \\ |U\cap W|=x, |U'\cap W|=y, \\ |U\cap U'\setminus W|=i, |U\cap U'\cap W|=j}} \mathbb{E}_{\mathbb{P}'}[\mathbf{1}\{G[U] \text{ is clique}\}\mathbf{1}\{G[U'] \text{ is clique}\}] \\
&= 1 + 2 \sum_{0 \le x < k} \sum_{\substack{U \subseteq V: \\ |U|=k, \\ |U\cap W|=x}} p^{\binom{k}{2}-\binom{x}{2}} \\
&\quad + \sum_{0 \le x < k} \sum_{0 \le y < k} \sum_{\substack{i \le \min\{k-x,k-y\}, \\ j \le \min\{x,y\}}} \sum_{\substack{U,U' \subseteq V: \\ |U|=|U'|=k, \\ |U\cap W|=x, |U'\cap W|=y, \\ |U\cap U'\setminus W|=i, |U\cap U'\cap W|=j}} p^{2\binom{k}{2}-\binom{x}{2}-\binom{y}{2}-\binom{i}{2}-ij} \\
&\le 1 + 2 \sum_{0 \le x < k} k^x n^{k-x} p^{\binom{k}{2}-\binom{x}{2}} \\
&\quad + \sum_{0 \le x < k} \sum_{0 \le y < k} \sum_{\substack{0 \le i \le \min\{k-x,k-y\}, \\ 0 \le j \le \min\{x,y\}}} k^{x+y-j} n^i n^{k-x-i} n^{k-y-i} p^{2\binom{k}{2}-\binom{x}{2}-\binom{y}{2}-\binom{i}{2}-ij}
\end{aligned}
$$

$$
\begin{aligned}
&\le 1 + 2 \sum_{0 \le x < k} k^x n^{k-x} p^{\binom{k}{2}-\binom{x}{2}} \\
&\quad + 2 \sum_{0 \le x \le y < k} \sum_{\substack{0 \le i \le k-y, \\ 0 \le j \le x}} k^{x+y-j} n^{2k-x-y-i} p^{2\binom{k}{2}-\binom{x}{2}-\binom{y}{2}-\binom{i}{2}-ij} \\
&\le 1 + C_k \max_{0 \le x < k} n^{k-x} p^{\binom{k}{2}-\binom{x}{2}} \qquad (10.2) \\
&\quad + C_k \max_{\substack{0 \le x \le y < k, \\ 0 \le i \le k-y, \\ 0 \le j \le x}} n^{2k-x-y-i} p^{2\binom{k}{2}-\binom{x}{2}-\binom{y}{2}-\binom{i}{2}-ij}, \qquad (10.3)
\end{aligned}
$$

where C_k in the last inequality is a constant depending on the constant k. Let us separately examine the terms $n^{k-x} p^{\binom{k}{2}-\binom{x}{2}}$ and $n^{2k-x-y-i} p^{2\binom{k}{2}-\binom{x}{2}-\binom{y}{2}-\binom{i}{2}-ij}$ appearing in (10.2) and (10.3). Recall $p = n^{-\gamma} = o(n^{-2/(k-1)})$.

For any $0 \le x < k$, we have

$$
n^{k-x} p^{\binom{k}{2}-\binom{x}{2}} = \left(n^k p^{\binom{k}{2}}\right)^{1-x/k} p^{x(k-x)/2} = o(1).
$$

For any $0 \le x \le y < k$, $0 \le i \le k-y$, and $0 \le j \le x$, we have

$$
\begin{aligned}
&n^{2k-x-y-i} p^{2\binom{k}{2}-\binom{x}{2}-\binom{y}{2}-\binom{i}{2}-ij} \\
&\quad = \left(n^k p^{\binom{k}{2}}\right)^{2-(x+y+i)/k} p^{x(k-x)/2+y(k-y)/2+i(k-i-2j)/2}. \qquad (10.4)
\end{aligned}
$$

Note that the first exponent $2 - (x+y+i)/k$ is strictly positive as $x + y + i \le x + k < 2k$, and the second exponent satisfies

$$
\begin{aligned}
&\frac{x(k-x)}{2} + \frac{y(k-y)}{2} + \frac{i(k-i-2j)}{2} \\
&\quad \ge \frac{x(k-x)}{2} + \frac{y(k-y)}{2} + \frac{i(k-i-2x)}{2} \\
&\quad \ge \frac{x(k-x)}{2} + \frac{y(k-y)}{2} + \frac{i(k-(k-y)-2x)}{2} \\
&\quad \ge \frac{x(k-x)}{2} + \frac{y(k-y)}{2} + \frac{-ix}{2} \\
&\quad \ge \frac{x(k-x)}{2} + \frac{y(k-y)}{2} + \frac{-(k-y)x}{2} \\
&\quad \ge \frac{x(k-x)}{2} + \frac{(y-x)(k-y)}{2} \\
&\quad \ge 0,
\end{aligned}
$$

where we repeatedly apply $0 \le i \le k-y$, $0 \le j \le x$, and $0 \le x \le y < k$. Therefore, we conclude that the expression in (10.4) is $o(1)$. Now that we know both terms

in (10.2) and (10.3) are $o(1)$, we use these bounds and conclude that

$$\begin{aligned}\mathbb{E}_{\mathbb{P}}[f^2] - 1 &\le C_k \max_{0\le x<k} n^{k-x}\, p^{\binom{k}{2}-\binom{x}{2}} \\ &\quad + C_k \max_{\substack{0\le x\le y<k,\\ 0\le i\le k-y,\\ 0\le j\le x}} n^{2k-x-y-i}\, p^{2\binom{k}{2}-\binom{x}{2}-\binom{y}{2}-\binom{i}{2}-ij} \\ &\le o(1).\end{aligned}$$

Since moreover $\mathbb{E}_{\mathbb{P}}[f] \ge 1$, we know $\mathrm{Var}_{\mathbb{P}}[f] = o(1)$. As $\mathrm{Var}_{\mathbb{Q}}[f], \mathrm{Var}_{\mathbb{P}}[f] = o(1)$ and $|\mathbb{E}_{\mathbb{P}}[f] - \mathbb{E}_{\mathbb{Q}}[f]| \ge 1 - o(1)$, we conclude that f achieves strong separation. ∎

Proof of Lemma 5.2. Let $k = C\sqrt{n}$ where $C > 0$ is a constant. Consider the planted subgraph detection task with a clique of size k planted in $G(n, \frac{1}{2})$. We want to show under this setting, it holds that (1) counting stars fails to achieve strong separation and (2) there exists a degree-$O(\log n)$ polynomial test that strongly separates the two hypotheses.

For (1), let us consider the signed count polynomial $f_{\mathcal{S}}$ where $\mathcal{S}$ is a star shape. When $\mathcal{S}$ is a t-star for any $t \ge 2$, $|\operatorname{Aut}(\mathcal{S})| = t!$. By Proposition 7.9, we may calculate the advantage of counting t-stars for detecting a planted clique of size k in $G(n, 1/2)$ as

$$\begin{aligned}\big(\mathrm{Adv}(f_{\mathcal{S}})\big)^2 &= \frac{M^2_{\mathcal{S},K_k}}{|\operatorname{Aut}(\mathcal{S})| \cdot M_{\mathcal{S}}} \\ &= \big(1+o(1)\big) \cdot \frac{\left(\sum_{i\in V(K_k)} (d_i)_{(t)}\right)^2}{n^{1+t}\cdot t!} \\ &\le \big(1+o(1)\big) \cdot \frac{(k\cdot k^t)^2 \cdot e^t}{n^{1+t}\cdot t^t} \le \frac{k^2}{n}\left(\frac{e\cdot k^2}{t\cdot n}\right)^t,\end{aligned}$$

which is bounded by $O(1)$ for any $t \ge 2$ when $k = C\sqrt{n}$ for a constant C. It is also easy to check that counting edges does not achieve strong separation in this case. We thus conclude that counting stars fails to strongly separate $\mathbb{P}$ under this setting for any t.

Now we turn to prove (2). The seminal work of [1] proved that a spectral method successfully detects a planted k-clique with high probability when $k = C\sqrt{n}$ for a large enough constant $C > 0$, and it is known that such spectral method can be well approximated by degree-$O(\log n)$ polynomials (see [11, 20]). However, these results do not address the aspect of strong separation. Here, we will show that some degree-$O(\log n)$ polynomial does achieve strong separation when $k = C\sqrt{n}$ for a large enough constant $C > 0$, following a slight variant of the polynomial that approximates the trace method.

For convenience, in the following discussion we will use M to denote the $\{+1,-1\}$ adjacency matrix of G drawn from the null distribution $\mathbb{Q}$ or the planted distribution $\mathbb{P}$, where an entry (i, j) is 1 if $\{i, j\}$ is an edge in G, -1 if $\{i, j\}$ is not an edge, and 0 on the diagonal $i = j$. Let $l = B \log n$, where B is a large enough constant. Consider the following polynomial

$$f(M) = \sum_{\substack{(i_1,\ldots,i_l):\\ i_t \in [n], \text{ all distinct}}} M_{i_1,i_2} M_{i_2,i_3} \ldots M_{i_{l-1},i_l} M_{i_l,i_1}$$

of M, which is a degree-l polynomial in the entries of M. Moreover, we will call $\overline{i} = (i_1, \ldots, i_l)$ a (simple) closed path (of length l on l vertices), and use

$$M^{\overline{i}} := M_{i_1,i_2} M_{i_2,i_3} \ldots M_{i_{l-1},i_l} M_{i_l,i_1}$$

to denote this product that corresponds to the closed path $\overline{i}$.

Under the null distribution $\mathbb{Q}$, we have

$$\begin{aligned}
\mathbb{E}_{\mathbb{Q}}[f(M)] &= 0, \\
\mathbb{E}_{\mathbb{Q}}[f(M)^2] &= \sum_{\text{closed paths } \overline{i},\overline{j}} \mathbb{E}_{\mathbb{Q}}[M^{\overline{i}} M^{\overline{j}}] \\
&= \sum_{\text{closed paths } \overline{i},\overline{j}} \mathbf{1}\{E(\overline{i}) = E(\overline{j})\} \\
&= n(n-1)\ldots(n-l+1) \cdot 2l \\
&= \big(1 + o(1)\big) 2l n^l.
\end{aligned}$$

Let us now consider the planted distribution $\mathbb{P}$. Since the planted clique can be specified by a subset of vertices the clique is planted on, we will denote this subset of vertices as $\mathbf{W}$, where we use bold letter to emphasize it is uniformly random among all subsets of size k. Under $\mathbb{P}$, we have

$$\begin{aligned}
\mathbb{E}_{\mathbb{P}}[f(M)] &= \sum_{\text{closed path } \overline{i}} \mathbb{E}_{\mathbb{P}}[M^{\overline{i}}] \\
&= \sum_{\text{closed path } \overline{i}} \mathbb{E}_{\mathbb{P}}[\mathbf{1}\{i_1, \ldots, i_l \in \mathbf{W}\}],
\end{aligned}$$

since conditioning on the planted clique, $\mathbb{E}_{\mathbb{P}}[M^{\overline{i}} | \mathbf{W}] = 0$ whenever $\overline{i}$ is not fully contained in the planted clique, and 1 otherwise. We further observe that the probability that any fixed closed path $\overline{i}$ of length l is contained in $\mathbf{W}$ is simply

$$\frac{k(k-1)\ldots(k-l+1)}{n(n-1)\ldots(n-l+1)} = \big(1 + o(1)\big) \cdot \Big(\frac{k}{n}\Big)^l,$$

which can also be verified using Lemma 7.4. As a result,

$$\begin{aligned}\mathbb{E}_{\mathbb{P}}[f(M)] &= \sum_{\text{closed path } \bar{i}} \mathbb{E}_{\mathbb{P}}[\mathbf{1}\{i_1,\dots,i_l \in \mathbf{W}\}] \\ &= \sum_{\text{closed path } \bar{i}} \mathbb{P}(i_1,\dots,i_l \in \mathbf{W}) \\ &= \big(1+o(1)\big)n(n-1)\dots(n-l+1)\cdot\Big(\frac{k}{n}\Big)^l \\ &= \big(1+o(1)\big)k^l.\end{aligned}$$

Finally, we turn to the second moment of $f(M)$ under $\mathbb{P}$. We have

$$\begin{aligned}\mathbb{E}_{\mathbb{P}}[f(M)^2] &= \sum_{\text{closed paths } \bar{i},\bar{j}} \mathbb{E}_{\mathbb{P}}[M^{\bar{i}}M^{\bar{j}}] \\ &= \sum_{\text{closed paths } \bar{i},\bar{j}} \mathbb{E}_{\mathbb{P}}\big[\mathbf{1}\{V(\bar{i}\triangle\bar{j})\subseteq \mathbf{W}\}\big] \\ &= \big(1+o(1)\big)\sum_{\text{closed paths } \bar{i},\bar{j}} \Big(\frac{k}{n}\Big)^{|V(\bar{i}\triangle\bar{j})|},\end{aligned}$$

following similar reasoning as before, where $\bar{i}\triangle\bar{j}$ denotes the symmetric difference shape of the two closed paths $\bar{i}$ and $\bar{j}$, which in particular, as we recall from Definition 6.8, does not contain any isolated vertex. Note that for the case of vertex-disjoint $\bar{i}$ and $\bar{j}$, there are approximately n^{2l} many such terms in the sum, each contributing approximately $(k/n)^{2l}$, and thus the total contribution is close to k^{2l} which matches the square of the first moment of $f(M)$ under $\mathbb{P}$. If we can show that the rest of the contribution from the other terms corresponding to pairs of $\bar{i}$ and $\bar{j}$ that are not vertex-disjoint[7] is much smaller than k^{2l}, then we would be done.

Let us further break down the situation when $\bar{i}$ and $\bar{j}$ are not vertex-disjoint into cases. In the following definitions, the indices in the constraints are cyclic modulo l (e.g., $i_{t'-1}$ refers to i_l if $t'=1$). For a pair of closed paths $\bar{i}$ and $\bar{j}$, define

$$\begin{aligned}&A := \{t\in[l] : \forall t'\in[l], j_t \neq i_{t'}\},\\ &B := \{t\in[l] : \exists t'\in[l], \text{s.t. } j_t = i_{t'}, j_{t-1}\neq i_{t'-1}, j_{t+1}\neq i_{t'+1}\},\\ &S := \{\{j_t, j_{t+1}\}, t\in[l] : \exists t'\in[l], \text{s.t. } \{j_t,j_{t+1}\} = \{i_{t'}, i_{t'+1}\}\},\\ &a := |A|, \quad b := |B|, \quad s := |S|,\end{aligned}$$

[7] In fact, one can even follow the argument used in the proof of our main theorem to only focus on the pairs of $\bar{i}$ and $\bar{j}$ that are not edge-disjoint. Nevertheless, the argument we present here suffices for the proof.

where A is the index set of vertices in $\bar{j}$ not shared with $\bar{i}$, B is the index set of vertices in $\bar{j}$ shared with $\bar{i}$ that do not participate in any shared edges, and S is the set of edges in $\bar{j}$ shared with $\bar{i}$. We moreover observe that if we restrict the attention to the set S of shared edges, they form a number of connected components, and we denote this number as c. Formally,

$$\begin{aligned} C &:= \{t \in [l] : \{j_t, j_{t+1}\} \in S, \{j_{t-1}, j_t\} \notin S\}, \\ c &:= |C|, \\ CC &:= \{(j_t, j_{t+1}, \ldots, j_{t+s_t}) : t \in C, s_t \text{ is the maximum such that} \\ &\qquad \forall t \le r \le t + s_t - 1, \{j_r, j_{r+1}\} \in S\}, \end{aligned}$$

where C is the set of the starting indices in the closed path $\bar{j}$ of the connected components of the set S of shared edges (if $\bar{i}$ and $\bar{j}$ do not overlap completely), and CC is the collection of connected components (sub-paths of $\bar{j}$) of S. Let us state as a fact that whenever $s = |S| < l$, the parameters a, b, s, c satisfy the identity $l = a + b + s + c$, which is easy to verify.

Now, we claim that $|V(\bar{i} \triangle \bar{j})| = 2l - b - 2s$. To see this, let us consider which vertices in $V(\bar{i}) \cup V(\bar{j})$ belong to $V(\bar{i} \triangle \bar{j})$. If $s = |S| = l$, then the two closed walks overlap completely, and the claim obviously holds. So let us consider otherwise.

- First, all vertices in $V(\bar{i}) \cup V(\bar{j})$ that do not correspond to any shared vertex or participate in any shared edges belong to $V(\bar{i} \triangle \bar{j})$, and there are
$$2(l - b - (s + c)) = 2a$$
such vertices.
- Second, each pair of shared vertices from $\bar{i}$ and $\bar{j}$ that do not participate in shared edges (corresponding to vertices in $\bar{j}$ indexed by B) contributes exactly 1 vertex to $V(\bar{i} \triangle \bar{j})$, and there are b of them.
- Third, every connected component of shared edges in S contributes 2 vertices, since each connected component is a sub-path shared by $\bar{i}$ and $\bar{j}$, and only the two endpoints of the sub-path survive the symmetric difference operation. This gives a total of $2c$ vertices.

From the analysis above, we have $|V(\bar{i} \triangle \bar{j})| = 2a + b + 2c = 2l - b - 2s$.

Moreover, we may estimate the number of pairs of $\bar{i}$ and $\bar{j}$ with the specific choice of parameters a, b, s, c in the following way:

- Without loss of generality, we enumerate the number of $\bar{i}$ as
$$n(n-1)\ldots(n-l+1) \le n^l.$$
- Next, we enumerate the vertices indexed in A in the order they are traversed in $\bar{j}$, creating $\le n^a$ choices.

- Then, we enumerate the vertices in $\overline{j}$ indexed in B, by enumerating the indices t and the matching indices t' of $\overline{i}$, creating $\leq (l \times l)^b = l^{2b}$ choices.
- Finally, we enumerate the vertices that participate in the shared edges according to the connected components the shared edges form. For each of the c connected components, we enumerate the starting index t in $\overline{j}$, the matching index t' in $\overline{i}$, the size s_t of this connected component, and the direction (whether they follow the forward direction $j_t = j_{t'}$, $j_{t+1} = i_{t'+1}$, $j_{t+2} = i_{t'+2}, \dots$ or the backward direction $j_t = j_{t'}$, $j_{t+1} = i_{t'-1}$, $j_{t+2} = i_{t'-2}, \dots$) of this component of shared edges, creating $\leq (l \times l \times l \times 2)^c = (2l^3)^c$ choices.

It is not hard to see one can uniquely recover a pair of $\overline{i}$ and $\overline{j}$ using the information above. Thus, the total number of pairs of $\overline{i}$ and $\overline{j}$ with the specific choice of parameters a, b, s, c is bounded by

$$N(a,b,s,c) \leq n^l \cdot n^a \cdot l^{2b} \cdot (2l^3)^c = 2^c l^{2b+3c} n^{l+a}.$$

Now we are ready to prove a bound on the second moment of $f(M)$ under $\mathbb{P}$:

$$\begin{aligned}
\mathbb{E}_{\mathbb{P}}[f(M)^2] &= \big(1+o(1)\big) \sum_{\text{paths } \overline{i},\overline{j}} \Big(\frac{k}{n}\Big)^{|V(\overline{i} \triangle \overline{j})|} \\
&= \big(1+o(1)\big) \sum_{a,b,s,c} \sum_{\substack{\text{paths } \overline{i},\overline{j}: \\ \overline{i},\overline{j} \text{ satisfies the parameters } a,b,s,c}} \Big(\frac{k}{n}\Big)^{2l-b-2s} \\
&= \big(1+o(1)\big) \sum_{a,b,s,c} N(a,b,s,c) \Big(\frac{k}{n}\Big)^{2l-b-2s},
\end{aligned}$$

note that when $\overline{i}, \overline{j}$ are vertex-disjoint, the parameters are $(a = l, b = 0, s = 0, c = 0)$

$$\leq \big(1+o(1)\big) \Bigg[n^{2l} \cdot \Big(\frac{k}{n}\Big)^{2l} + \sum_{\substack{a,b,s,c: \\ a<l}} 2^c l^{2b+3c} n^{l+a} \Big(\frac{k}{n}\Big)^{2l-b-2s} \Bigg].$$

Let us examine one term in the sum above:

$$\begin{aligned}
2^c l^{2b+3c} n^{l+a} \Big(\frac{k}{n}\Big)^{2l-b-2s} &= 2^c l^{2b+3c} k^{2l-b-2s} n^{a+b+2s-l} \\
&= 2^c l^{2b+3c} k^{2l-b-2s} n^{s-c} \\
&= k^{2l} \Big(\frac{l^2}{k}\Big)^b \Big(\frac{2l^3}{n}\Big)^c \Big(\frac{n}{k^2}\Big)^s,
\end{aligned}$$

where we use the previously stated identity $l = a + b + s + c$ in the second to last line. Moreover, when $\overline{i}$ and $\overline{j}$ are not vertex-disjoint (i.e., $a < l$), either $b > 0$

or $c > 0$. In either situation, when $k = C\sqrt{n}$ for a large constant C and $l = \Theta(\log n)$, the expression $k^{2l}(l^2/k)^b(2l^3/n)^c(n/k^2)^s$ above is $o(k^{2l}/n^{0.49})$. Since there are at most $l^4 = o(n^{0.01})$ choices of parameters a, b, s, c that satisfies $a < l$, and each term in the sum contributes $o(k^{2l}/n^{0.49})$, we conclude that

$$\begin{aligned}\mathbb{E}_{\mathbb{P}}[f(M)^2] &\le \big(1+o(1)\big)\left[n^{2l}\cdot\left(\frac{k}{n}\right)^{2l} + \sum_{\substack{a,b,s,c:\\ a<l}} 2^c l^{2b+3c} n^{l+a}\left(\frac{k}{n}\right)^{2l-b-2s}\right]\\ &\le \big(1+o(1)\big)[k^{2l} + k^{2l}/n^{0.48}]\\ &\le \big(1+o(1)\big)k^{2l}.\end{aligned}$$

Together with the first moment under $\mathbb{P}$, we get that $\mathrm{Var}_{\mathbb{P}}[f(M)] = o(k^{2l})$. From the moment computation for $\mathbb{Q}$, we also have

$$\begin{aligned}\mathrm{Var}_{\mathbb{Q}}[f(M)] &= O(ln^l) = o(k^{2l}),\\ |\mathbb{E}_{\mathbb{P}}[f(M)] - \mathbb{E}_{\mathbb{Q}}[f(M)]| &= \big(1-o(1)\big)k^l.\end{aligned}$$

We have thus verified that $f(M)$ achieves strong separation for detecting a planted clique of size k in $G(n, 1/2)$. ∎

11. Conclusion and future directions

In this paper, we initiate the unified study of the computational thresholds in detecting *arbitrary* planted subgraph structures in $G(n, p)$. We give a complete characterization of the strong separation power of *constant* degree polynomials in the regime of $p = \Omega(1)$. In particular, we reveal that under these assumptions, it is always optimal to count stars among all constant-degree polynomials.

Our work suggests many future directions.

(1) At a conceptual level, we believe our results make a strong case that studying "unified" planted random graph models, containing as special cases multiple well-studied models, is very beneficial. In our case, it was the generality of the studied model (Definition 1.1) that allowed us to reveal the (perhaps surprising) constant-degree optimality of the star counts. To the best of our knowledge, this optimality has not been observed before for any specific case of planted H. It is interesting what other common structural computational properties are satisfied by all planted subgraph models.

(2) In terms of specific directions, we consider a very interesting project to study what the optimal degree-D polynomial is when either $D = \omega(1)$ or $p = o(1)$ that our established star-count optimality fails (see Section 5). Moreover, our proof techniques

carefully leverage the graph structure and do not trivially extend to hypergraphs. It is a nice question whether and how these phenomena generalized to planted hypergraph settings.

(3) The recent work [23] suggests the constant-degree optimality of counting trees (equivalently of Approximate Message Passing) in terms of *recovering* a hidden spike in the spiked Wigner model (with a "dense" prior). While no trivial connection with our work is possible (we study the *detection version* between Bernoulli graph models), it is an exciting direction to study possible connections between the star-optimality from our work and the tree-optimality from [23].

(4) It is also interesting to explore whether similar unified phenomena hold in planted random hypergraph or even more general settings.

A. Proof of auxiliary lemmas

Proof of Proposition 7.5. Recall that the fact that

$$\chi_S(X) = \prod_{\{i,j\}\in E(S)} \frac{X_{i,j} - p}{\sqrt{p(1-p)}}$$

for $S \subseteq \binom{V}{2} : |S| \le D$ form an orthonormal basis of the multilinear polynomials in $\mathbb{R}[X]_{\le D}$ with respect to $\langle \cdot, \cdot \rangle_{\mathbb{Q}}$. Thus, we may expand any polynomial $f \in \mathbb{R}[X]_{\le D}$ in this basis as

$$f(X) = \sum_{S\subseteq\binom{V}{2}:|S|\le D} \langle f, \chi_S\rangle_{\mathbb{Q}} \chi_S = \sum_{S\subseteq\binom{V}{2}:|S|\le D} \hat{f}_S \chi_S,$$

where $\hat{f}_S := \langle f, \chi_S\rangle_{\mathbb{Q}}$ are the Fourier coefficients of f. Thus, we may compute

$$\begin{aligned}
\mathrm{Adv}^{\le D} &= \max_{f\in\mathbb{R}[X]_{\le D}} \frac{\mathbb{E}_{\mathbb{P}}[f]}{\sqrt{\mathbb{E}_{\mathbb{Q}}[f^2]}} = \max_{\substack{\{\hat{f}_S\} \\ S\subseteq\binom{V}{2}:|S|\le D}} \frac{\mathbb{E}_{\mathbb{P}}[\sum_S \hat{f}_S \chi_S]}{\sqrt{\mathbb{E}_{\mathbb{Q}}[(\sum_S \hat{f}_S \chi_S)^2]}} \\
&= \max_{\substack{\{\hat{f}_S\} \\ S\subseteq\binom{V}{2}:|S|\le D}} \frac{\sum_S \hat{f}_S \mathbb{E}_{\mathbb{P}}[\chi_S]}{\sqrt{\sum_{S,S'} \hat{f}_S \hat{f}_{S'} \mathbb{E}_{\mathbb{Q}}[\chi_S \chi_{S'}]}} \\
&= \max_{\substack{\{\hat{f}_S\} \\ S\subseteq\binom{V}{2}:|S|\le D}} \frac{\sum_S \hat{f}_S \mathbb{E}_{\mathbb{P}}[\chi_S]}{\sqrt{\sum_S \hat{f}_S^2}} = \sqrt{\sum_{S\subseteq\binom{V}{2}:|S|\le D} \mathbb{E}_{\mathbb{P}}[\chi_S]^2}.
\end{aligned}$$

We may compute the expectation $\mathbb{E}_{\mathbb{P}}[\chi_S]$ by first conditioning on $\mathbf{H}$ and then taking the expectation over random $\mathbf{H}$:

$$\mathbb{E}_{\mathbb{P}}[\chi_S] = \mathbb{E}_{\mathbf{H}}\mathbb{E}_{\mathbb{P}}[\chi_S|\mathbf{H}].$$

Conditioning on a fixed $\mathbf{H}$, $\mathbb{E}_{\mathbb{P}}[\chi_S|\mathbf{H}] = 0$ whenever S is not fully contained in $\mathbf{H}$, and $\mathbb{E}_{\mathbb{P}}[\chi_S|\mathbf{H}] = ((1-p)/p)^{|S|/2}$ if $S \subseteq \mathbf{H}$. Therefore, we have

$$\mathbb{E}_{\mathbb{P}}[\chi_S|\mathbf{H}] = \mathbf{1}(S \subseteq \mathbf{H})\Big(\frac{1-p}{p}\Big)^{|S|/2}.$$

Thus,

$$\begin{aligned}\mathbb{E}_{\mathbb{P}}[\chi_S] &= \mathbb{E}_{\mathbf{H}}\mathbf{1}(S \subseteq \mathbf{H})\Big(\frac{1-p}{p}\Big)^{|S|/2} \\ &= \mathbb{P}(S \subseteq \mathbf{H})\Big(\frac{1-p}{p}\Big)^{|S|/2} \\ &= \frac{M_{S,H}}{M_S}\Big(\frac{1-p}{p}\Big)^{|S|/2}, \end{aligned} \tag{A.1}$$

where the last line uses Lemma 7.4. Plugging this expression back in (A), we get

$$(\mathrm{Adv}^{\le D})^2 = \sum_{S \subseteq \binom{V}{2} : |S| \le D} \frac{M_{S,H}^2}{M_S^2}\Big(\frac{1-p}{p}\Big)^{|S|}.$$

Finally, we group the summation over subsets $S \subseteq \binom{V}{2} : |S| \le D$ according to the isomorphism classes of the shapes $\mathcal{S}$. For each shape $\mathcal{S}$, there are $M_{\mathcal{S}}/|\operatorname{Aut}(\mathcal{S})|$ subsets of $\binom{V}{2}$ (unlabelled copies) that are isomorphic to $\mathcal{S}$, and for isomorphic subsets, the expressions in the summation are equal. As a result, we may rewrite the summation as

$$\begin{aligned}(\mathrm{Adv}^{\le D})^2 &= \sum_{\mathcal{S} \in \mathbb{G}_{\le D}} \frac{M_{\mathcal{S}}}{|\operatorname{Aut}(\mathcal{S})|} \cdot \frac{M_{\mathcal{S},H}^2}{M_{\mathcal{S}}^2}\Big(\frac{1-p}{p}\Big)^{|\mathcal{S}|} \\ &= \sum_{\mathcal{S} \in \mathbb{G}_{\le D}} \frac{M_{\mathcal{S},H}^2}{M_{\mathcal{S}} \cdot |\operatorname{Aut}(\mathcal{S})|}\Big(\frac{1-p}{p}\Big)^{|\mathcal{S}|}. \end{aligned}$$

■

Proof of Lemma 7.7. For the enumeration of $M_{\mathcal{S},H}$, we first fix vertex $i \in V(H)$ and count the number of copies of $\mathcal{K}_{1,t}$ that have its root at i. Suppose the degree of i is d_i, then the number of copies of $\mathcal{K}_{1,t}$ rooted at i is $(d_i)_{(t)}$. The total number of copies of $\mathcal{K}_{1,t}$ in H is obtained by summing over all the vertex $i \in V(H)$:

$$M_{\mathcal{S},H} = \sum_{i \in V(H)} (d_i)_{(t)}.$$

■

Proof of Lemma 7.8. Clearly, one side of the inequality is obvious

$$\sum_{i\in[k]} (d_i)_{(t)} \le \sum_{i\in[k]} d_i^t,$$

so we focus on the other inequality.

We have for all $i \in [k]$,

$$\begin{aligned}(d_i)_{(t)} &\ge d_i^t\Big(1-\frac{t}{d_i}\Big)^t 1(d_i \ge t)\\ &\ge d_i^t\Big(1-\frac{t^2}{d_i}\Big)\mathbf{1}\{d_i \ge t\} \ge d_i^t - t^t - t^2 d_i^{t-1}.\end{aligned}$$

Hence,

$$\frac{\sum_{i\in[k]}(d_i)_{(t)}}{\sum_{i\in[k]} d_i^t} \ge 1 - \frac{k\,t^t}{\sum_{i\in[k]} d_i^t} - t^2\frac{\sum_{i\in[k]} d_i^{t-1}}{\sum_{i\in[k]} d_i^t}. \tag{A.2}$$

Clearly, by our assumption,

$$\frac{k\,t^t}{\sum_{i\in[k]} d_i^t} = o(1). \tag{A.3}$$

Also by the Hölder inequality,

$$\begin{aligned}t^2\frac{\sum_{i\in[k]} d_i^{t-1}}{\sum_{i\in[k]} d_i^t} &\le t^2\frac{(\sum_{i\in[k]} d_i^t)^{1-1/t}k^{1/t}}{\sum_{i\in[k]} d_i^t}\\ &= t^2\Bigg(\frac{k}{\sum_{i\in[k]} d_i^t}\Bigg)^{1/t} = o(1).\end{aligned} \tag{A.4}$$

Inserting (A.3) and (A.4) into (A.2), we obtain

$$\frac{\sum_{i\in[k]}(d_i)_{(t)}}{\sum_{i\in[k]} d_i^t} \ge 1 - o(1),$$

as desired. ∎

Proof of Proposition 7.9. Recall that the Walsh–Fourier basis χ_S form an orthonormal basis with respect to $\mathbb{Q}$. Thus,

$$\begin{aligned}\mathbb{E}_{\mathbb{Q}}[\chi_S] &= 0 \quad \text{if } S \ne \varnothing,\\ \mathbb{E}_{\mathbb{Q}}[\chi_S^2] &= 1,\\ \mathbb{E}_{\mathbb{Q}}[\chi_S\chi'_S] &= 0 \quad \text{if } S \ne S',\end{aligned}$$

and we may compute the first and the second moments of $f_{\mathcal{S}}$ under $\mathbb{Q}$ as

$$\begin{aligned}\mathbb{E}_{\mathbb{Q}}[f_{\mathcal{S}}] &= \sum_{S\subseteq\binom{V}{2}:S\cong\mathcal{S}} \mathbb{E}_{\mathbb{Q}}[\chi_S] = 0,\\ \mathbb{E}_{\mathbb{Q}}[f_{\mathcal{S}}^2] &= \sum_{S,S'\subseteq\binom{V}{2}:S,S'\cong\mathcal{S}} \mathbb{E}_{\mathbb{Q}}[\chi_S\chi_{S'}]\\ &= \sum_{S\subseteq\binom{V}{2}:S\cong\mathcal{S}} \mathbb{E}_{\mathbb{Q}}[\chi_S^2] = \frac{M_{\mathcal{S}}}{|\operatorname{Aut}(\mathcal{S})|}.\end{aligned}$$

Finally, we calculate the first moment of $f_{\mathcal{S}}$ under $\mathbb{P}$. Using the expectation of χ_S under $\mathbb{P}$ in (A.1), it holds that

$$\begin{aligned}\mathbb{E}_{\mathbb{P}}[f_{\mathcal{S}}] &= \sum_{S\subseteq\binom{V}{2}:S\cong\mathcal{S}} \mathbb{E}_{\mathbb{P}}[\chi_S]\\ &= \sum_{S\subseteq\binom{V}{2}:S\cong\mathcal{S}} \frac{M_{S,H}}{M_S}\Big(\frac{1-p}{p}\Big)^{|S|/2} = \frac{M_{\mathcal{S},H}}{|\operatorname{Aut}(\mathcal{S})|}\Big(\frac{1-p}{p}\Big)^{|\mathcal{S}|/2}.\end{aligned}$$

■

The advantage of $f_{\mathcal{S}}$ follows from the definition

$$\operatorname{Adv}(f_{\mathcal{S}}) = \frac{\mathbb{E}_{\mathbb{P}}[f_{\mathcal{S}}]}{\sqrt{\mathbb{E}_{\mathbb{Q}}[f_{\mathcal{S}}^2]}} = \frac{M_{\mathcal{S},H}}{M_{\mathcal{S}}^{1/2}\cdot|\operatorname{Aut}(\mathcal{S})|^{1/2}}\Big(\frac{1-p}{p}\Big)^{|\mathcal{S}|/2}.$$

B. Deferred lemmas and proof of lemmas in applications and counterexamples

Lemma B.1. *Let $k \le n$ and $0 < q \le 1$ be such that $kq = \omega(\log k)$. With probability $1 - o(1)$, for $H \sim G(k,q)$, its maximum degree is $(1 \pm o(1))\cdot(k-1)q$, and its number of edges is $(1 \pm o(1))\cdot\binom{k}{2}q$.*

Proof of Lemma B.1. Let $X \sim \text{Binomial}(k-1, q)$, where k, q are given as in the statement. We use the multiplicative Chernoff inequality to obtain the tail bound for X as follows:

$$\begin{aligned}\mathbb{P}\big[|X - \mathbb{E}[X]| \ge \delta\mathbb{E}[X]\big] &= \mathbb{P}\big[|X-(k-1)q| \ge \delta\cdot(k-1)q\big]\\ &\le 2\exp\Big(-\frac{\delta^2(k-1)q}{3}\Big),\end{aligned}$$

where $0 < \delta < 1$ is a parameter to be determined.

Now let $H \sim G(k,q)$, and denote the degree of a vertex $i \in V(H)$ as d_i. Note $d_i \sim \text{Binomial}(k-1,q)$ for every $i \in V(H)$. Using union bound, we compute an upper bound on the probability that the maximum degree of H is too large as

$$\mathbb{P}\Big[\max_{i\in V(H)} d_i \geq (1+\delta)\cdot(k-1)q\Big] \leq k\cdot\mathbb{P}\big[X \geq (1+\delta)\cdot(k-1)q\big] \leq 2k\cdot\exp\left(-\frac{\delta^2(k-1)q}{3}\right). \tag{B.1}$$

In the same way, we can also bound the probability that the minimum degree is too small:

$$\mathbb{P}\Big[\min_{i\in V(H)} d_i \leq (1-\delta)\cdot(k-1)q\Big] \leq 2k\cdot\exp\left(-\frac{\delta^2(k-1)q}{3}\right). \tag{B.2}$$

Since $kq \geq \omega(\log k)$, we may set $\delta^2 = C(\log k)/kq$ for a large enough constant C that ensures both bounds (B.1) and (B.2) are $o(1)$. This implies, with high probability, that all degrees d_i of $H \sim G(k,q)$ is concentrated in the range $(1\pm o(1))\cdot(k-1)q$, and consequently, the maximum degree of $H \sim G(k,q)$ is $(1\pm o(1))\cdot(k-1)q$ and the number of edges is $(1\pm o(1))\cdot\binom{k}{2}\cdot q$. ∎

C. Explanation of Remark 3.2

Let us consider the advantage of the signed count of a triangle $\mathcal{K}_4$. We claim that it diverges to infinity in the setting of Remark 3.2, where in the planted model $\mathbb{P}$ we have $G(n^\gamma, n^{-\alpha})$ planted in a $G(n, n^{-\beta})$ with $\alpha = 5/16, \gamma = 1/4, \beta = 1$.

Indeed, we may verify

$$\mathrm{Adv}(f_{\mathcal{K}_4})^2 = \frac{\mathbb{E}_{\mathbb{P}}[f_{\mathcal{K}_4}]^2}{\mathbb{E}_{\mathbb{Q}}[f_{\mathcal{K}_4}^2]} \geq \frac{\big(\sum_{S\subseteq\binom{n}{2}:S\cong\mathcal{K}_4}\mathbb{E}_{\mathbb{P}}[\chi_S]\big)^2}{n^{|V(\mathcal{K}_4)|}/|\operatorname{Aut}(\mathcal{K}_4)|},$$

by Proposition 7.9, we have

$$\geq \big(1-o(1)\big)\cdot\frac{n^{|V(\mathcal{K}_4)|}}{|\operatorname{Aut}(\mathcal{K}_4)|}\cdot\mathbb{E}_{\mathbb{P}}[\chi_S]^2$$

for an arbitrary $S \subseteq \binom{[n]}{2}$ such that $S \cong \mathcal{K}_4$,

$$\geq \frac{1}{C}\cdot n^4\left(\mathbb{P}(S\subseteq H)\Big(\frac{1-n^{-\beta}}{n^{-\beta}}\Big)^{|S|/2}\right)^2$$
$$\geq \frac{1}{C'}\cdot n^4\left(\Big(\frac{n^\gamma}{n}\Big)^4 (n^{-\alpha})^{\binom{4}{2}}(n^{\beta})^{\binom{4}{2}/2}\right)^2$$

$$\geq \frac{1}{C'}\frac{n^{(\beta-2\alpha)\binom{4}{2}}}{n^{(1-2\gamma)4}}$$
$$= \frac{1}{C'}\frac{n^{9/4}}{n^2} = \Omega(n^{1/4}) = \omega(1),$$

where C, C' are universal constants.

Acknowledgments. The authors are thankful to Alex Wein and Tim Kunisky for useful comments.

Funding. X. Y. was partially supported by a Simons Investigator award from the Simons Foundation to Daniel Spielman. P. Z. is supported by a Yale University Fund to Amin Karbasi.

References

[1] N. Alon, M. Krivelevich, and B. Sudakov, Finding a large hidden clique in a random graph. *Random Structures Algorithms* **13** (1998), no. 3-4, 457–466 Zbl 0959.05082 MR 1662795

[2] E. Arias-Castro and N. Verzelen, Community detection in dense random networks. *Ann. Statist.* **42** (2014), no. 3, 940–969 Zbl 1305.62035 MR 3210992

[3] V. Bagaria, J. Ding, D. Tse, Y. Wu, and J. Xu, Hidden Hamiltonian cycle recovery via linear programming. *Oper. Res.* **68** (2020), no. 1, 53–70 Zbl 1445.90007 MR 4059492

[4] B. Barak, S. Hopkins, J. Kelner, P. K. Kothari, A. Moitra, and A. Potechin, A nearly tight sum-of-squares lower bound for the planted clique problem. *SIAM J. Comput.* **48** (2019), no. 2, 687–735 Zbl 1421.68056 MR 3945259

[5] Q. Berthet and P. Rigollet, Optimal detection of sparse principal components in high dimension. *Ann. Statist.* **41** (2013), no. 4, 1780–1815 Zbl 1277.62155 MR 3127849

[6] M. Brennan, G. Bresler, and W. Huleihel, Reducibility and computational lower bounds for problems with planted sparse structure. In *Proceedings of the 31st Conference On Learning Theory*, pp. 48–166, Proceedings of Machine Learning Research 75, PMLR, 2018

[7] M. Chertkov, L. Kroc, F. Krzakala, M. Vergassola, and L. Zdeborová, Inference in particle tracking experiments by passing messages between images. *Proc. Natl. Acad. Sci. USA* **107** (2010), no. 17, 7663–7668

[8] A. Coja-Oghlan, O. Gebhard, M. Hahn-Klimroth, A. S. Wein, and I. Zadik, Statistical and computational phase transitions in group testing. In *Proceedings of the 35th Conference on Learning Theory*, pp. 4764–4781, Proceedings of Machine Learning Research 178, PMLR, 2022

[9] A. Dhawan, C. Mao, and A. S. Wein, Detection of dense subhypergraphs by low-degree polynomials. *Random Structures Algorithms* **66** (2025), no. 1, article no. e21279 Zbl 1557.05112 MR 4853082

[10] J. Ding, Y. Wu, J. Xu, and D. Yang, The planted matching problem: Sharp threshold and infinite-order phase transition. *Probab. Theory Related Fields* **187** (2023), no. 1-2, 1–71 Zbl 1540.60016 MR 4634336

[11] D. Gamarnik, A. Jagannath, and A. S. Wein, Low-degree hardness of random optimization problems. In *2020 IEEE 61st Annual Symposium on Foundations of Computer Science*, pp. 131–140, IEEE Computer Soc., Los Alamitos, CA, 2020 Zbl 1475.68016 MR 4232029

[12] D. Gamarnik and I. Zadik, Sparse high-dimensional linear regression. Estimating squared error and a phase transition. *Ann. Statist.* **50** (2022), no. 2, 880–903 Zbl 1486.62200 MR 4404922

[13] B. Hajek, Y. Wu, and J. Xu, Computational lower bounds for community detection on random graphs. In *Proceedings of the 28th Conference on Learning Theory*, pp. 899–928, Proceedings of Machine Learning Research 40, PMLR, 2015

[14] S. Hopkins, Statistical inference and the sum of squares method. PhD thesis, Cornell University, 2018

[15] W. Huleihel, Inferring hidden structures in random graphs. *IEEE Trans. Signal Inform. Process. Netw.* **8** (2022), 855–867 MR 4497466

[16] M. Jerrum, Large cliques elude the Metropolis process. *Random Structures Algorithms* **3** (1992), no. 4, 347–359 Zbl 0754.60018 MR 1179827

[17] C. Jones, A. Potechin, G. Rajendran, M. Tulsiani, and J. Xu, Sum-of-squares lower bounds for sparse independent set. In *2021 IEEE 62nd Annual Symposium on Foundations of Computer Science*, pp. 406–416, IEEE Computer Soc., Los Alamitos, CA, 2022 Zbl 1502.68015 MR 4399701

[18] J. Kahn and G. Kalai, Thresholds and expectation thresholds. *Combin. Probab. Comput.* **16** (2007), no. 3, 495–502 Zbl 1118.05093 MR 2312440

[19] D. Kunisky and J. Niles-Weed, Strong recovery of geometric planted matchings. In *Proceedings of the 2022 Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pp. 834–876, SIAM, Philadelphia, PA, 2022 Zbl 07883617 MR 4415073

[20] D. Kunisky, A. S. Wein, and A. S. Bandeira, Notes on computational hardness of hypothesis testing: Predictions using the low-degree likelihood ratio. In *Mathematical analysis, its applications and computation*, pp. 1–50, Springer Proc. Math. Stat. 385, Springer, Cham, 2022 Zbl 07632760 MR 4461037

[21] L. Massoulié, L. Stephan, and D. Towsley, Planting trees in graphs, and finding them back. In *Proceedings of the 32nd Conference on Learning Theory*, pp. 2341–2371, Proceedings of Machine Learning Research 99, PMLR, 2019

[22] M. Moharrami, C. Moore, and J. Xu, The planted matching problem: Phase transitions and exact results. *Ann. Appl. Probab.* **31** (2021), no. 6, 2663–2720 Zbl 1485.90115 MR 4350971

[23] A. Montanari and A. S. Wein, Equivalence of approximate message passing and low-degree polynomials in rank-one matrix estimation. *Probab. Theory Related Fields* **191** (2025), no. 1-2, 181–233 Zbl 07990323 MR 4869255

[24] R. Montgomery, Spanning trees in random graphs. *Adv. Math.* **356** (2019), article no. 106793 Zbl 1421.05080 MR 3998769

[25] E. Mossel, J. Niles-Weed, Y. Sohn, N. Sun, and I. Zadik, Sharp thresholds in inference of planted subgraphs. In *Proceedings of the 36th Conference on Learning Theory*, pp. 5573–5577, Proceedings of Machine Learning Research 195, PMLR, 2023

[26] E. Mossel, J. Niles-Weed, N. Sun, and I. Zadik, A second moment proof of the spread lemma. 2022, arXiv:2209.11347

[27] E. Mossel, J. Niles-Weed, N. Sun, and I. Zadik, On the second Kahn–Kalai conjecture. 2022, arXiv:2209.03326v1

[28] J. Park and H. T. Pham, A proof of the Kahn–Kalai conjecture. *J. Amer. Math. Soc.* **37** (2024), no. 1, 235–243 Zbl 1522.05442 MR 4654612

[29] L. Pósa, Hamiltonian circuits in random graphs. *Discrete Math.* **14** (1976), no. 4, 359–364 Zbl 0322.05127 MR 0389666

[30] A. Rotenberg, W. Huleihel, and O. Shayevitz, Planted bipartite graph detection. *IEEE Trans. Inform. Theory* **70** (2024), no. 6, 4319–4334 Zbl 1555.05210 MR 4751648

[31] T. Schramm and A. S. Wein, Computational barriers to estimation from low-degree polynomials. *Ann. Statist.* **50** (2022), no. 3, 1833–1858 Zbl 1539.62183 MR 4441142

[32] N. Verzelen and E. Arias-Castro, Community detection in sparse random networks. *Ann. Appl. Probab.* **25** (2015), no. 6, 3465–3510 Zbl 1326.05145 MR 3404642

[33] A. S. Wein, Optimal low-degree hardness of maximum independent set. *Math. Stat. Learn.* **4** (2021), no. 3-4, 221–251 Zbl 1530.68205 MR 4383734

Received 9 October 2024; revised 30 April 2025.

Xifan Yu
Department of Computer Science, Yale University, 51 Prospect Street, New Haven, CT 06511, USA; xifan.yu@yale.edu

Ilias Zadik
Department of Statistics and Data Science, Yale University, 219 Prospect Street, New Haven, CT 06511, USA; ilias.zadik@yale.edu

Peiyuan Zhang
Department of Electrical and Computer Engineering, Yale University, 17 Hillhouse Avenue, New Haven, CT 06511, USA; peiyuan.zhang@yale.edu