PTF Testing Lower Bounds for Planted Subgraphs

September 20, 2026

Abstract

Let Γ_n be a deterministic simple graph on at most n vertices. Embed it uniformly into $G(n, 1/2)$ and force its edges to be present, without prescribing any nonedges. We prove that all polynomial threshold tests of degree at most $c_\kappa \log n / \log\log n$ have vanishing distinguishing advantage if $2|E(\Gamma_n)| \leq n^{1-2\kappa}$ and the maximum degree of Γ_n is at most $n^{1/2-\kappa}$, for a fixed $0 < \kappa < 1/2$. One may take $c_\kappa = \min\{\kappa, 1/100\}$. A more general theorem assumes only bounds on embeddings of small connected graphs, or equivalently of small trees; a degree-moment criterion is also proved. The argument gives a uniform planting-coupling estimate and then handles arbitrary coefficients and arbitrarily small threshold margins by an explicit restriction tree. All combinatorial and restriction arguments are included. The only analytic theorem invoked without proof is the Gaussian Carbery–Wright inequality.

Note. This work was fully AI generated using GPT 6-Astra and the Codex harness, prompted by Zhangsong Li without any mathematical input.

1 Model, conditions, and main results

1.1 Monotone edge planting

All graphs in this note are finite, undirected, and simple. Let $\Gamma = \Gamma_n$ be a deterministic graph with $|V(\Gamma)| \leq n$, and put

$$m = |E(\Gamma)|, \qquad \Delta_\Gamma = \max_{x \in V(\Gamma)} \deg_\Gamma(x), \qquad E_n = \binom{[n]}{2}.$$

The maximum degree of an edgeless graph is understood to be zero. Under the null law $\mathbb{Q}_n$, the adjacency coordinates $(A_e)_{e \in E_n}$ are independent Bernoulli variables with parameter $1/2$.

Under the planted law $\mathbb{P}_{\Gamma,n}$, choose a uniformly random injection $\pi : V(\Gamma) \hookrightarrow [n]$. Set $A_e = 1$ for every $e \in \pi(E(\Gamma))$ and sample all other coordinates independently with parameter $1/2$. Thus this is *monotone edge planting*, not induced planting: a nonedge of Γ is not forced to be absent. Adding isolated vertices to the template does not change the distribution.

For an integer $d \geq 0$, define

$$\mathrm{Adv}_{n,\Gamma}(d) = \sup_{\deg f \leq d} |\mathbb{P}_{\Gamma,n}(f(A) > 0) - \mathbb{Q}_n(f(A) > 0)|.$$

The supremum is over all real polynomials of the adjacency entries. In particular, coefficients may depend arbitrarily on n and Γ, and no coefficient, symmetry, or margin assumption is imposed. All logarithms are natural.

Theorem 1.1 (Edge-count and maximum-degree criterion)**.** *Fix $0 < \kappa < 1/2$ and let $c_\kappa = \min\{\kappa, 1/100\}$. Suppose that, for all sufficiently large n,*

$$2|E(\Gamma_n)| \leq n^{1-2\kappa}, \qquad \Delta_{\Gamma_n} \leq n^{1/2-\kappa}. \tag{C}$$

If the nonnegative integers d_n satisfy

$$d_n \le c_\kappa \frac{\log n}{\log\log n},$$

then $\mathrm{Adv}_{n,\Gamma_n}(d_n) \to 0$. *The convergence is uniform over all templates satisfying* (C) *and all test polynomials in the displayed degree range.*

1.2 A small-pattern embedding condition

For a graph H, define

$$\mathrm{inj}(H,\Gamma) = \#\left\{\phi : V(H) \hookrightarrow V(\Gamma) : \{\phi(u),\phi(v)\} \in E(\Gamma) \text{ for every } \{u,v\} \in E(H)\right\}.$$

These are injective edge-preserving maps; they are not divided by automorphisms. Nonedges need not be preserved. The notation $\hom(H,\Gamma)$ denotes the analogous number of edge-preserving maps without the injectivity requirement.

Theorem 1.2 (Embedding-count criterion)**.** *Fix $0 < \kappa < 1/2$ and $c_\kappa = \min\{\kappa, 1/100\}$. Let $d = d_n \ge 1$ satisfy $d \le c_\kappa \log n/\log\log n$. Suppose that, for all sufficiently large n,*

$$\mathrm{inj}(H,\Gamma_n) \le n^{(1/2-\kappa)|V(H)|} \quad \textit{for every connected } H \textit{ with } 1 \le |E(H)| \le 2d. \tag{E}$$

Then $\mathrm{Adv}_{n,\Gamma_n}(d_n) \to 0$, *uniformly over the indicated templates and polynomials. It is equivalent to require* (E) *only for trees with between one and $2d$ edges.*

Corollary 1.3 (Degree-moment criterion)**.** *The conclusion of Theorem 1.2 holds if, instead of* (E), *one assumes*

$$\sum_{x \in V(\Gamma_n)} \deg_{\Gamma_n}(x)^t \le n^{(t+1)(1/2-\kappa)} \qquad \textit{for every integer } 1 \le t \le 2d_n. \tag{M}$$

The cases $d_n = 0$ and $E(\Gamma_n) = \varnothing$ have zero advantage and require no proof. They may be removed from any sequence when establishing convergence. We prove the implication

$$\text{(C)} \implies \text{(M)} \implies \text{(E)}$$

next, as well as a direct proof that (C) implies (E). The proof of Theorem 1.2 will then complete all three results.

1.3 Proof strategy

In Rademacher coordinates $X_e = 2A_e - 1$, planting independently chooses an edge set $I = \pi(E(\Gamma))$ and overwrites its coordinates by $+1$. The embedding assumption bounds $\Pr(U \subseteq I)$ for every small edge set U. A weighted Fourier-matrix argument converts those bounds into

$$\mathbb{E}[(f(Y) - f(X))^2] \le n^{-4\kappa+o(1)}\,\mathrm{Var}(f(X)).$$

This estimate controls polynomial values, but not signs near zero. We therefore construct a coordinate decision tree whose actual restricted polynomial is, at most leaves, either regular or dominated by its constant term. A regular leaf is handled by a proved Boolean small-ball estimate; a mean-dominated leaf is handled by the coupling and Chebyshev's inequality. The depth is $\exp(O_\eta(d \log d))$, and the planting avoids the queried coordinates with high probability in the stated degree range. The conditioning at a leaf is described explicitly; no conditioning on the planting avoiding that leaf is used.

1.4 Related work and scope

The distinction between polynomial-value separation and polynomial threshold testing is essential here. In the same monotone planted-subgraph model, Yu, Zadik, and Zhang [5] characterize constant-degree polynomial *strong separation*, in which the difference of means dominates both standard deviations, and show that signed star counts are optimal for that criterion. Their degree-profile characterization motivates comparison with (M). The conclusion proved here is different: throughout our sufficient parameter range, even the acceptance-probability advantage of an arbitrary polynomial threshold test tends to zero, and the permitted degree grows with n. We do not claim a complete characterization of PTF testing for all templates.

Hsieh, Kane, Kothari, Li, Mohanty, and Tiegel [4] obtain rigorous consequences of low-degree likelihood-ratio bounds for noisy distributions, including fully permutation-symmetric Boolean vectors and specified classes of Gaussian polynomial statistics. Those results do not directly give the theorem here for unsmoothed Boolean graph data and arbitrary test polynomials. In particular, vertex-permutation invariance of a random graph is weaker than invariance under all permutations of its edge coordinates. Diakonikolas, Kane, Liu, and Pittas [3] establish PTF testing lower bounds for non-Gaussian component analysis and related problems; their discussion identifies planted clique as a separate direction not covered by their results.

The regularity framework and Gaussian anti-concentration input used below are established tools [2, 1]. The role of the present argument is to combine an explicit graph-mask coupling estimate with a transfer statement for the *actual* restricted polynomial, retaining uniformity in all coefficients and threshold margins. This is a model-specific sufficient theorem, not a general implication from low-degree moment indistinguishability to PTF indistinguishability, and not a lower bound against all polynomial-time algorithms.

2 From graph structure to small-pattern containment

2.1 The graph-theoretic implications

Lemma 2.1 (Connected embeddings from edges and degrees)**.** *For every connected graph H with $v \geq 2$ vertices and every template Γ with $m \geq 1$,*

$$\mathrm{inj}(H, \Gamma) \leq 2m\, \Delta_\Gamma^{v-2}.$$

Consequently, (C) *implies* (E) *for every degree cutoff d.*

Proof. Choose a spanning tree of H, distinguish an edge of that tree, and order all remaining vertices so that the parent of each precedes it. There are $2m$ choices for the ordered image of the distinguished edge. Each subsequent vertex has at most Δ_Γ possible images adjacent to its already mapped parent. Ignoring the remaining edges of H and ignoring collisions can only enlarge the count. This proves the bound. Under (C),

$$2m\, \Delta_\Gamma^{v-2} \leq n^{1-2\kappa} n^{(1/2-\kappa)(v-2)} = n^{(1/2-\kappa)v}.$$

For an edgeless template, $\mathrm{inj}(H, \Gamma) = 0$, so (E) holds directly. □

If the embedding bound holds for all connected H in (E), it holds for the stated trees. Conversely, choose a spanning tree T of a connected H with at most $2d$ edges. It has the same vertex set, between one and $2d$ edges, and $\mathrm{inj}(H, \Gamma) \leq \mathrm{inj}(T, \Gamma)$. This proves the equivalence asserted in Theorem 1.2.

Lemma 2.2 (Tree homomorphisms from degree moments)**.** *If T is a tree with $t \geq 1$ edges, then*

$$\mathrm{hom}(T, \Gamma) \leq \sum_{x \in V(\Gamma)} \deg_\Gamma(x)^t. \tag{2.1}$$

Thus (M) *implies* (E)*, and* (C) *implies* (M)*.*

Proof. An edgeless template gives zero on both sides. Suppose $m > 0$, and write $d_\Gamma(x) = \deg_\Gamma(x)$ for this proof. If $t = 1$, both sides of (2.1) equal $2m$.

Suppose $t \geq 2$. Root T at a vertex o. Draw its image Z_o from the distribution

$$\nu(x) = \frac{d_\Gamma(x)}{2m}.$$

For each child, independently conditional on all previously sampled vertices, draw its image uniformly among the neighbors of its parent's image. This constructs a random homomorphism Z from T into Γ. All sampled images have positive degree. Moreover, every Z_u has marginal law ν: if its parent has law ν, then

$$\Pr(Z_u = y) = \sum_{x:\{x,y\}\in E(\Gamma)} \frac{d_\Gamma(x)}{2m} \frac{1}{d_\Gamma(x)} = \frac{d_\Gamma(y)}{2m}.$$

For $u \in V(T)$, put $a_u = \deg_T(u) - 1$. These numbers are nonnegative and sum to $t - 1$. For a fixed homomorphism ϕ, multiplication of the sampling probabilities gives

$$\Pr(Z = \phi) = \frac{1}{2m} \prod_{u\in V(T)} d_\Gamma(\phi(u))^{-a_u}.$$

At the root, its initial degree factor cancels one of its child-choice denominators; at any other vertex, its number of children is a_u. Therefore

$$\hom(T, \Gamma) = 2m\, \mathbb{E} \prod_{u\in V(T)} d_\Gamma(Z_u)^{a_u}.$$

Hölder's inequality, with exponents $(t - 1)/a_u$ for the positive a_u, gives

$$\begin{aligned} \mathbb{E} \prod_u d_\Gamma(Z_u)^{a_u} &\leq \prod_{u:a_u>0} \left(\mathbb{E} d_\Gamma(Z_u)^{t-1}\right)^{a_u/(t-1)} \\ &= \frac{1}{2m} \sum_x d_\Gamma(x)^t. \end{aligned}$$

Independence of the different Z_u is not needed for this application of Hölder. This proves (2.1).

For a connected H, take a spanning tree T with $t = |V(H)| - 1$ edges. If $1 \leq |E(H)| \leq 2d$, then $1 \leq t \leq 2d$, and

$$\operatorname{inj}(H, \Gamma) \leq \hom(T, \Gamma) \leq \sum_x d_\Gamma(x)^t \leq n^{(t+1)(1/2-\kappa)} = n^{(1/2-\kappa)|V(H)|}.$$

Finally,

$$\sum_x d_\Gamma(x)^t \leq \Delta_\Gamma^{t-1} \sum_x d_\Gamma(x) = 2m\, \Delta_\Gamma^{t-1} \leq n^{(t+1)(1/2-\kappa)}$$

under (C). If Γ is edgeless, the same conclusion follows directly, without any convention involving 0^0. $\square$

2.2 The planting mask and its containment probabilities

Let $X = (X_e)_{e\in E_n}$ be independent uniform signs. Independently choose the uniform injection π, set $I = \pi(E(\Gamma))$, and put

$$Y_e = \begin{cases} 1, & e \in I, \\ X_e, & e \notin I. \end{cases} \tag{2.2}$$

Thus X has the null law in Rademacher coordinates and Y has the planted law. The affine change from adjacency entries to signs preserves the degree bound; replacing X_e^2 by 1 multilinearizes any polynomial without changing its values on either input.

For an edge set $U \subseteq E_n$, let $V(U)$ be its incident vertex set, $v(U) = |V(U)|$, and let $H_U = (V(U), U)$. In particular, H_U has no isolated vertices. Set $v(\varnothing) = 0$, and define $(n)_v = n(n-1)\cdots(n-v+1)$, with $(n)_0 = 1$.

Lemma 2.3 (Containment identity and bound)**.** *For every nonempty $U \subseteq E_n$,*

$$\Pr(U \subseteq I) = \frac{\operatorname{inj}(H_U, \Gamma)}{(n)_{v(U)}}. \tag{2.3}$$

If (E) *holds and $n \geq 8d$, then, for $r = 2n^{-\kappa}$,*

$$\Pr(U \subseteq I) \leq \left(\frac{r}{\sqrt{n}}\right)^{v(U)} \qquad (1 \leq |U| \leq 2d). \tag{2.4}$$

Under the same embedding assumption, for every edge $e \in E_n$,

$$\alpha := \Pr(e \in I) = \frac{2m}{n(n-1)} \leq \frac{n^{-1-2\kappa}}{1 - 1/n}. \tag{2.5}$$

Proof. Pad Γ with isolated vertices to obtain a graph on $[n]$, and apply a uniformly random permutation. Its restriction to the original vertices is a uniform injection, so the resulting planted mask has the prescribed law. The inverse images of the specified vertices $V(U)$ form a uniform injection into $[n]$, with $(n)_{v(U)}$ possible values. Since H_U has no isolated vertices, no edge-preserving map can use an added isolated vertex. Thus exactly $\operatorname{inj}(H_U, \Gamma)$ of these injections preserve the edges of U, proving (2.3).

Let $H_1, \ldots, H_\ell$ be the connected components of H_U. The restriction of an injection on H_U gives an injection on each component; allowing images from distinct components to intersect only enlarges the number. Hence

$$\operatorname{inj}(H_U, \Gamma) \leq \prod_{j=1}^{\ell} \operatorname{inj}(H_j, \Gamma).$$

For $1 \leq |U| \leq 2d$, each component has between one and $2d$ edges. By (E), the last product is at most $n^{(1/2-\kappa)v(U)}$. Also $v(U) \leq 2|U| \leq 4d \leq n/2$, so every factor of $(n)_{v(U)}$ is at least $n/2$. Substituting in (2.3) gives

$$\Pr(U \subseteq I) \leq \frac{n^{(1/2-\kappa)v(U)}}{(n/2)^{v(U)}} = \left(2n^{-1/2-\kappa}\right)^{v(U)},$$

which is (2.4).

Finally, $\operatorname{inj}(K_2, \Gamma) = 2m$. Apply (2.3) to a single edge, and use (E) for $H = K_2$, to obtain (2.5). $\square$

3 Fourier preliminaries and exact restriction identities

We first give notation and two elementary ingredients on a generic cube $\{-1,1\}^M$. A multilinear polynomial has a unique expansion

$$f(x) = \sum_{S \subseteq [M]} a_S \chi_S(x), \qquad \chi_S(x) = \prod_{i \in S} x_i,$$

where $a_S = 0$ for $|S| > d$. Empty products equal one. Orthogonality gives

$$\mathbb{E}f(X) = a_\varnothing, \qquad \mathbb{E}f(X)^2 = \sum_S a_S^2, \qquad \mathrm{Var}(f(X)) = \sum_{S\neq\varnothing} a_S^2. \tag{3.1}$$

These identities also hold for independent coordinates each of which is either a Rademacher or a standard Gaussian: distinct multilinear monomials are orthogonal, and every such monomial has squared norm one.

Define the coordinate and total influences by

$$\mathrm{Inf}_i(f) = \sum_{S\ni i} a_S^2, \qquad \mathrm{Inf}(f) = \sum_i \mathrm{Inf}_i(f) = \sum_{S\neq\varnothing} |S| a_S^2.$$

Consequently,

$$\mathrm{Var}(f) \le \mathrm{Inf}(f) \le d\,\mathrm{Var}(f). \tag{3.2}$$

A nonconstant polynomial is *τ-regular* if

$$\max_i \mathrm{Inf}_i(f) \le \tau\,\mathrm{Inf}(f).$$

This is a condition on the polynomial representation itself, not just its threshold function.

3.1 A fourth-moment bound, including mixed inputs

Lemma 3.1. *Let $Z_1, \dots, Z_m$ be independent, each either Rademacher or standard Gaussian. For every multilinear $g(z) = \sum_S b_S \chi_S(z)$,*

$$\|g(Z)\|_4^2 \le \sum_S 3^{|S|} b_S^2.$$

In particular, if $\deg g \le d$, *then*

$$\|g(Z)\|_4 \le 3^{d/2} \|g(Z)\|_2. \tag{3.3}$$

Proof. Induct on the number of coordinates. For zero coordinates the assertion is an equality. Write $g(z) = a(z_1, \dots, z_{m-1}) + z_m b(z_1, \dots, z_{m-1})$. Both types of coordinate have first and third moments zero, second moment one, and fourth moment at most three. Conditional on the other coordinates,

$$\mathbb{E}_{Z_m} g(Z)^4 \le a^4 + 6a^2b^2 + 3b^4 \le (a^2 + 3b^2)^2.$$

Taking expectation over the remaining coordinates and applying the triangle inequality in L^2 gives

$$\|g\|_4^2 \le \|a^2 + 3b^2\|_2 \le \|a\|_4^2 + 3\|b\|_4^2.$$

Apply the induction hypothesis to a and b. The extra factor three for b accounts exactly for the presence of coordinate m in its monomials. This proves the weighted bound. Finally, $3^{|S|} \le 3^d$ and (3.1) imply (3.3). □

We will also use the elementary Paley–Zygmund estimate

$$\Pr(W \ge \tfrac{1}{2}\mathbb{E}W) \ge \frac{(\mathbb{E}W)^2}{4\mathbb{E}W^2}, \qquad W \ge 0, \quad \mathbb{E}W > 0. \tag{3.4}$$

Indeed, the expectation of W on this event is at least $\mathbb{E}W/2$, and Cauchy–Schwarz bounds that expectation by $(\mathbb{E}W^2)^{1/2}$ times the square root of the event probability.

3.2 Exact identities under coordinate restrictions

Let $H \subseteq [M]$ be a set of queried coordinates, and assign them a uniform vector $\rho \in \{-1,1\}^H$. Write $p_\rho(x) = f(\rho, x)$ for the actual restricted polynomial on H^c. Its coefficient of χ_T, for $T \subseteq H^c$, is

$$b_T(\rho) = \sum_{R \subseteq H} a_{R \cup T} \chi_R(\rho). \tag{3.5}$$

Thus $\deg b_T \le d - |T|$ whenever it is nonzero, and

$$\mathbb{E}_\rho b_T(\rho)^2 = \sum_{R \subseteq H} a_{R \cup T}^2.$$

For $j \notin H$, summing this identity over $T \ni j$ yields

$$\mathbb{E}_\rho \operatorname{Inf}_j(p_\rho) = \operatorname{Inf}_j(f). \tag{3.6}$$

Writing $\mu(\rho) = \mathbb{E}[p_\rho]$ and $V(\rho) = \operatorname{Var}(p_\rho)$, we also obtain

$$\mathbb{E}_\rho V(\rho) = \sum_{S : S \cap H^c \neq \varnothing} a_S^2 \le \sum_{j \notin H} \operatorname{Inf}_j(f), \tag{3.7}$$

$$\mathbb{E}_\rho \mu(\rho)^2 + \mathbb{E}_\rho V(\rho) = \mathbb{E} f(X)^2. \tag{3.8}$$

All means and variances at a fixed restriction are with respect to independent uniform values of the unqueried coordinates.

4 A uniform coupling bound for general planted masks

In this section I may be *any* random subset of E_n independent of X. No symmetry or clique assumption on its law is made. We assume only the containment estimate specified below and use the overwriting rule (2.2).

Lemma 4.1 (Elementary graph count). *For all integers $d \ge 1$ and $v \ge 1$, the number of graphs on a specified v-vertex set with at most $2d$ edges is at most $e^{2v\sqrt{d}}$.*

Proof. Put $M = \binom{v}{2}$. If $v \le 4\sqrt{d}$, count all graphs:

$$2^M \le \exp((\log 2) v^2/2) \le e^{2v\sqrt{d}}.$$

If $v > 4\sqrt{d}$, then $2d < M/2$. For $1 \le \ell \le M/2$ put $z = \ell/M$. Since $z \le 1/2$, all binomial probability weights $z^j(1-z)^{M-j}$ with $j \le \ell$ are at least $z^\ell(1-z)^{M-\ell}$. Thus

$$\sum_{j=0}^{\ell} \binom{M}{j} \le z^{-\ell}(1-z)^{-(M-\ell)} \le (eM/\ell)^\ell,$$

using $-(1-z)\log(1-z) \le z$. With $\ell = 2d$ and $w = v/(2\sqrt{d}) > 2$, the logarithm of this bound is at most

$$2d \log \frac{ev^2}{4d} = 2d(1 + 2\log w) \le 4dw = 2v\sqrt{d}.$$

The last inequality follows from $1 + 2\log w \le 2w$ for $w \ge 2$. □

Theorem 4.2 (General planting-coupling bound). *Let $d \ge 1$, let $r \ge n^{-1/2}$, and suppose that I is independent of the uniform Rademacher vector X and satisfies*

$$\Pr(U \subseteq I) \le \left(\frac{r}{\sqrt{n}}\right)^{v(U)} \qquad \textit{for every nonempty } U \subseteq E_n \textit{ with } |U| \le 2d. \tag{4.1}$$

Set

$$B_d = 4(d+1)e^{2\sqrt{d}}, \qquad A_d = 8(d+1)B_d^4 = 2048(d+1)^5 e^{8\sqrt{d}}.$$

If $B_d r \le 1/2$, then every polynomial f of degree at most d satisfies

$$\mathbb{E}[(f(Y) - f(X))^2] \le A_d r^4 \operatorname{Var}(f(X)). \tag{4.2}$$

Proof. Multilinearize f and write $f = \sum_{|S|\le d} a_S \chi_S$. Define the symmetric matrix

$$N_{S,T} = \mathbb{E}[(\chi_S(Y) - \chi_S(X))(\chi_T(Y) - \chi_T(X))].$$

Then the left side of (4.2) equals $a^\top N a$. We prove a weighted row-sum bound for N.

Step 1: compute all types of entries. For a fixed mask, $\chi_S(Y) = \chi_{S\setminus I}(X)$. Orthogonality gives

$$\begin{aligned}
\mathbb{E}_X[\chi_S(Y)\chi_T(Y)] &= \mathbf{1}_{\{S\triangle T\subseteq I\}},\\
\mathbb{E}_X[\chi_S(Y)\chi_T(X)] &= \mathbf{1}_{\{S\setminus I = T\}},\\
\mathbb{E}_X[\chi_S(X)\chi_T(Y)] &= \mathbf{1}_{\{T\setminus I = S\}},\\
\mathbb{E}_X[\chi_S(X)\chi_T(X)] &= \mathbf{1}_{\{S=T\}}.
\end{aligned}$$

Consequently,

$$N_{S,S} = 2\Pr(S\cap I \ne \varnothing). \tag{4.3}$$

If neither $S\subseteq T$ nor $T\subseteq S$, the two mixed indicators vanish, and

$$N_{S,T} = \Pr(S\triangle T \subseteq I). \tag{4.4}$$

If $T \subsetneq S$, then $\{S\setminus I = T\}$ is the intersection of $\{S\setminus T\subseteq I\}$ with $\{T\cap I = \varnothing\}$. Thus

$$N_{S,T} = \Pr(S\setminus T\subseteq I,\ T\cap I\ne\varnothing). \tag{4.5}$$

The remaining nested case follows by symmetry. Every entry of N is nonnegative. Its row and column indexed by $\varnothing$ vanish.

Step 2: a weighted counting rule. Let $w_S = n^{-v(S)/2}$. Fix a nonempty S and write, for a possible T,

$$a = |V(S)\setminus V(T)|, \qquad b = |V(T)\setminus V(S)|.$$

Then $w_T/w_S = n^{(a-b)/2}$. Suppose an auxiliary edge set H uniquely determines T, has at most $2d$ edges and v incident vertices, and its vertices outside $V(S)$ are exactly the b new vertices of T. The number of possible H with a specified b is at most

$$n^b(2d)^{v-b}e^{2v\sqrt{d}}. \tag{4.6}$$

Indeed, choose the new vertex labels in at most n^b ways, choose the other vertices from $V(S)$ in at most $(2d)^{v-b}$ ways, and use Lemma 4.1. The count may include extra choices that do not define a valid T; this only increases the upper bound.

Suppose, moreover, that all valid choices satisfy $a+b\le v-\ell$. Then (4.1) gives

$$\Pr(H\subseteq I)\frac{w_T}{w_S} \le \left(\frac{r}{\sqrt{n}}\right)^v n^{(v-\ell-2b)/2} = n^{-b-\ell/2}r^v.$$

Multiplying by (4.6) and summing over b gives the bound

$$n^{-\ell/2}r^v e^{2v\sqrt{d}}\sum_{b=0}^{v}(2d)^{v-b} \le n^{-\ell/2}(B_d r)^v. \tag{4.7}$$

Here $\sum_{b=0}^{v}(2d)^{v-b} \le (v+1)(2d)^v \le [4(d+1)]^v$, because $v+1\le 2^v$ for $v\ge 1$. The same rule applies when an edge $e\in S$ is held fixed and H determines T once e is specified.

Step 3: incomparable edge sets. Take $H = S \triangle T$, which determines $T = S \triangle H$ and has at most $2d$ edges. Every vertex lost from S or new in T is incident to H, so $a + b \le v(H)$. Incomparability implies that H contains an edge of $S \setminus T$ and an edge of $T \setminus S$, and hence $v(H) \ge 3$.

If $v(H) = 3$, the two edges just described share an endpoint. This endpoint is incident to H and belongs to both $V(S)$ and $V(T)$. It is therefore not among the $a + b$ lost and new vertices, so $a + b \le 2 = v(H) - 1$. Apply (4.7) with $\ell = 1$ when $v = 3$ and $\ell = 0$ when $v \ge 4$. By (4.4), the total contribution is at most

$$n^{-1/2}(B_d r)^3 + \sum_{v \ge 4} (B_d r)^v = B_d^3 n^{-1/2} r^3 + \frac{(B_d r)^4}{1 - B_d r}. \tag{4.8}$$

Only finite ranges of v can actually occur; extending the sum to infinity is an upper bound.

Step 4: strictly nested edge sets. For either inclusion direction, (4.5) and a union bound imply

$$N_{S,T} \le \sum_{e \in S \cap T} \Pr((S \triangle T) \cup \{e\} \subseteq I).$$

Fix a possible common edge $e \in S$; there are at most d choices. Put $H = (S \triangle T) \cup \{e\}$. For this fixed e, it determines

$$T = S \triangle (H \setminus \{e\}).$$

The new vertices of H are exactly the new vertices of T. Since the edge sets are strictly nested and e is common, H has at most $d + 1 \le 2d$ edges. It contains at least two distinct edges, so $v(H) \ge 3$. Both endpoints of e belong to $V(S) \cap V(T)$ and are incident to H. Thus $a + b \le v(H) - 2$. Apply (4.7) with $\ell = 2$, and sum over e and $v \ge 3$. The total nested contribution is at most

$$dn^{-1} \sum_{v \ge 3} (B_d r)^v = \frac{dn^{-1}(B_d r)^3}{1 - B_d r}. \tag{4.9}$$

If the common edge set is empty, the corresponding matrix entry is zero.

Step 5: diagonal terms and the weighted Schur estimate. By (4.3), a union bound, and the single-edge case of (4.1),

$$N_{S,S} \le 2d\, r^2/n.$$

Combining this with (4.8) and (4.9), and using $B_d r \le 1/2$, gives

$$\begin{aligned} \sum_T N_{S,T} \frac{w_T}{w_S} &\le 2dr^2/n + B_d^3 n^{-1/2} r^3 + 2B_d^4 r^4 + 2dB_d^3 n^{-1} r^3 \\ &\le (2d + B_d^3 + 2B_d^4 + 2dB_d^3) r^4 \le A_d r^4. \end{aligned}$$

Here $r \ge n^{-1/2}$ implies $r^2/n \le r^4$, $n^{-1/2} r^3 \le r^4$, and $n^{-1} r^3 \le r^4$.

For every S, T,

$$2|a_S a_T| \le a_S^2 \frac{w_T}{w_S} + a_T^2 \frac{w_S}{w_T}.$$

Using symmetry and entrywise nonnegativity of N, and omitting its zero constant row and column, we conclude

$$a^\top N a \le \sum_{S,T \ne \varnothing} N_{S,T} |a_S a_T| \le A_d r^4 \sum_{S \ne \varnothing} a_S^2.$$

Parseval proves (4.2). $\square$

Corollary 4.3 (Coupling under the graph condition)**.** *Assume* (E)*. If $n \geq 8d$, $r = 2n^{-\kappa}$, and $B_d r \leq 1/2$, then* (4.2) *holds for the planting of* Γ*. In particular, for fixed κ and $d = O(\log n)$, subject to* (E) *at that degree,*

$$\mathbb{E}[(f(Y) - f(X))^2] \leq n^{-4\kappa+o(1)} \operatorname{Var}(f(X)) \qquad (\deg f \leq d),$$

uniformly over the templates and polynomials.

Proof. Lemma 2.3 provides (4.1). Since $0 < \kappa < 1/2$, the choice $r = 2n^{-\kappa}$ satisfies $r \geq n^{-1/2}$. Apply Theorem 4.2. Finally,

$$\log A_d = O(\sqrt{d} + \log(d+1)), \qquad \log B_d = O(\sqrt{d} + \log(d+1)).$$

For $d = O(\log n)$ these are $o(\log n)$, so $B_d r = o(1)$ and $A_d r^4 = 16 A_d n^{-4\kappa} \leq n^{-4\kappa+o(1)}$. $\square$

Lemma 4.4 (Disagreement with a margin)**.** *Suppose a fixed polynomial f satisfies $\mathbb{E}[(f(Y) - f(X))^2] \leq \delta^2 \operatorname{Var}(f(X))$. If $\sigma^2 = \operatorname{Var}(f(X)) > 0$, then for every $t > 0$,*

$$\Pr\big(\mathbf{1}_{\{f(X)>0\}} \neq \mathbf{1}_{\{f(Y)>0\}}\big) \leq \Pr(|f(X)| \leq t\sigma) + \delta^2/t^2.$$

Proof. If $|f(X)| > t\sigma$ and $|f(Y) - f(X)| < t\sigma$, the two values have the same strict sign. Bound the complementary events by a union bound and Markov's inequality applied to the squared difference. $\square$

The small-ball term cannot be omitted: a Boolean polynomial can have an atom at zero or values on arbitrarily small scales. The next two sections provide the necessary uniform treatment.

5 A decision tree for the actual restricted polynomial

Fix $d \geq 1$, $0 < \eta < 1/10$, and $0 < \tau < 1$. A successful leaf will mean that its polynomial is constant, τ-regular, or satisfies

$$\operatorname{Var}(p_\rho) \leq \eta (\mathbb{E} p_\rho)^2. \tag{5.1}$$

The critical-index organization follows the regularity framework of [2]. We prove the precise variant needed here: success must concern the actual restricted polynomial, not merely a different threshold function approximating it under the null law. A deterministic tie-breaking rule is used whenever influences are sorted.

Lemma 5.1. *Define*

$$q = \frac{1}{4 \cdot 9^d}, \qquad \beta = \frac{\tau^2}{32 \cdot 81^d}, \qquad L = \left\lceil \beta^{-1} \log \frac{8d}{\eta q} \right\rceil,$$

$$R = \left\lceil \frac{2}{q} \log \frac{1}{\eta} \right\rceil, \qquad h = LR.$$

Every degree-at-most-d polynomial on a Boolean cube admits a deterministic coordinate decision tree of depth at most h such that a uniform input reaches a successful leaf with probability at least $1 - \eta$. For fixed η,

$$\log h \leq 2 \log(1/\tau) + O_\eta(d + \log(d+1)). \tag{5.2}$$

Proof. We first show that at any current nonconstant node, at most L additional queries produce success with conditional probability at least $q/2$. At a successful node no queries are needed. If at most L coordinates remain, query all of them and obtain constants. Otherwise, for the estimates in this round, divide the current polynomial by its positive standard deviation, so that its variance is one. This normalization is used only in the analysis: the tree still stores the actual restrictions of the original polynomial. The choices of coordinates, regularity, (5.1), and signs are invariant under this positive rescaling. Let M be the number of remaining coordinates. Order the remaining coordinates so that their influences decrease. Write $J_i = \mathrm{Inf}_i(f)$ and

$$T_j = \sum_{i>j} J_i.$$

By (3.2), $T_0 \leq d$. Define the critical index k as the smallest $j \in \{0, \dots, M\}$ with

$$J_{j+1} \leq \beta T_j,$$

where $J_{M+1} = 0$ and $T_M = 0$. Thus the index always exists. For $j < k$, one has $J_{j+1} > \beta T_j$, hence

$$T_{j+1} = T_j - J_{j+1} < (1-\beta)T_j.$$

Inductively,

$$T_j \leq d(1-\beta)^j \leq de^{-\beta j}, \qquad j \leq k. \tag{5.3}$$

Case 1: $k = 0$. Then $J_1 \leq \beta \,\mathrm{Inf}(f)$, and all other influences are no larger. Since $\beta \leq \tau$, the current polynomial is already τ-regular.

Case 2: $k \geq L$. Query the first L coordinates and let ρ denote their uniform answers. Put $\mu(\rho) = \mathbb{E}p_\rho$ and $V(\rho) = \mathrm{Var}(p_\rho)$. By (3.7), (5.3), and the definition of L,

$$\mathbb{E}_\rho V(\rho) \leq T_L \leq de^{-\beta L} \leq \frac{\eta q}{8}.$$

Because the original variance is one, (3.8) gives

$$\mathbb{E}_\rho \mu(\rho)^2 = 1 + (\mathbb{E}f)^2 - \mathbb{E}_\rho V(\rho) \geq \frac{1}{2}.$$

The function μ is a polynomial of degree at most d in ρ. Lemma 3.1 and (3.4), applied to $W = \mu^2$, yield

$$\Pr_\rho\left(\mu^2 \geq \tfrac{1}{2}\mathbb{E}_\rho \mu^2\right) \geq \frac{1}{4 \cdot 9^d} = q.$$

On this event, $\mu^2 \geq 1/4$. Markov's inequality gives

$$\Pr_\rho(V > \eta/4) \leq \frac{4\mathbb{E}_\rho V}{\eta} \leq q/2.$$

The intersection therefore has probability at least $q/2$, and on it $V \leq \eta/4 \leq \eta\mu^2$. Thus the actual restriction satisfies (5.1).

Case 3: $0 < k < L$. Query the first k coordinates. Put $B = T_k$. If $B = 0$, the polynomial depends only on those queried coordinates, and every restriction is constant. Suppose $B > 0$. For $j > k$, define

$$I_j(\rho) = \mathrm{Inf}_j(p_\rho), \qquad Z(\rho) = \sum_{j>k} I_j(\rho).$$

Identity (3.6) gives $\mathbb{E}_\rho I_j = J_j$ and $\mathbb{E}_\rho Z = B$. In the notation of (3.5),

$$I_j = \sum_{T \ni j} b_T^2, \qquad Z = \sum_{T \neq \varnothing} |T| b_T^2.$$

The triangle inequality in L^2 and Lemma 3.1 imply

$$\|I_j\|_2 \le \sum_{T \ni j} \|b_T\|_4^2 \le 3^d \sum_{T \ni j} \mathbb{E} b_T^2 = 3^d J_j,$$

and likewise $\|Z\|_2 \le 3^d B$. It follows from (3.4) that

$$\Pr_\rho(Z \ge B/2) \ge q.$$

The critical-index condition and the ordering give $J_j \le \beta B$ for every $j > k$. A union bound and Markov's inequality applied to I_j^2 now give

$$\begin{aligned}
\Pr_\rho\left(\max_{j>k} I_j > \frac{\tau B}{2}\right) &\le \frac{4}{\tau^2 B^2} \sum_{j>k} \mathbb{E}_\rho I_j^2 \\
&\le \frac{4 \cdot 9^d}{\tau^2 B^2} \sum_{j>k} J_j^2 \\
&\le \frac{4 \cdot 9^d \beta}{\tau^2} = \frac{1}{8 \cdot 9^d} = \frac{q}{2}.
\end{aligned}$$

Hence, with probability at least $q/2$, both $Z \ge B/2$ and $\max_j I_j \le \tau B/2$ hold. Then $\max_j I_j \le \tau Z$, so the actual restricted polynomial is τ-regular.

Iteration and depth. Apply this procedure at every unsuccessful leaf, freezing every successful leaf. Repeat at most R times. Conditional on any unsuccessful history, the unqueried coordinates are still independent uniform bits, so the next round succeeds with probability at least $q/2$. The unsuccessful probability after R rounds is therefore at most

$$(1 - q/2)^R \le e^{-qR/2} \le \eta.$$

Each round queries at most L coordinates along any path, proving the depth bound $h = LR$.

For completeness, the quantities inside both ceilings are greater than one, so

$$h \le \left(2\beta^{-1} \log \frac{8d}{\eta q}\right)\left(\frac{4}{q} \log \frac{1}{\eta}\right).$$

Now $\log \beta^{-1} = 2\log(1/\tau) + d \log 81 + \log 32$, $\log q^{-1} = d \log 9 + \log 4$, and

$$\log\log \frac{8d}{\eta q} = O_\eta(\log(d+1)).$$

Taking logarithms proves (5.2). $\square$

6 Small-ball estimates for regular polynomials

We use the following specialization of Carbery–Wright [1, Theorem 8]: for a nonzero degree-at-most-d polynomial f and a standard Gaussian vector G,

$$\Pr(|f(G)| \le s\|f(G)\|_2) \le C_{\mathrm{CW}} d\, s^{1/d}, \qquad s > 0, \tag{6.1}$$

with an absolute constant C_{CW}. This is obtained from their log-concave-measure inequality by taking the norm parameter to be $2d$. The same Gaussian formulation is recorded in [2, Theorem 8]. No centering assumption is required.

The Rademacher-to-Gaussian step needed below can be proved directly, instead of invoking a quantitative invariance theorem as a black box.

Lemma 6.1 (Smooth coordinate replacement). *Let X be uniform on $\{-1,1\}^M$, let G be a standard Gaussian vector, and let f be multilinear of degree at most d. If $\varphi \in C^4(\mathbb{R})$ has bounded fourth derivative, then*

$$|\mathbb{E}\varphi(f(X)) - \mathbb{E}\varphi(f(G))| \le \frac{9^d}{6}\|\varphi^{(4)}\|_\infty \sum_i \mathrm{Inf}_i(f)^2. \tag{6.2}$$

Proof. Replace the coordinates one at a time. At the replacement of coordinate i, condition on all other coordinates, which form an independent mixture of Gaussian and Rademacher variables. Multilinearity gives $f = U + z_i V$, with U and V independent of z_i. Taylor's formula gives

$$\varphi(U + zV) = \sum_{j=0}^{3} \frac{\varphi^{(j)}(U)}{j!} z^j V^j + R_z, \qquad |R_z| \le \frac{\|\varphi^{(4)}\|_\infty}{24}|zV|^4.$$

The first three moments of a Rademacher and a standard Gaussian agree. Their fourth moments sum to four. Therefore the absolute error at this replacement is at most

$$\frac{\|\varphi^{(4)}\|_\infty}{6}\mathbb{E}V^4.$$

The polynomial V has degree at most $d-1$, and orthogonality under the mixed input gives $\mathbb{E}V^2 = \mathrm{Inf}_i(f)$. Lemma 3.1 implies

$$\mathbb{E}V^4 \le 9^{d-1}\,\mathrm{Inf}_i(f)^2 \le 9^d\,\mathrm{Inf}_i(f)^2.$$

Summing the individual replacement errors proves (6.2). □

Lemma 6.2 (Regular small-ball bound). *There is an absolute $C_0 \ge 1$ such that every nonconstant τ-regular multilinear polynomial f of degree at most d satisfies, for $\sigma^2 = \mathrm{Var}(f(X))$ and every $t > 0$,*

$$\Pr(|f(X)| \le t\sigma) \le C_0 d\left[t^{1/d} + (d\tau)^{1/(8d)}\right]. \tag{6.3}$$

Proof. Divide f by σ, without subtracting its mean, so that its variance is one. Let $u = d\tau$. By regularity and (3.2),

$$\max_i \mathrm{Inf}_i(f) \le u, \qquad \sum_i \mathrm{Inf}_i(f) \le d, \qquad \sum_i \mathrm{Inf}_i(f)^2 \le du.$$

We first establish a smoothed bound. For any $s > 0$, choose a smooth function φ satisfying

$$0 \le \varphi \le 1, \qquad \varphi = 1 \text{ on } [-t,t], \qquad \varphi = 0 \text{ outside } [-t-s, t+s], \qquad \|\varphi^{(4)}\|_\infty \le C_2 s^{-4},$$

where C_2 is absolute. To construct it, take a fixed nonnegative smooth density ψ supported in $[-1/2, 1/2]$, set $\psi_s(x) = s^{-1}\psi(x/s)$, and convolve ψ_s with the indicator of $[-t-s/2, t+s/2]$. The fourth derivative is bounded by $\|\psi_s^{(4)}\|_1 = s^{-4}\|\psi^{(4)}\|_1$.

Lemma 6.1 gives

$$\Pr(|f(X)| \le t) \le \mathbb{E}\varphi(f(X)) \le \Pr(|f(G)| \le t+s) + C_1\, 9^d du\, s^{-4}.$$

Orthogonality implies $\mathrm{Var}(f(G)) = 1$ and hence $\|f(G)\|_2 \ge 1$. Thus (6.1) yields

$$\Pr(|f(X)| \le t) \le C_{\mathrm{CW}} d(t^{1/d} + s^{1/d}) + C_1\, 9^d du\, s^{-4}, \tag{6.4}$$

using $(t+s)^{1/d} \le t^{1/d} + s^{1/d}$.

If $u \ge 9^{-4d}$, then $u^{1/(8d)} \ge 1/3$, and (6.3) is trivial after taking $C_0 \ge 3$. Otherwise choose $s = u^{1/8}$. Then

$$s^{1/d} = u^{1/(8d)}, \qquad 9^d u\, s^{-4} = 9^d u^{1/2} \le u^{1/4} \le u^{1/(8d)}.$$

The middle inequality uses $u \le 9^{-4d}$. Substitution in (6.4), followed by enlargement of an absolute constant, proves (6.3). □

Remark 6.3. The estimate controls the full closed interval $[-t\sigma, t\sigma]$, including any atom at zero of $f(X)$. Its proof does not assume that the Boolean distribution is continuous. The nonzero mean of f also causes no loss because the Gaussian comparison uses $\|f(G)\|_2 \geq \sigma$.

7 Adaptive restrictions and a general PTF transfer inequality

The next proposition is formulated for any independent overwrite mask, so its dependence on the planted graph is completely isolated in α and δ.

Proposition 7.1 (Uniform threshold transfer)**.** *Let X be uniform on $\{-1,1\}^{E_n}$, let I be independent of X, and let Y be given by* (2.2)*. Suppose that*

$$\alpha = \max_{e \in E_n} \Pr(e \in I), \qquad \mathbb{E}[(g(Y) - g(X))^2] \leq \delta^2 \operatorname{Var}(g(X)) \quad \textit{for every } \deg g \leq d.$$

For $h = h(d, \eta, \tau)$ from Lemma 5.1 and every $t > 0$, every degree-at-most-d polynomial f satisfies

$$\begin{aligned} &|\Pr(f(Y) > 0) - \Pr(f(X) > 0)| \\ &\quad \leq h\alpha + \eta + \max\left\{ C_0 d \left[t^{1/d} + (d\tau)^{1/(8d)} \right] + \frac{\delta^2}{t^2},\ (2 + 2\delta + \delta^2)\eta \right\}. \end{aligned} \tag{7.1}$$

The same right-hand side bounds the coupling probability that the two threshold indicators disagree.

Proof. Construct the deterministic decision tree of Lemma 5.1 for f and run it on X. Let ρ be the leaf reached, $F_\rho \subseteq E_n$ the coordinates queried along the path, and x^ρ their answers. The tree never queries an already fixed coordinate. Its actual restricted polynomial is

$$f_\rho(z) = f(x^\rho, z), \qquad z \in \{-1,1\}^{E_n \setminus F_\rho},$$

and $|F_\rho| \leq h$.

Step 1: charge all changed queried coordinates separately. The set F_ρ is a function of X alone. Conditional on X, independence of I and a union bound give

$$\Pr(I \cap F_\rho \neq \varnothing) \leq \mathbb{E} \sum_{e \in F_\rho} \Pr(e \in I) \leq h\alpha. \tag{7.2}$$

On the complementary event, planting changes none of the answers along the path reached by X. Induction along this path shows that Y reaches the same leaf. In particular,

$$f(X) = f_\rho(X_{F_\rho^c}), \qquad f(Y) = f_\rho(Y_{F_\rho^c})$$

on that event. The first identity always holds. Consequently,

$$\begin{aligned} &\Pr\left(\mathbf{1}_{\{f(X)>0\}} \neq \mathbf{1}_{\{f(Y)>0\}}\right) \\ &\quad \leq h\alpha + \sum_\rho \Pr(X \text{ reaches } \rho) \\ &\qquad\qquad \cdot \Pr\left(\mathbf{1}_{\{f_\rho(X_{F_\rho^c})>0\}} \neq \mathbf{1}_{\{f_\rho(Y_{F_\rho^c})>0\}} \,\middle|\, X \text{ reaches } \rho\right). \end{aligned} \tag{7.3}$$

This is a pointwise event inclusion followed by averaging. No assumption on the distribution of leaves reached by Y is required.

Step 2: justify use of the coupling at each leaf. For a fixed leaf ρ, the event that X reaches it is exactly the cylinder event $X_{F_\rho} = x^\rho$. Indeed, those answers reproduce every deterministic choice and branch on the path. Conditional on this event, the unqueried coordinates are still independent uniform signs. The mask I retains its original distribution and is independent of them, because the conditioning event depends only on X.

View f_ρ as a polynomial on all of E_n that ignores the coordinates in F_ρ. Its degree is at most d. Adding arbitrary independent signs in the ignored coordinates does not change its values or its variance. Hence, conditional on the leaf,

$$\mathbb{E}\big[(f_\rho(Y_{F_\rho^c}) - f_\rho(X_{F_\rho^c}))^2 \mid X \text{ reaches } \rho\big] \le \delta^2 \operatorname{Var}(f_\rho).$$

Here the variance is with respect to uniform unqueried coordinates. We condition *only* on the leaf reached by X, not additionally on $I \cap F_\rho = \varnothing$. Such extra conditioning would change the mask law. Its cost has already been paid in (7.2).

Step 3: control every type of leaf. Unsuccessful leaves have total null probability at most η, and their conditional disagreement probability is at most one. Constant leaves contribute zero. Indeed, the uniform law has full support on the remaining cube, so a zero-variance restriction is constant at every point of that cube, including every overwritten input.

At a nonconstant regular leaf, combine Lemma 4.4 and Lemma 6.2. The conditional disagreement probability is at most

$$C_0 d\left[t^{1/d} + (d\tau)^{1/(8d)}\right] + \delta^2/t^2.$$

At a mean-dominated leaf, write $\mu = \mathbb{E} f_\rho$ and $\sigma^2 = \operatorname{Var}(f_\rho) \le \eta\mu^2$. If $\mu = 0$, then $\sigma = 0$, and the polynomial is constant zero. Otherwise the conditional coupling estimate and the triangle inequality in L^2 give

$$\|f_\rho(Y) - \mu\|_2 \le \|f_\rho(Y) - f_\rho(X)\|_2 + \|f_\rho(X) - \mu\|_2 \le (1+\delta)\sigma.$$

If either threshold indicator disagrees with $\mathbf{1}_{\{\mu>0\}}$, its polynomial value differs from μ by at least $|\mu|$. Therefore the conditional probability that the two indicators disagree is at most

$$\frac{\sigma^2}{\mu^2} + \frac{(1+\delta)^2\sigma^2}{\mu^2} \le (2 + 2\delta + \delta^2)\eta.$$

If a leaf satisfies both successful criteria, either corresponding bound is valid. Averaging in (7.3) proves the assertion about coupling disagreement. Finally,

$$|\Pr(f(Y) > 0) - \Pr(f(X) > 0)| \le \Pr\big(\mathbf{1}_{\{f(Y)>0\}} \ne \mathbf{1}_{\{f(X)>0\}}\big).$$

The upper bound is independent of f, which proves the uniform assertion. □

8 Parameter choices and completion of the graph theorems

Proof of Theorem 1.2. As noted earlier, $d = 0$ or an edgeless template gives zero advantage. Work along the remaining indices with $d \ge 1$. Fix $0 < \eta < 1/10$, independently of n, and use the absolute constant $C_0 \ge 1$ from Lemma 6.2. Set

$$t = \left(\frac{\eta}{4C_0 d}\right)^d, \qquad \tau = \frac{1}{d}\left(\frac{\eta}{4C_0 d}\right)^{8d}.$$

Then $0 < \tau < 1$ and

$$C_0 d\left[t^{1/d} + (d\tau)^{1/(8d)}\right] = \eta/2.$$

The logarithmic costs are

$$\log t^{-2} = 2d \log d + O_\eta(d),$$
$$\log(1/\tau) = \log d + 8d \log d + O_\eta(d).$$

Thus Lemma 5.1 implies

$$\log h \leq 16d \log d + O_\eta(d + \log(d+1)).$$

Write $c = c_\kappa$ and assume $1 \leq d \leq c \log n / \log \log n$. Uniformly over this range,

$$d \log d \leq (c + o(1)) \log n, \qquad d + \log(d+1) = o(\log n).$$

These estimates include bounded d. In particular,

$$h \leq n^{16c+o(1)}, \qquad t^{-2} \leq n^{2c+o(1)}.$$

For all sufficiently large n, we have $n \geq 8d$. Lemma 2.3 and Corollary 4.3 apply with $r = 2n^{-\kappa}$, since $r \geq n^{-1/2}$ and $B_d r = o(1)$. Take

$$\delta^2 = A_d r^4 = 16 A_d n^{-4\kappa} \leq n^{-4\kappa+o(1)}.$$

Also (2.5) gives

$$\alpha = \frac{2m}{n(n-1)} \leq n^{-1-2\kappa+o(1)}.$$

Consequently,

$$h\alpha \leq n^{16c-1-2\kappa+o(1)}, \tag{8.1}$$
$$\delta^2/t^2 \leq n^{-4\kappa+2c+o(1)}. \tag{8.2}$$

For $c = \min\{\kappa, 1/100\}$,

$$16c - 1 - 2\kappa \leq -0.84 - 2\kappa < 0, \qquad -4\kappa + 2c \leq -2\kappa < 0.$$

Thus (8.1) and (8.2) both tend to zero; also $\delta \to 0$.

Apply Proposition 7.1, and take the supremum over degree-at-most-d polynomials. It follows that

$$\limsup_{n\to\infty} \mathrm{Adv}_{n,\Gamma_n}(d_n) \leq \eta + \max\{\eta/2, 2\eta\} = 3\eta.$$

The estimate holds for every fixed $\eta > 0$, so letting $\eta \downarrow 0$ proves the conclusion. Every finite-n bound depends on the template only through the common assumptions, and never on the coefficients of the test. Therefore the convergence is uniform as claimed. The equivalence with the tree condition was proved in Section 2. □

Proof of Theorem 1.1 and Corollary 1.3. Lemma 2.1 shows that (C) implies (E). Lemma 2.2 shows that (M) implies (E). Apply Theorem 1.2 in each case. The previously separated cases of constant tests and edgeless templates are immediate. □

Remark 8.1 (Constants and degree range). The value 1/100 is a convenient safe constant, not an optimization. The displayed proof permits any fixed $c > 0$ with $2c < 4\kappa$ and $16c < 1 + 2\kappa$. Under the embedding condition, the polynomial-value coupling estimate also holds at degree $O(\log n)$, but the threshold transfer incurs factors $\exp(O_\eta(d \log d))$. The theorem proved here concerns the stated $\log n/ \log \log n$ degree range. A small coupling error by itself is not a justification for a larger PTF degree range.

9 Examples and interpretation of the graph conditions

9.1 Cliques, including the specified example

For the clique $\Gamma = K_K$,

$$2|E(\Gamma)| = K(K-1) \le K^2, \qquad \Delta_\Gamma = K-1.$$

Hence any $K \le n^{1/2-\kappa}$ satisfies (C). In particular, if $0 < \epsilon < 1/10$ and

$$K = \lfloor n^{1/10-\epsilon} \rfloor,$$

then we may take $\kappa = 2/5 + \epsilon \in (0, 1/2)$. All polynomial threshold tests of degree at most $(1/100)\log n/\log\log n$ consequently have vanishing advantage. The exponent $1/10 - \epsilon$ is the specified example; the theorem in fact includes cliques as large as $n^{1/2-\kappa}$ for every fixed $\kappa > 0$. If $\epsilon \ge 1/10$, the example has at most one vertex for large n and is trivial.

For any fixed $0 < a < 1/2$, take $K = \lfloor n^a \rfloor$ and $\kappa = 1/2 - a$. In this case unrestricted testing succeeds: the planted graph always contains a K-clique, whereas the null probability of this event is at most

$$\binom{n}{K} 2^{-\binom{K}{2}} \le \exp\left(K \log \frac{en}{K} - \frac{K(K-1)}{2} \log 2 \right) \longrightarrow 0.$$

Thus the clique specialization separates unrestricted testing from the PTF class in Theorem 1.1; its conclusion is not explained by information-theoretic indistinguishability.

9.2 Sparse templates with more than square-root many vertices

Let Γ_n be a matching with $\lfloor n^{9/10} \rfloor$ edges. It uses at most n vertices for all sufficiently large n. Its maximum degree is one and its edge count satisfies (C) for every fixed $0 < \kappa < 1/20$. Thus Theorem 1.1 applies even though the template has substantially more than $\sqrt{n}$ vertices. The same conclusion holds for any uniformly bounded-degree template with $O(n^{9/10})$ vertices, by absorbing the fixed multiplicative constants into the strict exponent slack.

The matching example illustrates the range of admissible template supports. It does not exhibit a separation between unrestricted testing and PTF testing. In fact, if Γ is a matching with $m = o(n)$ edges, even unrestricted tests have vanishing advantage. To see this, let $L = d\mathbb{P}_{\Gamma,n}/d\mathbb{Q}_n$ and let I, I' be independent planted masks. The likelihood identity $L(X) = \mathbb{E}_I \prod_{e\in I}(1 + X_e)$ gives

$$\mathbb{E}_{\mathbb{Q}_n} L^2 = \mathbb{E}_{I,I'} 2^{|I\cap I'|} = \sum_{k=0}^{m} \binom{m}{k} \frac{2^k (m)_k}{(n)_{2k}} \le \left(1 + \frac{2m}{(n-2m)^2}\right)^m \le \exp\left(\frac{2m^2}{(n-2m)^2}\right) = 1 + o(1).$$

Here $n > 2m$ for all sufficiently large n, and the equality follows by expanding $2^{|I\cap I'|} = \sum_{U\subseteq I} \mathbf{1}_{\{U\subseteq I'\}}$ and applying the containment identity to k disjoint edges. Cauchy–Schwarz then bounds the total variation distance by $\frac{1}{2}(\mathbb{E}_{\mathbb{Q}_n} L^2 - 1)^{1/2} = o(1)$.

As another example, let Γ_n be a disjoint union of $\lfloor n^{3/10} \rfloor$ cliques, each on $\lfloor n^{3/10} \rfloor$ vertices. Then

$$|V(\Gamma_n)| \asymp n^{3/5}, \qquad |E(\Gamma_n)| \asymp n^{9/10}, \qquad \Delta_{\Gamma_n} \asymp n^{3/10}.$$

Again (C) holds for every fixed $0 < \kappa < 1/20$ when n is sufficiently large. Unlike the matching example, this example also separates unrestricted testing from the stated PTF class: the planted graph contains a clique of size $\lfloor n^{3/10} \rfloor$, while the preceding first-moment bound makes the null probability of such a clique tend to zero. Its template support has size $n^{3/5+o(1)}$, exceeding $\sqrt{n}$.

9.3 Why neither numerical scale should be omitted

The following observations explain the roles of edge count and maximum degree; they do not assert necessity at the exact boundary of (C).

If $m/n \to \infty$, the total edge count Z is already a successful affine-threshold statistic. With $M = \binom{n}{2}$,

$$Z \sim \mathrm{Bin}(M, 1/2) \quad \text{under } \mathbb{Q}_n, \qquad Z \sim m + \mathrm{Bin}(M - m, 1/2) \quad \text{under } \mathbb{P}_{\Gamma,n}.$$

The latter law does not depend on the chosen embedding because exactly m distinct edges are forced. The means differ by $m/2$ and both variances are at most $M/4$. Chebyshev's inequality, with threshold $M/2 + m/4$, bounds the sum of the two error probabilities by $8M/m^2 = o(1)$.

Edge count alone is not sufficient. Suppose some template vertex has degree at least $n^{1/2+\beta}$, for a fixed $\beta > 0$. Define

$$Z_i = \frac{1}{\sqrt{n-1}} \sum_{j \neq i} (2A_{ij} - 1), \qquad T_r = \sum_{i=1}^{n} Z_i^{2r},$$

where r is a fixed positive integer with $2r\beta > 1$. Under the null, expansion of the even moment shows $\mathbb{E}Z_i^{2r} \le C_r$: only index tuples in which each edge coordinate appears an even number of times contribute, and there are at most $(2r)!(n-1)^r$ such tuples. Hence $\mathbb{E}_{\mathbb{Q}_n} T_r \le C_r n$.

Conditional on the embedding, at the image i of the high-degree vertex, Z_i has mean at least n^β and variance at most one. Chebyshev gives $Z_i \ge n^\beta/2$ with probability at least $1 - 4n^{-2\beta}$. Therefore

$$\mathbb{Q}_n(T_r > 2^{-2r-1} n^{2r\beta}) \le 2^{2r+1} C_r n^{1-2r\beta} \to 0,$$

while the same event has probability tending to one under the planted law. This is a degree-$2r$ polynomial threshold test. For example, a star with $\lfloor n^{3/4} \rfloor$ leaves has $o(n)$ edges but is detected by such a degree-six test.

Thus $m \approx n$ and $\Delta_\Gamma \approx \sqrt{n}$ mark genuine elementary detection scales. Our sufficient theorem keeps fixed polynomial slack below both scales. The tree-embedding and degree-moment formulations can be used when more information about the template is available.

References

[1] A. Carbery and J. Wright. Distributional and L^q norm inequalities for polynomials over convex bodies in $\mathbb{R}^n$. *Mathematical Research Letters* **8** (2001), 233–248.

[2] I. Diakonikolas, R. A. Servedio, L.-Y. Tan, and A. Wan. A regularity lemma, and low-weight approximators, for low-degree polynomial threshold functions. arXiv:0909.4727.

[3] I. Diakonikolas, D. M. Kane, S. Liu, and T. Pittas. PTF testing lower bounds for non-Gaussian component analysis. arXiv:2511.19398 (2025).

[4] J.-T. Hsieh, D. M. Kane, P. K. Kothari, J. Li, S. Mohanty, and S. Tiegel. Rigorous implications of the low-degree heuristic. arXiv:2601.05850 (2026).

[5] X. Yu, I. Zadik, and P. Zhang. Counting stars is constant-degree optimal for detecting any planted subgraph. *Mathematical Statistics and Learning* **8** (2025), no. 1/2, 105–164.