Improved Lower Bounds on the Randomized Complexity of Graph Properties*

Amit Chakrabarti and Subhash Khot

Department of Computer Science, Princeton University
35 Olden Street, Princeton NJ 08544, USA.
E-mail: {amitc,khot}@cs.princeton.edu
FAX: +1-(609)-258-1771

Abstract. We prove a lower bound of $\Omega(n^{4/3}\log^{1/3} n)$ on the randomized decision tree complexity of any nontrivial monotone n-vertex bipartite graph property, thereby improving the previous bound of $\Omega(n^{4/3})$ due to Hajnal [H91]. Our proof works by improving a probabilistic argument in that paper, which also improves a graph packing lemma proved there. By a result of Gröger [G92] our complexity lower bound carries over from bipartite to general monotone n-vertex graph properties. Graph packing being a well-studied subject in its own right, our improved packing lemma and the probabilistic technique used to prove it, may be of independent interest.

Keywords: Decision tree complexity, monotone graph properties, randomized complexity, randomized algorithms, graph packing, probabilistic method.

1 Introduction

Consider the problem of deciding whether or not a given input graph G has a certain (isomorphism invariant) property P. The graph is given by an oracle which answers queries of the form "is (x, y) an edge of G?" A *decision tree algorithm* for P makes a sequence of such queries to the oracle, where each query may depend upon the information obtained from the previous ones, until sufficient information about G has been obtained to decide whether or not P holds for G, whereupon it either accepts or rejects. Let $\mathcal{A}_P$ denote the set of decision tree algorithms for P and for $A \in \mathcal{A}_P$, let $\operatorname{cost}(A, G)$ denote the number of queries that A asks on input G. The quantity $\mathcal{C}(P) = \min_A \max_G \operatorname{cost}(A, G)$ is called the *deterministic decision tree complexity*, or simply the deterministic complexity of P.

A *randomized decision tree algorithm* for P is a probability distribution $\mathcal{D}$ over $\mathcal{A}_P$, and its cost (on input G) is the expectation of $\operatorname{cost}(A, G)$ with A drawn from $\mathcal{D}$:

$$\operatorname{cost}^R(\mathcal{D}, G) = \sum_{A \in \mathcal{A}_P} \Pr_{\mathcal{D}}[A] \operatorname{cost}(A, G) \ .$$

* This work was supported in part by NSF Grant CCR-96-23768, ARO Grant DAAH04-96-1-0181, and NEC Research Institute.

The *randomized decision tree complexity*, or simply the randomized complexity, of P is defined to be

$$\mathcal{C}^R(P) = \min_{\mathcal{D}} \max_{G} \operatorname{cost}^R(\mathcal{D}, G) \ .$$

An n-vertex graph property is said to be *nontrivial* if there is at least one n-vertex graph which has the property and at least one which does not. It is said to be *monotone* if addition of edges does not destroy the property. Let $\mathcal{P}_n$ denote the set of all nontrivial monotone n-vertex graph properties. We shall need to consider the analogously defined notion of randomized complexity for *bipartite* graph properties; let $\mathcal{P}_{n,n}$ denote the set of all nontrivial monotone (n,n)-bipartite graph properties.

Clearly any n-vertex graph property P satisfies $\mathcal{C}^R(P) \leq \mathcal{C}(P) \leq \binom{n}{2} = O(n^2)$. A classic result of Rivest and Vuillemin [RV76] shows that any $P \in \mathcal{P}_n$ has $\mathcal{C}(P) = \Omega(n^2)$, which settles the deterministic complexity of monotone properties up to a constant.[1] We remark that monotonicity is crucial for this result; there are examples of nontrivial non-monotone graph properties with $\mathcal{C}(P) = O(n)$.

The situation is far less satisfactory for randomized complexity. The first non-linear lower bound on $\mathcal{C}^R(P)$, for general $P \in \mathcal{P}_n$, was an $\Omega(n \log^{1/12} n)$ bound proven by Yao [Y87]. This was subsequently improved by Valerie King [K88] to $\Omega(n^{5/4})$ and later by Hajnal [H91] to $\Omega(n^{4/3})$. The only other significant work in the area is due to Gröger [G92] who established lower bounds stronger than Hajnal's for certain special classes of graph properties.

No property in $\mathcal{P}_n$ is known to have randomized complexity below $n^2/4$. Closing this gap between Hajnal's lower bound and this upper bound is one of the most important open problems concerning the complexity of graph properties. It is remarkable that this quarter-century-old problem has yielded so few results.

In this paper we take a small step by proving

Theorem 1.1 (Main Theorem). *Any property* $P \in \mathcal{P}_n$ *satisfies* $\mathcal{C}^R(P) = \Omega(n^{4/3} \log^{1/3} n)$. [2]

Our proof will rely on an important theorem from the pioneering work of Yao [Y87], as well as on a framework developed by Hajnal [H91]. In this framework we associate with a graph property a special pair of graphs which cannot be "packed" together. We then argue that if the property has low randomized complexity, then certain degree upper bounds can be proven for these special graphs. Finally, we use these degree bounds to prove that the special graphs can be "packed", thereby arriving at a contradiction.

The notion of graph packing, which we shall formally define later, is a well-studied subject in its own right [B78]. A packing lemma (Lemma 2.8) which we

[1] However, in the world of deterministic complexity, a far more interesting conjecture is that any $P \in \mathcal{P}_n$ has $\mathcal{C}(P) = \binom{n}{2}$ exactly. Remarkably, this conjecture remains open to this day.

[2] Throughout this paper $\log x$ denotes the logarithm of x to the base 2. The natural logarithm of x is denoted by $\ln x$.

establish in this paper is therefore of independent interest since it improves a packing theorem due to Hajnal and Szegedy [HS92].

The rest of the paper is organized as follows: in Section 2, we define some preliminary notions, describe the framework alluded to above and prove Theorem 1.1 assuming our graph packing lemma. In Section 3 we state and prove a technical lemma which is then used in Section 4 to prove the packing lemma we need, thereby completing the proof of Theorem 1.1. We conclude with some remarks in Section 5.

2 Preliminaries and Proof Outline

The first important step is to change the objects of study from graph properties to *bipartite* graph properties. A result of Gröger [G92] lets us do just that.

Theorem 2.1. *Let $f(n)$ be a function satisfying $f(n) = O(n^{3/2})$ and suppose any $P \in \mathcal{P}_{n,n}$ satisfies $\mathcal{C}^R(P) = \Omega(f(n))$. Then any $Q \in \mathcal{P}_n$ satisfies $\mathcal{C}^R(Q) = \Omega(f(n))$.*

Proof. This is a simple restatement of Theorem 3.5 of [G92]. □

For the purposes of proving a lower bound of $\Omega(n^{4/3}\log^{1/3} n)$, we may therefore safely concentrate on monotone bipartite graph properties alone. We now need some definitions.

Definition 2.2 (Basic definitions). *An (m,n)-bipartite graph G is a graph whose vertices can be partitioned into two independent sets, denoted $V_L(G)$ and $V_R(G)$ respectively, of sizes m and n respectively. The edge set of G is denoted $E(G)$. For such a graph we define*

$$\Delta_L(G) = \max_{v \in V_L(G)} \deg_G(v) \ , \qquad \delta_L(G) = \frac{1}{|V_L(G)|} \sum_{v \in V_L(G)} \deg_G(v) \ = \frac{|E(G)|}{|V_L(G)|} \ .$$

$\Delta_R(G)$ and $\delta_R(G)$ are defined similarly. When $|V_L(G)| = |V_R(G)|$ we define $\delta(G) = \delta_L(G) = \delta_R(G)$. We define $\bar{G}$ to be the (m,n)-bipartite graph with the same bipartition and with edge set $V_L(G) \times V_R(G) - E(G)$.

Definition 2.3 (Sparseness). *The bipartite graph G is said to be L-*sparse *if $V_L(G)$ contains at least $\frac{1}{2}|V_L(G)|$ isolated vertices, i.e. vertices of degree 0. The notion of R-*sparseness *is defined analogously.*

Let $P \in \mathcal{P}_{n,n}$. An (n,n)-bipartite graph G is called a *minterm* of P if G satisfies P but removing any edge from G yields a graph which does not. Suppose we associate with G an n-tuple $(d_1, d_2, \ldots, d_n)$ with $d_1 \geq \ldots \geq d_n$ where the d_i are the degrees of the vertices in $V_L(G)$; we then say that G is an *L-first minterm* of P if it is a minterm and its associated n-tuple is lexicographically smallest amongst all minterms. We say that G is an *L-first sparse minterm* of

P if it is a minterm, is L-sparse and its associated n-tuple is lexicographically smallest amongst all L-sparse minterms. We define *R-first minterms* and *R-first sparse minterms* analogously.

Finally, we define the *dual* of a property $P \in \mathcal{P}_{n,n}$ to be the property $P^* \in \mathcal{P}_{n,n}$ such that a graph G satisfies P^* iff $\bar{G}$ does not satisfy P.

Lemma 2.4. *For any $P \in \mathcal{P}_{n,n}$ either P or P^* has an R-sparse minterm.*

Proof. Let G be an edge-maximal R-sparse (n,n)-bipartite graph. Then $\bar{G}$ is isomorphic to G; therefore G must satisfy either P or P^*. □

It is easy to see that any decision tree algorithm for P can be converted into one for P^*; this gives $\mathcal{C}^R(P) = \mathcal{C}^R(P^*)$. Therefore from now on we shall assume WLOG that G has an R-sparse minterm. The next theorem summarizes the key result of Yao [Y87] and an extension of the result due to Hajnal [H91].

Theorem 2.5 ([Y87,H91]). *For $P \in \mathcal{P}_{n,n}$, the following hold*
(1) *If G is a minterm of P then $\mathcal{C}^R(P) = \Omega(|E(G)|)$.*
(2) *If G is either an L-first minterm or an L-first sparse minterm, then*

$$\mathcal{C}^R(P) = \Omega(n\Delta_L(G)/\delta_L(G)) \ ,$$

and a similar statement holds for R-first minterms and R-first sparse minterms.
□

2.1 Graph Packing

We now introduce the key graph theoretic concept which we shall need. Let us say that graphs G and H *can be packed* if there is a way to identify their vertices without identifying any edge of G with an edge of H. Such an identification, when it exists, shall be called a *packing* of G and H. To see the relevance of this concept, consider the case when G and H are minterms of P and P^* respectively, for some property $P \in \mathcal{P}_{n,n}$. To say that G and H can be packed is equivalent to saying that G is isomorphic to a aubgraph of $\bar{H}$. Now from monotonicity and the definition of dual properties one can see that this gives rise to a contradiction.

These ideas are formalized in the next definition[3] and the following theorem.

Definition 2.6. *Let G and H be (m,n)-bipartite graphs. A* packing *of G and H is a pair of bijections $\varphi_L : V_L(G) \to V_L(H)$ and $\varphi_R : V_R(G) \to V_R(H)$ such that for any $x \in V_L(G)$ and $y \in V_R(G)$, either $(x,y) \notin E(G)$ or $(\varphi_L(x), \varphi_R(y)) \notin E(H)$. We say that G and H* can be packed *if there exists such a packing.*

Theorem 2.7 ([Y87]). *For $P \in \mathcal{P}_{n,n}$, let G be a minterm of P and H be a minterm of P^*. Then G and H cannot be packed.* □

[3] We have defined the notion of packing only for bipartite graphs here because that is all we need. In the literature, packing has been studied both for general graphs as well as bipartite graphs.

2.2 Outline of the Proof of the Main Theorem

We are now ready to outline the proof of Theorem 1.1. Let $P \in \mathcal{P}_{n,n}$ and let $q = q(n)$ be a parameter to be fixed later. We wish to prove that $\mathcal{C}^R(P) = \Omega(nq)$. Suppose this is not the case. Let G be an R-first sparse minterm of P and H be an L-first minterm of P^*. By part (1) of Theorem 2.5 the following conditions hold:

$$\delta(G) \leq q \ , \qquad \delta(H) \leq q \ .$$

Using these in part (2) of Theorem 2.5 gives us the following additional conditions:

$$\Delta_R(G) \leq q^2 \ , \qquad \Delta_L(H) \leq q^2 \ .$$

What we would like to show is that for an appropriate choice of q, these conditions imply that G and H can be packed. Then by Theorem 2.7 we would have a contradiction.

The above framework is the same as that used by Hajnal [H91]. Our improvement is in the parameters of the packing lemma. Our improved lemma says:

Lemma 2.8 (Packing Lemma). *Set $q = (\frac{1}{16}\varepsilon n \log n)^{1/3}$. Let G and H be (n,n)-bipartite graphs with $\delta(G) \leq q$, $\delta(H) \leq q$, $\Delta_R(G) \leq q^2$ and $\Delta_L(H) \leq q^2$. Furthermore, suppose G is R-sparse. Then if ε is a small enough constant, G and H can be packed.*

Remark. This is a stronger result than that of Hajnal and Szegedy [HS92]; this makes the lemma interesting on its own.

As outlined above, proving this lemma will establish that $\mathcal{C}^R(P) = \Omega(nq) = \Omega(n^{4/3} \log^{1/3} n)$ with the above choice of q. This will prove the Main Theorem 1.1. The rest of the paper will therefore be devoted to proving this lemma.

3 A Technical Lemma

The proof of our improved packing lemma will depend on a probabilistic fact: specifically, a tail inequality. We shall be interested in considering set systems with the property that a small random sample of the ground set is unlikely to hit too many of the sets in the system. More precisely, the property is that the number of sets *missed* by the sample is only a constant factor below the expected number, with high probability. The purpose of this section is to establish that certain simple conditions, if satisfied by the set system, guarantee this type of property. As it turns out, it suffices to upper bound the size of each set, and the maximum and average number of sets containing each element of the ground set; of course we also need a large enough set system over a large enough ground set. We shall call a set system *favourable* — with the appropriate parameters — if it satisfies these conditions.

Assumption 3.1 *Throughout this section we assume that* n *is large enough and that* $r(n)$*,* $s(n)$ *and* $t(n)$ *are integer valued functions in* $o(n) \cap \omega(1)$*.*

Definition 3.2. *Let* V *be a finite set and let* $\mathcal{F} \subseteq 2^V$ *be a set system on ground set* V*. We say that* $\mathcal{F}$ *is* $(n, r(n), s(n), \bar{s}(n))$-favourable *if*

$$|V| = n \ , \qquad |\mathcal{F}| \geq n \ ,$$

$$\forall F \in \mathcal{F} \quad |F| \leq r(n) \ ,$$

$$\forall v \in V \quad |\{F \in \mathcal{F} : v \in F\}| \leq s(n) \ , \quad \text{and}$$

$$\frac{1}{n} \sum_{v \in V} |\{F \in \mathcal{F} : v \in F\}| \leq \bar{s}(n) \ .$$

Now consider a set system $\mathcal{F}$ on ground set V, and a function $t(n)$. Let S be a random subset of V of size $t(n)$, chosen uniformly from all subsets of size $t(n)$. Our Technical Lemma is concerned with the behaviour of the following random variable:

$$X(\mathcal{F}; t(n)) = |\{F \in \mathcal{F} : F \cap S = \emptyset\}| \ . \tag{1}$$

Lemma 3.3 (Technical Lemma). *Let* $\mathcal{F}$ *be* $(n, r(n), s(n), \bar{s}(n))$*-favourable. Suppose* $r(n)t(n) \leq \frac{1}{4}\varepsilon n \log n$*, and* $t(n)s(n)\bar{s}(n) \leq n^{2-3\varepsilon}$ *for some constant* $\varepsilon > 0$*. Then we have*

$$\Pr\left[X(\mathcal{F}; t(n)) < \frac{1}{2} n^{1-\varepsilon}\right] \leq \frac{1}{n^2} \ .$$

Example. It may help to first think about a concrete example of a favourable set system and what the lemma says about it. Consider the ground set $V = \{1, 2, \ldots, n\}$ and let $\mathcal{F}$ be the collection of all n intervals of size $3n^{1/4}$ (say) with wrap-around (i.e. n and 1 are consecutive). Every interval is of size $3n^{1/4}$ and each point of the ground set belongs to $3n^{1/4}$ intervals. Therefore this $\mathcal{F}$ is $(n, 3n^{1/4}, 3n^{1/4}, 3n^{1/4})$-favourable. A straightforward calculation shows that a random subset of V of size $2n^{3/4}$ (say) is expected to be disjoint from $\Omega(n)$ of these intervals. From the above lemma we can conclude, in particular, that it is disjoint from $\Omega(n^{0.99})$ intervals, with "high" probability.

In order to prove Lemma 3.3, we define another random variable Y which "behaves like" $X(\mathcal{F}; t(n))$ and is easier to handle. Let us number the elements of V from 1 to n and set $p = 2t(n)/n$. Let $Z_1, \ldots, Z_n$ be i.i.d. boolean random variables with $\Pr[Z_i = 1] = p$. Let $S' \subseteq V$ be the random subset given by $S' = \{i : Z_i = 1\}$ and define

$$Y = |\{F \in \mathcal{F} : F \cap S' = \emptyset\}| \ . \tag{2}$$

The next lemma connects Y with X.

Lemma 3.4. *For any* α *we have* $\Pr[X(\mathcal{F}; t(n)) < \alpha] \leq 2 \cdot \Pr[Y < \alpha]$.

Proof. We proceed as in [H91]. For $0 \le k \le n$ let $\pi_k = \Pr[X(\mathcal{F};k) < \alpha]$. Observe that $\pi_0 \le \pi_1 \le \cdots \le \pi_n$. Let $A = \lfloor \frac{1}{2} np \rfloor$ and $B = \lfloor \frac{3}{2} np \rfloor$. We have

$$\Pr[Y < \alpha] = \sum_{k=0}^{n} \binom{n}{k} p^k (1-p)^{n-k} \pi_k \ge \pi_A \sum_{k=A}^{B} \binom{n}{k} p^k (1-p)^{n-k} \ge \frac{1}{2}\pi_A \ .$$

Noting that $A = t(n)$ completes the proof. □

Lemma 3.5. *Under the hypotheses of the Technical Lemma,* $E[Y] \ge n^{1-\varepsilon}$.

Proof. We have $E[Y] = \sum_{F \in \mathcal{F}} (1-p)^{|F|} \ge n(1-p)^{r(n)}$ using the fact that $\mathcal{F}$ is $(n, r(n), s(n), \bar{s}(n))$-favourable. For any constant $\alpha > 1$, we have $(1-p)^{1/p} \ge e^{-\alpha}$ for large enough n. Therefore

$$E[Y] \ge n \left((1-p)^{1/p} \right)^{2t(n)r(n)/n} \ge n^{1-\alpha\varepsilon/2\ln 2} \ ,$$

since by hypothesis $t(n)r(n) \le \frac{1}{4}\varepsilon n \log n$. Choosing $\alpha = 2\ln 2$ completes the proof. □

Now that we know that Y has high expectation, the main task is to prove that it does not fall far below its expectation too often. To this end we would like to consider an exposure martingale corresponding to Y that is obtained by revealing the values of Z_i one at a time. For $i = 0, 1, \ldots, n$, we define random variables $Y_i = Y_i(Z_1, \ldots, Z_i)$:

$$Y_i(z_1, \ldots, z_i) = E[Y | Z_1 = z_1, \ldots, Z_i = z_i] \ , \qquad \forall z_1, \ldots, z_i \in \{0,1\} \tag{3}$$

where the expectation is taken over $Z_{i+1}, \ldots, Z_n$. It is clear that $Y_{i-1}(z_1, \ldots, z_{i-1}) = (1-p)Y_i(z_1, \ldots, z_{i-1}, 0) + pY_i(z_1, \ldots, z_{i-1}, 1)$. Therefore, defining another set of random variables $D_i = D_i(Z_1, \ldots, Z_{i-1})$ by

$$\begin{aligned} D_i(z_1, \ldots, z_{i-1}) := Y_i(z_1, \ldots, z_{i-1}, 0) - Y_i(z_1, \ldots, z_{i-1}, 1) \ , \\ \forall z_1, \ldots, z_{i-1} \in \{0,1\} \end{aligned} \tag{4}$$

gives

$$\begin{aligned} Y_i(z_1, \ldots, z_i) &= \begin{cases} Y_{i-1}(z_1, \ldots, z_{i-1}) + pD_i(z_1, \ldots, z_{i-1}) \ , & \text{if } Z_i = 0 \\ Y_{i-1}(z_1, \ldots, z_{i-1}) - (1-p)D_i(z_1, \ldots, z_{i-1}) \ , & \text{if } Z_i = 1 \end{cases} \\ &= Y_{i-1} + (p(1-Z_i) - (1-p)Z_i)D_i \\ &= Y_{i-1} + (p - Z_i)D_i \ , \end{aligned}$$

whence

$$Y = E[Y] + \sum_{i=1}^{n} (p - Z_i) D_i \ . \tag{5}$$

To bound the random variables D_i, note that for any fixed $z_1, \ldots, z_{i-1}$, the quantity D_i is a convex combination of the quantities

$$Y(\underbrace{z_1, \ldots, z_{i-1}}_{\text{fixed}}, 0, \underbrace{z_{i+1}, \ldots, z_n}_{\text{variable}}) - Y(z_1, \ldots, z_{i-1}, 1, z_{i+1}, \ldots, z_n) \tag{6}$$

where $(z_{i+1},\ldots,z_n)$ varies over all tuples in $\{0,1\}^{n-i}$. From the definition of Y in (2), it is clear that each of the quantities (6) lies between 0 and d_i where

$$d_i := |\{F \in \mathcal{F} : i \in F\}| \ .$$

Therefore $0 \le D_i \le d_i$.

Lemma 3.6. *With $Z_1,\ldots,Z_n$ as above, let $c_1,\ldots,c_n$ be positive integers, let $C_i = C_i(Z_1,\ldots,Z_{i-1})$ be real functions satisfying $0 \le C_i \le c_i$ and let $\lambda > 0$ be an arbitrary real. Define $\Delta = \max_{i=1}^n c_i$ and $\delta = \max\{\lambda^2, \frac{1}{n}\sum_{i=1}^n c_i\}$. Then, if $pn \ge \Delta$ we have*

$$\Pr\left[\ \sum_{i=1}^{n}(p - Z_i)C_i \ < \ -\lambda\sqrt{pn\delta\Delta}\,\log n\ \right] \le n^{-\lambda^2/6} \ .$$

Remark. The lemma is interesting because (1) we are summing *dependent* random variables and (2) it is a martingale inequality that does not follow from Azuma's inequality.

Proof. To simplify the proof, we assume that the C_i's are integer-valued; this assumption can easily be removed by discretizing with denominator n^2 and rescaling. For $1 \le i \le n$, $1 \le j \le d_i$, define random variables W_{ij} as follows

$$W_{ij} = \begin{cases} p - Z_i \ , & \text{if } j \le C_i \\ 0 \ , & \text{if } j > C_i \ . \end{cases}$$

The key observation is that the nonzero W_{ij}'s for distinct i are independent, because the Z_i's are. Therefore, if we set $\ell = n\delta/\Delta$, the W_{ij}'s can be arranged in an $\ell \times \Delta$ matrix such that the nonzero random variables in each column are independent.[4] Fix a column and discard the zero entries; suppose m entries remain and sum to S. Standard Chernoff bounds (e.g. see [ASE00], Theorems A.11 and A.12) imply that for any $a > 0, b > 1$:

$$\Pr[S < -a] < \exp\left(-\frac{a^2}{2pm} + \frac{a^3}{2p^2m^2}\right) \ , \quad \text{and} \tag{7}$$

$$\Pr[S < -(b-1)pm] < \left(e^{b-1}b^{-b}\right)^{pm} \ . \tag{8}$$

Set $a = \lambda\sqrt{p\ell}\log n$ and $b = 1 + a/pm$. Suppose $a \le \frac{2}{3}pm$. From (7) we immediately obtain

$$\Pr[S < -\lambda\sqrt{p\ell}\log n] < \exp(-\frac{1}{6}\lambda^2\log^2 n) \le \exp(-\frac{1}{6}\lambda^2\log n) \ .$$

Next suppose $a > \frac{2}{3}pm$. Since the real function $f(x) = (x-1)^{-1}x\ln x$ is increasing for $x \ge 1$, from (8) we get

$$\Pr[S < -\lambda\sqrt{p\ell}\log n] < \exp\left(-\lambda\sqrt{\frac{pn\delta}{\Delta}}\,(f(5/3)-1)\log n\right) < \exp(-\frac{1}{6}\lambda^2\log n) \ ,$$

[4] We pad the matrix with extra zero entries, if necessary, to end up with the required shape of $\ell \times \Delta$.

where the last inequality from the facts that $pn \geq \Delta$ and that $\delta \geq \lambda^2$. Thus, in each case we have $\Pr[S < -\lambda\sqrt{p\ell}\log n] < n^{-\lambda^2/6}$.

To finish the proof we simply note that $\sum_{i=1}^{n}(p - Z_i)C_i$ is the sum of Δ such random variables S. □

We are now ready to prove our technical lemma.

Proof. **(of Lemma 3.3)** Let us apply Lemma 3.6 with $c_i = d_i$ and $C_i = D_i$. This choice of parameters gives $\Delta \leq s(n)$ and $\delta \leq \bar{s}(n)$. Because of the way we defined Y in (2), increasing p can only increase the quantity $\Pr[Y < \alpha]$, for any α; thus we may safely assume than $pn \geq \Delta$. Recalling that $p = 2t(n)/n$ we get

$$\Pr\left[Y < E[Y] - \lambda\sqrt{2t(n)s(n)\bar{s}(n)}\right] < n^{-\lambda^2/6} \ .$$

Recall that by hypothesis $t(n)s(n)\bar{s}(n) \leq n^{2-3\varepsilon}$. Using Lemmas 3.4 and 3.5 we then get

$$\Pr\left[X < n^{1-\varepsilon} - \lambda\sqrt{2n^{2-3\varepsilon}}\right] < 2n^{-\lambda^2/6} \ .$$

Setting $\lambda > \sqrt{12}$ yields $\Pr[X < \frac{1}{2}n^{1-\varepsilon}] < n^{-2}$ as desired, when n is large enough. □

4 Proof of the Packing Lemma

We now return to proving our improved packing lemma (Lemma 2.8). Recall that from the hypotheses we already have the following degree conditions on the bipartite graphs G and H we wish to pack:

$$\delta(G) \leq q \ , \quad \delta(H) \leq q \ , \quad \Delta_R(G) \leq q^2 \ , \quad \Delta_L(H) \leq q^2 \ , \tag{9}$$

where we have set

$$q = \left(\frac{\varepsilon}{16} n \log n\right)^{1/3} \ , \tag{10}$$

where ε is a small constant to be fixed later. We shall assume throughout this section that n is large enough.

Definition 4.1. *For a subset W of the vertex set of a graph and integer $k \leq |W|$, let $\mathcal{N}(W)$ denote the neighbourhood of W. Let $\mathrm{top}(W,k)$ and $\mathrm{bot}(W,k)$ denote the subsets of W consisting of, respectively, the k highest and k lowest degree vertices in W. For a vertex x, let $\mathcal{N}(x)$ be defined as $\mathcal{N}(\{x\})$.*

Following Hajnal [H91], our first step will be to modify G and H suitably so that even stronger degree conditions hold. Let $k = \min\{n/2, n/4\delta(H)\}$. From the hypotheses, we know that $V_R(G)$ has at least $n/2$ isolated vertices; let V_1 be a set of size $n/2$ of these. Let $V_0 = \mathrm{top}(V_L(G), k)$, $V_2 = \mathrm{bot}(V_L(H), k)$ and $V_3 = \mathcal{N}(V_2) \cup \mathrm{top}(V_R(H), \frac{n}{2} - |\mathcal{N}(V_2)|)$. Let us define graphs G' and H' as follows:

$$G' = G - (V_0 \cup V_1) \ ; \qquad H' = H - (V_2 \cup V_3) \ .$$

It follows from the construction above that if G' and H' can be packed then so can G and H. This is because having packed G' and H' we may *arbitrarily* identify the vertices in V_0 with those in V_2 and the vertices in V_1 with those in V_3. Now, to show that G' and H' can be packed, we shall need the degree conditions guaranteed by the following lemma.

Lemma 4.2. *G' and H' are $(n-k, n/2)$-bipartite graphs with the following properties:*

$$\delta_L(G') \le q \ , \qquad \delta_R(G') \le q \ , \qquad \delta_L(H') \le q \ , \qquad \delta_R(H') \le q \ ,$$

$$\Delta_L(G') \le 4q^2 \ , \qquad \Delta_R(G') \le q^2 \ ,$$

$$\Delta_L(H') \le q^2 \ , \qquad \Delta_R(H') \le 4q \ .$$

Proof. The first four inequalities are obvious from (9), as are the bounds on $\Delta_R(G')$ and $\Delta_L(H')$. By construction $|\mathcal{N}(V_2)| \le \sum_{v \in V_2} \deg_H(v) \le n/4$; therefore V_3 contains at least $n/4$ of the highest degree vertices in $V_R(H)$. Since these vertices are removed to obtain H' we have $\Delta_R(H') \le 4\delta(H) \le 4q$. Similarly, we have $\Delta_L(G') \le 4\delta(H)\delta(G) \le 4q^2$. □

We prove that G' and H' can be packed using the probabilistic method: let $\varphi_L : V_L(G') \to V_L(H')$ be a random bijection; we shall show that with positive probability there exists a bijection $\varphi_R : V_R(G') \to V_R(H')$ such that (φ_L, φ_R) is a packing. Let $\Gamma = \Gamma(\varphi_L)$ be a bipartite graph on vertex set $(V_R(G), V_R(H))$ defined as follows: for $x \in V_R(G'), y \in V_R(H')$ we have $(x, y) \in E(\Gamma)$ iff $\varphi_L(\mathcal{N}(x)) \cap \mathcal{N}(y) = \emptyset$. It is clear that the required bijection φ_R exists iff the graph Γ has a perfect matching. Our task now is to show that the (random) bipartite graph Γ has a perfect matching with positive probability. The most straightforward way of doing this is to obtain lower bounds on the degrees of vertices in Γ and then apply König's Theorem.

The next two lemmas establish such lower bounds. It is important to note that unlike [H91] we exploit the asymmetry between G' and H' in a crucial way; the degree lower bound for H' is proved along lines similar to [H91], whereas for G' we need the power of our Technical Lemma.

Lemma 4.3. *With probability greater than $\frac{1}{2}$, for every vertex $y \in V_R(H')$ we have $\deg_\Gamma(y) \ge \frac{n}{2} - 8q^2$.*

Proof. Let $y \in V_R(H')$ be arbitrary and let $S(y) = \varphi_L^{-1}(\mathcal{N}(y))$. Then $S(y)$ is a random subset of $V_L(G')$ of size at most $\Delta_R(H') = 4q$. Since $\delta(G') \le q$, we have

$$E\left[\sum_{v \in S(y)} \deg_{G'}(v) \right] \le 4q^2 \ .$$

This bound on the expectation implies a high probability result proven through Chernoff bounds in exactly the same manner as Lemma 5.4 in [H91]; we need to have $\Delta_L(G') = O(n/\log n)$, but this is indeed the case by Lemma 4.2 and our choice of q in (10). Therefore, we get

$$\Pr\left[\sum_{v \in S(y)} \deg_{G'}(v) > 8q^2\right] < \frac{1}{2n} \ .$$

Thus $\Pr[\,|\mathcal{N}(S(y))| > 8q^2] < \frac{1}{2n}$ and so $\Pr[\deg_\Gamma(y) \geq \frac{n}{2} - 8q^2] < \frac{1}{2n}$. Since $|V_R(H')| \leq n$, the lemma follows. □

Lemma 4.4. *With probability greater than $\frac{1}{2}$, for every vertex $x \in V_R(G')$ we have $\deg_\Gamma(x) \geq 8q^2$.*

Proof. Fix $\varepsilon = \frac{1}{10}$. Let $x \in V_R(G')$ be arbitrary and let $T(x) = \varphi_L(\mathcal{N}(x))$. Consider the set system $\mathcal{H} = \{\mathcal{N}(y) : y \in V_R(H')\}$ on ground set $V_L(H')$. By Lemma 4.2 we see that $\mathcal{H}$ is $(\frac{n}{2}, 4q, q^2, q)$-favourable and by (10) we have $4q \cdot q^2 = \frac{1}{4}\varepsilon n \log n$. Now $T(x)$ is a random subset of $V_L(H')$ of size $|T(x)| \leq \Delta_R(G') \leq q^2$ and $q^2 \cdot q^2 \cdot q = q^5 \leq n^{2-3\varepsilon}$. Therefore, $\mathcal{H}$ and $T(x)$ satisfy the hypotheses of the Technical Lemma 3.3.

Applying the Technical Lemma, we see that the number of sets in $\mathcal{H}$ that $T(x)$ is disjoint from falls below $\frac{1}{2}n^{1-\varepsilon}$ with probability at most $\frac{1}{n^2} < \frac{1}{2n}$. In other words $\Pr[\deg_\Gamma(x) < \frac{1}{2}n^{1-\varepsilon}] < \frac{1}{2n}$. Noting that $\frac{1}{2}n^{1-\varepsilon} \geq 8q^2$ gives us the desired result. □

We now have all the pieces we need to prove the packing lemma.

Proof. **(Of the Packing Lemma)** From Lemma 4.3 and Lemma 4.4 we see that if the bijection φ_L is chosen at random, then with positive probability the following event occurs:

$$\forall x \in V_R(G') \quad \deg_\Gamma(x) \geq \frac{n}{2} - 8q^2 \ , \quad \text{and} \quad \forall y \in V_R(H') \quad \deg_\Gamma(y) \geq 8q^2 \ .$$

Since Γ is an $(n/2, n/2)$-bipartite graph, by König's Theorem, this event is a sufficient condition for Γ to have a perfect matching. Therefore, there exists a bijection φ_L such that Γ has a perfect matching; thus there exists a packing of G' and H'. By the discussion preceding Lemma 4.2 we see that G and H can be packed. □

5 Concluding Remarks

The complexity of graph properties has been studied for a quarter of a century now. Yet the most basic conjecture, namely an $\Omega(n^2)$ randomized decision tree complexity lower bound for monotone properties, remains open to this day. As mentioned before, the number of results leading towards a settlement of this conjecture have been very few.

Nine years have passed since the best previously known lower bound was established. We believe that this makes our result, a slight improvement of the bound, significant for injecting new life into this problem.

To improve the lower bound further it appears necessary to break out of Hajnal's framework. Our Technical Lemma is not constrained by this framework — instead, Hajnal's framework constrains the parameters we are forced to apply the lemma with — and we hope that it will be useful in further work on this problem.

6 Acknowledgments

We are grateful to Professor Andrew Yao for introducing us to this fascinating problem. We would like to thank Professor Bernard Chazelle for several important comments and suggestions.

References

[ASE00] Alon, N., Spencer, J. H. *The probabilistic method, Second Edition*, Wiley-Interscience Series, 2000, Appendix A.

[B78] Bollobás, B. *Extremal Graph Theory*, Academic Press, 1978, Chapter 8.

[G92] Gröger, H. D. *On the randomized complexity of monotone graph properties*, Acta Cybernetica, **10** (1992), 119–127.

[H91] Hajnal, P. *An* $\Omega(n^{\frac{4}{3}})$ *lower bound on the randomized complexity of graph properties*, Combinatorica, **11** (1991), 131–143.

[HS92] Hajnal, P., Szegedy, M. *On packing bipartite graphs*, Combinatorica, **12** (1992), 295–301.

[K88] King, V. *Lower bounds on the complexity of graph properties*, Proc. 20th ACM STOC (1988), 468–476.

[RV76] Rivest, R.L., Vuillemin, J. *On recognizing graph properties from adjacency matrices*, Theoret. Comput. Sci., **3** (1976), 371–384.

[Y87] Yao, A.C. *Lower bounds to randomized algorithms for graph properties*, Proc. 28th IEEE FOCS (1987), 393–400.