# Coloring d-Degenerate Graphs Equitably

A.V. Kostochka\* K. Nakprasit<sup>†</sup> S.V. Pemmaraju<sup>‡</sup>

#### Abstract

An equitable coloring of a graph is a proper vertex coloring such that the sizes of any two color classes differ by at most 1. A d-degenerate graph is a graph G in which every induced subgraph has a vertex with degree at most d. It is well known that trees are 1-degenerate, outerplanar graphs are 2-degenerate, and planar graphs are 5-degenerate. The results in this paper concern the problem of equitable coloring of d-degenerate graphs with few colors.

Our first result shows that every d-degenerate graph can be equitably partitioned into three (d-1)-degenerate graphs. Repeated application of this result implies that every n-vertex d-degenerate graph G with  $\Delta(G) \leq n/3^d$  can be equitably  $3^d$ -colored. Then we show that every n-vertex d-degenerate graph G with  $\Delta(G) \leq n/15$  can be equitably k-colored for any  $k \geq 16d$ . The proof of this bound is constructive and implies an O(d)-factor approximation algorithm for equitable coloring with fewest colors any n-vertex d-degenerate graph G with  $\Delta(G) \leq n/15$ . We then extend this to an O(d)-factor approximation algorithm for equitably coloring any d-degenerate graph. Among the implications of this result is the first O(1)-factor approximation algorithm for equitable coloring planar graphs with minimum number of colors. These results have applications in improved Chernoff-Hoeffding bounds for sums of random variables with limited dependence and to partitioning problems such as MAX p-SECTION.

### 1 Introduction

An equitable coloring of a graph is a proper vertex coloring such that the sizes of any two color classes differ by at most 1. Equitable colorings naturally arise in some scheduling, partitioning, and load balancing problems [16]. In contrast with ordinary coloring, a graph may have an equitable k-coloring (i.e., an equitable coloring with k colors) but have no equitable (k + 1)-coloring. The equitable chromatic number of a graph G, denoted  $\chi_{eq}(G)$ , is the smallest k such that G is equitably k-colorable.

The deep result of Hajnal and Szemerédi [4] from 1970 says that for every graph G and any  $k \geq \Delta(G) + 1$ , G has an equitable k-coloring. In its "complementary" form this results concerns the decomposition of a sufficiently dense graph into cliques of equal size and this has been used in a number of applications of Szemerédi's Regularity Lemma [8]. The Hajnal-Szemerédi theorem implies the upper bound  $\chi_{eq}(G) \leq \Delta(G) + 1$ . The bound of the Hajnal-Szemerédi theorem is sharp, but it can be improved for some important classes of graphs. Consider the class of trees. For the star graph,  $S_n$ , that is an n-vertex tree in which one vertex is adjacent to the rest of the vertices,  $\chi_{eq}(S_n) = 1 + \lceil (n-1)/2 \rceil$ . This turns out to be the worst case for trees because of Meyer's bound  $\lceil 13 \rceil \chi_{eq}(T) \leq 1 + \lceil \Delta(T)/2 \rceil$  for every tree T. Following Meyer's result, one direction of research in equitable colorings has been to obtain upper bounds better than  $1 + \Delta(G)$  on  $\chi_{eq}(G)$  for G in various classes of graphs not containing  $K_{\Delta+1}$ . Such bounds are known for bipartite graphs  $(\Delta)$ , proved by

<sup>\*</sup>Department of Mathematics, The University of Illinois, Urbana, IL 61801, and Institute of Mathematics, Novosibirsk, Russia, kostochk@math.uiuc.edu; this was partially supported by the NSF grant DMS-0099608 and by the grants 02-01-00039 and 00-01-00916 of the Russian Foundation for Basic Research.

<sup>†</sup>Department of Mathematics, The University of Illinois, Urbana, IL 61801, nakprasit@math.uiuc.edu

<sup>&</sup>lt;sup>‡</sup>Department of Computer Science, The University of Iowa, sriram@cs.uiowa.edu; this was partially supported by the NSF grant DMS-0213305

Lih and Wu [12]), outerplanar graphs  $(1 + \lceil \Delta/2 \rceil)$ , proved in [9]), and planar graphs  $(3 + \Delta/2)$  for sufficiently large  $\Delta$ , proved in [10]). Lih [11] provides a nice survey of this direction of research.

Another interesting and important direction of research for equitable colorings was initiated by Bollobás and Guy [3]. They showed that while  $1 + \lceil \Delta/2 \rceil$  is a tight upper bound on the equitable chromatic number of trees, "most" trees can be equitably 3-colored. Specifically, their result implies that any n-vertex forest F with  $\Delta(F) \leq n/3$  can be equitably 3-colored. This result seems to uncover a fundamental phenomenon in equitable colorings: apart from some "star-like" graphs, most of graphs admit equitable colorings with few colors. Another example of this phenomenon appears in [15], where the Bollobás-Guy result was extended to outerplanar graphs. Specifically, [15] showed that any n-vertex outerplanar graph G with  $\Delta(G) \leq n/6$  can be equitably 6-colored.

In this paper we show that this phenomenon is widely pervasive. A d-degenerate graph is a graph G in which every induced subgraph has a vertex with degree at most d. It is well known that forests are exactly 1-degenerate graphs, outerplanar graphs are 2-degenerate, and planar graphs are 5-degenerate. By the definition, the vertices of every d-degenerate graph can be ordered  $v_1, \ldots, v_n$ in such a way that for every i > 2, vertex  $v_i$  has at most d neighbors  $v_i$  with j < i. A d-degenerate graph can be colored (not necessarily, equitably) with d+1 colors using a greedy algorithm. This upper bound is tight because  $K_{d+1}$  is a d-degenerate graph. Defining the chromatic number of a class of graphs as the maximum chromatic number of a graph in the class, we can then assert that the chromatic number of the class of d-degenerate graphs is d+1. Roughly speaking, the main result in this paper shows that the equitable chromatic number of the class of d-degenerate graphs with "star-like" graphs deleted, is O(d). Furthermore, using techniques from the proof of this result, we show how to equitably color any d-degenerate graph (including the "star-like" graphs that were excluded earlier), with the number of colors used being O(d) times the equitable chromatic number. For any fixed d (for example, for planar graphs we would set d=5), this gives a constant-factor approximation algorithm for equitably coloring any d-degenerate graph using minimum number of colors. The reader should note that equitable versions of NP-complete minimum coloring problems are also NP-complete (a simple reduction involving adding isolated vertices suffices). So for example, determining if a given planar graph with maximum vertex degree 4 has an equitable coloring with 3 or fewer colors is NP-complete.

One motivation for understanding equitable colorings comes from the use of equitable colorings in the derivation of deviation bounds for sums of dependent random variables that exhibit limited dependence [14]. The basic idea there was to color equitably the dependency graph of the variables involved in the sum and use the Chernoff-Hoeffding bound for the variables associated with each color class. This connection was independently observed by Ruciński and others [6] and in [7], Janson and Ruciński compare bounds obtained by this "break-up" method to those obtained by other more sophisticated techniques. Subsequently, Janson [5] explores further the use of the Hajnal-Szemerédi theorem to obtain equitable colorings with applications to U-statistics, random strings, and random graphs. In all of these applications, the fewer colors we use, the better the deviation bound is. In fact, Pemmaraju [14] notes that if the dependency graph of a set of random variables can be colored with a constant number of colors, then the deviation bounds we get for the sum of these random variables are essentially what we would have obtained had we assumed independence among the random variables and used Chernoff-Hoeffding bounds. Since the upper bounds on the equitable chromatic number proved in this paper for d-degenerate graphs are better than those provided by the Hajnal-Szemerédi theorem, we get correspondingly better deviation bounds. Consider for example the case of planar graphs. We show that "most" planar graphs can be equitably colored with a constant number of colors and furthermore we show that if we are allowed to delete a constant number of vertices from any planar graph, then the remaining graph can be colored with a constant number of colors. This means that for a set of random variables whose dependency graph is planar, we obtain deviation bounds on their sum that are almost as good as those obtained had the random variables been mutually independent.

MAX p-SECTION is the problem that takes as input an edge weighted graph G and a positive integer p and produces a partition of V(G) into p subsets of equal size such that the total weight of edges connecting different parts is maximized. Recently Andersson [1] used semidefinite pro-

gramming to obtain a  $\left(\frac{p-1}{p} + \Theta(p^{-3})\right)$ -factor approximation algorithm for MAX p-SECTION. Our results imply alternate combinatorial algorithms for the MAX-p-SECTION problem for d-degenerate graphs and might lead to a better approximation factor for planar graphs.

We have also been able to extend some of the results in this paper to list colorings. An list analogue of equitable colorings was introduced in [2].

Now, we describe the main results more precisely. An equitable k-partition of a graph G is the collection of subgraphs  $\{G[V_1], G[V_2], ..., G[V_k]\}$  of G induced by the vertex partition  $\{V_1, V_2, ..., V_k\}$  of V(G) where, for any pair  $V_i$  and  $V_j$ , the sizes of  $V_i$  and  $V_j$  differ by at most 1. Our first result is the following.

**Theorem 1.** Let  $k \geq 3$  and  $d \geq 2$ . Then every d-degenerate graph has an equitable k-partition into (d-1)-degenerate graphs.

This is an extension of the Bollobás-Guy result [3] which essentially asserts the same for d=1 and k=3. By fixing k=3 in the above theorem and recursively applying it for  $d \geq 2$  and applying the Bollobás-Guy Theorem for d=1, we get the following result.

Corollary 1. For  $d, n \ge 1$ , every d-degenerate, n-vertex graph G with  $\Delta(G) \le n/3^d$  can be equitably  $3^d$ -colored.

This guarantees a  $3^d$ -equitable coloring for d-degenerate graphs that are not "star-like", where "star-like" refers to any n-vertex graph G with  $\Delta(G) > n/3^d$ . In our main result, we strengthen Corollary 1 to the following.

**Theorem 2.** For  $d, n \geq 1$ , every d-degenerate, n-vertex graph G with  $\Delta \leq n/15$  is equitably k-colorable for any  $k \geq 16d$ .

The proof of this result is constructive and provides an O(d)-factor approximation algorithm for equitable coloring with fewest colors of each d-degenerate n-vertex graph G with  $\Delta \leq n/15$ . Finally, we extend this to all d-degenerate graphs and show the following.

**Theorem 3.** There exists a polynomial time algorithm that for every equitably s-colorable d-degenerate graph G produces an equitable k-coloring of G for each  $k \geq 31ds$ .

# 2 Equitable partitions of d-degenerate graphs

Every d-degenerate graph G admits a d-degenerate vertex ordering, that is, an ordering  $v_1, v_2, \ldots, v_n$  of the vertices of G such that each vertex  $v_i$  has at most d neighbors in  $\{v_1, v_2, \ldots, v_{i-1}\}$ . A d-degenerate ordering of G can be constructed by picking a vertex v with smallest degree and appending it to a degenerate ordering of G - v. It is easy to see that any d-degenerate graph G can be partitioned into two (d-1)-degenerate graphs: construct a degenerate ordering and color the vertices in this order, red or blue, using the rule that a vertex v is colored red if it has less than d red neighbors; otherwise color v blue. While this procedure leads to a partition into (d-1)-degenerate graphs, this partition needs not be equitable. In fact, the only partition of a star graph (which is 1-degenerate) into two independent sets (which are 0-degenerate) is the one in which one part has one vertex and the other has the rest. In this section we show that if  $d \geq 2$  and we allow for a third part, we can provide equitability. This extends the Bollobás-Guy result [3] to arbitrary  $d \geq 2$  and also gives a tool to get equitable colorings that use few colors. Specifically, we will prove the following fact.

**Theorem 1.** Let  $k \geq 3$  and  $d \geq 2$ . Then every d-degenerate graph can be equitably partitioned into k (d-1)-degenerate graphs.

*Proof.* We prove the result by contradiction, assuming that the above claim is false. Let G be a smallest (with respect to the number of vertices) counterexample to the theorem. Let n = |V(G)|. Then n > dk, because otherwise, any equitable vertex partition is good enough. A simple observation that forms the basis of the proof is the following.

Claim 1. Let  $v_1, v_2, \ldots, v_m$  be a d-degenerate vertex ordering of a d-degenerate graph H. If  $H - v_m$  has a k-partition  $(W_1, \ldots, W_k)$  where every  $W_i$  induces a (d-1)-degenerate subgraph, then among  $W_1 + v_m, \ldots, W_k + v_m$  at most one is not (d-1)-degenerate. Furthermore, if  $W_i + v_m$  is not (d-1)-degenerate, then  $v_m$  has d neighbors and  $W_i$  contains all d neighbors of  $v_m$ .

*Proof.* By the definition of a d-degenerate vertex ordering, the degree of  $v_m$  is at most d. If  $W_i$  has fewer than d neighbors of  $v_m$ , then we can append  $v_m$  to a (d-1)-degenerate ordering of  $W_i$ .  $\square$ 

#### Claim 2. The minimum degree of G is d and n is divisible by k.

Proof. Suppose that  $n=k\cdot s+r$ , where  $1\leq r\leq k$ . We can choose a degenerate ordering of G such that the last vertex in the ordering,  $v_n$ , is a vertex of minimum degree. By the minimality of G, there exists an equitable k-partition  $(W_1,\ldots,W_k)$  of  $V(G)-v_n$  into sets inducing (d-1)-degenerate graphs. Note that exactly r-1 of these sets have size s+1 and remaining k-r+1 sets are of size s. Since  $k-r+1\geq 1$ , there is at least one  $W_i$  of size s. If  $\deg_G(v_n)\leq d-1$ , then adding  $v_n$  to any set  $W_i$  of size s creates the desired equitable k-partition of G. This contradicts the choice of G and so we have that  $\deg_G(v_n)\geq d$ .

If k does not divide n, then we have r < k. This implies that there are  $k - r + 1 \ge 2$  sets of size s and by Claim 1, we can add  $v_n$  to at least one of these sets of size s. Again, this contradicts the choice of G as a minimal counterexample and implies that k divides n.

Given a vertex ordering  $R = \{v_1, \ldots, v_n\}$  of a graph H and an edge  $e = v_i v_j \in E(H)$ , we denote  $l_R(e) = i$  and  $r_R(e) = j$  if i < j. Among all d-degenerate orderings of V(G) choose a special ordering  $U = (u_1, \ldots, u_n)$  where the maximum index  $l_U(e)$  of an edge  $e \in E(G)$  is maximized. Let  $i_0$  be the maximum of  $l_U(e)$  over all the edges in the special ordering U. For convenience, we use  $U_i$  to denote the set  $\{u_i, u_{i+1}, \ldots, u_n\}$  for each  $i, 1 \le i \le n$ .

Claim 3. The vertex  $u_{i_0}$  is adjacent to  $u_i$  for every  $i_0 < i \le n$ , and the set  $U_{i_0+1}$  is independent.

Proof. The second part of the claim is directly implied by the definition of  $i_0$ . Suppose that some  $u_j$ , for some  $j > i_0$  is not adjacent to  $u_{i_0}$ . Then all the neighbors of  $u_j$  are in  $V(G) - U_{i_0}$ . So moving  $u_j$  from its current position to just before  $u_{i_0}$  creates another d-degenerate ordering of V(G). In this ordering the maximum index of the left end of an edge is  $i_0 + 1$ , a contradiction to the choice of the special ordering U.

Now we are ready to prove the theorem.

CASE 1.  $i_0 \geq n-k+1$ . Let  $G' = G - U_{n-k+1}$ . By the minimality of G, V(G') has an equitable partition  $(W_1, \ldots, W_k)$  into sets inducing (d-1)-degenerate graphs. Now we attempt to consecutively add  $u_{n-k+1}, u_{n-k+2}, \ldots, u_n$  (in this order) so that (a) we add one vertex to every set and (b) every new set still induces a (d-1)-degenerate graph. For vertices  $u_{n-k+1}, u_{n-k+2}, \ldots, u_{n-1}$  we can do this by Claim 1. Suppose that after adding vertices  $u_{n-k+1}, u_{n-k+2}, \ldots, u_{n-1}, W_i$  is the only set to which no vertex has been added. The trick with  $u_n$  is that one of its neighbors is  $u_{i_0}$  which has already been added to a set different from  $W_i$ . Thus  $u_n$  has at most (d-1) neighbors in  $W_i$  and therefore after adding  $u_n$  to  $W_i$ , it still induces a (d-1)-degenerate graph.

CASE 2.  $i_0 \leq n-k$ . Let  $G'' = G - U_{i_0}$ . By the minimality of G, V(G'') has an equitable partition  $(W_1, \ldots, W_k)$  into sets inducing (d-1)-degenerate graphs. For  $i > i_0$  call a set  $W_\ell$ ,  $1 \leq \ell \leq k$  i-incompatible, if all d-1 neighbors of  $u_i$  different from  $u_{i_0}$  are in  $W_\ell$ . By Claim 1, for every  $i > i_0$ , there could be at most one i-incompatible set. However, a set  $W_\ell$  may be i-incompatible for several i. By Claim 1,  $u_{i_0}$  can be added to any one of at least k-1 sets among the  $W_i$ 's. Let  $S = \{W_i \mid 1 \leq i \leq k \text{ and } u_{i_0} \text{ can be added to } W_i\}$ . There exists some set  $W_{\ell'} \in S$  such that  $W_{\ell'}$  is i-incompatible with at most  $(n-i_0)/|S|$  values of  $i > i_0$ . Since  $k \geq 3$ ,  $|S| \geq 2$  and so  $(n-i_0)/|S| \leq (n-i_0)/2$ . Now add  $u_{i_0}$  to  $W_{\ell'}$ . Any  $u_i$ ,  $i > i_0$ , for which  $W_{\ell'}$  is i-incompatible, can be added to any set other than  $W_{\ell'}$ . So distribute such  $u_i$ 's among sets other than  $W_{\ell'}$  so that the sizes of new sets don't exceed s = n/k. The remaining  $u_i$ 's can be added to any set. Thus, we add these in an arbitrary way so that the size of every  $W_\ell$  becomes s = n/k.

**Algorithm.** The algorithm implied by the above proof is sketched here; the correctness of the algorithm follows from the proof. An equitable k-partition of a given n-vertex graph G is constructed recursively. If G contains a vertex of degree less than d or if n is not divisible by k, we construct a d-degenerate ordering of G, and assuming that v is the last vertex in this ordering, construct an equitable k-partition of G-v and then add v to one of the k sets. Otherwise, we construct a special d-degenerate ordering U of G, referred to in the proof, as follows. Let  $L_0$  be the set of vertices in G with degree at most d. If  $L_0$  contains a pair of adjacent vertices, say u and v, then U is obtained by constructing an arbitrary d-degenerate ordering of G-u-v and appending u and v to this. Otherwise, let  $L_1$  be the set of vertices in  $G-L_0$  with degree at most d. By the definition, every vertex in  $L_1$  has a neighbor in  $L_0$ . Find a vertex  $v \in L_1$  with fewest neighbors in  $L_0$ . Let S denote the set of neighbors of v in  $L_0$ . U is obtained by constructing an arbitrary d-degenerate ordering of G-v-S and appending v followed by vertices in S to this. Once U is constructed, we determine whether Case 1 (respectively, Case 2) of the proof applies and accordingly construct an equitable k-partition of  $G' = G - U_{n-k+1}$  (respectively,  $G'' = G - U_{i_0}$ ) and add vertices in  $U_{n-k+1}$ (respectively,  $U_{i_0}$ ) to the sets in the partition. It is easy to see that  $O(n^2)$  time suffices for algorithm, though it seems likely that with more care this can be implemented in subquadratic time.

Theorem 1 and the equitable k-partitioning algorithm described above can be used to get an equitable coloring of a given d-degenerate graph G. Let  $S = \{G\}$ . In each stage, replace each of the graphs in S by the 3 subgraphs obtained by equitably 3-partitioning it. It is easy to verify that after t stages, S contains  $3^t$  (d-t)-degenerate subgraphs, each containing either  $\lfloor n/3^t \rfloor$  or  $\lfloor n/3^t \rfloor$  vertices. So this is an equitable  $3^t$ -partition of G into (d-t)-degenerate graphs. Repeat this for t=d-1 stages to get  $3^{d-1}$  1-degenerate graphs (forests) and then use the Bollobás-Guy Theorem [3] to 3-color equitably each of these forests, to get an equitable  $3^d$ -coloring of G. Since the Bollobás-Guy Theorem can only be applied to an n-vertex forest F with  $\Delta(F) \leq n/3$ , we require that the maximum degree of any of the subgraphs obtained after d-1 stages be at most  $n/3^d$ . Stipulating that  $\Delta(G) \leq n/3^d$  ensures this condition and we obtain the following.

Corollary 1. For  $d, n \ge 1$ , every d-degenerate, n-vertex graph G with  $\Delta(G) \le n/3^d$  can be equitably  $3^d$ -colored.

## 3 Coloring d-degenerate graphs with O(d) colors

The significance of Corollary 1 is that it extends to various classes of graphs (including planar graphs) what is known for trees and outerplanar graphs: a constant number of colors suffice to color equitably each graph in the class, apart from some "star-like" graphs. In this section, we sharpen Corollary 1 by reducing the number of colors used from  $3^d$  to O(d) and by claiming far fewer graphs being "star-like".

Given a graph G, we say that an ordering  $v_1, v_2, \ldots, v_n$  of the vertices of G is *greedy* if for every  $i, v_i$  has the highest degree in  $G - v_1 - \cdots - v_{i-1}$ .

**Theorem 2.** Every d-degenerate graph with maximum degree at most  $\Delta$  is equitably k-colorable when  $k \geq 16d$  and  $n \geq 15\Delta$ .

*Proof.* Let G be a d-degenerate graph with vertex set V of size n and edge set E(G). Let  $k(t-1) < n \le kt$  and  $k \ge 16d$ .

CASE 1. t < 15. This is the simple case and its proof can be found in the appendix.

CASE 2.  $t \ge 16$ . This is the more interesting and significantly more difficult case. Let  $t = \beta_1 4^m + \beta_2 4^{m-1} + \ldots + \beta_{m+1}$  where  $\beta_j$  is an integer,  $0 \le \beta_j \le 3$ . For  $i = 1, 2, \ldots, m+1$ , define  $l_i = \beta_1 4^{i-1} + \beta_2 4^{i-2} + \ldots + \beta_i$ . For notational convenience let  $l_0 = 0$ . We have that  $l_i = 4l_{i-1} + \beta_i$  for each  $i = 1, 2, \ldots, m+1$  and also that  $t = l_{m+1}$ .

We now partition V(G) into sets  $C_1, C_2, \ldots, C_{m+1}$  and color the vertices in  $C_i$  in the *i*th phase of the algorithm. We use the values of  $l_1, l_2, \ldots, l_m$  to control the sizes of these sets. For notational convenience set  $A_0 = B_0 = C_0 = \emptyset$ . For each  $i = 1, 2, \ldots, m$ , we construct sets  $A_i$  and  $B_i$  and set  $C_i = A_i \cup B_i$ . We use  $C_i'$  to denote the vertices in the sets constructed thus far. In other words, for

each  $i=0,1,\ldots,m+1$ , we let  $C_i'$  denote  $\cup_{j=0}^i C_j$ . For each  $i=1,2,\ldots,m$ ,  $A_i$  is constructed by selecting vertices in  $G-C_{i-1}'$  as follows. Arrange the vertices of of  $G-C_{i-1}'$  in greedy ordering and let  $A_i$  be the first  $(l_i-l_{i-1})k$  vertices in this ordering.  $B_i$  is selected from vertices in  $G-C_{i-1}'-A_i$  as follows. Initially set  $B_i=\emptyset$  and while there is a vertex  $w\in G-C_{i-1}'-A_i-B_i$  that has at least 13d neighbors in  $A_i\cup B_i\cup C_{i-1}'$ , add w to  $B_i$ . Repeat this process until every vertex  $w\in G-C_{i-1}'-A_i-B_i$  has fewer than 13d neighbors in  $C_{i-1}'\cup A_i\cup B_i$ . This completes the construction of  $A_i$  and  $B_i$  and we simply set  $C_i=A_i\cup B_i$ . After constructing  $C_1,C_2,\ldots,C_m$ , we set  $C_{m+1}=V(G)-C_m'$ .

Now let  $b_i = |B_i|$  for each i = 0, 1, 2, ..., m and let e(H) denote the number of edges in a graph H. It follows from our construction that for each i = 0, 1, ..., m,  $e(G[C_i]) \ge 13d \sum_{j=0}^i b_j$ . On the other hand  $G[C_i]$  is a d-degenerate graph and has  $l_i k + \sum_{j=0}^i b_j$  vertices, and so  $e(G[C_i]) < (l_i k + \sum_{j=0}^i b_j)d$ . It follows that  $\sum_{j=0}^i b_j < \frac{l_i k}{12}$ , or in other words, for each i = 1, ..., m,

$$|C_i'| < \frac{13}{12} l_i k. \tag{1}$$

Since  $C'_{m+1} = V(G)$  we also know that  $|C'_{m+1}| \le tk = l_{m+1}k$ .

We will color  $C_1$  with k colors in such a way that each color class has at most  $\lceil \frac{7}{6}l_1 \rceil$  vertices. We color vertices in  $C_1$  one by one in a degenerate order. Hence when we color vertex  $u \in C_1$ , there are at least (k-d) color classes that do not contain neighbors of u. Since

$$|C_1| < \frac{13l_1k}{12} \le \frac{13l_1k}{12} \frac{16(k-d)}{15k} < \frac{7}{6}l_1(k-d),$$

there exists a color class M of size less than  $\frac{7}{6}l_1$  that does not contain neighbors of u. We color u with color M.

We now show how to color the rest of the sets  $C_2, C_3, \ldots, C_{m+1}$ . For  $2 \le i \le m+1$ , in the *i*th phase we start with G such that all vertices in  $C'_{i-1}$  have been colored. In this phase we will color the vertices in  $C_i$  in a degenerate order in such a way that: (i) Every color class is of size at most  $L_i$ , where  $L_i = \lceil \frac{7}{6}l_i \rceil$  for  $2 \le i \le m$ , and  $L_{m+1} = t$ ; (ii) The vertices in  $C'_{i-1}$  will not be recolored.

Claim 4. For every  $i \ge 2$ ,  $L_{i-1}/L_i \le 2/5$ .

*Proof.* Recall that  $l_i \geq 4l_{i-1}$  for every  $i \geq 2$ . If i = m + 1, then  $L_i = l_i = t \geq 16$ . Therefore,

$$\frac{L_m}{L_{m+1}} = \frac{\lceil 7l_m/6 \rceil}{t} \le \frac{7l_m/6 + 5/6}{t} \le \frac{7}{6 \cdot 4} + \frac{5/6}{16} = \frac{11}{32} < 0.4.$$

If  $2 \le i \le m$ , then  $L_i = \lceil \frac{7l_i}{6} \rceil$ . If  $l_{i-1} \ge 2$ , then  $l_i \ge 8$  and

$$\frac{L_{i-1}}{L_i} \le \frac{7l_{i-1}/6 + 5/6}{7l_i/6} \le \frac{1}{4} + \frac{5/6}{7 \cdot 8/6} = \frac{19}{56}.$$

Finally, if  $l_{i-1} = 1$ , then  $L_{i-1} = 2$  and  $L_i \ge 5$ . This proves the claim.

Suppose we want to color a vertex v. Let  $M_1, \ldots, M_k$  be the current color classes. Let  $Y_0$  denote the set of color classes of cardinality less than  $L_i$ . If some  $M_j \in Y_0$  contains no neighbors of v, then we color v with  $M_j$  and work with the next vertex. Otherwise, let  $Y_0$ -candidate be a vertex  $w \in V - C'_{i-1}$  such that there exists a color class  $M(w) \in Y_0$ , with  $w \notin M(w)$  and  $N_G(w) \cap M(w) = \emptyset$ . Let  $Y_1$  be the set of color classes containing a  $Y_0$ -candidate. If a member  $M_j$  of  $Y_1$  does not contain a neighbor of v, then we color v with  $M_j$  and recolor some  $Y_0$ -candidate  $w \in M_j$  with M(w). For  $h \geq 1$ , let a  $Y_h$ -candidate be a vertex  $w \in C_i - \bigcup_{M \in Y_0 \cup \ldots \cup Y_h} M$  such that there exists  $M(w) \in Y_h$  with  $N_G(w) \cap M(w) = \emptyset$ . Let  $Y_{h+1}$  be the set of color classes containing a  $Y_h$ -candidate. If a member  $M_j$  of  $Y_{h+1}$  does not contain a neighbor of v, then we color v with  $M_j$  and similarly to the

above recolor a sequence of candidates. Finally, let  $Y = \bigcup_{j=0}^{\infty} Y_j$  and y = |Y|. Then by the above Y possesses the following properties:

- (a) every member of Y contains a neighbor of v,
- (b) every vertex  $u \in C_i \bigcup_{M \in Y} M$  has a neighbor in every  $M \in Y$  (otherwise the color class of u would be in Y).

We will prove that there is at least one color class M in Y that does not contain neighbors of v. Suppose this is not the case.

Observe that each vertex  $u \in C_i$  has at most 13d neighbors in  $C'_{i-1}$  (by the construction of  $B_{i-1}$ ), and at the moment of coloring has at most d neighbors among colored earlier vertices of  $C_i$  (since vertices are considered in a degenerate order). So when we color a vertex  $u \in C_i$ , there are at most (13+1)d color classes that have neighbors of u. By property (a) of  $Y, y \leq 14d$ .

Claim 5. y < 8d/7.

Proof. Let  $S = \bigcup_{M \in Y} M$  and  $T = C_i - S$ . By property (b) of Y, at least y|T| edges connect T with S. Since G is d-degenerate, we conclude that y|T| < d(|S| + |T|), i.e., that (y - d)|T| < d|S|. Clearly,  $|S| \le yL_i$ . By the definition of  $Y_0$ ,  $|T| \ge (k - y)(L_i - L_{i-1})$ .

By Claim 4,  $\frac{L_i - L_{i-1}}{L_i} \ge 1 - \frac{2}{5} = \frac{3}{5}$ , for every  $i \ge 2$ . Therefore,

$$(y-d)(k-y)\frac{3}{5} < dy.$$

Since  $k \ge 16d$ , the last inequality yields that  $(y-d)(16d-y)\frac{3}{5} < dy$ . This implies the following inequality for  $\gamma = y/d$ :

$$\gamma^2 - \frac{46}{3}\gamma + 16 > 0.$$

Therefore, either  $\gamma > (23 + \sqrt{385})/3 \sim 14.207...$  or  $\gamma < (23 - \sqrt{385})/3 \sim 1.1261... < 8/7$ . The former is impossible, since  $y \le 14d$ , thus the latter holds. This proves the claim.

SUBCASE 2.1:  $2 \le i \le m$ . The total number of colored vertices is at least  $L_i(k-y)$  which by Claim 5 is greater than

$$\left\lceil \frac{7l_i}{6} \right\rceil \left( k - \frac{8d}{7} \right) \ge \frac{7l_i}{6} \frac{13k}{14} = \frac{13l_i k}{12}.$$

This contradicts (1) for j = i - 1.

SUBCASE 2.2: i = m + 1. Let  $D_i$  be the highest degree in  $G[V - C_i']$ .

Claim 6. 
$$l_1\Delta + (l_2 - l_1)D_1 + (l_3 - l_2)D_2 + \ldots + (l_{m+1} - l_m)D_m < 2.75\Delta + 4.25dt$$
.

*Proof.* The proof of this claim is somewhat technical and appears in the appendix.

Let  $M_1 \in Y_0$ . By construction, every  $M_j$  contains at most  $L_i$  vertices in  $C'_i$ . So the number of neighbors of  $M_1$  is at most

$$\begin{split} &L_1\Delta + (L_2 - L_1)D_1 + \ldots + (L_{m+1} - L_m)D_m = \\ &= \left\lceil \frac{7l_1}{6} \right\rceil \Delta + \left( \left\lceil \frac{7l_2}{6} \right\rceil - \left\lceil \frac{7l_1}{6} \right\rceil \right)D_1 + \ldots + \left( t - \left\lceil \frac{7l_m}{6} \right\rceil \right)D_m \\ &= \left\lceil \frac{7l_1}{6} \right\rceil (\Delta - D_1) + \left\lceil \frac{7l_2}{6} \right\rceil (D_1 - D_2) + \ldots + \left\lceil \frac{7l_m}{6} \right\rceil (\Delta_{m-1} - D_m) + tD_m \\ &< \frac{7l_1}{6}(\Delta - D_1) + \frac{5}{6}(\Delta - D_1) + \frac{7l_2}{6}(D_1 - D_2) + \frac{5}{6}(D_1 - D_2) + \ldots \\ &+ \frac{7l_m}{6}(D_{m-1} - D_m) + \frac{5}{6}(D_{m-1} - D_m) + tD_m \\ &\leq \left( \frac{7l_1}{6} + \frac{5}{6} \right)\Delta + \frac{7}{6}\left( (l_2 - l_1)D_1 + (l_3 - l_2)D_2 + \ldots + (l_{m+1} - l_m)D_m \right). \end{split}$$

On the other hand, by property (b) of Y, the number of neighbors of  $M_1$  is at least  $(k-y)(t-L_m)$ . Note that

$$t - L_m = t - \left\lceil \frac{7l_m}{6} \right\rceil \ge t \left( 1 - \frac{\frac{7l_m}{6} + \frac{5}{6}}{t} \right) = t \left( 1 - \frac{7}{4 \cdot 6} - \frac{5}{6 \cdot 16} \right) = \frac{21}{32}t.$$

Hence by Claim 6 we have  $(k-y)(t-L_m) \ge (k-8d/7)\frac{21}{32}t$ . Comparing this with the upper bound above and applying Claim 6 we get

$$\left(k - \frac{8d}{7}\right) \frac{21}{32}t \le \frac{5}{6}\Delta + \frac{7}{6} (2.75\Delta + 4.25dt).$$

Since  $\Delta \leq n/15 \leq kt/15$ , this reduces to

$$\left(k - \frac{8d}{7}\right) \frac{21}{32} \le \frac{5}{6 \cdot 15} k + \frac{7}{6} \left(\frac{2.75}{15} k + 4.25d\right),$$

which gives

$$\left(\frac{21}{32} - \frac{1}{18} - \frac{7}{6} \frac{2.75}{15}\right)k \le \left(\frac{21}{32} \frac{8}{7} + \frac{7 \cdot 4.25}{6}\right)d.$$

It follows that

$$\frac{k}{d} \le \frac{68.5}{12} \cdot \frac{1440}{557} = \frac{8220}{557} < 15,$$

a contradiction to  $k \geq 16d$ . This proves the theorem.

Algorithm. The above proof implies a simple algorithm for equitably k-coloring any n-vertex d-degenerate graph with  $\Delta(G) \leq n/15$ . We first partition of V(G) into sets  $C_i$ ,  $1 \leq i \leq m+1$  as described in the first part of the proof. Then for each  $i=1,2,\ldots,m+1$ , we attempt to color vertices of  $C_i$  in d-degenerate order. It is possible that in the process some vertices may have to be recolored, but these recolorings are restricted to the set being currently colored, namely  $C_i$ . The algorithm is clearly polynomial time and it can be implemented in  $O(n^3)$  time; we do not give details here.

# 4 Constant-factor approximation algorithm

The algorithm above can be (trivially) thought of as providing an O(d)-factor approximation algorithm for equitably coloring, with fewest colors, an n-vertex, a d-degenerate graph with  $\Delta(G) \leq n/15$ . In this section, we extend this to an O(d)-factor algorithm for equitably coloring any d-degenerate graph. This implies that first known O(1)-factor algorithm for planar graphs. The main result in this section is the following.

**Theorem 4.** Every n-vertex d-degenerate graph G with maximum degree at most  $\Delta$  is equitably k-colorable for any  $k, \ k \geq \max\left\{62d, \ 31d\frac{n}{n-\Delta+1}\right\}$ .

*Proof.* Let G be an n-vertex d-degenerate graph. Let  $G_0 = G$ , h = 30d - 1 and for j = 1, ..., h, let  $w_j$  be a vertex of the maximum degree in  $G_{j-1}$  and  $G_j = G_{j-1} - w_j$ .

Claim 7. For every  $v \in V(G_h)$ ,  $\deg_{G_h}(v) < n/30$ .

Proof. If  $\deg_{G_h}(v) \geq n/30$  for some  $v \in V(G_h)$ , then it is also the case that  $\deg_{G_{j-1}}(w_j) \geq n/30$  for every  $j = 1, \ldots, 30d-1$ , and hence  $|E(G)| \geq 30d(n/30) = dn$ . This is a contradiction since any n-vertex d-degenerate graph has fewer than dn edges.

Claim 8. There are pairwise disjoint independent sets  $M_1, M_2, \ldots, M_h$  such that for every  $j, 1 \le j \le h$ , (i)  $w_j \in \bigcup_{s=1}^j M_s$ , (ii)  $\lfloor n/k \rfloor \le |M_j| \le \lceil n/k \rceil$ , and (iii)  $nj/k \le \sum_{s=1}^j |M_s| < 1 + nj/k$ .

*Proof.* Let  $X_1 = V(G) - w_1 - N_G(w_1)$ . Clearly,  $|X_1| \ge n - \Delta - 1$ . Since G is d-degenerate,  $X_1$  contains an independent set  $M_1'$  of size at least  $\frac{X_1}{d+1} \ge \frac{n-\Delta-1}{d+1}$ . Since

$$\frac{n}{k} \le \frac{n - \Delta + 1}{31d} < \frac{n - \Delta}{d + 1},$$

 $|M_1'| > \frac{n}{k} - \frac{1}{d+1}$ . Hence, we can choose a subset  $M_1''$  of  $M_1'$  of size  $\lceil \frac{n}{k} \rceil - 1$  and let  $M_1 = M_1'' + w_1$ . By construction,  $M_1$  satisfies properties (i)–(iii) for j = 1.

Suppose we have constructed  $M_1, M_2, \ldots, M_{j-1}$  satisfying (i)–(iii) for some  $j \leq h$ . Let  $x_j = w_j$  if  $w_j \notin \bigcup_{s=1}^{j-1} M_s$ , and let  $x_j$  be any vertex outside  $\bigcup_{s=1}^{j-1} M_s$  otherwise. Let  $X_j = V(G) - \bigcup_{s=1}^{j-1} M_s - x_j - N_G(x_j)$ . Since G is d-degenerate,  $X_j$  contains an independent set  $M'_j$  of size at least  $\frac{X_j}{d+1}$ . Suppose that  $|M'_j| < -1 + n/k$ . In view of (iii), this means that

$$\frac{n-1-(j-1)n/k-1-\Delta}{d+1}<\frac{n}{k}-1.$$

For n > k and  $d \ge 1$ , the last inequality yields  $n - \Delta + 1 < \frac{(j+d)n}{k} + 1 < \frac{31dn}{k}$ . But this contradicts the choice of k. Thus, we can choose a subset of  $M'_j$  that together with  $x_j$  forms an independent set  $M''_j$  of size  $\lceil n/k \rceil$ . If

$$|M_j''| + \sum_{s=1}^{j-1} |M_s| < \frac{jn}{k} + 1,$$

then we let  $M_j = M_j''$ , otherwise we get  $M_j$  by deleting a vertex  $v \neq x_j$  from  $M_j''$ . Note that in the latter case,  $\lfloor n/k \rfloor \neq \lceil n/k \rceil$ , and thus (i)–(iii) hold in both cases. This proves the claim.

Let G' be the graph obtained by deleting vertices in  $M_1 \cup M_2 \cup ... M_h$  from G and let V' = V(G'). Claim 9.  $|V'| \ge 16n/31$ .

*Proof.* By (iii) of Claim 8,  $|V'| \ge n - (30d - 1)n/k - 1 \ge n - 30dn/k$ . Since  $k \ge 62d$ , we get  $|V'| \ge 32n/62$ .

By Claims 7 and 9,

$$\frac{|V'|}{\Delta(G')} \ge \frac{32n}{62} \cdot \frac{30}{n} > 15.$$

Since  $k-h \ge 62d-30d=32d$ , by Theorem 1, G' is equitably (k-h)-colorable. Hence G is equitably k-colorable. This proves the theorem.

Corollary 2. Every d-degenerate graph with n vertices and maximum degree at most 1 + n/2 is equitably k-colorable when  $k \geq 62d$ .

**Theorem 3.** There exists a polynomial time algorithm that, given a d-degenerate graph G with  $\chi_{eq}(G) \leq s$ , can equitably k-color G with for any  $k, k \geq 31ds$ .

*Proof.* Assume that a graph G on n vertices with maximum degree  $\Delta$  admits an equitable coloring  $\phi$  with s colors. Let  $v \in V(G)$  have degree  $\Delta$ . The color class of v contains at most  $n-\Delta$  vertices. Thus no other color class can contain more than  $n-\Delta+1$  vertices. Hence,

$$s > \frac{n}{n - \Delta + 1}. (2)$$

Also, if G has at least one edge,  $s \ge 2$ . If  $\Delta \le 1 + n/2$ , then by Corollary 4 G can be equitably k-colored for any  $k \ge 62d$ . Since  $62d \ge 31ds$ , G can be equitably k-colored for any  $k \ge 31ds$ . If

 $\Delta > 1 + n/2$ , then  $31d\frac{n}{n-\Delta+1} > 62d$  and therefore by Theorem 4, G be equitably k-colored for any  $k \geq 31d\frac{n}{n-\Delta+1}$ . It follows from inequality 2 that G can be equitably k-colored for any  $k \geq 31ds$ .

The fact that such an equitable k-coloring can be constructed in polynomial time is implied by the proof of Theorem 4. The algorithm is sketched here. First identify the high degree vertices  $w_1, w_2, \ldots, w_h$  in G and construct the color classes  $M_1, M_2, \ldots, M_h$ . Constructing these color classes uses as a subroutine an algorithm that constructs an independent set of size at least m/(d+1) in a given m-vertex, d-degenerate graph. It is easy to see that the following greedy algorithm suffices for this task: pick a minimum degree vertex, delete the vertex and its neighbors, and repeat until no vertices are left. Once the color classes  $M_1, M_2, \ldots, M_h$  are constructed and the colored vertices are deleted, we are left with a graph whose maximum vertex degree is less than n/30. We color the vertices in this graph using the algorithm from the previous section. This phase dominates the running time of the algorithm and hence we have an  $O(n^3)$  algorithm.

### References

- [1] G. Andersson. An approximation algorithm for max p-section. In Proceedings of the 16th Annual Symposium on Theoretical Aspects of Computer Science, Volume 1563 of Lecture Notes in Computer Science, pages 237–247, 1999.
- [2] M.J. Pelsmajer A.V. Kostochka and D.B. West. A list analogue of equitable coloring. Submitted, 2002.
- [3] B. Bollobás and R.K. Guy. Equitable and proportional coloring of trees. *Journal of Combinatorial Theory, Series B*, 34:177–186, 1983.
- [4] A. Hajnal and E. Szeméredi. Proof of a conjecture of Erdös. In P. Erdös, A. Rényi, and V.T. Sós, editors, Combinatorial Theory and Its Applications, Vol II, Volume 4 of Colloquia Mathematica Societatis János Bolyai, pages 601–623. North-Holland, 1970.
- [5] S. Janson. Large deviations for sums of partly dependent random variables. Isaac Newton Institute preprint NI02024-CMP, 2002.
- [6] S. Janson, T. Luczak, and A. Ruciński. Random Graphs. Wiley-Interscience, 2000.
- [7] S. Janson and A. Ruciński. The infamous upper tail. Random Structures and Algorithms, 20(3), 2002.
- [8] J. Komlos and M. Simonovits. Szemerédi's regularity lemma and its applications in graph theory. DIMACS Technical Report 96-10, 1996.
- [9] A.V. Kostochka. Equitable colorings of outerplanar graphs. To appear in Discrete Math, 2001.
- [10] A.V. Kostochka and K. Nakprasit. Equitable colorings of d-degenerate graphs. To appear in Combinatorics, Probability, and Computing, 2002.
- [11] K. Lih. The equitable coloring of graphs. In D.-Z. Du and P.M. Pardalos, editors, The Handbook of Combinatorial Optimization, Vol. 3, pages 543-566. Kluwer, Boston, 1998.
- [12] K. Lih and P. Wu. On equitable coloring of bipartite graphs. Discrete Mathematics, 151:155-160, 1996.
- [13] W. Meyer. Equitable coloring. American Mathematical Monthly, 80:143-149, 1973.
- [14] S.V. Pemmaraju. Equitable colorings extend chernoff-hoeffding bounds. In Proceedings of the 5th International Workshop on Randomization and Approximation Techniques in Computer Science (APPROX-RANDOM 2001), pages 285-296, 2001.
- [15] S.V. Pemmaraju. Coloring outerplanar graphs equitably. In review at Journal of Combinatorial Theory, Series B, 2002.
- [16] A. Tucker. Perfect graphs and an application to optimizing municipal services. SIAM Review, 15:585–590, 1973.

## Appendix.

#### Proof of CASE 1 of Theorem 1

CASE 1.  $t \leq 15$ . We will color the vertices one by one in a d-degenerate order  $v_1, \ldots, v_n$  (with some recolorings). Suppose we cannot color vertex  $v_i$ . Let Z be the set of color classes containing neighbors of  $v_i$ . Since G is d-degenerate,  $|Z| \leq d$ . If a color class  $M \notin Z$  has fewer than t vertices, then we can color  $v_i$  with M. Since  $n \leq kt$ , there is a color class  $M_0 \in Z$  with at most t-1 vertices. If a vertex w in a color class  $M \notin Z$  has no neighbors in  $M_0$ , then we can recolor w with  $M_0$  and color  $v_i$  with M. Thus, every of (k-|Z|)t colored vertices outside of Z has a neighbor in  $M_0$ . Therefore,

$$(t-1)\Delta \ge (k-d)t$$
.

Since  $n > 15\Delta$ , we have

$$(t-1)\frac{n}{15} \ge \frac{15}{16}kt \ge \frac{15}{16}n$$

and hence  $t-1 > 15^2/16 > 14$ , a contradiction to the choice of t.

#### Technical claim in the proof of Theorem 1

#### Claim 10.

$$l_1\Delta + (l_2 - l_1)D_1 + (l_3 - l_2)D_2 + \dots + (l_{m+1} - l_m)D_m \le 2.75\Delta + 4.25dt.$$

Proof. Observe that

$$|E(G)| \geq \sum_{\substack{1 \leq i \leq m \\ 1 \leq j \leq l_i \, k}} \deg_{V-C'_{i-1}-\{v^i_1, \ldots, v^i_{j-1}\}}(v^i_j) \ldots$$

By the definition of  $A_i$ , for  $v_j^i \in A_i$ ,  $\deg_{G[V-C'_{i-1}-\{v_1^i,...,v_{j-1}^i\}]}(v_j^i) \ge D_{i+1}$ , and  $|A_i| = (l_i - l_{i-1})k$ . Thus,

$$|E(G)| \ge k(l_1D_1 + (l_2 - l_1)D_2 + (l_3 - l_2)D_3 + \dots + (l_m - l_{m-1})D_m).$$

Since  $|E(G)| < dn \le dtk$ , we have

$$l_1 D_1 + (l_2 - l_1) D_2 + (l_3 - l_2) D_3 + \ldots + (l_m - l_{m-1}) D_m < dt.$$
(3)

Note that

$$\frac{l_{i+1} - l_i}{l_i - l_{i-1}} = \frac{4l_i + \beta_{i+1} - l_i}{4l_i - 1 + \beta_i - l_{i-1}} \le \frac{3(4l_{i-1} + \beta_i) + 3}{3l_{i-1} + \beta_i} = 4 + \frac{3 - \beta_i}{3l_{i-1} + \beta_i} \le 4 + \frac{1}{l_{i-1}}.$$

For  $i \geq 3$ , we obtain  $l_{i+1} - l_i \leq (4 + \frac{1}{4})(l_i - l_{i-1})$ . Also  $(l_2 - l_1) - 4.25l_1 = \beta_2 - 1.25l_1$ . Therefore,

$$4.25 (l_1D_1 + (l_2 - l_1)D_2 + (l_3 - l_2)D_3 + \ldots + (l_m - l_{m-1})D_m) \ge$$

$$\geq (l_2 - l_1)D_1 + (l_3 - l_2)D_2 + \ldots + (l_{m+1} - l_m)D_m + (1.25l_1 - \beta_2)D_1.$$

Comparing with (3), we get

$$(l_2 - l_1)D_1 + (l_3 - l_2)D_2 + \dots + (l_{m+1} - l_m)D_m < 4.25dt + \beta_2 D_1 - 1.25l_1 D_1.$$

Hence

$$l_1 \Delta + (l_2 - l_1)D_1 + (l_3 - l_2)D_2 + \dots + (l_{m+1} - l_m)D_m \le l_1 \Delta + 4.25dt + \beta_2 D_1 - \frac{5}{4}l_1 D_1$$

$$\le 4.25dt + \beta_2 D_1 - \frac{1}{4}l_1 D_1 \le 4.25dt + (3 - \frac{1}{4})D_1 \le 4.25dt + 2.75\Delta.$$

This proves the claim.