Bipartite Coverings and the Chromatic Number

Dhruv Mubayi
Department of Mathematics,
Statistics, and Computer Science
University of Illinois
Chicago, IL 60607, USA
mubayi@math.uic.edu

Sundar Vishwanathan
Department of Computer Science
Indian Institute of Technology
Mumbai
India 400076
sundar@cse.iitb.ernet.in

Abstract

Consider a graph G with chromatic number k and a collection of complete bipartite graphs, or bicliques, that cover the edges of G. We prove the following two results:

• If the bipartite graphs form a partition of the edges of G, then their number is at least $2^{\sqrt{\log_2 k}}$. This is the first improvement of the easy lower bound of $\log_2 k$, while the Alon-Saks-Seymour conjecture states that this can be improved to $k-1$.

• The sum of the orders of the bipartite graphs in the cover is at least $(1-o(1))k\log_2 k$. This generalizes, in asymptotic form, a result of Katona and Szemerédi who proved that the minimum is $k\log_2 k$ when G is a clique.

1 Introduction

It is a well-known fact that the minimum number of bipartite graphs needed to cover the edges of a graph G is $\lceil \log \chi(G) \rceil$, where $\chi(G)$ is the chromatic number of G (all logs are to the base 2). Two classical theorems study related questions. One is the Graham-Pollak theorem [1] which states that the minimum number of complete bipartite graphs needed to partition $E(K_k)$ is $k-1$. Another is the Katona-Szemerédi theorem [4], which states that the minimum of the sum of the orders of a collection of complete bipartite graphs that cover $E(K_k)$ is $k \log k$. Both of these results are best possible.

An obvious way to generalize these theorems is to ask whether the same results hold for any G with chromatic number k.

Conjecture 1 (Alon-Saks-Seymour) *The minimum number of complete bipartite graphs needed to partition the edge set of a graph G with chromatic number k is $k-1$.*

Note that every graph has a partition of this size, simply by taking a proper coloring $V_1, \ldots V_k$ and letting the ith bipartite graph be $(V_i, \cup_{j>i} V_j)$.

Another motivation for Conjecture 1 is that the non-bipartite analogue is an old conjecture of Erdős-Faber-Lovász. The Erdős-Faber-Lovász conjecture remains open although it has been proved asymptotically by Kahn [3]. Conjecture 1 seems much harder than the Erdős-Faber-Lovász conjecture, indeed, as far as we know there are no nontrivial results towards it except the folklore lower bound of $\log_2 k$ which doesn't even use the fact that we have a partition. Our first result improves this to a superlogarithmic bound for k large.

Theorem 2 *The number of complete bipartite graphs needed to partition the edge set of a graph G with chromatic number k is at least $2^{\sqrt{2 \log k}(1+o(1))}$.*

Motivated by Conjecture 1, we make the following conjecture that generalizes the Katona-Szemerédi theorem.

Conjecture 3 *Let G be a graph with chromatic number k. The sum of the orders of any collection of complete bipartite graphs that cover the edge set of G is at least $k \log k$.*

We prove Conjecture 3 with $k \log k$ replaced by $(1 - o(1))k \log k$.

Theorem 4 *Let G be a graph with chromatic number k, where k is sufficiently large. The sum of the orders of any collection of complete bipartite graphs that cover the edge set of G is at least*

$$k \log k - k \log \log k - k \log \log \log k.$$

The next two sections contain the proofs of Theorems 2 and 4.

2 The Alon-Saks-Seymour Conjecture

It is more convenient to phrase and prove our result in inverse form. Let G be a disjoint union of m complete bipartite graphs $(A_i, B_i), 1 \leq i \leq m$. The Alon-Saks-Seymour conjecture then states that the chromatic number of G is at most $m+1$.

We prove the following theorem which immediately implies Theorem 2.

Theorem 5 *Let G be a disjoint union of m complete bipartite graphs. Then $\chi(G) \leq m^{\frac{1+\log m}{2}}(1+o(1))$.*

Proof. We will begin with a proof of a worse bound. We will first show that $\chi(G) \leq m^{\log m}(1+o(1))$. A color will be an ordered tuple of length at most $\log m$, with each element a positive integer of value at most m. We will construct this tuple in stages. In the ith stage we will fill in the ith co-ordinate. Note that the length of the tuple may vary with vertices.

With each vertex v, at stage i, we will associate a set $S(i,v) \subset V(G)$. The set $S(i,v)$ will contain all vertices which have the same color sequence, so far, as v (in particular, $v \in S(i,v)$ for all i).

A bipartite graph (A_j, B_j) is said to *cut* a subset of vertices S if $S \cap A_j \neq \emptyset$ and $S \cap B_j \neq \emptyset$.

Consider two bipartite graphs (A_k, B_k) and (A_l, B_l) from our collection. Since they are edge disjoint, (A_l, B_l) cuts either A_k or B_k, but not both.

Fix a vertex v. We set $S(0,v) := V(G)$. The assignment for the $i+1$st stage is as follows. Suppose we have defined $S(i,v)$. Let $\mathcal{F}(i,v)$ denote the set of all bipartite graphs that cut $S(i,v)$. For each bipartite graph $(A_j, B_j) \in \mathcal{F}(i,v)$ for which $v \in A_j \cup B_j$, let C_j be the set among A_j, B_j that contains v and let D_j be the set among A_j, B_j that omits v. For a vertex v, check if there is a bipartite graph $(A_j, B_j) \in \mathcal{F}(i,v)$ such that $v \in A_j \cup B_j$ and one of the following two conditions are satisfied:

- The number of bipartite graphs in $\mathcal{F}(i,v)$ that cut C_j is smaller than the number that cut D_j. OR
- The number of bipartite graphs in $\mathcal{F}(i,v)$ that cut C_j is equal to the number that cut D_j and $C_j = A_j$.

If there is such a j, then the $i + 1$st co-ordinate of the color of v is j and $S(i + 1, v) = S(i, v) \cap C_j$. If there are many candidates for j, pick one arbitrarily.

If there is no such (A_j, B_j), then the coloring of v ceases and the vertex will not be considered in subsequent stages. In other words, the final color of vertex v will be a sequence of length i.

Note that in this process every vertex is assigned a color except vertices that were not assigned a color in the very first step. We will show below that no two vertices that are assigned a color are adjacent. The same argument shows that the vertices that do not get assigned a color in the first step form an independent set. These vertices are all assigned a special color which is swallowed up in the $o(1)$ term.

The following technical lemma establishes the statements needed to prove correctness and a bound on the number of colors used.

Lemma 6 *For each vertex v, the set $S(i, v)$ is determined by the color sequence $x_1, \ldots, x_i$ assigned to the vertex v. It will be independent of the vertex v. Note that if the color sequence stops before i then $S(i, v)$ is not defined. Also, the number of bipartite graphs that cut $S(i, v)$ is at most $m/2^i$.*

Proof. The proof is by induction on i. Both statements are trivially true for $i = 0$. For the inductive step, assume that $S(i, v)$ is determined by $x_1, \ldots, x_i$ and at most $m/2^i$ bipartite graphs cut $S(i, v)$. If v ceases to be colored then we are done. Now suppose that v is colored with $x_{i+1} = t$ in step $i + 1$. Then $(A_t, B_t) \in \mathcal{F}(i, v)$ and $v \in A_t \cup B_t$. As before, define C_t and D_t. Because v is colored in this step, the number of bipartite graphs in $\mathcal{F}(i, v)$ that cut C_t is either smaller than the number which cut D_t or they are equal and $C_t = A_t$. Knowing $S(i, v)$ and t we can determine which of the cases we are in and we can determine $S(i+1, v) = C_t \cap S(i, v)$. Notice that C_t can be determined by looking at $S(i, v)$ and t alone and is independent of the vertex v.

Also, since the number of bipartite graphs that cut C_t is at most half the number that cut $S(i, v)$ the second assertion follows. ∎

We argue first that the coloring is proper. Assume for a contradiction that two adjacent vertices v and w are assigned the same color sequence. Suppose the sequence is of length i. Then by the previous lemma $S(i, v) = S(i, w)$. There has to be one bipartite graph, say (A_p, B_p), such that $v \in A_p$ and $w \in B_p$. If the number of bipartite graphs in $\mathcal{F}(i, v)$ that cut A_p is less than the number that cut B_p then v will be given color p in the $i+1$st step. If the number of bipartite graphs in $\mathcal{F}(i, v)$ that cut A_p is equal to the number that cut B_p then since $C_p = A_p$, again v will be given color p in the $i+1$st step. Consequently, the number of bipartite graphs in $\mathcal{F}(i, v)$ that cut B_p is smaller than the number that cut A_p and hence w will be given color p. In all three cases, at least one of v or w will be given a color contradicting our assumption that both sequences are of length i. This argument also shows that vertices which were not assigned a color in the first step form an independent set. The coloring stops when $\mathcal{F}(i, v)$ is empty for every vertex and that happens after $\log m$ steps from the lemma.

A simple observation helps in reducing this bound by a square-root factor. At each stage, the colorings of the $S(i, v)$s are independent. Hence the colors only matter within the vertices in each of these sets. The number of bipartite graphs that cut $S(i, v)$ is at most $m/2^i$. We *renumber* these bipartite graphs from 1 to $m/2^i$. Hence the labels in the ith stage will be restricted to this set. The total number of colors used, of length i is at most $m \cdot \frac{m}{2} \cdots \frac{m}{2^i}$. The number for $i < m$ is swallowed up in the $o(1)$ term and the value for $i = m$ simplifies to the main term in the bound given. ∎

3 Generalizing the Katona-Szemerédi Theorem

In this section we prove Theorem 4. Given a graph G, let $b(G)$ denote the minimum, over all collections of bipartite graphs that cover the edges of G, of the sum of the orders of these bipartite graphs.

One proof of the Katona-Szemerédi theorem is due to Hansel [2] and the same proof yields the following lemma which is part of folklore.

Lemma 7 *Let $G = (V, E)$ be an n vertex graph with independence number*

α. Then $\alpha \geq \frac{n}{2^{b(G)/n}}$.

The lemma is proved by considering a bipartite covering achieving $b(G)$, deleting at random one of the parts of each bipartite graph, and computing a lower bound on the expected number of vertices that remain. It is easy to see that these remaining vertices form an independent set, and hence one obtains a lower bound on the independence number.

Let $k = \chi(G)$. We may assume that $n \leq k \log k$, since we are done otherwise. Let $G = G_0$. Starting with G_0, repeatedly remove independent sets of size given by Hansel's lemma as long as the number of vertices is at least k. Let the graphs we get be $G_0, G_1, \ldots, G_t$. Let $|V(G_i)| = n_i$ and $\beta = \max_i 2^{b(G_i)/n_i}$. Let this maximum be achieved for $i = p$. From the definition, we see that $n_{i+1} \leq n_i(1 - \frac{1}{2^{b(G_i)/n_i}})$. Hence $n_t \leq n(1 - 1/\beta)^t < ne^{-t/\beta} < n2^{-t/\beta}$ and together with $n_t \geq k$ we obtain

$$t \leq \beta \log(n/k).$$

There are two cases to consider. First suppose that $t \geq k/\log k$. Then from the above two inequalities we obtain

$$2^{b(G_p)/n_p} \log(n/k) \geq k/\log k.$$

Taking logs and using the facts that $n \leq k \log k$ and $n_p \geq k$ we get

$$b(G_p) \geq k(\log k - \log\log k - \log\log\log k).$$

We now consider the case that $t < k/\log k$. Let G' be the graph obtained after removing an independent set from G_t. By definition of t we have $|V(G')| < k$. Also $\chi(G') \geq k(1 - 1/\log k)$. Since the color classes of size one in an optimal coloring form a clique, this implies that G' has a clique of size at least $k(1 - 2/\log k)$. Using the fact that k is sufficiently large, $\log(1-x) > -2x$ for x sufficiently small and applying the Katona-Szemerédi theorem, we get

$$b(G') \geq \left(k - \frac{2k}{\log k}\right) \log\left\{k\left(1 - \frac{2}{\log k}\right)\right\} > \left(k - \frac{2k}{\log k}\right)\left(\log k - \frac{4}{\log k}\right)$$

$$> k\log k - 3k > k\log k - k\log\log k - k\log\log\log k.$$

Since $b(G) \geq b(G')$, the proof is complete. ▮

Note that in the proof $b(G_i)$ could use different covers, but with sizes smaller than the one induced by $b(G_0)$. One can get better lower order terms by adjusting the threshold between the two cases.

4 Acknowledgments

We thank the referees for helpful comments that improved the presentation. The research of Dhruv Mubayi was supported in part by NSF grant DMS 0653946.

References

[1] R. L. Graham, H. O. Pollak, On the addressing problem for loop switching. Bell System Tech. J. 50 1971 2495–2519

[2] G. Hansel, Nombre minimal de contacts de fermeture ncessaires pour raliser une fonction boolenne symtrique de n variables. (French) C. R. Acad. Sci. Paris 258 1964 6037–6040.

[3] J. Kahn, Coloring nearly-disjoint hypergraphs with $n + o(n)$ colors. J. Combin. Theory Ser. A 59 (1992), no. 1, 31–39.

[4] G. Katona, E. Szemerédi, On a problem of graph theory. Studia Sci. Math. Hungar 2 1967 23–28.