Multicolored forests in complete bipartite graphs ⋆

Richard A. Brualdi and Susan Hollingsworth

Department of Mathematics, University of Wisconsin-Madison, Madison, WI 53706-1388

Abstract

If the edges of a graph G are colored using k colors, we consider the *color distribution* for this coloring $\vec{a} = (a_1, a_2, \ldots, a_k)$, in which a_i denotes the number of edges of color i for $i = 1, 2, \ldots, k$. We find inequalities and majorization conditions on color distributions of the complete bipartite graph $K_{n,n}$ which guarantee the existence of multicolored subgraphs: in particular, multicolored forests and trees. We end with a conjecture on partitions of $K_{n,n}$ into multicolored trees.

Key words: complete bipartite graph, majorization, multicolored subgraphs, trees

1 Definitions

If the edges of a graph G are colored using k colors (we will not insist on a *proper* coloring), we may consider the *color distribution* for this coloring, $\vec{a} = (a_1, a_2, \ldots, a_k)$, in which a_i denotes the number of edges of color i for $i = 1, 2, \ldots, k$.

Let $\vec{p} = (p_1, p_2, \ldots, p_k)$ and $\vec{q} = (q_1, q_2, \ldots, q_k)$ be two sequences of nonnegative integers. Then we say $\vec{p}$ *majorizes* $\vec{q}$, denoted $\vec{p} \succeq \vec{q}$, provided that when the subscripts are re-ordered so that $p_1 \leq p_2 \leq \cdots \leq p_k$ and $q_1 \leq q_2 \leq \cdots \leq q_k$, we have

$$p_i + p_{i+1} + \cdots + p_k \geq q_i + q_{i+1} + \cdots + q_k \qquad \text{for } i = 1, 2, \ldots, k, \qquad (1)$$

⋆ Research partially supported by NSF Grant DMS-9424346.

with equality when $i = 1$. In what follows, we will be writing the inequalities in (1) in the equivalent form

$$p_1 + p_2 + \cdots + p_i \leq q_1 + q_2 + \cdots + q_i \qquad \text{for } i = 1, 2, \ldots, k. \tag{2}$$

Observe that two color distributions for a given graph G need not be sequences of the same length. We shall sometimes wish to apply the majorization order to color distributions: in cases in which the sequences are of different lengths, it is to be understood that the shorter sequence has been "padded" with zeroes so that the two sequences may be compared.

Given an edge-coloring of a graph G, we shall say that a subgraph H of G is *multicolored* provided no two of the edges of H share the same color.

2 Main results

Now we shall concentrate our attention on the "balanced" complete bipartite graph $K_{n,n}$.

Given a coloring of $K_{n,n}$, it is possible to partition $K_{n,n}$ into multicolored subgraphs in various ways. Observing that

$$1 + 3 + \cdots + (2n-3) + (2n-1) = n^2,$$

the number of edges of $K_{n,n}$, leads us to consider colorings of $K_{n,n}$ which allow a partition into multicolored subgraphs of sizes 1,3, ... ,2n-3,2n-1.

Theorem 1 *If $G = K_{n,n}$ is colored with color distribution $(a_1, a_2, \ldots, a_p)$, then G admits a partition into multicolored subgraphs of sizes $2n-1, 2n-3, \ldots, 3, 1$ if and only if*

$$(a_1, a_2, \ldots, a_p) \preceq (1, 1, 2, 2, \ldots, n-1, n-1, n). \tag{3}$$

Moreover, if G is colored with color distribution $(a_1, a_2, \ldots, a_p)$ satisfying (3), then it is possible to partition G into multicolored forests *of sizes $2n-1, 2n-3, \ldots, 3, 1$.*

If we represent $K_{n,n}$ by an n by n matrix (the bipartite adjacency matrix), we can view the colors and forests as a sort of orthogonal decomposition reminiscent of orthogonal latin squares. Figure 1 shows some examples for $n = 4$.

For color distributions which do not satisfy the majorization condition in (3), we can still find a multicolored forest of size $2n-1$ (which is, in fact, a spanning

$$\begin{bmatrix} 4 & 3 & 4 & 2 \\ 4 & 3 & 2 & 1 \\ 4 & 3 & 2' & 1' \\ 3' & 3' & 3' & 2' \end{bmatrix} \quad \text{is orthogonal to} \quad \begin{bmatrix} A & A & B & A \\ D & B & B & A \\ C & C & A & A \\ A & C & B & B \end{bmatrix}$$

(a) This coloring of $K_{4,4}$ is orthogonal to this partition of $K_{4,4}$ into forests (in this case actually trees).

$$\begin{bmatrix} 4 & 2 & 3 & 3' \\ 3' & 4 & 2 & 1 \\ 3 & 2' & 4 & 1' \\ 2' & 3 & 3' & 4 \end{bmatrix} \quad \text{is orthogonal to} \quad \begin{bmatrix} C & A & A & B \\ A & B & B & A \\ B & A & D & A \\ B & C & C & A \end{bmatrix}$$

(b) This coloring of $K_{4,4}$ (in this case actually proper) is orthogonal to this partition of $K_{4,4}$ into forests.

Fig. 1. Orthogonal decompositions

tree for $K_{n,n}$), provided that certain conditions are met:

Theorem 2 *If $(a_1, a_2, \ldots, a_{2n-1})$ is a color distribution for $G = K_{n,n}$, with subscripts ordered so that $a_1 \leq a_2 \leq \cdots \leq a_{2n-1}$, then every coloring of G with this distribution contains a multicolored spanning tree if and only if for every integer k with $k \leq 2n-1$,*

$$\sum_{i=1}^{k} a_i > k^2/4. \tag{4}$$

3 Proofs of Theorems 1 and 2

For the sake of notational simplicity, in the proofs which follow we will assume that all color distributions have been re-indexed as necessary so that the terms are listed in non-decreasing order.

We take advantage of the technique of [**?**] and [**?**], which utilize a theorem of Rado [**?**]:

Theorem 3 (Rado) *Let M be a matroid on a finite set E with rank function*

ρ and let $(A_1, A_2, \ldots, A_m)$ be a family of subsets of E. Then there exists an independent set $F = \{e_1, \ldots, e_m\}$ such that $e_i \in A_i$ for each $i = 1, 2, \ldots, m$ if and only if

$$\rho(\bigcup_{i \in I} A_i) \geq |I| \quad (\emptyset \neq I \subseteq \{1, \ldots, m\}).$$

If G is a graph with v vertices and edge set E and M is the polygon matroid of G, then the rank function is given by

$$\rho(X) = v - t \quad (X \subseteq E)$$

where t is the number of connected components of the spanning subgraph of G with edge set X.

We shall also be using the following lemma:

Lemma 4 *Let G be a bipartite graph with p vertices and $t \geq p-k+1$ connected components.*

- *If k is even, G contains at most $k^2/4$ edges.*
- *If k is odd, G contains at most $(k^2-1)/4$ edges.*

PROOF. We begin by counting the number of edges in the connected components of G. For this purpose, we may assume that G has a single nontrivial component, the other components being isolated vertices. For, suppose there are two nontrivial components C_1 and C_2, with bipartitions

$$C_i = A_i \cup B_i, \quad |A_i| = a_i, |B_i| = b_i \quad \text{for } i = 1, 2.$$

We may assume that $a_1 \geq b_1$, $a_2 \geq b_2$, and $a_1 \geq a_2$. Now modify the components as follows:

- Choose a vertex v in B_2. Let $\mathcal{N}$ denote the neighbors of v. Remove all edges incident to v, thereby isolating it.
- Let $r = |\mathcal{N}|$, and observe that $r \leq a_2 \leq a_1$. Choose r vertices in A_1, and create r new edges matching up these vertices to the vertices in $\mathcal{N}$.

This operation preserves both the number of edges and the number of bipartite connected components.

Since $t \geq p - k + 1$, there are at least $p - k$ isolated vertices, hence at most k vertices in the large connected component, hence at most $(k/2)(k/2) = k^2/4$ edges.

In the case that k is odd, by a parallel argument we see that G can contain at most $(k^2-1)/4$ edges. ■

Finally, we shall be using the following theorem of Gale and Ryser [**?**]:

Theorem 5 (Gale, Ryser) *Let $R=(r_1,r_2,\ldots,r_m)$ and $S=(s_1,s_2,\ldots,s_n)$ be two vectors whose components are nonnegative integers. For $i=1,2,\ldots,m$, let $\delta_i=(1,1,\ldots,1,0,0,\ldots,0)$ be a vector of n components with 1's in the first r_i positions and 0's elsewhere. Then there exists an $m\times n$ $(0,1)$-matrix with row sum vector R and column sum vector S if and only if*

$$S \preceq \overline{S},$$

where $\overline{S}$ is the column sum vector of the matrix $\overline{A}=\begin{bmatrix}\delta_1\\ \delta_2\\ \vdots\\ \delta_m\end{bmatrix}$.

We now have the necessary machinery to prove Theorems 1 and 2.

Proof (Theorem 1). Suppose G is colored with color distribution $\vec{a}=(a_1,a_2,\ldots,a_p)$ and that G has been partitioned into multicolored subgraphs of sizes $2n-1,2n-3,\ldots,3,1$. Consider the $n\times(2n-1)$ $(0,1)$-matrix with rows indexed by subgraphs and columns by colors, in which the (i,j)th entry is 1 iff the i^{th} subgraph has an edge of color j. Then the row sum vector is $R=(2n-1,2n-3,\ldots,3,1)$, the column sum vector is $S=\vec{a}$, and the vector $\overline{S}$ of Theorem 5 is $(n,n-1,n-1,\ldots,2,2,1,1)$. The "only if" statement of Theorem 1 hence follows from Theorem 5.

Now, let us assume that G is colored with color distribution $\vec{a}$ satisfying (3). We will show that G may be partitioned into multicolored forests of the desired sizes. For $i=1,2,\ldots,2n-1$, let M_i denote the set of edges of color i. We establish the existence of edge-disjoint multicolored forests $F_1,F_2,\ldots,F_n$, where F_i contains $2i-1$ edges. We begin by showing the existence of F_n. By Rado's theorem it suffices to show that in the polygon matroid of G, the rank of U_I is at least k, where $U_I=\bigcup_{i\in I}M_i$, for each $k=1,2,\ldots 2n-1$ and each $I\subseteq\{1,2,\ldots,2n-1\}$ with $|I|=k$. Fix k and I, and let t be the number of connected components of G_I, the spanning subgraph of G with edge set U_I.

Our objective is to show that $2n-t\geq k$. Assume to the contrary that $t\geq 2n-k+1$. Let e denote the number of edges in G_I.

Assume first that k is even. Since the edges of G_I comprise some of the M_i's, and since $\vec{a}$ satisfies (3), we have $e \geq k^2/4 + k/2$, because

$$(a_1 + a_2 + \cdots + a_k) \leq 2\left(1 + 2 + \cdots + \frac{k}{2}\right) = 2\frac{\frac{k}{2}\left(\frac{k}{2}+1\right)}{2} = \frac{k^2}{4} + \frac{k}{2}.$$

On the other hand, according to Lemma 4, $e \leq k^2/4$, a contradiction.

In the case that k is odd, by a similar calculation we have

$$e \geq \frac{k-1}{2}\frac{k+1}{2} + \frac{k+1}{2},$$

again contradicting Lemma 4.

Now consider $M'_1, M'_2, \ldots, M'_{2n-3}$, where $M'_i = M_i \backslash F_n$. (Notice that $M_{2n-2} \backslash F_n$ and $M_{2n-1} \setminus F_n$ are both empty since each of M_{2n-2}, M_{2n-1} consisted of only one edge.)

By applying Rado's theorem again to $U'_I = \bigcup_{i \in I} M'_i$ for $k = 1, 2, \ldots, 2n-3$ and $I \subseteq \{1, 2, \ldots, 2n-3\}$ with $|I| = k$, we discover that there is a multicolored forest F_{n-1}.

Continuing in this way, we obtain all the multicolored forests. ■

Proof (Theorem 2). We first show that inequality (4) is necessary. Suppose $(c_1, c_2, \ldots, c_{2n-1})$ is a color distribution for G, and suppose there are positive integers r, s with $\sum_{i=1}^{r+s} \leq rs$. Choose a biclique $B = K_{r,s}$ in G. B contains rs edges, so it is possible to color its edges so that for $i = 1, 2, \ldots, r+s$, color i is applied to exactly c_i edges of B. Now the rest of G may be colored arbitrarily so that the color distribution is $(c_1, c_2, \ldots, c_{2n-1})$.

Because there are only $2n-1$ colors, in order to form a multicolored spanning tree we must use one edge of each color. Thus we must choose $r+s$ edges from B; but as B has only $r+s$ vertices, we cannot avoid forming a cycle.

Now, to show that inequality (4) is sufficient, we suppose that $(c_1, c_2, \ldots, c_{2n-1})$ is a color distribution satisfying (4), and we color G arbitrarily with this color distribution. We apply Rado's theorem to the polygon matroid of G to show that there must be a multicolored spanning tree.

For $i = 1, 2, \ldots, 2n-1$, let C_i be the collection of edges of color i. We show that the rank of U_I is at least k, where $U_I = \bigcup_{i \in I} C_i$, for each $k = 1, 2, \ldots 2n-1$

$$\begin{bmatrix} 5 & 4 & 3 & 2 & 1 \\ 4' & 5 & 4 & 3 & 2 \\ 3' & 4' & 5 & 4 & 3 \\ 2' & 3' & 4' & 5 & 4 \\ 1' & 2' & 3' & 4' & 5 \end{bmatrix} \quad \text{is orthogonal to} \quad \begin{bmatrix} A & A & A & A & A \\ A & B & B & B & B \\ A & B & C & C & C \\ A & B & C & D & D \\ A & B & C & D & E \end{bmatrix}$$

Fig. 2. This proper coloring of $K_{5,5}$ is orthogonal to this decomposition of $K_{5,5}$ into trees.

and each $I \subseteq \{1, 2, \ldots, 2n-1\}$ with $|I| = k$. Fix k and I, and let t be the number of connected components of G_I, the spanning subgraph of G with edge set U_I.

Our objective is to show that $2n - t \geq k$. Assume to the contrary that $t \geq 2n - k + 1$. Let e denote the number of edges in G_I.

First suppose that k is even. Take $r = s = k/2$. Then according to (4), $e \geq k^2/4+1$. On the other hand, according to Lemma 4, $e \leq k^2/4$, a contradiction.

Similarly, for k odd we have $e \geq (k^2 - 1)/4+1$ by (4), but also $e \leq (k^2 - 1)/4$ by Lemma 4, again a contradiction. ■

4 Concluding remarks

We believe that the second statement of Theorem 1 can be strengthened to the following:

Conjecture 6 *If $G = K_{n,n}$ is colored by a* proper *coloring with color distribution $(a_1, a_2, \ldots, a_p)$ satisfying (3), then it is possible to partition G into multicolored* trees *of sizes $2n-1, 2n-3, \ldots, 3, 1$.*

Examples of such orthogonal decompositions are easy to come by. The example in Figure 2 extends to a family of decompositions of this sort.

The stronger hypothesis is necessary, since colorings which are not proper need not allow a decomposition into multicolored trees. Figure 3 shows representatives from a family of such colorings.

$$\begin{bmatrix} 3 & 2 & 2' \\ 3 & 2 & 2' \\ 3 & 1 & 1' \end{bmatrix}$$

(a) coloring of $K_{3,3}$

$$\begin{bmatrix} 4 & 3 & 2 & 2' \\ 4 & 3 & 2 & 2' \\ 4 & 3 & 1 & 1' \\ 4 & 3' & 3' & 3' \end{bmatrix}$$

(b) coloring of $K_{4,4}$

$$\begin{bmatrix} 5 & 4 & 3 & 2 & 2' \\ 5 & 4 & 3 & 2 & 2' \\ 5 & 4 & 3 & 1 & 1' \\ 5 & 4 & 3' & 3' & 3' \\ 5 & 4' & 4' & 4' & 4' \end{bmatrix}$$

(c) coloring of $K_{5,5}$

Fig. 3. A family of colorings of $K_{n,n}$ which do not permit a decomposition into multicolored trees.