

COMMUNICATION

DECOMPOSITION OF PRODUCT GRAPHS INTO COMPLETE BIPARTITE SUBGRAPHS

Bruce REZNICK*, Prasoon TIWARI, and Douglas B. WEST

University of Illinois, Urbana, Illinois 61801, U.S.A.

Communicated by G.-C. Rota

Received 17 May 1985

Let $\tau(G)$ be the minimum number of complete bipartite subgraphs needed to partition the edges of G . Let $G_{\vec{n}}$ be the weak product of cliques, $K_{n_1} \times \cdots \times K_{n_r}$. This graph has vertices $\{\vec{x} : 0 \leq x_i < n_i\}$, with edges between those vectors that differ in each coordinate.

Theorem:

$$\tau(G_{\vec{n}}) = \sum_{|S| \text{ even}} \prod_{i \notin S} (n_i - 1).$$

1. Introduction

Graph decomposition is the expression of a graph G as a disjoint union of subgraphs, which can also be viewed as partitioning the edges of G . When the subgraphs that may be used are restricted to some family \mathbf{F} , we study the **F**-decomposition number, which is the minimum number of subgraphs from \mathbf{F} in a decomposition of G . (We consider only undirected simple graphs: no loops or multiple edges.) For example, the minimum number of forests needed to partition the edges of G is its “arboricity”. Decompositions in which each subgraph has bounded maximum degree have also been studied. Let \mathbf{F} be the collection of all complete bipartite graphs $K_{m,n}$, we study the **F**-decomposition number $\tau(G)$. Note that listing the subgraphs in a minimum complete bipartite decomposition of G will generally give a shorter representation of G than the adjacency list representation.

We refer to $K_{1,n}$ as a *star* “centered” on the single vertex of high degree. The number of vertices in the complement of a maximum independent set gives an upper bound on $\tau(G)$, since the edges can be partitioned using stars centered on

* Supported by the National Science Foundation and a fellowship from the Alfred P. Sloan Foundation

those vertices. Erdős conjectured that this gives the best result for almost all graphs. We consider a highly structured class of graphs in which the optimum decomposition is considerably better.

Graham and Pollak [1] proved algebraically that for K_n the optimum decomposition is that mentioned above, using $n - 1$ stars. Tverberg [5] gave a shorter proof using a system of homogeneous linear equations, also found independently by Lovász [3]. G.W. Peck [4] recently obtained a similar short proof using rank of matrices; no purely combinatorial short proof is known. We determine τ for more general graphs $G_{\vec{n}}$ determined as follows by an integer vector $\vec{n} = n_1, \dots, n_r$. Provide a vertex in $G_{\vec{n}}$ for each integer vector \vec{x} with $0 \leq x_i < n_i$, and place edges between vertices whose vectors differ in every coordinate. (We refer to the vertices and the corresponding vectors interchangeably.)

We follow the algebraic argument of Graham and Pollak to obtain the correct lower bound; the more difficult part is constructing the decomposition that achieves this. Letting S denote arbitrary subsets of the indices, we obtain $\tau(G_{\vec{n}}) = \sum_{|S| \text{ even}} \prod_{i \notin S} (n_i - 1)$. The initial term, $\prod (n_i - 1)$, corresponds in the construction to stars centered at each vertex that is nonzero in every component; the edges between vertices that have zeros can be partitioned more efficiently than by using stars.

Write \sim to denote adjacency of vertices. In general, the *weak product* $G \times H$ of G and H has vertices $\{(x, y)\}$, with $(x, y) \sim (x', y')$ if $x \sim x'$ in G and $y \sim y'$ in H . Note that $G_{\vec{n}} = K_{n_1} \times \dots \times K_{n_r}$.

2. The lower bound

For any graph G , let $A(G)$ be its symmetric 0, 1 adjacency matrix, and let p and q be the number of positive and negative eigenvalues of A , respectively (counted with multiplicity). As proved by Graham and Pollak [1], we claim $\tau(G) \geq \max\{p, q\}$; we follow their proof.

The approach is algebraic; note that the quadratic form associated with $A(G)$ can be written in terms of products of variables associated with adjacent vertices: $Q_G(z_1, \dots, z_n) = z^T A z = 2 \sum_{i \sim j} z_i z_j$. If G is bipartite, with partite sets X and Y , then $Q_G = 2 \sum_{i \in X} z_i \sum_{j \in Y} z_j$. Thus, for an arbitrary G , $\tau(G)$ is the minimum m for which we can write $Q_G(\vec{z}) = 2 \sum_{k=1}^m \sum_{i \in X_k} z_i \sum_{j \in Y_k} z_j$.

Consider an arbitrary matrix M with p positive and q negative eigenvalues, and let Q be its associated quadratic form. Suppose Q is written as a linear combination of purely squared variables, i.e. $Q(\vec{z}) = \sum \lambda_i w_i^2$, where $w_i = \sum_{j=1}^n d_{ij} z_j$. If r of $\{\lambda_i\}$ are positive and s of $\{\lambda_i\}$ are negative, then Sylvester's Law of Inertia (see [2], for example) states that $r \geq p$ and $s \geq q$. In particular, suppose $\{X_k, Y_k\}$ are $\tau(G)$ complete bipartite subgraphs partitioning G . Then $Q_G(\vec{z}) = 2 \sum_{k=1}^{\tau(G)} \bar{u}_k \bar{v}_k = \sum_{k=1}^{\tau(G)} \frac{1}{2} (\bar{u}_k + \bar{v}_k)^2 - \frac{1}{2} (\bar{u}_k - \bar{v}_k)^2$, where $\bar{u}_k = \sum_{i \in X_k} z_i$ and $\bar{v}_k = \sum_{j \in Y_k} z_j$. Hence $\tau(G) \geq \max\{p, q\}$.

For weak products of graphs, the eigenvectors can be obtained from the eigenvectors of the factors. For any weak product $G \times H$, the adjacency matrix $A(G \times H)$ can be partitioned into blocks i, j corresponding to vertex pairs x_i, x_j in G such that that i, j th block equals the i, j th entry of $A(G)$ times the matrix $A(H)$. In other words, $A(G \times H)$ is the *Kronecker product* $A(G) \otimes A(H)$. Suppose M, N are matrices with eigenvalues λ, μ corresponding to eigenvectors \bar{u}, \bar{v} , and let \bar{w} be the vector formed by concatenating $u_1\bar{v}, u_2\bar{v}, \dots$. Then $(M \otimes N)\bar{w} = \lambda\mu\bar{w}$. If collections of \bar{u} 's and \bar{v} 's are independent, then the \bar{w} 's so formed are also independent. Hence, a full set of eigenvectors for M and N yields a full set of eigenvectors for $M \otimes N$.

Now consider the graphs $G_{\bar{n}}$. We have $A(G_{\bar{n}}) = A(K_{n_1}) \otimes \dots \otimes A(K_{n_t})$. Since $A(K_r)$ has eigenvalues $\{r-1, -1\}$ with multiplicities $(1, r-1)$, it follows that to every subset $S \subseteq \{1, \dots, t\}$ we may associate the eigenvalue $(-1)^{|S|} \prod_{i \notin S} (n_i - 1)$ with multiplicity $\prod_{i \in S} (n_i - 1)$. Therefore, the multiplicity of positive and negative eigenvalues are $p = \sum_{|S| \text{ even}} \prod_{i \in S} (n_i - 1)$ and $q = \sum_{|S| \text{ odd}} \prod_{i \in S} (n_i - 1)$. Note that $p + q = \prod n_i$ and $p - q = \sum_S (-1)^{|S|} \prod_{i \in S} (n_i - 1) = (-1)^t \prod (n_i - 2)$. Hence the maximum of p and q is obtained by summing over products *leaving out* an even number of coordinates, or in other words $\tau(G_{\bar{n}}) \geq \sum_{|S| \text{ even}} \prod_{i \notin S} (n_i - 1)$.

3. The construction

Theorem. $\tau(G_{\bar{n}}) = \sum_{|S| \text{ even}} \prod_{i \notin S} (n_i - 1)$.

Proof. It remains only to provide a construction for the upper bound. As mentioned in the introduction, use $\prod (n_i - 1)$ stars to partition the edges incident to vertices with no zero coordinate. For the remaining edges, we define $\prod_{i \notin S} (n_i - 1)$ bipartite subgraphs for each set of coordinates S of even size.

Each of these bipartite subgraphs can be described as follows. In each coordinate, the vertices in one part have a fixed value, and the vertices in the other part take on any value other than that. Thus each has $\prod (n_i - 1)$ edges. It remains to specify the fixed value and which part has it, for each coordinate.

We encode the subgraphs with a symbol for each coordinate in each part, making a distinction between when the fixed value is 0 or some other value. This leads to four symbols:

- O – Every vertex in this part has this coordinate 0.
- X – Every vertex in this part has this coordinate any nonzero value.
- C – Every vertex in this part has this coordinate some nonzero constant.
- ⊕ – Every vertex in this part has this coordinate any value differing from the corresponding nonzero C (in particular, allowing 0).

Thus, the two parts of a subgraph can be described by strings with a symbol for each coordinate; an example with $t = 11$ appears below. Note that O and X

always appear opposite each other in a given coordinate, and Θ and C always appear opposite each other. The nonzero constant that Θ must avoid is of course the nonzero constant occurring as the corresponding C. Subscripts are placed on the C's and Θ 's to emphasize that the value of the special constant may differ from coordinate to coordinate.

$$\begin{array}{cccccccccccc} \Theta_1 & \Theta_2 & X & C_4 & O & \Theta_6 & X & O & X & C_{10} & O \\ C_1 & C_2 & O & \Theta_4 & X & C_6 & O & X & O & \Theta_{10} & X \end{array}$$

Now fix a set of coordinates S having even size. The subgraphs corresponding to S all have X's and O's in the positions belonging to S for the encoding of each part. Viewing the coordinates cyclically (i.e., modulo t), these alternate between XO and OX. To encode the remaining coordinates, use Θ if the cyclically preceding member of S in this part received O and the subsequent member received X. Use C if the cyclically preceding member of S received X. This rule was used in the example above. The alternation of X and O guarantees that the same graph results regardless of which part is considered.

Note that a C appears in one part or the other for each $i \notin S$. Since a different subgraph arises for each nonzero choice of $\{C_i\}$, this encoding yields $\prod_{i \notin S} (n_i - 1)$ bipartite subgraphs corresponding to S .

To verify that this is a decomposition, it suffices to show that each edge of $G_{\bar{n}}$ appears in exactly one of the subgraphs. We need only consider edges with at least one zero on each endpoint, since the others are supplied by the stars. The following algorithm retrieves the encoding for the subgraph containing $\bar{x}\bar{y}$; applying it to the edge between $\bar{x} = 03210010210$ and $\bar{y} = 14021302001$ yields the encoding illustrated above, with $C_i = 1, 4, 1, 3, 1$ for $i = 1, 2, 4, 6, 10$.

It suffices to determine the location of the O's; the X's appear opposite them, and the Θ 's and C's are distributed by the cyclic rule above. View the coordinates cyclically. If $x_i = 0$, then coordinate i receives O in the encoding of part X if and only if the next 0 that appears cyclically belongs to \bar{y} . Note that all the positions with 0's in \bar{x} or \bar{y} that do not receive O's by this rule receive Θ 's. Similarly, the coordinates encoded as C correspond to nonzero values. Thus the edge $\bar{x}\bar{y}$ does belong to the subgraph described. Furthermore, of the subgraphs defined above, this is the only one containing this edge, because placing the O's and X's in any other positions or any other amounts would place an O in a position where the vertex is nonzero or a C in a position where the vertex is zero. \square

We should note that it does not always hold that $\tau(G \times H)$ is the minimum of its number of positive and negative eigenvalues, even when this does hold for G and H . A small example where it fails is the product of two 5-cycles.

References

- [1] R.L. Graham and H.O. Pollak, On embedding graphs in squashed cubes, Springer Lect. Notes in Math. 303 (1973) 99–110.

- [2] S. Lipshutz, *Linear Algebra* (McGraw–Hill, New York, 1968) 265.
- [3] L. Lovász, unpublished.
- [4] G.W. Peck, A new proof of a theorem of Graham and Pollak, *Discrete Math.* 49 (1984) 327–328.
- [5] H. Tverberg, On the decomposition of K_n into complete bipartite subgraphs, *J. Graph Theory* 6 (1982) 493–494.