On Clique Coverings of Complete Multipartite Graphs

Akbar Davoodi[*]   Dániel Gerbner[†]   Abhishek Methuku[†]   Máté Vizer[†]

September 6, 2018

Abstract

A *clique covering* of a graph G is a set of cliques of G such that any edge of G is contained in one of these cliques, and the *weight* of a clique covering is the sum of the sizes of the cliques in it. The *sigma clique cover number* $scc(G)$ of a graph G, is defined as the smallest possible weight of a clique covering of G.

Let $K_t(d)$ denote the complete t-partite graph with each part of size d. We prove that for any fixed $d \geq 2$, we have

$$\lim_{t \to \infty} scc(K_t(d)) = \frac{d}{2} t \log t.$$

This disproves a conjecture of Davoodi, Javadi and Omoomi [4].

1   Introduction

Let $\mathcal{F}$ be a family of not necessarily distinct sets $S_1, S_2, \ldots, S_n$. We define the *intersection graph* of $\mathcal{F}$ in the following way: Put a vertex v_i corresponding to the set S_i, for every $1 \leq i \leq n$, and two vertices v_i and v_j ($i \neq j$) are adjacent if and only if the corresponding sets intersect. In other words, the edge set of this graph is $\{\{v_i, v_j\} : S_i \cap S_j \neq \emptyset, \ i \neq j\}$. Intersection graphs have many applications in real life problems (e.g. see [11]).

On the other hand, given an n-vertex graph G with vertex set $\{v_1, v_2, \ldots, v_n\}$, one can find a family of sets, $S_1, S_2, \ldots, S_n$ such that for every i, j with $i \neq j$, v_i and v_j are adjacent in G if and only if $S_i \cap S_j \neq \emptyset$ (for example each S_i could be defined as the set of edges incident to the vertex v_i). This is called a *set intersection representation* of G. In other words, a set intersection representation of G is a function $\mathcal{R} : V(G) \to \mathcal{P}(L)$, where $\mathcal{P}(L)$ is the family

---

[*]School of Mathematics, Institute for Research in Fundamental Sciences (IPM), P.O.Box: 19395-5746, Tehran, Iran, E-mail: `davoodi@ipm.ir`

[†]MTA Alfréd Rényi Institute of Mathematics, Budapest, Hungary, E-mails: `gerbner@renyi.hu`, `vizermate@gmail.com`, `abhishekmethuku@gmail.com`

of all subsets of a set L of labels such that for every two distinct vertices $v_i, v_j \in V(G)$, they are adjacent if and only if $\mathcal{R}(v_i) \cap \mathcal{R}(v_j) \neq \emptyset$. . Note that for a fixed graph there may be several different set intersection representations and there are many related parameters that can be studied. Often the goal is to give a set intersection representation with minimum number of labels. This number is the so-called *intersection number* of G, denoted by $i(G)$, see for example [6]. Another goal is to minimize $\max_{v \in V(G)} |\mathcal{R}(v)|$ over all the set intersection representations $\mathcal{R}$. This minimum is called the *local clique cover number* of G and is denoted by $\mathrm{lcc}(G)$ (see [2, 9]).

In this article, we are interested in a third parameter, where the goal is to minimize $\sum_{v \in V(G)} |\mathcal{R}(v)|$ over all the set intersection representations $\mathcal{R}$. This minimum is called the *sigma clique cover number* or the *edge clique cover sum* of G, and is denoted by $\mathrm{scc}(G)$ (see [4, 10]).

Let us explain the connection between set intersection representations and clique coverings. A *clique* of size k is a complete graph on k vertices. A *clique covering* of a graph G is a set $\mathcal{C} = \{C_1, \ldots, C_s\}$ of cliques of G such that any edge of G is contained in one of the cliques in $\mathcal{C}$. A *clique partition* is a clique covering where every edge of G is contained in exactly one of the cliques in $\mathcal{C}$. The *clique cover number* $cc(G)$ is the smallest number of cliques in a clique covering of G, while the *clique partition number* $cp(G)$ is the smallest number of cliques in a clique partition of G.

Note that there is a one-to-one correspondence between the clique coverings of G and the set intersection representations of G, as follows. Let G be a given graph and let $\mathcal{C}$ be a clique covering of G. We consider the set of labels $L = \mathcal{C}$. For every vertex $v \in V(G)$, let $\mathcal{R}(v)$ be the set of cliques containing v. Now, if two vertices x and y are adjacent, then by definition of clique covering, there is a clique in $\mathcal{C}$ covering the edge xy. This clique appears in both sets $\mathcal{R}(x)$ and $\mathcal{R}(y)$. This shows $i(G) \leq cc(G)$. For the other direction, observe that for every label $l \in L$ in a set intersection representation of G, the set of vertices of G containing l forms a clique in G, because these vertices are pairwise adjacent by definition. This shows $i(G) \geq cc(G)$. Therefore, $i(G) = cc(G)$. (For more details see [11].)

In the same vein, being a clique partition is equivalent to have the condition that for every two sets S_i and S_j with $i \neq j$, we have $|S_i \cap S_j| \leq 1$ in the language of set systems.

Erdős, Goodman and Pósa [6] showed that for any graph G with n vertices, there exists a set S with $\lfloor \frac{n^2}{4} \rfloor$ elements and a family $S_1, S_2, \ldots, S_n$ of (not necessarily distinct) subsets of S such that two vertices v_i and v_j are adjacent in G if and only if the sets S_i and S_j (corresponding to v_i and v_j, respectively) intersect. In other words they proved that $i(G) \leq \lfloor \frac{n^2}{4} \rfloor$. Furthermore, $\lfloor \frac{n^2}{4} \rfloor$ cannot be replaced by a smaller number as shown by the complete balanced bipartite graph on n vertices.

In this paper, we consider a weighted version of the above parameters. The *weight* of a set of cliques $\mathcal{C}$ is the sum of the sizes of the cliques in $\mathcal{C}$. Equivalently, the weight is $\sum_{i=1}^{n} |S_i|$, where $S_1, S_2, \ldots, S_n$ is the corresponding set intersection representation of the graph. Therefore, *sigma clique cover number* $\mathrm{scc}(G)$ is the smallest weight of a clique covering

of G. In other words,

$$\text{scc}(G) = \min_{\mathcal{C}: \text{ Clique Covering}} \sum_{c \in \mathcal{C}} |c| = \min_{S_1, S_2, \ldots, S_n} \sum_{i=1}^{n} |S_i|, \tag{1}$$

where the first minimum is taken over all possible clique coverings $\mathcal{C}$ of G, while the latter minimum is taken over all set intersection representations $S_1, S_2, \ldots, S_n$ of G. Similarly, the *sigma clique partition number*, denoted by $\text{scp}(G)$, is the smallest weight of a clique partition (see for example [5]) of G. That is,

$$\text{scp}(G) = \min_{\mathcal{C}: \text{ Clique Partition}} \sum_{c \in \mathcal{C}} |c| = \min_{S_1, S_2, \ldots, S_n} \sum_{i=1}^{n} |S_i|, \tag{2}$$

where the first minimum is taken over all possible clique partitions $\mathcal{C}$ of G, while the latter minimum is taken over all set intersection representations $S_1, S_2, \ldots, S_n$ in which every pair of S_i's intersect in at most one element. This extra condition is equivalent to saying that every edge of G is covered by a unique clique in the clique covering.

Therefore, it is obvious that for every graph G, we have $\text{scc}(G) \leq \text{scp}(G)$. In the 5th Hungarian Combinatorial Colloquium, Katona and Tarján raised the following conjecture. If G is a graph on n vertices, then we have $\text{scc}(G) \leq \lfloor \frac{n^2}{2} \rfloor$. This conjecture was proved independently by Győri and Kostochka [8], Kahn [10] and Chung [3]. Moreover in [8] and [3], it is proved that $\text{scp}(G) \leq \lfloor \frac{n^2}{2} \rfloor$ and the only extremal construction is the Turán bipartite graph. Note that this result is a stronger statement than the above mentioned result of Erdős, Goodman and Pósa [6]. Recently, this upper bound for scc was improved for a large class of graphs. Using the probabilistic method, Davoodi, Javadi and Omoomi in [4] showed that if G is a graph on n vertices with no isolated vertices and $\Delta(\overline{G}) = d - 1$, then we have

$$scc(G) \leq (e^2 + 1)nd \left\lceil \ln \left(\frac{n-1}{d-1} \right) \right\rceil. \tag{3}$$

In [4] they raised the following question to see if this bound is sharp.

Question 1. *For positive integers n, d, how large can the sigma clique cover number of an n−vertex graph be, if the maximum degree of its complement is $d - 1$?*

Complete Multipartite Graphs

In [4] Davoodi, Javadi and Omoomi investigated the sigma clique cover number of complete multipartite graphs as they were conjectured to be examples where the upper bound (3) for *scc* is sharp (at least for large n). They proved the following.

Theorem 2 (Davoodi, Javadi and Omoomi [4]). *For positive integers n, d with $n \geq 2d$, let G be a complete multipartite graph on n vertices with at least two parts of size d and the other parts of size at most d. Then $scc(G) \geq nd$. Moreover, if d is a prime power and $n \leq d(d+1)$, then $scc(G) = scp(G) = nd$.*

3

Let $K_t(d)$ denote the complete t-partite graph with each part of size d. We know that $scc(K_t(d)) \leq cd^2 t \log t$ for some constant c, by (3). On the other hand, by Theorem 2, we have $scc(K_t(d)) = d^2 t$ when $t \leq (d+1)$ and d is a prime power.

Conjecture 3 (Davoodi, Javadi and Omoomi [4])**.** *There exists a function f and a constant c, such that for every positive integers t and d, if $t \geq f(d)$, then*

$$scc(K_t(d)) \geq cd^2 t \log t.$$

Let us state an equivalent reformulation.

Conjecture 4 (Davoodi, Javadi and Omoomi [4])**.** *There exists a function f and a constant $c > 0$ such that if $d \geq 2$, $t \geq f(d)$ and $\mathcal{F} = \{(A_i^1, A_i^2, \ldots, A_i^d) \ : \ 1 \leq i \leq t\}$ with the property that $A_i^j \cap A_{i'}^{j'} = \emptyset$ if and only if $i = i'$ and $j \neq j'$. Then we have*

$$\sum_{i,j} |A_i^j| \geq cd^2 t \log t.$$

Note that the affirmative answer to this conjecture would imply that the upper bound (3) for scc is best possible up to a constant factor, at least for sufficiently large n. However, we will disprove this conjecture.

Qualitatively independent partitions

A partition of a set into d classes is called a *d-partition*. Two partitions P and P' of an n-element set are called *qualitatively independent* if every class of P has a non-empty intersection with every class of P'. This definition has an intuitive meaning in probability theory. Two partitions can be generated by two independent random variables if and only if the partitions are qualitatively independent.

Note that if we have a family of d-partitions $\mathcal{F}_i = \{(A_i^1, A_i^2, \ldots, A_i^d)\}$ $(1 \leq i \leq t)$ such that any two are qualitatively independent, then $A_i^j \cap A_{i'}^{j'} = \emptyset$ if and only if $i = i'$ and $j \neq j'$. This is the property of the families prescribed in Conjecture 4. On the other hand, if we are given families $\mathcal{F}_i = \{(A_i^1, A_i^2, \ldots, A_i^d)\}$ $(1 \leq i \leq t)$ with $A_i^j \cap A_{i'}^{j'} = \emptyset$ if and only if $i = i'$ and $j \neq j'$, then let $X := \bigcup_{i=1}^t \bigcup_{j=1}^d A_i^j$ and $B_i^d := X \setminus \bigcup_{j=1}^{d-1} A_i^j$. Then the families $\mathcal{F}_i' = \{(A_i^1, A_i^2, \ldots, A_i^{d-1}, B_i^d)\}$ $(1 \leq i \leq t)$ are pairwise qualitatively independent. In other words, for each i we extend the partial partition given by $\mathcal{F}_i$ to a full partition by adding the unused elements of X to A_i^d.

Gargano, Körner and Vaccaro [7] studied the following. Let $N(n, d)$ be the largest cardinality of a family of d-partitions of an n-set under the restriction that any two partitions in the family are qualitatively independent.

Theorem 5 (Gargano, Körner and Vaccaro [7])**.** *For every d we have*

$$\lim_{n \to \infty} \frac{1}{n} \log N(n, d) = \frac{2}{d}.$$

4

Note that Gargano, Körner and Vaccaro stated their result with lim sup instead of lim, but it is not hard to see that actually it holds with lim.

Our results

We disprove Conjecture 4 by proving the following theorem:

Theorem 6. *For any $d \geq 2$ and any $\epsilon > 0$, there exist a constant $C_d < \frac{1}{2} + \epsilon$ and $t(\epsilon)$ such that for all $t \geq t(\epsilon)$ there is a family $\mathcal{F}_t = \{(A_i^1, A_i^2, \ldots, A_i^d) : 1 \leq i \leq t\}$ with the property that $A_i^j \cap A_{i'}^{j'} = \emptyset$ if and only if $i = i'$ and $j \neq j'$, and*

$$\sum_{A_i^j \in \mathcal{F}_t} |A_i^j| \leq C_d dt \log t.$$

We also prove the following lower bound:

Theorem 7. *Let $d, t \geq 2$ and $\mathcal{F} = \{(A_i^1, A_i^2, \ldots, A_i^d) : 1 \leq i \leq t\}$ such that $A_i^j \cap A_{i'}^{j'} = \emptyset$ if and only if $i = i'$ and $j \neq j'$. Then we have*

$$\frac{d}{2} t \log t \leq \sum_{A_i^j \in \mathcal{F}} |A_i^j|.$$

By Theorem 6 and Theorem 7 we get:

Corollary 8. *For any $d \geq 2$ we have*

$$\lim_{t \to \infty} scc(K_t(d)) = \frac{d}{2} t \log t.$$

2   Proofs

Proof of Theorem 6

By Theorem 5, we have that for any $\epsilon_1 > 0$ there is an $n(\epsilon_1)$ such that for all $n \geq n(\epsilon_1)$ there is a qualitatively independent family $\mathcal{F}_n(\epsilon_1) = \{(A_i^1, A_i^2, \ldots, A_i^d) : 1 \leq i \leq N(n, d)\}$ on an n-set with the property that

$$\frac{2}{d} - \epsilon_1 \leq \frac{1}{n} \log(N(n, d)).$$

Or equivalently we have $(n \geq n(\epsilon_1))$

$$e^{\frac{2n}{d} - \epsilon_1 n} \leq N(n, d).$$

For any fixed d, choose $\epsilon_1 < \frac{2}{d}$ and for $n \geq n(\epsilon_1)$ let $t := N(n, d)$. Note that

$$\sum_{A_i^j \in \mathcal{F}_n(\epsilon_1)} |A_i^j| = tn,$$

thus we have

$$\sum_{A_i^j \in \mathcal{F}_n(\epsilon_1)} |A_i^j| \leq \frac{1}{2 - d\epsilon_1} dt \log t.$$

Therefore, if we choose ϵ_1 small enough compared to ϵ, then we are done.

$\square$

Proof of Theorem 7

Let $A = [A_{i,j}]$ be a t by d matrix, where $A_{i,j}$ is corresponding to the set A_i^j defined in the statement of Theorem 7. For every fixed two columns of this matrix, two sets intersect if and only if they are not in a same row of the matrix (equivalently they do not belong to a same part of the multipartite graph). Therefore, the sets corresponding to every two columns satisfies Bollobás's Two Families Theorem [1]. Thus, for two consecutive columns we have

$$\sum_{i=1}^{t} \binom{|A_i^j| + |A_i^{j+1}|}{|A_{i,j}|}^{-1} \leq 1.$$

To prove the lower bound we apply Bollobás's theorem d times. First, for the first and second columns, then for the second and the third columns and finally for the last and the first columns. To simplify the notation, let $k_{i,j} := |A_i^j|$ for $1 \leq i \leq t$, $1 \leq j \leq d$ and $k_{i,d+1} := k_{i,1}$. In this way, we have the following inequality

$$\sum_{j=1}^{d} \sum_{i=1}^{t} \binom{k_{i,j} + k_{i,j+1}}{k_{i,j}}^{-1} \leq d.$$

Let $f(m) = \binom{m}{m/2}^{-1}$ for every even integer m, and let $f(x)$ be its linear extension in $\mathbb{R}_{>0}$. Then $f(x)$ is convex and note that for every integers a_i and b_i, $\binom{a_i+b_i}{a_i}^{-1} \geq \binom{a_i+b_i}{\lfloor \frac{a_i+b_i}{2} \rfloor}^{-1} \geq f(a_i + b_i)$. Thus we have

$$\sum_{j=1}^{d} \sum_{i=1}^{t} f(k_{i,j} + k_{i,j+1}) \leq d.$$

On the other hand, by Jensen's inequality,

$$f\left(\frac{2 \sum_{i,j} k_{i,j}}{td}\right) \leq \frac{1}{td} \sum_{j=1}^{d} \sum_{i=1}^{t} f(k_{i,j} + k_{i,j+1}).$$

6

Hence, $f\left(\frac{2\sum_{i,j}k_{i,j}}{td}\right) \leq \frac{1}{t}$, and by $\binom{n}{r} \leq 2^n$ we conclude that

$$\sum_{i,j}k_{i,j} \geq \frac{td}{2}\log t.$$

$\square$

3   Concluding remarks

Let us note that Question 1 remains open. Recall that it was conjectured that $K_t(d)$ was the natural candidate for showing the sharpness of (3). However, Theorem 6 implies that this is not the case. This suggests that perhaps (3) can be improved, but Theorem 7 shows that it cannot be improved by more than a factor of d. If on the other hand (3) is sharp (for n large enough), then one needs to find another candidate graph G with $\Delta(\overline{G}) = d - 1$ and large sigma clique cover number.

Acknowledgment

References

[1] B. Bollobás. On generalized graphs. *Acta Mathematica Hungarica*, **16**(3–4):447–452, 1965.

[2] Cs. Bujtás, A. Davoodi, E. Győri, and Zs. Tuza. Clique coverings and claw-free graphs. *arXiv*:1608.07686.

[3] F. R. K. Chung. On the decomposition of graphs. *SIAM J. Algebraic Discrete Methods* **2**(1), 1–12, 1981.

[4] A. Davoodi, R. Javadi, B. Omoomi. Edge clique covering sum of graphs. *Acta Mathematica Hungarica* **149**(1), 82–91, 2016.

[5] A. Davoodi, R. Javadi, B. Omoomi. Pairwise balanced designs and sigma clique partitions. *Discrete Math.* **339**(5), 1450–1458, 2016.

[6] P. Erdős, A. W. Goodman, L. Pósa. The representation of a graph by set intersections. *Canad. J. Math.*, **18**, 106–112, 1966.

[7] L. Gargano, J. Körner, U. Vaccaro. Sperner capacities. *Graphs and Combinatorics*, **9**(1), 31–46, 1993.

[8] E. Győri, A. V. Kostochka. On a problem of G. O. H. Katona and T. Tarján. *Acta Math. Acad. Sci. Hungar.*, **34**(3-4), 321–327, 1979.

[9] R. Javadi, Z. Maleki, B. Omoomi. Local clique covering of claw-free graphs. *J. Graph Theory*, **81**(1), 92–104, 2016.

[10] J. Kahn. Proof of a conjecture of Katona and Tarján. *Period. Math. Hungar.*, **12**(1), 81–82, 1981.

[11] T. A. McKee, F. R. McMorris. Topics in intersection graph theory. *Vol. 2. SIAM*, 1999.