Clique Partitions
of
Dense Graphs

October 2, 2006

Michael Cavers
Dept of Combinatorics & Optimization
University of Waterloo
200 University Avenue West
Waterloo, Ontario, Canada N2L 3G1

Jacques Verstraëte
Dept of Mathematics & Statistics
McGill University
805 Sherbrooke Street West
Montreal, Quebec, Canada H2A 3K6

Abstract

In this paper, we prove that for any forest $F \subset K_n$, the edges of $E(K_n)\backslash E(F)$ can be partitioned into $O(n \log n)$ cliques. This extends earlier results on clique partitions of the complement of a perfect matching and of a hamiltonian path in K_n. We also show that if a graph G has maximum degree $\triangle$, then the edges of $E(K_n)\backslash E(G)$ can be partitioned into roughly $n^{\frac{3}{2}}\triangle^{\frac{1}{2}}\log^2 n$ cliques provided there exist Steiner systems with certain parameters. Furthermore, as $n \to \infty$, almost every $\triangle$-regular graph G on n vertices has the property that $E(K_n)\backslash E(G)$ cannot be partitioned into fewer than about $\triangle^2[\frac{\log n}{\log \triangle} - 1]^2$ cliques.

1 Introduction

A *clique partition* of a graph G is a collection of complete subgraphs of G (called cliques) that partition the edge set of G. In this paper, we study the problem of finding clique partitions of $K_n\backslash F$, where $F \subset K_n$ is a forest or a graph of maximum degree $\triangle$. Here $K_n\backslash F$ refers to the graph on $V(K_n)$ consisting of all edges of K_n which are not in F, and is called the *complement of* F. We denote by $\mathrm{cp}(K_n\backslash F)$ the *clique partition number* of $K_n\backslash F$, which is the smallest number of cliques partitioning $E(K_n)\backslash E(F)$. Any further notation not defined here is found in Bondy and Murty [6]. Gregory, McGuinness and Wallis [7] proved that the complement of a perfect matching on n vertices can be partitioned into $O(n \log\log n)$ cliques. In the case that F is a forest, we prove the following theorem:

Theorem 1 *Let $F \subset K_n$ be a forest. Then* $\mathrm{cp}(K_n\backslash F) = O(n \log n)$.

The proof of Theorem 1 will be given in Section 2. We are not aware of any lower bounds for $\mathrm{cp}(K_n \backslash F)$ which are not of order of magnitude $O(n)$; we conjecture that there exist forests for which

$$\frac{\mathrm{cp}(K_n \backslash F)}{n} \to \infty.$$

In the second part of the paper, we are interested in estimating clique partition numbers of very dense graphs – we wish to find bounds on $\mathrm{cp}(K_n \backslash G)$ when G has maximum degree $\triangle$. An early result of Erdős and de Bruijn [4] shows that $\mathrm{cp}(K_n \backslash K_2) = n - 1$ for $n \geq 3$, with equality only for clique partitions consisting of $n-2$ complete graphs of order two and one complete graph of order $n-2$. Using projective planes, Wallis [8] showed if H is a graph with at most $\sqrt{n}$ vertices, then $K_n \backslash H$ can be partitioned into $O(n)$ cliques. A projective plane is a particular example of a family of sets called a Steiner system. We recall that a *Steiner* (n, k)*-system* is a family of k-element subsets of an n-element set such that each pair of elements appears in exactly one of the k-element sets. In other words, a Steiner (n, k)-system provides a clique partition of K_n into cliques of size k; in particular $\binom{n}{2}/\binom{k}{2}$ cliques are present in this partition. Conditional on the existence of Steiner systems with certain parameters, we give bounds on $\mathrm{cp}(K_n \backslash G)$ for graphs G with prescribed maximum degree $\triangle(G)$:

Theorem 2 *Let G be a graph on n vertices, let $k = \lfloor (\frac{n}{2\triangle})^{1/2} \rfloor$, and suppose there exists a Steiner (n, k)-system. Then, provided that $\triangle(G) = o(\frac{n}{\log^4 n})$ as n tends to infinity,*

$$\mathrm{cp}(K_n \backslash G) = O(n^{\frac{3}{2}} \triangle(G)^{\frac{1}{2}} \log^2 n).$$

Furthermore, if $\frac{n-\triangle}{\log n} \to \infty$, then as $n \to \infty$, almost every $\triangle$-regular graph G on n vertices has

$$\mathrm{cp}(K_n \backslash G) = \Omega\big(\triangle^2 [\tfrac{\log n}{\log \triangle} - 1]^2\big).$$

For the purpose of comparison, if $\triangle = n^{1-\epsilon}$ in Theorem 2, then the upper bound for $\mathrm{cp}(K_n \backslash G)$ is of order $n^{2-\frac{1}{2}\epsilon} \log^2 n$ whereas the lower bound is of order $n^{2-2\epsilon}$. It would be interesting to determine whether either of these bounds is tight in order of magnitude. We conclude with the following conjecture:

Conjecture 3 *Let G be a graph on n vertices with maximum degree $\triangle$. Then $\mathrm{cp}(K_n \backslash G) = O(\triangle n \log n)$. Furthermore, if $\triangle = o(n)$, then $\mathrm{cp}(K_n \backslash G) = o(n^2)$.*

2 Complements of Forests

To prove Theorem 1, we will restrict our attention to trees and show $\mathrm{cp}(K_n \backslash T) = O(n \log n)$ for any tree T on n vertices. Theorem 1 follows from this statement, since every forest $F \subset K_n$ is contained in a spanning tree $T \subset K_n$, and

$$\mathrm{cp}(K_n \backslash F) \leq \mathrm{cp}(K_n \backslash T) + n - 1.$$

To prove the claim $\mathrm{cp}(K_n \backslash T) = O(n \log n)$, we make use of the following definition.

Definition 4 A *tree partition* of a graph G is a collection of subtrees $\{T_1, T_2, \ldots, T_r\}$ of G such that every edge of G is in exactly one subtree:

$$G = \bigcup_{i=1}^{r} T_i,$$

and for all $i \neq j$, $|V(T_i) \cap V(T_j)| \leq 1$.

For a positive integer b, we say a tree partition $\{T_1, T_2, \ldots, T_r\}$ is *b-smooth* if for some k, $k \leq b|T_i| \leq bk$, for $i = 1, 2, \ldots, r$. The following lemma, which is easily seen to be best possible, will be used to prove Theorem 1.

Lemma 5 *Let T be a tree on n vertices and let $2 \leq k \leq n$. Then there exists a 3-smooth tree partition of T into at most $2n/k$ trees such that every tree in the partition has size at most k.*

Proof. Let T be a tree on n vertices. It is well known that there exists a 2-smooth tree-partition $\{T_1, T_2\}$ of T. To see this, take a tree partition $\{T_1, T_2\}$ of T so that $||V(T_1)| - |V(T_2)||$ is minimized, and assume T_1 and T_2 share vertex v. If $\frac{n}{3} \leq |T_i| \leq \frac{2n}{3}$, for $i = 1, 2$, then we are done. Without loss of generality, suppose that $|T_1| < \frac{n}{3}$. As $||V(T_1)| - |V(T_2)||$ is minimized, v is adjacent to at least two vertices of T_2. Form a tree partition $\{J_1, J_2\}$ of T_2, such that J_1 and J_2 share vertex v. Then

$$\frac{2n}{3} + 1 < |T_1 \cup J_1| + |T_1 \cup J_2| < \frac{4n}{3} + 1.$$

Then $\{T_1 \cup J_1, J_2\}$ or $\{T_1 \cup J_2, J_1\}$ is a 2-smooth tree partition of T.

To finish the proof, we construct a 3-smooth tree partition of T into trees of size at most k. Repeatedly take 2-smooth partitions of all trees of size more than k in the current tree-partition, starting with the trivial tree partition, $\{T\}$. This procedure gives a tree partition of T all of whose trees have size at most k and at least $k/3$, and average size at least $k/2$ as required. ∎

We now use Lemma 5 to prove Theorem 1.

Proof of Theorem 1. Define for $n \in \mathbb{N}$:

$$g(n) = \max\{\text{cp}(K_n \backslash T) : T \text{ is a tree on } n \text{ vertices}\}.$$

Let T be a tree on n vertices such that $g(n) = \text{cp}(K_n \backslash T)$. By Lemma 5, there exists a 3-smooth tree partition $\{T_1, T_2, \ldots, T_r\}$ of T such that $|T_i| \leq \lfloor\sqrt{n}\rfloor$ and $r \leq 2\sqrt{n}$. Without loss of generality, suppose that for $i = 2, 3, \ldots, r$,

$$V(T_i) \cap \left(\bigcup_{j=1}^{i-1} V(T_j)\right) = \{v_i\}.$$

Let t be the smallest integer such that $t \geq 7\sqrt{n}$ and there is a projective plane of order t. By results on the distribution of prime numbers [2], and the fact that there exists a projective plane of every prime order, we can take $t = 7\sqrt{n} + n^{\theta}$ for some $\theta < \frac{1}{2}$ and choose a projective plane of order t. We identify the t^2+t+1 points of the projective plane with the vertices of a complete graph K_{t^2+t+1}, and the blocks form a clique partition of this complete graph. We claim that we can embed the trees $T_1, T_2, \ldots, T_r$ in the cliques $B_1, B_2, \ldots, B_r$ in such a way that the union of these embedded trees is T.

First identify the vertices of T_1 with points from an arbitrary block, say B_1, of the projective plane, where the vertex v_2 is identified with some point w_2, and all other vertices of T_1 are identified arbitrarily with points from $B_1 \backslash \{w_2\}$. Suppose that for some $i : 2 \leq i \leq r$, we have already identified the vertices of T_{i-1} with the points of B_{i-1}, such that vertex v_i is identified with point w_i of some block B_j, where $j \leq i-1$. Pick a block B_i (different from $B_1, B_2, \ldots, B_{i-1}$) that contains the point w_i. There exists such a block as there are $t+1 \geq 7\sqrt{n}$ blocks containing the point w_i (and at most $r \leq 2\sqrt{n}$ blocks have been used). Identify the vertices of T_i with points from B_i such that v_i is identified with w_i, and all other vertices of T_i are identified arbitrarily with points from $B_i \backslash W$, where W is the set of points from $B_1 \cup B_2 \cup \cdots \cup B_{i-1}$ that intersect with B_i. Note that $|W| < r$, as B_i intersects every other block in at most one point. This identification can be done, as each block has $t+1 \geq 7\sqrt{n}$ points, each tree has at most $\lfloor \sqrt{n} \rfloor$ vertices, and removing at most $r-1 \leq 2\sqrt{n}$ points of block B_i leaves at least $\sqrt{n}$ points which can be identified with T_i. This defines the embedding of the tree T_i into B_i for $i = 1, 2, \ldots, r$, and the union of the embedded trees is clearly T.

Now delete points of the projective plane such that each block B_i has $\lfloor \sqrt{n} \rfloor$ points. Then:

$$\begin{aligned} g(n) &\leq t^2+t+1-r+\sum_{i=1}^{r} \operatorname{cp}(K_{|B_i|} \backslash T_i) \\ &= O(n) + \sum_{i=1}^{r} \operatorname{cp}(K_{\lfloor \sqrt{n} \rfloor} \backslash T_i) \\ &\leq O(n) + \sum_{i=1}^{r} g(\lfloor \sqrt{n} \rfloor) \end{aligned}$$

Defining $c(x) = g(\lfloor x \rfloor)$ for $x \in \mathbb{R}$ gives

$$c(x) \leq O(x) + 2\sqrt{x} \cdot c(\sqrt{x}).$$

Dividing through by x and setting $z = \log_2 \log_2 x$, and $h(z) = c(x)/x$ gives,

$$h(z) \leq O(1) + 2 \cdot h(z-1).$$

So, $h(z) = O(2^z)$, for x (and hence z) arbitrarily large. Hence, $c(x) = O(x \log x)$ implying that $g(n) = O(n \log n)$. ∎

3 Dense Graphs

In this section, we prove Theorem 2 using Steiner systems (see Cameron and van Lint [5]) and the probabilistic method (see Alon and Spencer [1]).

Recall that the blocks of a Steiner (n,k)-system correspond to a clique partition of K_n into $\binom{n}{2}/\binom{k}{2}$ cliques of size k. Necessary existence conditions for the existence of an $\mathcal{S}(n,k)$ are

$$n \equiv 1 \mod k-1,$$

$$n(n-1) \equiv 0 \mod k(k-1).$$

Wilson's theorem [3] says that the necessary conditions above for the existence of an $\mathcal{S}(n,k)$ are sufficient for almost all $n \in \mathbb{N}$. However, the proofs presented by Wilson do not give an explicit constant $n_0(k)$ such that an $\mathcal{S}(n,k)$ exists for all $n \geq n_0(k)$ satisfying the necessary conditions. Recently, Chang showed that $n_0(k) \leq \exp(\exp(k^{k^2}))$ (see page 800 in Beth, Jungnickel and Lenz [3]). It is therefore out of the reach of current research to determine for which $k \in \{1,2,\ldots,n\}$ a Steiner (n,k)-system exists.

Proof of Theorem 2. Both the proofs of the upper and lower bounds in Theorem 2 are probabilistic. First we show that for almost every $\triangle$-regular graph G on n vertices,

$$\mathrm{cp}(K_n \backslash G) = \Omega\big(\triangle^2 [\tfrac{\log n}{\log \triangle} - 1]^2\big).$$

Note that this statement is true whenever the right hand side is of order n, since $\mathrm{cp}(K_n \backslash G) = \Omega(n)$ for every $\triangle$-regular graph G. For this part of the proof, let $m = m(\triangle, n)$ be an integer such that as $n \to \infty$,

$$m = o\big(\triangle^2 [\tfrac{\log n}{\log \triangle} - 1]^2\big).$$

Note that $\triangle^2[\frac{\log n}{\log \triangle} - 1]^2 \to \infty$ since $\frac{n-\triangle}{\log n} \to \infty$. For each clique partition $\mathcal{C}$ of a graph G on n vertices, let $B(G,\mathcal{C})$ be the bipartite graph with parts $\mathcal{C}$ and $V(G)$ such that $v \in V(G)$ is joined to $C \in \mathcal{C}$ if $v \in C$. We observe that $B(G,\mathcal{C})$ has no cycles of length four and all vertices $C \in \mathcal{C}$ have degree at least two. Therefore

$$\sum_{C \in \mathcal{C}} \binom{d(C)}{2} \leq \binom{n}{2}.$$

This implies that if $|\mathcal{C}| = m$, then the number of edges in $B(G,\mathcal{C})$ is at most

$$\frac{m}{2} + m\Big(\frac{1}{4} + \frac{n^2}{m}\Big)^{\frac{1}{2}} < 2\sqrt{m}n$$

provided n is large enough.
Let $\beta(m,n,\triangle)$ be the total number of graphs $B(G,\mathcal{C})$ with $|\mathcal{C}| = m$. Then

$$\begin{aligned}\log \beta(m,n,\triangle) \leq \log \sum_{k < 2\sqrt{m}n} \binom{mn}{k} &= O(\sqrt{m}n \log m) \\ &= o(\triangle n \log \triangle \cdot [\tfrac{\log n}{\log \triangle} - 1])\end{aligned}$$

by definition of m. By the results on random regular graphs in Wormald [9], the logarithm of the number $\gamma(n,\triangle)$ of $\triangle$-regular graphs on n vertices is at least

$$\begin{aligned}\log \gamma(n,\triangle) \geq \log \left(\frac{(\triangle n)!}{2^{\frac{\triangle n}{2}} (\frac{\triangle n}{2})! (\triangle !)^n} \cdot \exp[-\Omega(\triangle^2)] \right) &= \Omega(\triangle n \log \tfrac{n}{\triangle}) \\ &= \Omega(\triangle n \log \triangle \cdot [\tfrac{\log n}{\log \triangle} - 1]).\end{aligned}$$

We conclude that for any m as defined above, and since since $\frac{n-\triangle}{\log n} \to \infty$,

$$\lim_{n\to\infty} \frac{\log \beta(m,n,\triangle)}{\log \gamma(n,\triangle)} = 0.$$

In particular, the proportion of $\triangle$-regular graphs $G \subset K_n$ such that $\mathrm{cp}(K_n \backslash G) = m$ tends to zero exponentially fast in n as $n \to \infty$. It follows that almost all $\triangle$-regular graphs G on n vertices have

$$\mathrm{cp}(K_n \backslash G) = \Omega\big(\triangle^2 [\tfrac{\log n}{\log \triangle} - 1]^2\big).$$

This completes the first part of the proof.

Now we prove the upper bound on $\mathrm{cp}(K_n \backslash G)$ given in Theorem 2. Suppose G, n, k satisfy the conditions of the theorem. Let $\mathcal{S} = (X, \mathcal{B})$ be a Steiner system with blocks of size k on n points. For a random permutation of the points, the probability that a fixed set of k points is a fixed block in $\mathcal{B}$ is exactly $1/\binom{n}{k}$. Take G to be a fixed graph on the same set of n points, with maximum degree $\triangle$. Let G_B denote the subgraph of G spanned by the edges contained in a block $B \in \mathcal{B}$. Consider the event $|E(G_B)| \geq r$, for some integer r. Pick a subgraph H_B of G_B with exactly r edges. If the maximum size of a matching in H_B is i for some positive integer $i \leq r$, and if there are s vertices of H_B which are unsaturated by a maximum matching, then

$$\max\left\{\frac{r - \binom{2i}{2}}{2i}, 0\right\} \leq s \leq r - i.$$

For convenience, let $s_i = (r - \binom{2i}{2})/2i$. Let $A_B(i,s)$ denote the event that the largest matching in H_B has size i and H_B has $s + 2i$ vertices. Fixing a matching M of size i in G, there are at most $(2i\triangle)^s$ ways to choose a set S of s vertices so that $V(M) \cup S = H_B$. Then there are $\binom{n-2i-s}{k-2i-s}$ ways to choose the vertices of $B \backslash V(H_B)$ from G. Therefore

$$\mathbb{P}[A_B(i,s)] \leq \frac{1}{\binom{n}{k}} \binom{\triangle n}{i} (2i\triangle)^s \binom{n-2i-s}{k-2i-s}.$$

Since $\{|E(G_B)| \geq r\} \subset \bigcup_{i,s} A_B(i,s)$, it follows that

$$\mathbb{P}[|E(G_B)| \geq r] \leq \frac{1}{\binom{n}{k}} \sum_{i=1}^{r} \sum_{\substack{s \geq s_i \\ s \geq 0}}^{r-i} \binom{\triangle n}{i} (2i\triangle)^s \binom{n-2i-s}{k-2i-s}.$$

To estimate the sums on the right, we use the inequality

$$\frac{\binom{a-t}{b-t}}{\binom{a}{b}} < \frac{b^t}{a^t}.$$

Let j be the largest integer such that

$$r - \binom{2j}{2} \geq 0$$

so that definitely $\sqrt{r/2}-1 \leq j \leq \sqrt{r/2}+1$. Recall $k = \lceil \sqrt{n}/\sqrt{2\triangle} \rceil$. Then

$$\begin{aligned}
\mathbb{P}[|E(G_B)| \geq r] &\leq \sum_{i=1}^{r} \sum_{\substack{s \geq s_i \\ s \geq 0}}^{r-i} \binom{\triangle n}{i} (2i\triangle)^s \frac{\binom{n-2i-s}{k-2i-s}}{\binom{n}{k}} \\
&< \sum_{i=1}^{r} \sum_{\substack{s \geq s_i \\ s \geq 0}}^{r-i} (\triangle n)^i (2r\triangle)^s \frac{k^{2i+s}}{n^{2i+s}} \\
&= \sum_{i=1}^{r} \left(\frac{\triangle k^2}{n}\right)^i \sum_{\substack{s \geq s_i \\ s \geq 0}}^{r-i} \left(\frac{2rk\triangle}{n}\right)^s \\
&\leq \sum_{i=1}^{j} \left(\frac{1}{2}\right)^i \sum_{s=s_i}^{r-i} \left(\frac{1}{2}\right)^s + \sum_{i=j+1}^{r} \left(\frac{1}{2}\right)^i \sum_{s=0}^{r-i} \left(\frac{1}{2}\right)^s \quad (*) \\
&\leq 2\sum_{i=1}^{j} \left(\frac{1}{2}\right)^i \left(\frac{1}{2}\right)^{s_i} + 4\left(\frac{1}{2}\right)^{j+1} \\
&= \sqrt{2}\sum_{i=1}^{j} \left(\frac{1}{2}\right)^{r/(2i)} + \left(\frac{1}{2}\right)^{j-1} \\
&\leq \sqrt{2}j\left(\frac{1}{2}\right)^{r/(2j)} + \left(\frac{1}{2}\right)^{j-1}.
\end{aligned}$$

To justify $(*)$, note that $\frac{2rk\triangle}{n} \leq \frac{1}{2}$ is true if we set $r = 2(\log_2(2n^2\triangle))^2$ and recall that

$$\frac{\triangle(\log_2 n)^4}{n} \to 0.$$

In addition, $\frac{\triangle k^2}{n} \leq \frac{1}{2}$ by our choice of k. These two statements verify $(*)$. We show that the above expression is less than $\frac{1}{|\mathcal{B}|}$. That is, we prove

$$\left[\sqrt{2}j\left(\frac{1}{2}\right)^{\frac{r}{2j}} + \left(\frac{1}{2}\right)^{j-1}\right] \frac{n(n-1)}{k(k-1)} < 1.$$

But, the left hand side of this is

$$\begin{aligned}
\left[\sqrt{2}j\left(\frac{1}{2}\right)^{\frac{r}{2j}} + \left(\frac{1}{2}\right)^{j-1}\right] \frac{n(n-1)}{k(k-1)} &= O\left(\frac{jn^2}{k^2}\left(\frac{1}{2}\right)^{\frac{r}{2j}}\right) \\
&= O\left(\triangle n \log_2(2n^2\triangle)\left(\frac{1}{2}\right)^{\frac{r}{2\log_2(2n^2\triangle)+1}}\right) \\
&= O\left(\triangle n \log_2(2n^2\triangle)\left(\frac{1}{2}\right)^{\log_2(2n^2\triangle)}\right) \\
&= O\left(\frac{\log_2(2n^2\triangle)}{n}\right) \\
&= O(\log_2 n/n).
\end{aligned}$$

So with positive probability $|E(G_B)| < r$ for all $B \in \mathcal{B}$, and

$$\mathrm{cp}(K_n \backslash G) < rk|\mathcal{B}| = O(n^{3/2} \triangle^{1/2} \log^2 n),$$

if n is sufficiently large. This completes the proof. ∎

References

[1] N. Alon and J.H. Spencer, *The Probabilistic Method*, Wiley, New York, (1992). (Second Edition, 2000).

[2] R.C. Baker, G. Harman, and J. Pintz, The difference between consecutive primes, *Proc. London Math. Soc.* **83**(3) (2001) 532-562.

[3] T. Beth, D. Jungnickel, and H. Lenz, *Design Theory, Volume* 2*, Second Edition*, Cambridge Univ. Press, New York, (1999).

[4] N.G. de Bruijn and P. Erdős, On a combinatorial problem, *Indag. Math.* **10** (1948) 421-423.

[5] P.J. Cameron and J.H. Van Lint, *Designs, graphs, codes and their links*, London Mathematical Society Student Texts 22, Cambridge Univ. Press, New York, (1991).

[6] J.A. Bondy and U.S.R. Murty, *Graph theory with applications*, MacMillan, London, (1972).

[7] D.A. Gregory, S. McGuinness, and W. Wallis, Clique partitions of the cocktail party graph, *Discrete Math.* **59** (1986) 267-273.

[8] W.D. Wallis, Finite planes and clique partitions, *Contemporary Mathematics*, **111** (1990) 279-285.

[9] N.C. Wormald, Models of random regular graphs. Surveys in Combinatorics, 1999, J.D. Lamb and D.A. Preece, eds. London Mathematical Society Lecture Note Series, vol 276, pp. 239-298. Cambridge University Press, Cambridge, 1999.