Regarding two questions about clique and biclique partitions

Dhruv Rohatgi[*]

MIT
Cambridge, MA, U.S.A.

drohatgi@mit.edu

John C. Urschel[†]

Institute for Advanced Study
Princeton, NJ, U.S.A.

urschel@mit.edu

Jake Wellens[‡]

Berkeley, CA, U.S.A.

narf1899@gmail.com

Abstract

For a graph G, let $\mathrm{cp}(G)$ denote the minimum number of cliques of G needed to cover the edges of G exactly once. Similarly, let $\mathrm{bp}_k(G)$ denote the minimum number of bicliques (i.e. complete bipartite subgraphs of G) needed to cover each edge of G exactly k times. We make progress on two open questions – one regarding the maximum possible value of $\mathrm{cp}(G) + \mathrm{cp}(\overline{G})$ (due to de Caen, Erdős, Pullman and Wormald) and the other regarding $\mathrm{bp}_k(K_n)$ (due to de Caen, Gregory and Pritikin). For the first, we improve the best-known upper and lower bounds on $\max_G \mathrm{cp}(G) + \mathrm{cp}(\overline{G})$. For the second question, we show that the classical lower bound is tight to leading order, i.e. $\mathrm{bp}_k(K_n) = (1 + o(1))n$.

Mathematics Subject Classifications: 05C35, 05C70

1 Introduction

For a fixed family of graphs $\mathcal{F}$, an $\mathcal{F}$-partition of a graph G is a collection $\mathcal{C} = \{H_1, \ldots, H_\ell\}$ of subgraphs $H_i \subset G$ such that each edge of G belongs to exactly one $H_i \in \mathcal{C}$, and each H_i is isomorphic to some graph in $\mathcal{F}$. When $\mathcal{F} = \{K_r\}_{r \geqslant 2}$, we refer to $\mathcal{F}$-partitions as *clique partitions*, and when $\mathcal{F} = \{K_{s,t}\}_{s,t \geqslant 1}$, the corresponding partitions are called *biclique partitions*. The size $|\mathcal{C}|$ of the smallest clique partition of G is called the *clique partition number* of G, denoted $\mathrm{cp}(G)$. The *biclique partition number* $\mathrm{bp}(G)$ is defined analogously. Both $\mathrm{cp}(G)$ and $\mathrm{bp}(G)$ (and their many variants) are NP-hard to compute

---

in general graphs, but have been studied extensively from a combinatorial perspective, in part because of their connections to various areas of computer science (see, e.g. [13]). In this paper, we consider two longstanding combinatorial questions related to these quantities.

1.1 Biclique partitions of K_n

In 1971, Graham and Pollak [9] showed that, for every $n \geqslant 2$,

$$\mathrm{bp}(K_n) = n - 1. \tag{1.1}$$

In particular, the edges of K_n can be partitioned into $n - 1$ stars

$$K_{1,n-1}, K_{1,n-2}, \ldots, K_{1,1}$$

centered at different vertices, while the corresponding lower bound holds by an elegant linear algebraic argument. The lower bound argument easily generalizes to give

$$\mathrm{bp}_k(K_n) \geqslant n - 1 \tag{1.2}$$

for any k, where $\mathrm{bp}_k(G)$ is the size of the smallest collection $\{H_1, \ldots, H_\ell\}$ of bicliques $H_i \subset G$ such that each edge of G belongs to H_i for exactly k different values of $i \in [\ell]$. As a matter of notation, such a collection is called a *k-biclique cover* of G. More generally, a $\{k_1, \ldots, k_t\}$-*biclique cover* of G is a collection $\{H_1, \ldots, H_\ell\}$ of bicliques $H_i \subset G$ such that for each edge of G there is some $k \in \{k_1, \ldots, k_t\}$ such that the edge belongs to exactly k of the bicliques. (See [5] for a variety of interesting results on $\mathrm{bp}_L(G)$ for certain lists L.)

In 1993, de Caen, Gregory and Pritikin conjectured that (1.2) is tight for sufficiently large n:

Conjecture 1 (de Caen et al. [3]). *For every positive integer k,*

$$bp_k(K_n) = n - 1$$

for all sufficiently large n.

The same authors prove their conjecture for each $k \leqslant 18$, using special constructions from design theory [3]. However, the best-known upper bound for general k is $\mathrm{bp}_k(K_n) = O(kn)$, obtained by simply compounding a small-k construction.

In Section 2, we show that, to leading order, Conjecture 1 is true.

Theorem 2. *For every positive integer k,*

$$\mathrm{bp}_k(K_n) = (1 + o(1))\, n.$$

More precisely, we construct a family of designs (inspired by classical ideas of Nisan and Wigderson [15]), that yields a k-covering of K_n by at most $n + 2kn^{3/4} + k\sqrt{n}$ complete bipartite subgraphs.

1.2 Clique partitions of G and $\overline{G}$

In 1986, de Caen, Erdős, Pullman and Wormald [4] investigated the maximum value of $\mathrm{cp}(G) + \mathrm{cp}(\overline{G})$ over the set $\mathcal{G}_n$ of all graphs G on n vertices, and proved that

$$\frac{7n^2}{25} + O(n) \leqslant \max_{G \in \mathcal{G}_n} \mathrm{cp}(G) + \mathrm{cp}(\overline{G}) \leqslant \frac{13n^2}{30} + O(n). \qquad (1.3)$$

They hypothesized that the lower bound $\frac{7n^2}{25}$ is nearly tight, and left closing the gap in (1.3) as an open problem.[1]

Problem 3 (de Caen et al. [4])**.** Find c such that

$$\max_{G \in \mathcal{G}_n} cp(G) + cp(\overline{G}) \sim cn^2.$$

In Section 3.1, we show that the family of graphs constructed in [4] can actually be modified to improve the lower bound in (1.3). Specifically, we show that the constant c in (3) must be at least $\frac{23}{82} > \frac{7}{25}$.

Theorem 4. *For infinitely many n, there exists a self-complementary graph $G \in \mathcal{G}_n$ with* $\mathrm{cp}(G) \geqslant \frac{23}{164}n^2 + o(n^2)$.

The upper bound in (1.3) essentially comes from greedily selecting edge-disjoint triangles from G and $\overline{G}$, forming clique partitions into K_3's and K_2's. Subsequent work on complementary triangle packings, first by Erdős et al. [7] and later by Keevash and Sudakov [14], improved significantly upon the greedy packing, with the latter authors showing the existence of a packing with $\frac{n^2}{12.89}$ edge-disjoint triangles. The resulting clique partitions (as observed by Bujtas et al. [2]) contain a total of $0.34481n^2 + o(n^2)$ cliques, improving the $0.4\overline{3}n^2$ upper bound in (1.3). However, partitions into triangles and edges can never push this bound below $0.3\overline{3}n^2$, as illustrated by $G = K_{n/2,n/2}$. In Section 3.2, we extend the ideas of Keevash and Sudakov to the complementary clique partition problem, improving (1.3) beyond the limits of triangle packings:

Theorem 5. *For all $G \in \mathcal{G}_n$,* $\mathrm{cp}(G) + \mathrm{cp}(\overline{G}) \leqslant 0.3186n^2 + o(n^2)$.

2 A k-biclique covering of K_n

Our goal in this section is to construct a collection of $(1 + o(1))n$ bicliques on a set of n vertices such that all $\binom{n}{2}$ edges belong to exactly k bicliques in the collection. We recall the definition of a *combinatorial design* in the sense of Nisan and Wigderson [15] from their classical paper on pseudorandom generators.

---

[1]In the same paper [4], the authors solve the corresponding problem for $\mathrm{cc}(G) + \mathrm{cc}(\overline{G})$, where $\mathrm{cc}(G)$ is the minimal number of cliques in G needed to cover every edge *at least* once, showing that $\max_{G \in \mathcal{G}_n} \mathrm{cc}(G) + \mathrm{cc}(\overline{G}) = \frac{n^2}{4}(1 + o(1))$. This is tight up to the $o(1)$ error by $K_{n/2,n/2}$, and the error term was later removed by Pyber [19] for $n > 2^{1500}$.

Definition 6. A family of sets $\{S_1, \ldots, S_n\}$ with $S_1, \ldots, S_n \subseteq [d]$ is a (n, d, t, m)-design if:

1. $|S_i| = m$ for all $i \in [n]$;

2. $|S_i \cap S_j| \leqslant t$ for all $i, j \in [n]$ with $i \neq j$.

We construct our designs in a way that differs from [15] and better suits our particular choice of parameters:

Lemma 7. *For any positive integers m and t, there exists some N such that an (n, d, t, m)-design with $d \leqslant 2mn^{1/(t+1)}$ exists for all $n \geqslant N$.*

Proof. Let N be large enough that there are at least m prime numbers in the interval $[n^{1/(t+1)}, 2n^{1/(t+1)}]$ for every $n \geqslant N$; this is possible by the Prime Number Theorem. Fix some $n \geqslant N$, and choose m distinct primes $p_1, \ldots, p_m$ in this interval. We will pick sets $S_1, \ldots, S_n$ from the disjoint union

$$U = \bigsqcup_{k=1}^{m} \mathbb{Z}/p_k\mathbb{Z}.$$

For $i \in [n]$, let S_i consist of m elements from U, one from each group. Specifically, for $k \in [m]$, pick element $i \pmod{p_k}$ from group $\mathbb{Z}/p_k\mathbb{Z}$.

It is clear that $|S_i| = m$ for all i, and that $d := |U| = \sum_{k=1}^{m} p_k \leqslant 2mn^{1/(t+1)}$. We claim that $|S_i \cap S_j| \leqslant t$ for all distinct $i, j \in [n]$. Indeed, suppose to the contrary that $|S_i \cap S_j| > t$ for some distinct $i, j \in [n]$. Then among the chosen primes, there are $t + 1$ primes $p_{l_1}, \ldots, p_{l_{t+1}}$ with $i \equiv j \pmod{p_{l_k}}$ for each $k \in [t+1]$. But then

$$\prod_{k=1}^{t+1} p_{l_k} \mid (i - j).$$

Since $i \neq j$, it follows that

$$|i - j| \geqslant \prod_{k=1}^{t+1} p_{l_k} \geqslant n,$$

a contradiction. $\qquad\square$

Remark 8. The above design is in fact optimal up to constant factors. Consider any (n, d, t, m)-design, where the sets are contained in a universe U of size d. For every $(t+1)$-element subset of U, there is at most one set among $S_1, \ldots, S_n$ that contains the subset. Since each S_i contains $\binom{m}{t+1}$ subsets of size $t+1$, we must have $\binom{d}{t+1} \geqslant n\binom{m}{t+1}$, so

$$d \geqslant \left(n\binom{m}{t+1}(t+1)!\right)^{1/(t+1)} \geqslant \frac{1}{e}n^{1/(t+1)}m.$$

We will only use the special case $(n, k\sqrt{n}, 1, \lfloor k/2 \rfloor)$ of Lemma 7, which we state explicitly below as a corollary.

Corollary 9. *For any positive integer k and all n sufficiently large, there is some $d \leqslant k\sqrt{n}$ and sets $S_1, \ldots, S_n \subseteq [d]$ with $|S_i| = \lfloor k/2 \rfloor$ for all i, and $|S_i \cap S_j| \leqslant 1$ for all $i \neq j$.*

We also require a result of Alon [1] on $\{1,2\}$-biclique coverings of K_n, which are collections of bicliques such that every edge of K_n belongs to either 1 or 2 of the bicliques in the collection. The size of the smallest such collection is denoted $\mathrm{bp}_{\{1,2\}}(K_n)$.

Fact 10 (Alon, [1]). [2] *For all n, $\mathrm{bp}_{\{1,2\}}(K_n) \leqslant 2\sqrt{n}$.*

Finally, we construct a k-biclique covering of K_n.

Theorem 11. *Let k be a positive integer. Then for all sufficiently large n,*

$$\mathrm{bp}_k(K_n) \leqslant n + 2kn^{3/4} + k\sqrt{n}.$$

Proof. Let n be large enough to apply Corollary 9. Let $S_1, \ldots, S_n \subseteq [d]$ be the sets in the resulting design, and let $p_1, \ldots, p_{\lfloor k/2 \rfloor}$ be the corresponding primes used in the proof of Lemma 7. Define bicliques $B_1, \ldots, B_d \subset K_n$ by letting B_i be the biclique between $\{j \in [n] \mid i \in S_j\}$ and $\{j \in [n] \mid i \notin S_j\}$. Then any edge $\{i,j\}$ is covered exactly $|S_i| + |S_j| - 2|S_i \cap S_j|$ times, and this number is equal to either $2\lfloor k/2 \rfloor$ or $2\lfloor k/2 \rfloor - 2$ (depending on whether $|S_i \cap S_j| = 0$ or 1).

If k is odd, every edge still needs to be covered either 1 or 3 more times. Let us define a *triple-edge* to be an edge $\{i,j\}$ with $|S_i \cap S_j| = 1$. An edge $\{i,j\}$ is a triple-edge if and only if there exists some index l and remainder r such that $i \equiv j \equiv r \pmod{p_l}$. We can define a clique $C_{l,r}$ consisting of all vertices i with $i \equiv r \pmod{p_l}$. Observe that every triple-edge is contained in exactly one such clique, and every such clique contains only triple-edges. To make progress, we will construct a $\{1,2\}$-biclique covering of each clique $C_{l,r}$. The number of cliques $C_{l,r}$ is at most $k\sqrt{n}$, and each has size at most $\sqrt{n}$, so by Fact 10, at most $k\sqrt{n} \cdot 2n^{1/4} = 2kn^{3/4}$ bicliques are needed to $\{1,2\}$-cover every clique $C_{l,r}$. Now every edge needs to be covered 1 or 2 more times.

If k is even, every edge needs to be covered only 0 or 2 more times, so we skip the above step. Finally, in either case, we'll "pad" the covering so that every edge is covered exactly k times. To do this, define bicliques $D_1, \ldots, D_n$ where D_i is the star centered at vertex i and containing edges to all vertices $j < i$ such that $\{i,j\}$ needs to be covered 1 or 2 more times, and to all vertices $j > i$ such that $\{i,j\}$ needs to be covered 2 more times.

This completes the construction. The total number of bicliques used is at most $n + 2kn^{3/4} + k\sqrt{n}$ (from the padding step, the $\{1,2\}$-covering step, and the initial design, respectively). $\square$

Remark 12. A key ingredient in the proof above is the $\{2k-2, 2k\}$-biclique covering of K_n using $2k\sqrt{n}$ bicliques. It is shown in [5] that $\sqrt{n/2}$ bicliques are necessary for this list covering, so the asymptotic dependence on n cannot be decreased.

---

[2]The value of $\mathrm{bp}_{\{1,2\}}(K_n)$ is not known for large n, although the lower bound $\mathrm{bp}_{\{1,2\}}(K_n) \geqslant \sqrt{n-1}$ always holds.
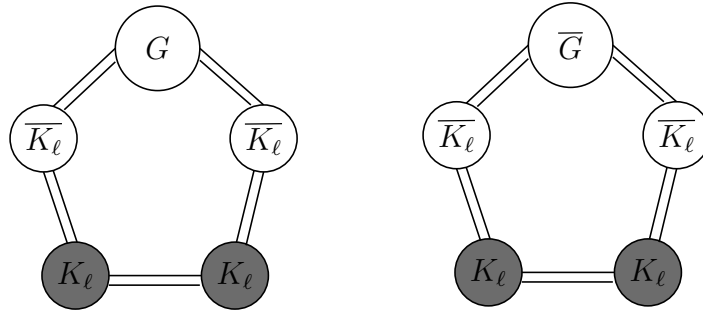
Figure 1: The graph $H_\ell(G)$ and its complement.

3 Clique partitions of a graph and its complement

3.1 Improving the lower bound

The construction in our proof of Theorem 4 is based on the original construction in [4], and the calculation of its clique partition number makes use of certain facts shown in [4] and [17]. Here we include the entire argument for the reader's convenience. Before proceeding with the construction, we need the following lemma, which has appeared in many places but perhaps first in Pullman and Donald [17]. Recall that the edge chromatic number $\chi'(G)$ of a graph G is the minimum number of colors needed to color the edges of G so that no two edges of the same color are incident to the same vertex. We use the notation $G \vee H$ to denote the graph on vertices $V(G) \sqcup V(H)$ formed by adding all edges between $V(G)$ and $V(H)$.

Lemma 13. *Let G be any graph with n vertices and e edges. Then $\mathrm{cp}(G \vee \overline{K_\ell}) \geqslant n\ell - e$. If $\chi'(G) \leqslant \ell$, then $\mathrm{cp}(G \vee \overline{K_\ell}) = n\ell - e$.*

Proof. Let $H = \overline{K_\ell}$ and let E_{G-H} be the set of all $n\ell$ edges between $V(G)$ and $V(H)$. Suppose $C_1, \ldots, C_r$ is a clique partition of $G \vee H$. Since C_i can have at most one vertex in H, it follows that $|E(C_i) \cap E(G)| \geqslant \binom{|E(C_i) \cap E_{G-H}| - 1}{2} \geqslant |E(C_i) \cap E_{G-H}| - 1$. Letting $S = \{i : E(C_i) \cap E_{G-H} \neq \emptyset\}$ and summing this inequality over S, we obtain

$$e \geqslant \sum_{i \in S} |E(C_i) \cap E(G)| \geqslant \sum_{i \in S} |E(C_i) \cap E_{G-H}| - |S| \geqslant n\ell - r, \tag{3.1}$$

which implies $\mathrm{cp}(G \vee H) \geqslant n\ell - e$. When $\chi'(G) \leqslant \ell$, we can assign each of the ℓ nodes in H to one of the ℓ color classes of a valid edge coloring in G, and obtain a collection of triangles of the form $\{v, x, y\}$, for $v \in H$ and $(x, y) \in E(G)$ that has been given color v in the edge coloring. No edge in E_{G-H} will be used twice precisely because no vertex in G is incident to two edges of the same color. This gives a collection of e edge-disjoint triangles that cover all the edges in G, and leaves at most $n\ell - 2e$ edges left to cover. Adding in those remaining edges yields a clique partition of size at most $n\ell - e$. $\square$

The construction: Let ℓ and m be any positive integers, and let G be any graph on m vertices. We define $H_\ell = H_\ell(G)$ to be the graph in Figure 1, where the double lines
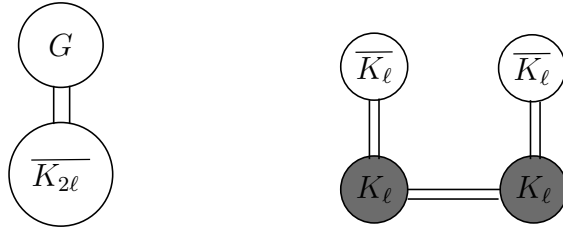
Figure 2: Decomposing H_ℓ into the edge-disjoint union of the two graphs $X_\ell(G) = G \vee \overline{K_{2\ell}}$ (left) and $Y_\ell = \overline{K_\ell} \vee K_\ell \vee K_\ell \vee \overline{K_\ell}$ (right).

are to be interpreted in the same way as the $\vee$ symbol, i.e. including all possible edges between the vertices on either end. Observe that $\overline{H_\ell(G)} \cong H_\ell(\overline{G})$, and that the edges of $H_\ell(G)$ can be split into $X_\ell(G) := G \vee \overline{K_{2\ell}}$ and $Y_\ell = \overline{K_\ell} \vee K_\ell \vee K_\ell \vee \overline{K_\ell}$, as depicted in Figure 2. Clearly $\chi'(G) \leqslant \chi'(K_m)$, which is at most m, since we can assign the numbers $0, 1, \ldots, m-1$ to each vertex and color the edge (i,j) by $i - j \mod m$. So if $m \leqslant 2\ell$, Lemma 13 implies that $\mathrm{cp}(X_\ell(G)) = m\ell - e(G)$. Therefore

$$
\begin{aligned}
\mathrm{cp}(H_\ell(G)) + \mathrm{cp}(\overline{H_\ell(G)}) &= \mathrm{cp}(H_\ell(G)) + \mathrm{cp}(H_\ell(\overline{G})) \\
&= \mathrm{cp}(X_\ell(G)) + \mathrm{cp}(X_\ell(\overline{G})) + 2\mathrm{cp}(Y_\ell) \\
&= 2m\ell - \binom{m}{2} + 2\mathrm{cp}(Y_\ell)
\end{aligned}
$$

for *any* graph G on $m \leqslant 2\ell$ vertices. (In fact, this still gives a lower bound on $\mathrm{cp}(H_\ell(G)) + \mathrm{cp}(\overline{H_\ell(G)})$ for any G and any m.) The term $\mathrm{cp}(Y_\ell)$ was computed in [4], and we include this calculation in the Appendix:

Lemma 14 (Lem. 2 and 3 in [4]). *For any ℓ, $\mathrm{cp}(Y_\ell) \geqslant \frac{7}{4}\ell^2 + O(\ell)$, and this is tight infinitely often.*

So for any G on m vertices, we have

$$
\mathrm{cp}(H_\ell(G)) + \mathrm{cp}(\overline{H_\ell(G)}) \geqslant 2m\ell - \binom{m}{2} + \frac{7}{2}\ell^2. \tag{3.2}
$$

Note that $H_\ell(G)$ has $n := m + 4\ell$ vertices, so when we maximize (3.2) in m while keeping n fixed, we find that the optimum occurs at $m = \frac{9}{8}\ell$. At this value of m, the lower bound is $(8 - \frac{81}{128})\ell^2 + O(\ell)$ for a graph on $\frac{41}{8}\ell$ vertices, implying that, for infinitely many n,

$$
\max_{G \in \mathcal{G}_n} \mathrm{cp}(G) + \mathrm{cp}(\overline{G}) \geqslant \frac{(8 - \frac{81}{128})}{(\frac{41}{8})^2} n^2 + O(n) = \frac{23}{82} n^2 + O(n).
$$

Note that if G is a self-complementary graph (i.e. $G \cong \overline{G}$), then $H_\ell(G)$ is also self-complementary. $\qquad\square$

3.2   Improving the upper bound

The problem of partitioning a graph G into as few cliques as possible is equivalent to the problem of *packing* disjoint copies of $K_3, K_4, \ldots, K_n$ inside of G in such a way as to maximize a certain linear objective function. Indeed, given a clique partition C of G, let C_i denote the number of cliques of size i in C, for $i = 2, \ldots, n$. Then $|C| = \sum_{i=2}^{n} C_i$ and $\sum_{i=2}^{n} \binom{i}{2} C_i = |E(G)|$, so

$$
\begin{aligned}
\mathrm{cp}(G) &= \min_C |C| \\
&= E(G) - \underbrace{\max_C \sum_{i \geqslant 3} \left(\binom{i}{2} - 1 \right) C_i}_{=:v(G)}.
\end{aligned}
\qquad (3.3)
$$

We will also consider *r-restricted clique packings/partitions*, in which the largest clique can have size at most r. We define $\mathrm{cp}(G, r)$ to be the minimum number of cliques of size at most r needed to partition the edges of G. Equivalently, $\mathrm{cp}(G, r) = E(G) - v_r(G)$, where

$$
v_r(G) := \max_C \sum_{i=3}^{r} \left(\binom{i}{2} - 1 \right) C_i.
\qquad (3.4)
$$

Clearly $\mathrm{cp}(G, r) \geqslant \mathrm{cp}(G)$, and one would expect the numbers $\mathrm{cp}(G, r)$ and $\mathrm{cp}(G)$ to be relatively close for large r. This is indeed the case, as we show in the following lemma.

Lemma 15. *For any $\epsilon > 0$, there exists an integer $r_0 = r_0(\epsilon)$ such that for any $r \geqslant r_0$ and any graph G on n vertices,*

$$
\mathrm{cp}(G, r) \leqslant \mathrm{cp}(G) + \epsilon \cdot n^2.
$$

Proof. We make use of the following fact, which is a straightforward consequence of Wilson's theorem [21]: for any fixed $t \geqslant 2$ and $\epsilon > 0$, there is an integer $m_0 = m_0(t, \epsilon)$ such that for all $m \geqslant m_0$, there is a partition of K_m into edge-disjoint copies of K_t and at most ϵm^2 leftover edges. Set $t = \frac{1}{2\epsilon}$ and $r_0 = m_0(t, \epsilon/5)$.

Let C be a clique partition with $|C| = \mathrm{cp}(G)$. For any $r \geqslant r_0$, we can obtain an r-restricted clique partition $\widetilde{C}$ from C as follows: keep each clique of size at most r, and, for each clique K_m with $m > r$, decompose it into at most $\binom{m}{2}/\binom{t}{2}$ copies of K_t and cover the remaining edges (of which there are at most $\frac{\epsilon}{5} \cdot m^2$) with K_2's. This gives a clique partition $\widetilde{C}$ of size

$$
\begin{aligned}
|\widetilde{C}| &\leqslant \sum_{i=2}^{r} C_i + \sum_{i > r} \left(\frac{\binom{i}{2}}{\binom{t}{2}} + \frac{\epsilon}{5} i^2 \right) C_i \\
&\leqslant \sum_{i=2}^{n} C_i + \epsilon \cdot \sum_{i=2}^{n} \binom{i}{2} C_i \\
&= |C| + \epsilon \cdot |E(G)|
\end{aligned}
$$

from which the lemma follows. $\square$

3.2.1 Fractional clique packings

For a fixed family $\mathcal{F}$ of graphs and any graph G, let $\binom{G}{\mathcal{F}}$ denote the set of (unlabeled, non-induced) subgraphs of G which are isomorphic to some $F \in \mathcal{F}$. Following Keevash and Sudakov [14] and Yuster [22], we say a function $\psi : \binom{G}{\mathcal{F}} \to [0,1]$ is a *fractional $\mathcal{F}$-packing of G* if for every edge $e \in E(G)$, we have

$$\sum_{e \in H \in \binom{G}{\mathcal{F}}} \psi(H) \leqslant 1.$$

We denote by $G_{\mathcal{F}}$ the polyhedron of all fractional $\mathcal{F}$-packings of G. As we are interested in the fractional analogue of clique packings, we will only be concerned with families of the form

$$\mathcal{F}_r := \{K_3, K_4, \ldots, K_r\}.$$

Let $\nu_r(G)$ be the value of the linear program

$$\max_{\psi \in G_{\mathcal{F}_r}} \sum_{H \in \binom{G}{\mathcal{F}_r}} \left(\binom{|H|}{2} - 1 \right) \psi(H). \tag{3.5}$$

When the objective function is simply $\sum_{H \in \binom{G}{\mathcal{F}}} \psi(H)$, and the family $\mathcal{F} = \{F\}$ is just a single graph, a theorem of Haxell and Rödl [11] implies that relaxing the domain of maximization from (integer) packings to fractional packings can only change the value of the optimum by $o(n^2)$. Subsequently, Yuster [22] extended this result to arbitrary families of graphs. For finite families (such as $\mathcal{F}_r$), Yuster's proof easily extends to arbitrary linear objective functions [23]. Therefore:

Theorem 16. *For any $r \geqslant 3$ and $G \in \mathcal{G}_n$,*

$$v_r(G) - \nu_r(G) = o(n^2).$$

The advantages of studying fractional clique packings rather than clique partitions are twofold. First, solving the linear program (3.5) is computationally feasible, unlike the corresponding integer program. Second, they can be *averaged*, which not only enables one to turn finite computations into asymptotic bounds, but also allows one to leverage the results of a search on n vertices to reduce the search space when looking for a minimizer on $n + 1$ vertices. This is the approach used by Keevash and Sudakov in [14], and the following averaging lemma (for a different LP) appears as their Lemma 2.1, with the same proof.

For each r, define

$$f_r(n) := \min_{G \in \mathcal{G}_n} \nu_r(G) + \nu_r(\overline{G}).$$

Lemma 17. *For any $r \geqslant 3$, the sequence $\frac{f_r(n)}{n(n-1)}$ is increasing in n.*

Proof. Let $G \in \mathcal{G}_{n+1}$, and let $G_1, \ldots, G_{n+1}$ be the induced subgraphs on the vertex subsets of size n. Let $\psi_i, \overline{\psi_i}$ be optimal fractional packings on G_i and $\overline{G_i}$. Since each edge of G (and $\overline{G}$) occurs in $n-1$ of the G_i, we have that

$$\psi := \frac{1}{n-1} \sum_{i=1}^{n+1} \psi_i, \qquad \overline{\psi} := \frac{1}{n-1} \sum_{i=1}^{n+1} \overline{\psi_i}$$

are fractional packings on G and $\overline{G}$ with combined objective value of at least $\frac{n+1}{n-1} f_r(n)$, and hence $\frac{f_r(n+1)}{(n+1)n} \geqslant \frac{(n+1)f_r(n)}{n(n-1)(n+1)} = \frac{f_r(n)}{n(n-1)}$, as claimed. $\qquad\square$

Since the sequence $\frac{f_r(n)}{n(n-1)}$ is obviously bounded above by $1/2$, it follows that it converges to a limit $c_r \in (0, 1/2)$. Since c_r is increasing in r, the sequence $\{c_r\}$ also converges to a limit that we will call c_∞.

Theorem 18.

$$\max_{G \in \mathcal{G}_n} \operatorname{cp}(G) + \operatorname{cp}(\overline{G}) \sim \left(\frac{1}{2} - c_\infty\right) n^2.$$

Proof. This essentially follows from Lemma 15 and Theorem 16. More explicitly, for any $\epsilon > 0$, we can pick r large enough so that $|\operatorname{cp}(G) - \operatorname{cp}(G, r)| < \epsilon n^2$ for any $G \in \mathcal{G}_n$, and $|c_r - c_\infty| < \epsilon$. Now pick n large enough so that $|v_r(G) - \nu_r(G)| < \epsilon n^2$ for any $G \in \mathcal{G}_n$ and $|f_r(n) - c_r n^2| < \epsilon n^2$. It follows that

$$\max_{G \in \mathcal{G}_n} \operatorname{cp}(G) + \operatorname{cp}(\overline{G}) \in \left(\frac{1}{2} - c_\infty \pm 8\epsilon\right) n^2$$

for n sufficiently large. $\qquad\square$

The same argument shows that $\max_{G \in \mathcal{G}_n} \operatorname{cp}(G, r) + \operatorname{cp}(\overline{G}, r) \sim \left(\frac{1}{2} - c_r\right) n^2$. Let us define $\alpha_r := \frac{1}{2} - c_r$, and $\alpha_\infty = \frac{1}{2} - c_\infty$. We seek an upper bound on α_∞, and since $\alpha_\infty \leqslant \alpha_r = \frac{1}{2} - c_r \leqslant \frac{1}{2} - \frac{f_r(n)}{n(n-1)}$ for any n, it suffices for our purposes to compute a lower bound on the value of $\frac{f_r(n)}{n(n-1)}$ for any particular pair of positive integers (r, n). For example, a modern computer can compute $f_4(8) = 6$ numerically by solving the LP (3.5) on every non-isomorphic graph on 8 vertices. This shows that $\alpha_\infty \leqslant \alpha_4 \leqslant \frac{1}{2} - \frac{6}{8 \cdot 7} = \frac{11}{28} \approx 0.3928$. This already beats the best bound one can get from purely Ramsey-based arguments[3],

---

[3]As was remarked in [4], one can begin with a maximal collection of edge disjoint K_r's (instead of triangles) in G and $\overline{G}$, and bound the number of remaining edges (using Turan's theorem) by $\xi_r n^2$, where $\xi_r := \frac{1}{2} - \frac{1}{2R(r,r)-2}$, and the iterate on the remaining edges with cliques of size K_{r-1}, etc. It is not hard to see that the bound one obtains is

$$\operatorname{cp}(G) + \operatorname{cp}(\overline{G}) \leqslant \left(\xi_3 + \frac{\xi_4 - \xi_3}{3} + \frac{\xi_5 - \xi_4}{6} + \cdots + \frac{\xi_r - \xi_{r-1}}{\binom{r-1}{2}} + \frac{\frac{1}{2} - \xi_r}{\binom{r}{2}}\right) n^2.$$

Even using the most optimistic (i.e. smallest) of the possible values for $R(k, k)$ for $k \geqslant 5$, this approach will not yield an upper bound better than $0.41n^2$.

although it does not beat the Keevash-Sudakov triangle packing bound. In the remainder of this section, we improve this bound in two ways: first, we show in Section 3.2.2, we can *combine* Ramsey-type arguments with estimates on $f_r(n_0)$ to yield better estimates on $f_r(n)$ for n much larger than n_0; second, in Section 3.2.3 we compute the exact value of $f_4(n)$ up to $n = 19$, using an algorithm of Keevash and Sudakov that is significantly more efficient than brute force search.

3.2.2 Ramsey-type improvements

In [14], it was observed that the averaging argument in Lemma 17 can be improved, in a sense, by using a different decomposition of G into smaller subgraphs based on a greedy packing as described in the introduction. In particular, given any bicoloring of K_{3n}, greedily select vertex-disjoint monochromatic triangles $T_1, \ldots, T_i$. The fact that $R(3,3) = 6$ guarantees that we can do this until 3 vertices remain, giving us $n - 1$ triangles $T_1, \ldots, T_{n-1}$, and one set of 3 vertices denoted T_n. Consider the 3^n colorings c of K_n obtained by picking one vertex in each T_i and the edges between them. Each coloring has some fractional packing ψ_c of weight at least $f_3(n)$, and since each edge between T_i and T_j for $i \neq j$ occurs in exactly 3^{n-2} of these, the average $3^{-(n-2)} \sum_c \psi_c$ is a valid fractional packing in K_{3n} of weight at least $9f_3(n)$. Since each of the monochromatic triangles $T_1, \ldots, T_{n-1}$ are edge disjoint from this packing, they can be included as well, yielding a lower bound

$$f_3(3n) \geqslant 9f_3(n) + 2(n - 1). \tag{3.6}$$

Since $R(4,4) = 18$, we can greedily find vertex disjoint monochromatic copies of K_4, $H_1, \ldots, H_{n-4}$, with 16 vertices remaining. From the remaining vertices, we can find edge disjoint monochromatic triangles $T_{n-3}, T_{n-2}, T_{n-1}, T_n$, which we join with the remaining four vertices to form $H_{n-3}, \ldots, H_n$, each of size four. Repeating the same process as above, we see that

$$f_4(4n) \geqslant 16f_4(n) + 5(n - 4) + 8. \tag{3.7}$$

For $r = 5$, we can use the bound $R(5,5) \leqslant 48$ to find $n - 9$ vertex-disjoint copies of K_5, with 45 vertices left over. We can then find $\lceil (45 - 18)/4 \rceil = 7$ copies of K_4, with 17 vertices left over, in which we can find 2 monochromatic triangles, and distribute the remaining vertices so that each of these 11 parts has size 5. Arguing as above, this then implies

$$f_5(5n) \geqslant 25f_5(n) + 9(n - 9) + 37. \tag{3.8}$$

We omit the details, but using similar arguments and the bounds on Ramsey numbers $R(6,6) \leqslant 165$ and $R(7,7) \leqslant 540$ yields the inequalities

$$f_6(6n) \quad \geqslant \quad 36f_6(n) + 14n - 151 \tag{3.9}$$
$$f_7(7n) \quad \geqslant \quad 49f_7(n) + 20n - 532. \tag{3.10}$$

According to András Gyárfás [10], Paul Erdős, sitting in the Atlanta Airport in 1995, asked his companions whether every bicoloring of the edges of $K_{R(k,k)}$ contains *two edge-disjoint monochromatic copies of K_k*. Ralph Faudree pointed out that this is not true,

at which point Erdős asked for the smallest number $n(k)$ for which any bicoloring of $K_{n(k)}$ *does* contain two edge-disjoint monochromatic K_k's. The next day, Faudree showed $n(3) = 7$, and some time later, Gyárfás showed $n(4) = 19$. For our purposes, however, we require *vertex-disjoint* monochromatic copies of K_r. In the appendix we give an argument, inspired by the proof of $n(4) = 19$ by Gyárfás, showing that $n = 20$ is sufficient to find two vertex-disjoint monochromatic K_4's, provided there is also a monochromatic K_5:

Lemma 19. *Any bicoloring of the edges of K_{20} with a monochromatic copy of K_5 contains two vertex-disjoint monochromatic copies of K_4.*

With this lemma in hand, we can obtain a slight improvement over (3.7):

Lemma 20. *For any $n \geqslant 12$, $f_4(4n) \geqslant 16f_4(n) + 5n - 9$.*

Proof. Consider any bicoloring of K_{4n}. Since $4n \geqslant 48 \geqslant R(5,5)$, there is some monochromatic copy of K_5 – call this subgraph N. While there are at least $R(4,4) = 18$ vertices in $K_{4n} \setminus N$, we can greedily select vertex-disjoint monochromatic copies of K_4 in $K_{4n} \setminus N$, $H_1, \ldots, H_{n-5}$. This leaves a set S of 15 remaining vertices. By Lemma 19, the coloring induced on $S \cup N \cong K_{20}$ has two vertex disjoint copies of K_4, which we call H_{n-4} and H_{n-3}. Removing the vertices in $H_{n-4} \cup H_{n-3}$ from $N \cup S$, we are left with 12 vertices, which must contain three vertex-disjoint monochromatic triangles T_1, T_2 and T_3. This leaves behind a set of three vertices $\{v_1, v_2, v_3\}$. Decomposing K_{4n} into the n blocks of size 4

$$H_1, \ldots, H_{n-3}, T_1 \cup \{v_1\}, T_2 \cup \{v_2\}, T_3 \cup \{v_3\},$$

we consider the 4^n edge-colorings c of K_n obtained by picking one vertex from each part. Each of these has a fractional clique packing ψ_c of size at least $f_4(n)$, and since each edge is used in $4^{-(n-2)}$ such ψ_c, we know that $4^{-(n-2)} \sum_c \psi_c$ is a valid packing in K_{4n}. Adding the copies of K_4 and K_3 inside the n individual blocks, we see that $f_4(4n) \geqslant 16f_4(n) + 5(n-3) + 6$. $\qquad\square$

3.2.3 Computer-aided calculations

We next describe a generalization of the algorithm used by Keevash and Sudakov in the case of triangle packings [14], which we call the KS extension method. For any finite family of graphs $\mathcal{F} = \{H_1, \ldots, H_r\}$, any graph $G \in \mathcal{G}_n$, and any vector $\Gamma \in \mathbb{R}^{\mathcal{F}}$, we let $\nu_{\mathcal{F},\Gamma}(G)$ be the value of the linear program

$$\max_{\psi \in G_{\mathcal{F}}} \sum_{H \in \binom{G}{\mathcal{F}}} \Gamma(H)\psi(H) \tag{3.11}$$

and define $\Lambda(\mathcal{F}, \Gamma, n) := \min_{G \in \mathcal{G}_n} \nu_{\mathcal{F},\Gamma}(G) + \nu_{\mathcal{F},\Gamma}(\overline{G})$. For any $\ell \in \mathbb{R}$, and any set L of graphs, define

$$\mathcal{L}(L, \ell) := \{G \in L : \nu_{\mathcal{F},\Gamma}(G) + \nu_{\mathcal{F},\Gamma}(\overline{G}) \leqslant \ell\},$$

and let $\Lambda_{\mathcal{F},\Gamma}(L) = \min_{G \in L} \nu_{\mathcal{F},\Gamma}(G') + \nu_{\mathcal{F},\Gamma}(\overline{G'})$. We also define $\text{ext}(L)$ be the set of *one-vertex extensions* of the graphs in L. The KS extension method is based on the following

observation: by Lemma 17, any graph $G' \in \mathcal{G}_{n+1}$ with $\nu_{\mathcal{F},\Gamma}(G') + \nu_{\mathcal{F},\Gamma}(\overline{G'}) \leqslant \frac{n+1}{n-1} \cdot \ell$ must be a one-vertex extension of some graph in $\mathcal{L}(\mathcal{G}_n, \ell)$. In other words, if $\{\ell_n\}_{n \in \mathbb{N}}$ is any sequence of numbers satisfying $\ell_{n+1} \geqslant \frac{n+1}{n-1}\ell_n$, then

$$\mathcal{L}(\mathcal{G}_{n+1}, \ell_{n+1}) \subseteq \mathrm{ext}(\mathcal{L}(\mathcal{G}_n, \ell_n)).$$

Let us refer to such sequences ℓ_n as *level sequences*.

Algorithm 1: KS Extension Method

---

1 $n \leftarrow n_0$
2 compute $L = \mathcal{L}(\mathcal{G}_n, \ell_n)$ (e.g. via exhaustive search)
3 **while** $L \neq \emptyset$ **do**
4 $\quad\quad \Lambda[n] = \Lambda_{\mathcal{F},\Gamma}(L)$
5 $\quad\quad S \leftarrow \mathrm{ext}(L)$
6 $\quad\quad L \leftarrow \mathcal{L}(S, \ell_{n+1})$
7 $\quad\quad \Lambda[n+1] = \ell_{n+1}$
8 $\quad\quad n \leftarrow n + 1$
9 **end**
10 **return** Λ

Note that the sequence ℓ_n used by Algorithm 1 does not have to be determined before runtime; as long as it is guaranteed to be a level sequence, this guarantees the loop invariant $\mathcal{L}(\mathcal{G}_n, \ell_n) \subseteq L$, and hence $\Lambda[n] \leqslant \Lambda_{\mathcal{F},\Gamma}(\mathcal{G}_n)$. In [14], they choose a parameter d (called the "search depth"), and define ℓ_n recursively by taking $\ell_{n_0} = +\infty$ and ℓ_{n+1} to be $\frac{n+1}{n-1} \cdot \alpha_n$, where α_n is either (a) the dth smallest value in the set $\{\nu_{\mathcal{F},\Gamma}(G') + \nu_{\mathcal{F},\Gamma}(\overline{G'}) : G \in \mathcal{L}(\mathcal{G}_n, \ell_n)\}$, if this set has at least d elements, or (b) ℓ_n, if the set has fewer than d elements. The role of d is to limit the number of graphs stored in the set L. If $d = \infty$, then Algorithm 1 has to solve the LP (3.11) on every graph up to size n in order to compute $\Lambda_{\mathcal{F},\Gamma}(\mathcal{G}_n)$, while if d is too small, then the while loop will terminate after a small number of iterations. We ran an implementation[4] of this method on a 24-core computing grid with $d = 11$, starting with an exhaustive search on $n_0 = 6$ vertices, and obtained the results summarized in Table 1. The last column in particular implies $f_4(20) > 64.725$, which implies $c_4 > 0.1703$. Using Lemma 20, and inequalities (3.8), (3.9), and (3.10) (in that order), we can obtain the bound $c_7 \geqslant 0.1814$, which implies

$$\max_{G \in \mathcal{G}_n} \mathrm{cp}(G) + \mathrm{cp}(\overline{G}) < 0.3186n^2 + o(n^2).$$

---

[4]There are other implementation details omitted from our pseudocode description of Algorithm 1 that also have significant impact on its runtime and memory usage, such as how and when to prune isomorphisms, which LP solver to use, which value of n_0 to exhaust from, and how to split work among processors. Our implementation is similar to the one used in [14], and we recommend reading their magma code, which can be found online at https://people.math.ethz.ch/~sudakovb/triangles-program.

i \ n	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1	2	4	6	8	11	15	19	23	27	33	39	45	51	57	> 64.725
2	4	5	7	9	12	16	20	24	28	34	40	46	52	58	*
3	5	6	8	10	12.5	16.5	20.5	24.5	28.5	34.5	40.5	*	*	*	*
4	6	7	9	11	13	17	21	25	29	34.75	40.75	*	*	*	*
5	7	8	9.5	12	14	17.5	21.25	25.25	29.25	35	*	*	*	*	*
6	8	$8.\overline{3}$	10	$12.\overline{3}$	14.5	18	21.5	25.5	29.5	35.25	*	*	*	*	*
7	$8.\overline{3}$	8.5	$10.\overline{3}$	12.5	$14.\overline{6}$	18.25	22	26	30	35.5	*	*	*	*	*
8	9	9	10.5	$12.\overline{6}$	14.8	$18.\overline{3}$	22.25	26.25	30.25	*	*	*	*	*	*
9	$10.\overline{6}$	9.5	10.6	$12.\overline{6}$	15	18.5	$22.\overline{3}$	$26.\overline{3}$	30.5	*	*	*	*	*	*
10	12.5	10	$10.\overline{6}$	12.8	15.5	$18.\overline{6}$	22.5	26.5	30.75	*	*	*	*	*	*
11	*	$10.\overline{3}$	10.8	13	15.6	18.75	*	*	31	*	*	*	*	*	*

Table 1: The lowest values of $\nu_4(G) + \nu_4(\overline{G})$ for $G \in \mathcal{G}_n$, $n = 6, \ldots, 19$, as found by the KS extension method. The level ℓ_{20} was $64.72527+$ when the algorithm terminated, which implies that $f_4(20) > 64.725$.

References

[1] N. Alon. Neighborly families of boxes and bipartite coverings. *Algorithms and Combinatorics*, Volume 14, 1997, pages 27–31.

[2] C. Bujtas, A. Davoodi, E. Győri, Z. Tuza. Clique Coverings and Claw-free Graphs, *European Journal of Combinatorics*, Volume 88, 2020.

[3] D. de Caen, D.A. Gregory, and D. Pritikin. Minimum biclique partitions of the complete multigraph and related designs. *Graphs, matrices, and designs*, 1993, pages 93–119.

[4] D. de Caen, P. Erdős, N. J. Pullmann, and N. C. Wormald. Extremal clique coverings of complementary graphs. *Combinatorica*, Volume 6 Issue 4, 1986, pages 309–314.

[5] S.M. Cioabă, M. Tait. Variations on a theme of Graham and Pollak. *Discrete Mathematics*, Volume 313, Issue 5, 2013, pages 665–676.

[6] P. Erdős. Some recent problems and results in graph theory. *Discrete Mathematics*, Volume 164, 1997, pages 81–85.

[7] P. Erdős, R. J. Faudree, R. J. Gould, M. S. Jacobson, and J. Lehel. Edge disjoint monochromatic triangles in 2-colored graphs, *Discrete Mathematics* Volume 231, 2001, pages 135–141.

[8] P. Erdős, A. W. Goodman, L. Pisa. The Representation of a Graph by Set Intersections, *Canadian Journal of Mathematics*, **18**, 1955, pages 105–112.

[9] R.L. Graham and H.O. Pollak. On the addressing problem for loop switching. *Bell System Technical J.*, Volume 50, Issue 8, 1971, pages 2495–2519.

[10] A. Gyárfás. Fruit Salad. *Electronic Journal of Combinatorics*, vol. 4, 1997, #R8.

[11] P. Haxell and V. Rödl. Integer and fractional packings in dense graphs, *Combinatorica* Volume 21, 2001, pages 13–38.

[12] H. Huang and B. Sudakov. A counterexample to the Alon-Saks-Seymour conjecture and related problems, *Combinatorica*, Volume 32, 2012, pages 205–219.

[13] S. Jukna, A.S. Kulikov. On covering graphs by complete bipartite subgraphs. *Discrete Mathematics*, Volume 309, Issue 10, 2009, pages 3399–3403.

[14] P. Keevash and B. Sudakov. Packing triangles in a graph and its complement, *Journal of Graph Theory*, Volume 47, Issue 3, 2004, pages 203–216.

[15] N. Nisan and A. Wigderson. Hardness vs Randomness. *Journal of Computer and System Sciences*, Volume 49, Issue 2, 1994, pages 149–167.

[16] J. Orlin. Contentment in graph theory : Covering graphs with cliques. *Indagationes Mathematicae* **39**, 1977, pages 406?424.

[17] N. J. Pullman and A. Donald. Clique Coverings of Graphs II – Complements of Cliques, *Utilitas Mathematicae* **19**, 1981, pages 207–213.

[18] N. J. Pullman, H. Shank and W. D. Wallis, Clique Coverings of Graphs V – Maximal Clique Partitions, *Bulletin of the Australian Mathematical Society* **25**, 1982, pages 337–356.

[19] L. Pyber. Clique covering of graphs. *Combinatorica* **6**, 1986, pages 393–398.

[20] J. Radhakrishnan, P. Sen, S. Vishwanathan, Depth-3 arithmetic for $S_n^2(X)$ and extensions of the Graham-Pollack theorem. *FST TCS 2000: Foundations of Software Technology and Theoretical Computer Science*, Springer, 2000, pages 176–187.

[21] R. M. Wilson. Decompositions of complete graphs into subgraphs isomorphic to a given graph. In *Proc. 5th British Combinatorial Conference, Congr. Numer.* XV, 1976, pages 647–659.

[22] R. Yuster. Integer and fractional packing of families of graphs. *Random Structures & Algorithms*, Vol. 26, 2005, pages 110–118.

[23] R. Yuster, personal communication. 2018.

Appendix

Proof of Lemma 14

Lemma 14. For any ℓ, $cp(Y_\ell) \geqslant \frac{7}{4}\ell^2 + O(\ell)$, and this is tight infinitely often.

Proof. Let $\mathcal{C} = \{C_1, \ldots, C_k\}$ be an optimal clique partition of Y_ℓ. Let us denote the left (according to Figure 2) copy of K_ℓ in Y_ℓ by A and the right copy by B. Suppose that $\mathcal{C}' = \{C_1, \ldots, C_t\}$, for some $t \leqslant k$, is the sub-collection of cliques which contain vertices in both A and B. Let E_A and E_B be the edges in $A \cap \mathcal{C}'$ and $B \cap \mathcal{C}'$, so that Y_ℓ is the edge disjoint union of $(A \setminus E_A) \vee \overline{K}_\ell$ and $(B \setminus E_B) \vee \overline{K}_\ell$ with $\mathcal{C}'$, and therefore

$$\mathrm{cp}(Y_\ell) \geqslant 2\ell^2 - 2\binom{\ell}{2} + |E_A| + |E_B| + t. \tag{3.12}$$

If clique C_i has a_i vertices in A and b_i vertices in B, then

$$\sum_i^t a_i b_i = \ell^2 \tag{3.13}$$

and

$$|E_A| + |E_B| + t = \sum_{i=1}^t \left(\binom{a_i}{2} + \binom{b_i}{2} + 1 \right). \tag{3.14}$$

Minimizing (3.14) over positive integers a_i, b_i subject to the constraint (3.13), we see the minimum occurs when $a_i = b_i = 2$, i.e. each $C_i \in \mathcal{C}'$ is a K_4 with two vertices in each of A and B. Therefore, at the minimum, $t = \ell^2/4$ and $|E_A| + |E_B| + t = 3\ell^2/4$, which gives

$$\mathrm{cp}(Y_\ell) \geqslant 2\ell^2 - 2\binom{\ell}{2} + 3\ell^2/4 = \frac{7}{4}\ell^2 + O(\ell),$$

as claimed. Tightness follows from Theorem 4 in [18], which essentially guarantees the existence of a decomposition of the edges between A and B into disjoint K_4's, whenever $\ell \geqslant 14$ is even. $\qquad\square$

Proof of Lemma 19

Lemma 19. Any bicoloring of the edges of K_{20} with a monochromatic copy of K_5 contains two vertex-disjoint monochromatic copies of K_4.

Proof. Suppose that we have a bicoloring of K_{20} with a red copy $N = \{n_1, \ldots, n_5\}$ of K_5. If there is a blue copy of K_4, then we are finished, because this blue copy and N cannot share an edge, and therefore share at most one vertex. We may now assume that all monochromatic copies of K_4 are red. We can address the case in which there exists a vertex v such that it is incident to at least nine red and blue edges each relatively quickly. We denote by R and B the cliques on the red and blue neighbors of v, respectively. Because the Ramsey number $R(3, 4) = 9$ and our graph has no blue copy of K_4, R must contain a red copy of K_3. Moreover, B cannot contain a blue copy of K_3, so B must contain a red copy of K_4. Adding v to the red copy of K_3 in R results in two vertex-disjoint red copies of K_4, one in $R \cup v$ and one in B. We may now assume that all vertices have at least eleven incident edges of the same color.

Consider the case in which some vertex v has two red and two blue edges adjacent to a red copy M of K_4. If v has at least eleven red edges, then it has at least nine red edges connected to $K_{20}\backslash(M \cup \{v\})$, which, by the same argument as above, implies $K_{20}\backslash M$ has a red copy of K_4. The same argument holds if v has at least eleven blue edges. We may now assume that no vertex has two red and two blue edges adjacent to a red copy of K_4.

From here, we consider two cases:

Case I: Suppose that there exists five vertices $V = \{v_1, \ldots, v_5\} \subset K_{20}\backslash N$, each with at least three red edges adjacent to N. Because no vertex has both two red and two blue edges adjacent to a red copy of K_4, each vertex of V has at least four red edges adjacent to N. In addition, because our graph has no blue copy of K_4 every set $V\backslash v_i$ has a red edge.

Suppose that some vertex of V, without loss of generality called v_1, has five red edges adjacent to N. Without loss of generality, $\{v_2, v_3\}$ is a red edge in $V\backslash v_1$. There are at most two blue edges from v_2 or v_3 to N; without loss of generality assume they are not incident to n_4 or n_5. Then the subsets $\{v_1, n_1, n_2, n_3\}$ and $\{v_2, v_3, n_4, n_5\}$ are both red copies of K_4. So we may now assume that each vertex in V has exactly four red edges adjacent to N.

Let $f(v_i)$ denote the unique vertex in N for which edge $\{v_i, f(v_i)\}$ is blue, and $f(V)$ denote the range of f. We consider several sub-cases, depending on the size of $|f(V)|$.

Suppose $|f(V)| > 2$. Let $\{v_1, v_2\}$ be (without loss of generality) a red edge in V. Both v_1 and v_2 have four red edges to N, so there are at least three vertices in N (without loss of generality n_1, n_2, and n_3) such that $\{v_i, n_j\}$ is a red edge for all $i \in \{1, 2\}$ and $j \in \{1, 2, 3\}$. By pigeonhole, $|f(V) \cap \{n_1, n_2, n_3\}| > 0$, so (without loss of generality) suppose that $f(v_3) = n_1$. Then $\{v_1, v_2, n_1, n_2\}$ and $\{v_3, n_3, n_4, n_5\}$ are vertex-disjoint red copies of K_4.

Suppose $|f(V)| = 2$. Without loss of generality, support that $f(V) = \{n_1, n_2\}$ and $|f^{-1}(n_1)| \geqslant 3$. Because there are no blue copies of K_4 in our graph, $f^{-1}(n_1)$ contains a red edge $\{v_i, v_j\}$, and the subsets $\{v_i, v_j, n_2, n_3\}$ and $\{v_k, n_1, n_4, n_5\}$ are vertex-disjoint red copies of K_4, where $v_k \in f^{-1}(n_2)$.

Suppose $|f(V)| = 1$. Without loss of generality, suppose $f(V) = \{n_1\}$. Then V does not contain a blue copy of K_3, otherwise our graph would contain a blue copy of K_4. If V contains a red copy $\{v_1, v_2, v_3\}$ of K_3, then $\{v_1, v_2, v_3, n_2\}$ and $\{v_4, n_3, n_4, n_5\}$ are two red copies of K_4, and we are done. If V does not contain a red or blue copy of K_3, then the red edges in V form a cycle of length five, and there are two vertex-disjoint red edges in V, denoted $\{v_i, v_j\}$ and $\{v_k, v_l\}$. In this case, the subsets $\{v_i, v_j, n_2, n_3\}$ and $\{v_k, v_l, n_4, n_5\}$ are both red copies of K_4.

Case II: Suppose that there exist at most four vertices in $K_{20}\backslash N$ with at least three red edges adjacent to N. Then there are at least eleven vertices in $K_{20}\backslash N$ with at least three blue edges adjacent to N. Because no vertex has two red and two blue edges

adjacent to a red copy of K_4, these vertices have at least four blue edges adjacent to N, and so there exists a vertex $n_i \in N$ with at least nine blue edges adjacent to $K_{20} \backslash N$. Therefore, $K_{20} \backslash N$ must contain a red copy of K_4.

This completes the proof. □