THE BICLIQUE COVERING NUMBER OF GRIDS

KRYSZTAL GUO, TONY HUYNH, AND MARCO MACCHIA

ABSTRACT. We determine the exact value of the biclique covering number for all grid graphs.

1. INTRODUCTION

Let $G$ be a graph. A biclique of $G$ is a complete bipartite subgraph. The biclique covering number of $G$, denoted $bc(G)$, is the minimum number of bicliques of $G$ required to cover the edges of $G$. The biclique covering number is studied in fields as diverse as polyhedral combinatorics [10, 2], biology [11], and communication complexity [12], where it is also known as bipartite dimension or rectangle covering number.

Computing the biclique covering number is a classic NP-hard problem. Indeed, deciding if $bc(G) \leq k$ appears as problem GT18 in Garey and Johnson [7]. It is also NP-hard to approximate within a factor of $n^{1-\epsilon}$, where $n$ is the number of vertices of $G$ (see [4]).

As such, there are very few classes of graphs for which we know the biclique covering number exactly. For example, it is well-known that the biclique covering number of the complete graph $K_n$ is $\lceil \log_2 n \rceil$ (see [5] for a proof). Note that the minimum number of bicliques that consist of maximal bicliques (under edge inclusion). For brevity, we call such biclique covers simply covers. For $G_{p,q}$, this implies that covers consist of elements which are isomorphic to $K_{1,3}, K_{1,4}$, or the 4-cycle. We first establish the following upper bound.

Lemma 2. For all integers $1 \leq p \leq q$, $bc(G_{p,q}) \leq \lfloor pq/2 \rfloor$.

Proof. Observe that $G_{p,q}$ is a bipartite graph with bipartition $(X, Y)$ where $|X| = \lfloor pq/2 \rfloor$ and $|Y| = \lceil pq/2 \rceil$. Therefore taking the set of stars centered at vertices in $X$ gives a cover of size $\lfloor pq/2 \rfloor$. See Figure 1. \hfill \Box

We now establish the following lower bounds.

Abstract. We determine the exact value of the biclique covering number for all grid graphs.
Lemma 3. For all integers $1 \leq p \leq q$,

$$\left\lfloor \frac{pq}{2} \right\rfloor \leq \text{bc}(G_{p,q}), \quad \text{if } p \text{ is odd.}$$

$$\frac{pq}{2} - 1 \leq \text{bc}(G_{p,q}), \quad \text{if } p \text{ is even.}$$

Proof. Let $1 \leq p \leq q$. We define a special subset $S(G_{p,q})$ of edges of $G_{p,q}$ inductively as follows. If $p \in \{1, 2\}$, we let $S(G_{p,q})$ be the set of all horizontal edges of $G_{p,q}$. If $p \geq 3$, let $\text{Out}$ be the outer cycle of $G_{p,q}$ and define $S(G_{p,q}) = E(\text{Out}) \cup S(G_{p,q} - V(\text{Out}))$. See Figure 2a and Figure 2b. An easy induction gives

(i) $|S(G_{p,q})| = pq - 1$, if $p$ is odd.

(ii) $|S(G_{p,q})| = pq - 2$, if $p$ is even.

On the other hand, every biclique of $G_{p,q}$ contains at most 2 edges of $S(G_{p,q})$. Therefore, bc($G_{p,q}$) $\geq \left\lfloor |S(G_{p,q})|/2 \right\rfloor$, which completes the proof. □

Next, we give values of $p$ and $q$ for which $\text{bc}(G_{p,q}) = \lfloor pq/2 \rfloor - 1$.

Lemma 4. $\text{bc}(G_{p,q}) = pq/2 - 1$ if $p$ is even and $q - 1 = k(p - 1) + 2\ell$ for some integers $0 \leq \ell < k$.

Proof. By Lemma 3, it suffices to construct a cover $C$ of size $pq/2 - 1$. Since $q - 1 = k(p - 1) + 2\ell$ for some integers $0 \leq \ell < k$, we can decompose $G_{p,q}$ as $k$ copies of $G_{p,p}$ and $\ell$ copies of $G_{p,3}$ such that each of the $\ell$ copies of $G_{p,3}$ is between two copies of $G_{p,p}$. Since $p$ is even, $G_{p,p}$ has two covers of size $p^2/2 - 1$, whose set of 4-cycles are the 4-cycles of the two diagonals of $G_{p,p}$, respectively. See Figure 3 for the case $p = 6$.

The cover $C$ is constructed as follows. For every two consecutive copies of $G_{p,p}$, we use one cover from Figure 3 on one copy and the other cover from Figure 3 on the other copy. Note that this results in $p^2/2 - 1$ bicliques that are in both copies. For every copy of $G_{p,3}$ between two copies of $G_{p,p}$ we again use the two covers of size $p^2/2 - 1$ on the two copies of $G_{p,p}$, and then we use $p^2/2 + 1$ bicliques in $G_{p,3}$ to cover the remaining edges, see Figure 4. The total construction when $p = 6$ and $q = 25$ is given in Figure 5. Observe that:

$$|C| = k \left( \frac{p^2}{2} - 1 \right) + \ell \left( \frac{p^2}{2} + 1 \right) - (k - \ell - 1) \left( \frac{p^2}{2} - 1 \right) = \frac{p^2}{2} (k(p - 1) + 2\ell + 1) - 1 = \frac{pq}{2} - 1. \quad \Box$$
We now establish that we may choose a cover with special properties, before completing the proof of the main theorem. Recall that $\text{Out}$ is the outer cycle of $G_{p,q}$. Let $\mathcal{C}$ be a cover of $G_{p,q}$. A boundary element of $\mathcal{C}$ is an element of $\mathcal{C}$ containing at least one edge of $\text{Out}$. A boundary 4-cycle is a boundary element that is a 4-cycle and a boundary star is a boundary element that is a star.

**Lemma 5.** For every cover of $G_{p,q}$, there exists a cover $\mathcal{C}$ of the same size with the following properties:

(i) all boundary stars of $\mathcal{C}$ are pairwise edge-disjoint;

(ii) no edge is contained in both a boundary 4-cycle and a boundary star of $\mathcal{C}$.

**Proof.** If two boundary stars meet in an edge, we replace them with a star and a 4-cycle which cover a superset of edges, as in Figure 6. If a boundary 4-cycle and a boundary star are not edge-disjoint, then we replace them with two 4-cycles which cover a superset of edges, as in Figure 7.
By repeatedly performing these two replacement rules, we eventually obtain a cover $C$ satisfying (i) and (ii).

We now complete the proof of Theorem 1 by establishing the following converse to Lemma 4.

**Lemma 6.** Let $1 \leq p \leq q$, with $p$ even. If $G_{p,q}$ has a cover of size $pq/2 - 1$, then $q - 1 = k(p - 1) + 2\ell$ for some integers $0 \leq \ell < k$.

**Proof.** The lemma clearly holds if $p = 2$, so we may assume $p \geq 4$. Let $C$ be a cover of $G_{p,q}$ of size $pq/2 - 1$ satisfying properties (i) and (ii) of Lemma 5.

We begin by defining some objects required for the proof. Let $\text{Out}$ be the outer cycle of $G_{p,q}$. The corners of $G_{p,q}$ are the vertices $(1, 1)$, $(1, p)$, $(q, 1)$, and $(q, p)$. Let $H$ be the subgraph of $G_{p,q}$ induced by the edges of the boundary 4-cycles of $C$. A fence is a connected component of $H$. The size of a fence is the number of boundary 4-cycles it contains. A link is a connected component of $\text{Out} \setminus E(H)$ containing at least two vertices and no corners. See Figure 8 for an illustration of fences and links.

Suppose $L$ is a link contained on the topmost path of $G_{p,q}$. Let $L_{\text{left}}$ be the path starting from the left end of $L$ by first proceeding down twice and then alternating between proceeding right and down. We define $L_{\text{right}}$ by interchanging left and right in the definition of $L_{\text{left}}$. The staircase generated by $L$ is the subgraph of $G_{p,q}$ contained in the region bounded by $L$, $L_{\text{left}}$, $L_{\text{right}}$, and (possibly) a subpath of the bottommost path of $G_{p,q}$, see Figure 9a and Figure 9b. Staircases for links contained in the leftmost, rightmost, and bottommost paths of $G_{p,q}$ are defined by rotating $G_{p,q}$ so that the link is on the topmost path, applying the above definition, and then rotating back.

The length of a path is its number of edges, and the length of a staircase is the length of the link that generates it. An edge is thick if it is covered by at least two bicliques in $C$. A pyramid $P$ is a staircase of length $2p - 4$ containing a size-2 fence $B$ such that the thick edge in $B$ is the only thick edge in $P$, see Figure 10. We call $B$ the tip of the pyramid.
Claim 7. Every staircase contains at least one thick edge.

Proof. Let \( L \) be a link of length \( k \) connecting two fences \( B \) and \( B' \) (see Figure 11a) and let \( S \) be the staircase generated by \( L \). Towards a contradiction, suppose \( S \) does not contain a thick edge. Since \( C \) satisfies properties (i) and (ii) of Lemma 5, \( k \) is even and \( P \) is covered by exactly \( k/2 \) boundary stars as in Figure 11b. Let \( E_1 \) be the set of edges in \( S \) with one end on \( L \) and not covered by a boundary element of \( C \) (blue edges in Figure 11b). Each edge \( e \) in \( E_1 \) must be covered by a \( K_{1,4} \), else \( e \) is thick, see Figure 11c. Let \( e_1 \) and \( e_2 \) be the two horizontal edges of \( S \) which intersect \((V(B) \cup V(B')) \setminus V(L)\) (blue edges in Figure 11c). Since neither \( e_1 \) nor \( e_2 \) are thick, they must be covered by \( 4 \)-cycles as in Figure 11d. By repeating this argument, we either find a thick edge (red edge in Figure 11e), or we reach the bottommost path of \( G_{p,q} \). Let \( F \) be the set of edges \( f \in E(S) \) such that \( f \) is covered by a \( K_{1,4} \) of \( C \) and \( f \) has exactly one endpoint on the bottommost path of \( G_{p,q} \). By the above argument, \( F \) is non-empty. Moreover, each edge in \( F \) is thick (see Figure 11f).

A double staircase is a pair \((S, S')\), where \( S \) and \( S' \) are distinct staircases such that there is exactly one thick edge in \( S \cup S' \). Suppose \( S \) has length \( 2a \) and \( S' \) has length \( 2b \). Observe that either \( a + b = p - 2 \) or \( a + b = p \) and \( C \) must cover \( S \cup S' \) as in Figures 12a and 12b, respectively.

We define the waste of \( C \) to be

\[
w(C) := 4|C| - |E(G_{p,q})|. 
\]  

(1)

Let \( \tau \) be the number of copies of \( K_{1,3} \) that appear in \( C \) and \( t_i \) be the number of edges of \( G_{p,q} \) that are covered exactly \( i \) times by \( C \). By double counting, we obtain:

\[
w(C) = \tau + t_2 + 2t_3 + 3t_4. 
\]  

(2)

Since \( |C| = pq/2 - 1 \) and \( |E(G_{p,q})| = 2pq - p - q \), (1) immediately gives the following:

\[
w(C) = p + q - 4. 
\]  

(3)

Note that a fence can have size at most \( N := 2p + 2q - 8 \). For \( i \in \{N\} \), let \( b_i \) be the number of fences of size \( i \) of \( C \).

Claim 8. \( b_N = 0 \).
Proof. If $b_N = 1$, then $C$ contains $2p + 2q - 8$ boundary 4-cycles. These boundary 4-cycles cover $6p + 6q - 24$ edges of $G_{p,q}$. Thus, $2pq - 7p - 7q + 24$ edges are not covered by boundary 4-cycles. It follows that

$$|C| \geq 2p + 2q - 8 + \frac{2pq - 7p - 7q + 24}{4} = \frac{pq}{2} + \frac{p}{4} + \frac{q}{4} - 2 \geq \frac{pq}{2},$$

where the last inequality follows since $4 \leq p \leq q$. However, this contradicts $|C| = \frac{pq}{2} - 1$. □

Let $\beta$ be the number of edges that are covered twice by boundary 4-cycles of $C$, and let $c \in [4]$ be the number of corners of $G_{p,q}$ covered by the set of fences. We now provide a lower bound on the waste of $C$.

Claim 9. $w(C) \geq \beta + \tau = p + q - 2 - \frac{c}{2} - \frac{b_1}{2} + \sum_{i=3}^{N-1} \left(\frac{i}{2} - 1\right)b_i$.

Proof. The inequality follows immediately from (2). For the equality, first observe $\beta = \sum_{i=1}^{N-1} (i-1)b_i$. Next, because each boundary star covers two edges of $\text{Out}$, each edge of $\text{Out}$ not covered by a fence contributes $\frac{1}{2}$
to \( \tau \). Since Out contains \( 2p + 2q - 4 \) edges and the set of fences cover \( c + \sum_{i=1}^{N-1} i b_i \) edges of Out, we have

\[
\beta + \tau = \sum_{i=1}^{N-1} (i-1)b_i + \frac{1}{2} \left( 2p + 2q - 4 - c - \sum_{i=1}^{N-1} i b_i \right)
\]

\[
= p + q - 2 - \frac{c}{2} - \frac{b_1}{2} + \sum_{i=3}^{N-1} \left( \frac{i}{2} - 1 \right) b_i.
\]

Claim 10. The set of staircases of \( C \) can be enumerated as \( P_1, \ldots, P_n, S_1, S'_1, \ldots, S_m, S'_m \) such that \( P_i \) is a pyramid for all \( i \in [n] \) and \( (S_i, S'_i) \) is a double staircase for all \( j \in [m] \). Moreover, \( b_2 = n, b_i = 0 \) for all \( i \geq 3 \), and \( w(C) = \tau + n + m \).

Proof. The number of staircases of \( C \) is at least \( \sum_{i=1}^{N-1} b_i + c - 4 \). At most \( b_2 \) fences can be the tips of pyramids, so at least \( b_1 + \sum_{i=3}^{N-1} b_i + c - 4 \) staircases are not pyramids. Each of these staircases contains a thick edge by Claim 7, and each thick edge can be in at most 2 staircases. Therefore, these staircases contribute at least \( (b_1 + \sum_{i=3}^{N-1} b_i + c - 4)/2 \) to the waste of \( C \) which is not counted in \( \beta + \tau \). By Claim 9,

\[
w(C) \geq p + q - 2 - \frac{c}{2} - \frac{b_1}{2} + \sum_{i=3}^{N-1} \left( \frac{i}{2} - 1 \right) b_i + \frac{1}{2} \left( b_1 + \sum_{i=3}^{N-1} b_i + c - 4 \right)
\]

\[
= p + q - 4 + \sum_{i=3}^{N-1} \left( \frac{i}{2} - 1 \right) b_i.
\]

By (3), \( w(C) = p + q - 4 \). Hence, we must have \( b_1 = 0 \) for \( i \geq 3 \) and equality throughout the above argument. In particular, there are exactly \( b_2 \) pyramids. The remaining staircases each contain exactly one thick edge, and each of these thick edges is in exactly two staircases. This gives the required enumeration of the staircases of \( C \) and implies \( w(C) = \tau + n + m \).

Claim 11. \( q - 1 = k(p - 1) + 2\ell \) for some integers \( 0 \leq \ell < k \).

Proof. Let \( L \) be the leftmost path of \( G_{p,q} \). Since \( q \geq p \), no link contained in \( L \) can generate a pyramid or be part of a double staircase. By Claim 10, no link is contained in \( L \). It follows that \( L \) intersects at most one fence. Since \( p \) is even, it follows that \( L \) intersects exactly one fence \( B_{\text{left}} \). Since all fences are of size 1 or 2 and \( p \) is even, \( B_{\text{left}} \) is of size 1 and \( L \setminus E(B_{\text{left}}) \) is the disjoint union of two even-length paths \( P_a \) and \( P_b \). We assume that \( P_a \) is above \( P_b \) and that \( P_a \) and \( P_b \) have lengths \( a \) and \( b \), respectively. By the same argument, the rightmost path of \( G_{p,q} \) intersects exactly one fence \( B_{\text{right}} \).

We first consider the case when \( B_{\text{left}} \) is a fence that does not contain a corner. Suppose \( L_1 \) is the first link along the topmost path of \( G_{p,q} \) and \( S_1 \) is the staircase generated by \( L_1 \). By Claim 10, there are no thick edges outside of staircases. Therefore, \( B_{\text{left}} \) and the boundary stars force the leftmost portion of \( G_{p,q} \) to be covered by \( C \) as in Figure 13a. This forces \( S_1 \) to be in a double staircase \( (S_1, S'_1) \), where \( S_1 \) has length \( 2a \) (see Figures 14a and 14b), or \( S_1 \) has length \( 2b + 2 \) and \( S'_1 \) has length \( 2a + 2 \) (see Figures 14c and 14d).

Suppose \( L_2 \) is the next link along the topmost path of \( G_{p,q} \). Because of the previous double staircase \( (S_1, S'_1) \), \( S_2 \) must be in a double staircase \( (S_2, S'_2) \) where \( S_2 \) has length \( 2a \) and \( S'_2 \) has length \( 2b \) (see Figures 14a and 14c), or \( S_2 \) has length \( 2a + 2 \) and \( S'_2 \) has length \( 2b + 2 \) (see Figures 14b and 14d). Repeating the argument, we obtain a sequence of double staircases \( (S_1, S'_1), \ldots, (S_{k-1}, S'_{k-1}) \), where \( (S_{k-1}, S'_{k-1}) \) is the last double staircase. Note that it is possible that this sequence is empty, corresponding to the case that there are no staircases and \( k - 1 = 0 \).

Again, because there are no thick edges outside of staircases, the fence \( B_{\text{right}} \) and the boundary stars force the rightmost portion of \( G_{p,q} \) to be covered as in Figure 13b. For all \( i \in [k - 1] \), let \( \ell_i \) and \( \ell'_i \) be the lengths of \( S_i \) and \( S'_i \) respectively. Let \( \ell \) be the number of times that \( \{\ell_i, \ell'_i\} = \{2a + 2, 2b + 2\} \). It follows that \( q - 1 = k(p - 1) + 2\ell \), as required.

The case when \( B_{\text{left}} \) contains a corner is treated exactly as above, resulting in the construction from Figure 5.

This completes the entire proof.

Acknowledgements. We thank Denis Cornaz for suggesting this problem and for carefully reading an early draft of this paper. We also thank Matthias Walter for help in computing \( bc(G_{p,q}) \) for small values of
p and q via SCIP [8]. This project is supported by ERC grant FOREFRONT (grant agreement no. 615640) funded by the European Research Council under the EU’s 7th Framework Programme (FP7/2007-2013).

References

1. Noga Alon, Neighborly families of boxes and bipartite coverings, The Mathematics of Paul Erdős II, Springer Berlin Heidelberg, Berlin, Heidelberg, 1997, pp. 27–31.
2. Manuel Aprile, Yuri Faenza, Samuel Fiorini, Tony Huynh, and Marco Macchia, Extension complexity of stable set polytopes of bipartite graphs, Graph-theoretic concepts in computer science, Lecture Notes in Comput. Sci., vol. 10520, Springer, Cham, 2017, pp. 75–87. MR 3746146
3. Sergei Bezzukov, Dalibor Fronček, Steven J. Rosenberg, and Petr Kovář, On biclique coverings, Discrete Math. 308 (2008), no. 2-3, 319–323. MR 2378030
4. Parinya Chalermsook, Sandy Heydrich, Eugenia Holm, and Andreas Karrenbauer, Nearly tight approximability results for minimum biclique cover and partition, Algorithms—ESA 2014, Lecture Notes in Comput. Sci., vol. 8737, Springer, Heidelberg, 2014, pp. 235–246. MR 3253135
5. Peter C. Fishburn and Peter L. Hammer, Bipartite dimensions and bipartite degrees of graphs, Discrete Math. 160 (1996), no. 1-3, 127–148. MR 1417566
6. Dalibor Fronček, Janja Jerebic, Sandi Klavžar, and Petr Kovář, Strong isometric dimension, biclique coverings, and Sperner’s theorem, Combin. Probab. Comput. 16 (2007), no. 2, 271–275. MR 2298814
7. Michael R. Garey and David S. Johnson, Computers and intractability, W. H. Freeman and Co., San Francisco, Calif., 1979, A guide to the theory of NP-completeness, A Series of Books in the Mathematical Sciences. MR 519066
8. Ambros Gleixner, Michael Bastubbe, Leon Eifler, Tristan Gally, Gerald Gamrath, Robert Lion Gottwald, Gregor Hendel, Christopher Hojný, Thorsten Koch, Marco E. Lübecke, Stephen J. Maher, Matthias Miltenberger, Benjamin Müller, Marc E. Pfetsch, Christian Puchert, Daniel Rehfeldt, Franziska Schlösser, Christoph Schubert, Felipe Serrano, Yuji Shinano, Jan Merlin Viernickel, Matthias Walter, Fabian Wegscheider, Jonas T. Witt, and Jakob Witzig, The SCIP Optimization Suite 6.0, Technical report, Optimization Online, July 2018.
9. R. L. Graham and H. O. Pollak, On the addressing problem for loop switching, Bell System Tech. J. 50 (1971), 2495–2519. MR 0289210
10. Volker Kaibel and Stefan Weltge, A short proof that the extension complexity of the correlation polytope grows exponentially, Discrete Comput. Geom. 53 (2015), no. 2, 397–401. MR 3316228
11. Dana S. Nau, George Markowsky, Max A. Woodbury, and D. Bernard Amos, A mathematical analysis of human leukocyte antigen serology, Math. Biosci. 40 (1978), no. 3-4, 243–270. MR 503830
12. A. A. Razborov, Applications of matrix methods to the theory of lower bounds in computational complexity, Combinatorica 10 (1990), no. 1, 81–93. MR 1075069
Figure 14