Theoretical Computer Science 636 (2016) 77–84

Contents lists available at ScienceDirect

Theoretical Computer Science

www.elsevier.com/locate/tcs

Tight lower bounds on the number of bicliques in false-twin-free graphs

Marina Groshaus [a,1], Leandro Montero [b,*]

[a] Departamento de Computación, Universidad de Buenos Aires, Buenos Aires, Argentina
[b] Laboratoire de Recherche en Informatique, Université Paris-Sud, Orsay CEDEX, France

ARTICLE INFO

Article history:
Received 14 November 2015
Received in revised form 7 May 2016
Accepted 19 May 2016
Available online 24 May 2016
Communicated by L.M. Kirousis

Keywords:
Bicliques
False-twin-free graphs
Lower bounds

ABSTRACT

A *biclique* is a maximal bipartite complete induced subgraph of G. Bicliques have been studied in the last years motivated by the large number of applications. In particular, enumeration of the maximal bicliques has been of interest in data analysis. Associated with this issue, bounds on the maximum number of bicliques were given. In this paper we study bounds on the minimun number of bicliques of a graph. Since adding false-twin vertices to G does not change the number of bicliques, we restrict to false-twin-free graphs. We give a tight lower bound on the minimum number bicliques for a subclass of $\{C_4, \text{false-twin}\}$-free graphs and for the class of $\{K_3, \text{false-twin}\}$-free graphs. Finally we discuss the problem for general graphs.

1. Introduction

Interest in the study of bicliques has increased recently motivated by the wide scope of applications [2,9,16,17,25,28]. For example, bicliques appear in automata and language theory, graph compression, artificial intelligence, biology and data mining. In particular, bicliques are studied in the contexts of: biclustering microarray data [4,26,27], optimizing phylogenetic tree reconstruction [23], identifying common gene-set associations [5], integrating diverse functional genomics data [3], analyzing proteome–transcriptome relationships [15] and discovering patterns in epidemiological research [21]. In genetics, bicliques represent subsets of genes and subsets of properties [4,18,26,27]. Also, bicliques appear in the reconstruction of phylogenetic trees [23].

In all these applications, bicliques represent the relation between different data types. Due to the big size of data used, one of the main problems consists of enumerating all the bicliques. Recall that the number of bicliques in a graph can be exponential [22]. Therefore, algorithms to list all bicliques of a graph have been of particular attention. Recall that the definition of bicliques can be different for each case. Some authors consider bicliques not induced, others, with bounded size of bipartition, etc. See for example [1,6–8,11,19–21,23,24,29,30].

In the context of induced bicliques, some bounds were given. Prisner studied various aspects [22]. He showed that the maximum number of bicliques in a bipartite graph on n vertices is $2^{\frac{n}{2}}$, that is attained by the *Cocktail-party* graphs. Also, he

* Corresponding author.
E-mail addresses: groshaus@dc.uba.ar (M. Groshaus), lmontero@lri.fr (L. Montero).

[1] Partially supported by UBACyT grant 20020100100754, PICT ANPCyT grant 2010-1970, CONICET PIP grant 11220100100310. Partially supported by Math-Amsud project 14 Math 06.

http://dx.doi.org/10.1016/j.tcs.2016.05.027
0304-3975/

gave a lower bound of $3^{\frac{n}{3}}$ and an upper bound of 1.6181^n on the maximum number of bicliques for general graphs. In [10], they improved this result giving a better upper bound of $\frac{1}{3^{1/3}-1}3^{\frac{n}{3}}$ on the number of bicliques of a graph.

In this work we focus on lower bounds on the minimum number of bicliques of a graph where by definition, bicliques are induced. Despite that in general it seems a "trivial" question, since any bipartite complete graph $K_{n,m}$ has just one biclique, it is not considered the fact that all vertices in each partition are false-twins. Since adding false-twin vertices to G does not change the number of bicliques [13], the bipartite complete graph $K_{n,m}$ can be thought as a biclique "equivalent" to $K_{1,1}$. Recall that many applications consider the bicliques as a whole group of items (represented on one side of the bipartition) having a common non-empty subset of characteristics (represented on the other side). So, it is natural to think that if there are groups of "twin objects" with the same characteristics, we could only maintain one representative object for each. Following this idea we introduce the problem of finding bounds for the minimum number of bicliques in a graph with no false-twin vertices. We recall that deleting false-twin vertices can be done in linear time using the modular decomposition [14].

Even though the bounds are polynomial, this approach can help to develop algorithms for listing the bicliques, not only because of the use itself of the bounds but also because of the ideas behind the proofs. Also these bounds can be used for heuristical algorithms.

Throughout this paper we discuss the general case and give bounds for the minimum number of bicliques of graphs in a subclass of $\{C_4, \text{false-twin}\}$-free graph and the class of $\{K_3, \text{false-twin}\}$-free graph. We prove that any graph in these classes on $n \geq 3$ and $n \geq 4$ vertices respectively, has at least $\lceil \frac{n}{2} \rceil$ bicliques. We show that these bounds are tight and finally we prove that the bound $\lceil \frac{n}{2} \rceil$ does not hold for general false-twin-free graphs. This work is the full improved version of a previous extended abstract [12].

This work is organized as follows. In Section 2 the notation is given. In Section 3 and Section 4 we study bicliques in false-twin-free graphs. We prove the lower bound for the minimum number of bicliques for a subclass of $\{C_4, \text{false-twin}\}$-free graphs and for the class of $\{K_3, \text{false-twin}\}$-free graphs. Finally, in Section 5 we discuss the problem for general graphs.

2. Preliminaries

Along the paper we restrict to undirected simple graphs. Let $G = (V, E)$ be a graph with vertex set V and edge set E, and let $n = |V|$ and $m = |E|$. A *subgraph* G' of G is a graph $G' = (V', E')$ where $V' \subseteq V$ and $E' \subseteq E$. A subgraph $G' = (V', E')$ of G is *induced* when for every pair of vertices $v, w \in G'$, $vw \in E'$ if and only if $vw \in E$. A graph $G = (V, E)$ is *bipartite* when $V = U \cup W$, $U \cap W = \emptyset$, and $E \subseteq U \times W$. Say that G is a *complete graph* when every possible edge belongs to E and say that G is *bipartite complete* when $E = U \times W$. A complete graph of n vertices is denoted K_n and a bipartite complete graph on n and n' vertices in each partition respectively, is denoted $K_{n,n'}$. A graph G is H-*free* if it does not contain H as an induced subgraph. An *independent set* of a graph is a set of vertices, no two of which are adjacent. A *clique* of G is a maximal complete induced subgraph, while a *biclique* is a maximal bipartite complete induced subgraph of G. The *open neighborhood* of a vertex $v \in V(G)$, denoted by $N(v)$, is the set of vertices adjacent to v while the *closed neighborhood* of v, denoted by $N[v]$, is $N(v) \cup \{v\}$. The *degree* of a vertex v, denoted by $d(v)$, is defined as $d(v) = |N(v)|$. The maximum degree among all vertices of G is denoted by $\Delta(G)$. Two vertices u, v are *false-twins* if $N(u) = N(v)$ and *true-twins* if $N[u] = N[v]$. A vertex v is *simplicial* if $\{v\} \cup N(v)$ is a clique. A *diamond* is the graph K_4 minus an edge. A *cycle* on k vertices, denoted by C_k, is a set of vertices $v_1 v_2 \ldots v_k \in V(G)$ such that $v_i \neq v_j$ for all $1 \leq i \neq j \leq k$, v_i is adjacent to v_{i+1} and v_k is adjacent to v_1 for all $1 \leq i \leq k-1$. We assume that all the graphs of this paper are connected.

As mentioned before, we want to keep only the representative vertices for each group of false-twins, in order to, for example reduce the size of the graph and bicliques considered. Formalizing this idea, we present the following definition as it is in [13]. Consider all maximal sets of false-twin vertices $Z_1, \ldots, Z_k$ and let $\{z_1, z_2, \ldots, z_k\}$ be the set of *representative vertices* such that $z_i \in Z_i$. The graph obtained by the deletion of all vertices of $Z_i \setminus \{z_i\}$, for $i = 1, \ldots, k$, is denoted $Tw(G)$. We mention that being false-twin can be thought as an equivalence relation that separates the graphs in equivalent classes.

3. Lower bounds in a subclass of {C_4, false-twin}-free graphs

We start defining a subclass of $\{C_4, \text{false-twin}\}$-free graphs. Observe that if G has no induced C_4, then all its bicliques are isomorphic to $K_{1,r}$, $r \geq 1$. We call these kinds of bicliques v-*star*, where v is the vertex in the partition of size one and is called the *center* of the star.

A vertex x is *alone* if x is a simplicial vertex and $N(x)$ has no simplicial vertices. Let G be a graph and let $\mathcal{A}$ be its set alone vertices. An *assignment* for $\mathcal{A}$ is an association for every alone vertex: each alone vertex x is associated to a vertex v and an edge vv', where v and v' belong to $N(x)$, such that either v is not dominated by v' or v and v' are true-twins. We say that $\mathcal{A}$ has a *good assignment* if the assignment verifies the following: If two different alone vertices in $\mathcal{A}$ have the same associated edge, then its endpoints are not true-twins and the associated vertices of the alone vertices are different. This implies that the edges incident to true-twins or to two vertices such that one is dominated by the other, can only be assigned to at most one alone vertex. See Figs. 1 and 2 for examples.

We study the subclass of $\{C_4, \text{false-twin}\}$-free graphs such that the set of alone vertices $\mathcal{A}$ has a good assignment. Observe that $\{C_4, \text{diamond}, \text{false-twin}\}$-free graphs are included in this class.

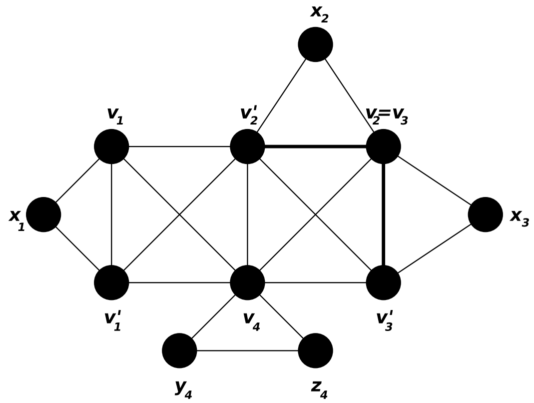

Fig. 1. $\{C_4, \text{false-twin}\}$-free graph that has a good assignment. Vertex x_1 has the vertex v_1 and an edge with true-twins as endpoints associated to it ($v_1v'_1$). Vertices x_2 and x_3 have the same vertex ($v_2 = v_3$) associated to them but the associated edges are different, $v_2v'_2$, $v_3v'_3$ respectively.

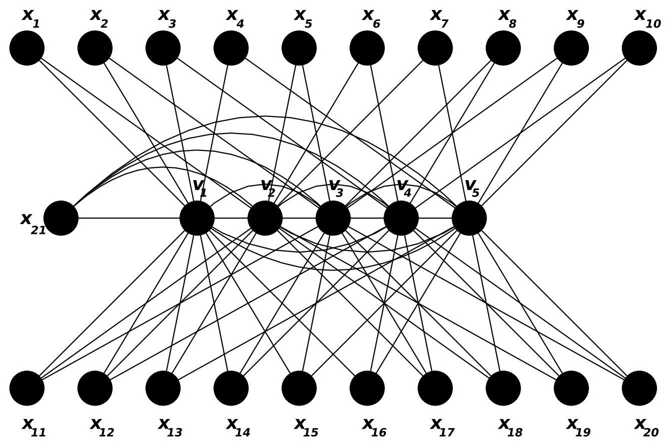

Fig. 2. $\{C_4, \text{false-twin}\}$-free graph that has not a good assignment. Vertices $x_1, \ldots, x_{21}$ are alone vertices and $v_1, \ldots, v_5$ are non-simplicial vertices. Therefore in any assignment at least two alone vertices will be associated to the same vertex and edge as each of the 10 edges in the clique $v_1, \ldots, v_5$ could be used twice, obtaining at most 20 possible assignments to associate to 21 vertices.

We need the following lemma.

Lemma 3.1. *Let G be a $\{C_4, \text{false-twin}\}$-free graph on $n \geq 3$ vertices, without vertices of degree one, such that the set $\mathcal{A}$ has a good assignment. Then G has at least n bicliques.*

Proof. Observe that if $G = K_n$, then the result holds as each edge is a biclique. Otherwise, we will assign each vertex to a different biclique. We will first give labels to some specific edges and associate them to some vertices. Also we will assign some vertices to some bicliques.

Let $C_1, C_2, \ldots, C_\ell$, $\ell \geq 1$, be the maximal sets of pairwise adjacent simplicial vertices in G. Observe that each C_i induces a complete subgraph.

Consider first the sets C_i such that $|C_i| \geq 3$. For each of these sets choose any non-simplicial vertex v adjacent to two vertices in C_i, say w_1, w_2. Clearly v exists as $G \neq K_n$. Label the edges vw_1, vw_2 with labels 1, 2 respectively. This pair of labeled edges are associated to v. Now as each of the $\frac{|C_i|(|C_i|-1)}{2}$ edges of C_i is a biclique in G and $\frac{|C_i|(|C_i|-1)}{2} \geq |C_i|$, for $|C_i| \geq 3$, we assign each of these vertices to a different biclique (edge) of the complete graph C_i.

Next consider the sets C_i such that $|C_i| = 2$. Assume $C_i = \{y_i, z_i\}$. For each set, choose any non-simplicial vertex v adjacent to y_i and z_i. Consider the pair of edges vy_i, vz_i and assign labels 1 and 2 respectively. This pair of labeled edges are associated to v. Assign vertex y_i to the biclique y_iz_i.

Finally, consider the sets C_i such that $|C_i| = 1$. Assume $C_i = \{x_i\}$. Recall that x_i is an alone vertex. By hypothesis, $\mathcal{A}$ has a good assignment. For each i, consider the vertex v and the edge vv' associated to x_i in the good assignment. If v and v' are true-twins, vv' is a biclique, and we assign vertex x_i to this biclique. Observe that since $\mathcal{A}$ has a good assignment, no two vertices are associated to the same edge (biclique). Otherwise, label the edge vx_i with label 1, and the edge vv' with label 2. This pair of labeled edges are associated to v.

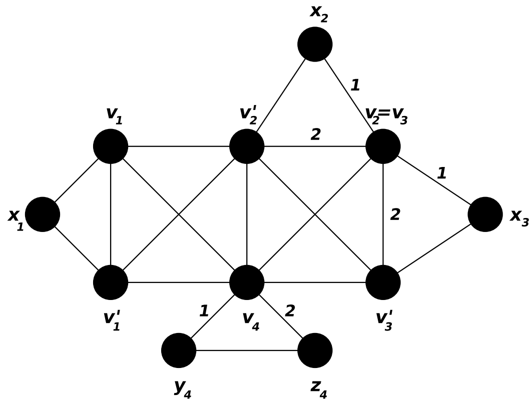

Fig. 3. Same graph as Fig. 1 with labeled edges.

Observe that each pair of edges labeled with 1 and 2 contain a common vertex and are also incident to two adjacent vertices. See Fig. 3.

Now, we assign the remaining vertices to different bicliques.

First consider non-simplicial vertices. Let v be a non-simplicial vertex. Observe that each v is the center of star biclique. Therefore if v has no labeled edges associated to it, then we assign v to any v-star biclique. Otherwise, consider any star biclique centered in v that contains all the associated edges with label 1. Recall that all these edges labeled with 1 are incident to v and to simplicial vertices which are all independent. Also, observe that any star containing all those edges labeled with 1 incident to v does not contain a labeled edge associated to v with label 2. So such a vertex v is associated with a star biclique centered at v, with all associated edges labeled 1.

Next, for each set of simplicial vertices $C_i = \{y_i, z_i\}$ with edges vy_i, vz_i labeled with 1, 2 respectively, we will assign z_i to a biclique. Consider any v-star biclique that contains the edge vz_i with label 2 and all the edges associated to v labeled with 1 (different from vy_i), if any. Observe that such a v-star biclique containing the edge vz_i always exists since, if v has no other labeled edges associated to it, as it is non-simplicial, then it must exist a vertex adjacent to v not adjacent to z_i. Assign vertex z_i to this biclique.

Finally, consider each $C_i = \{x_i\}$. Recall that x_i is an alone vertex. Let v and vv' be its corresponding vertex and edge in the assignment. Vertex v is not dominated by v' since the case when they are true-twins was solved before. Consider any v-star biclique that contains the labeled edge vv' with label 2 associated to v and no other labeled edge associated to v with label 2. Observe that such a star always exists since v is not dominated by v' and if a labeled edge with label 1 associated to v cannot be in the star, neither can be the corresponding labeled edge with label 2 of the same pair. We assign vertex x_i to this biclique.

We have assigned every vertex of G to a biclique of G. Observe that all these bicliques are different: each non-simplicial vertex v is assigned to a v-star biclique that contains no labeled edge associated to v with label 2. Each simplicial vertex is assigned to a biclique that is either an edge incident to two simplicial vertices, an edge incident to two non-simplicial true-twins vertices, or a star centered in a non-simplicial vertex that contains exactly one associated labeled edge with label 2. We conclude that G contains at least n bicliques. □

Theorem 3.2. *Let G be a $\{C_4, \text{false-twin}\}$-free graph on $n \geq 3$ vertices, such that G has at most k vertices of degree one and the set $\mathcal{A}$ of alone vertices has a good assignment. Then G has at least $n-k$ bicliques.*

Proof. Clearly, if G does not contain vertices of degree one, the result holds by Lemma 3.1. Let $v_1, v_2, \ldots, v_k$ be the vertices of degree one in G, $k \geq 1$, and let $x_1, x_2, \ldots, x_k$ be their unique neighbors respectively. We obtain from G a new graph G' as follows. For each vertex v_i, we add two new vertices u_i, w_i such that $\{v_i, u_i, w_i\}$ induces a K_3. Clearly G' is $\{C_4, \text{false-twin}\}$-free and it has $n+2k$ vertices where none of them has degree one and the set $\mathcal{A}$ of alone vertices has a good assignment. Then, by Lemma 3.1, G' has at least $n+2k$ bicliques. Now for each i, $\{u_i, w_i\}$, $\{v_i, u_i, x_i\}$ and $\{v_i, w_i, x_i\}$ are bicliques in G' that clearly, are not in G. Thus G' has exactly $3k$ bicliques more than G. Finally we obtain that G has at least $n+2k-3k = n-k$ bicliques as desired. □

Using the following Lemma, we can give a bound on the number $n-k$ of bicliques given by Theorem 3.2.

Lemma 3.3. *Let G be a false-twin-free graph on $n \geq 3$ vertices. Then G has at most $\lfloor \frac{n}{2} \rfloor$ vertices of degree one.*

Proof. By contrary, if G has more than $\lfloor \frac{n}{2} \rfloor$ vertices of degree one, then there must exist two of them having the same unique neighbor, that is, they are false-twin vertices. A contradiction. □

Combining last results, we obtain the main theorem of the section. It gives a tight lower bound on the number of bicliques for this subclass of $\{C_4, \text{false-twin}\}$-free graphs.

Theorem 3.4. *Let G be a $\{C_4, \text{false-twin}\}$-free graph on $n \geq 3$ vertices such that the set $\mathcal{A}$ of alone vertices has a good assignment. Then G has at least $\lceil \frac{n}{2} \rceil$ bicliques.*

Proof. The result follows from Lemma 3.3 and Theorem 3.2 since $n - k \geq n - \lfloor \frac{n}{2} \rfloor = \lceil \frac{n}{2} \rceil$ as desired. □

This bound is tight. For this, consider any cycle $C_k = v_1 v_2 \ldots v_k$, $k \geq 5$, and join each v_i with a new vertex x_i. Clearly this graph is $\{C_4, \text{diamond}, \text{false-twin}\}$-free, it has $n = 2k$ vertices and $\lceil \frac{n}{2} \rceil = k$ bicliques.

As a direct consequence of Theorem 3.4, we obtain the following corollaries.

Corollary 3.5. *Let T be a false-twin-free tree on $n \geq 4$ vertices. Then T has at least $\lceil \frac{n}{2} \rceil$ bicliques.*

Moreover,

Corollary 3.6. *For $\lceil \frac{n}{2} \rceil \leq k \leq n - 2$ and $n \geq 4$ there exists a false-twin-free tree T on n vertices and k bicliques.*

4. Lower bounds in $\{K_3, \text{false-twin}\}$-free graphs

In this section we study bounds on the minimum number of bicliques in $\{K_3, \text{false-twin}\}$-free graphs. We show first some useful lemmas.

Lemma 4.1. *Let G be a $\{K_3, \text{false-twin}\}$-free graph. Then every vertex is contained in a v-star biclique.*

Proof. Let v be a vertex. If $|N(v)| = 1$ then the result clearly follows. Suppose now that $|N(v)| > 1$. Now, since $N(v)$ is an independent set, $\{v\} \cup N(v)$ is contained in one biclique. If there is no vertex w such that $N(v) \subseteq N(w)$ then $\{v\} \cup N(v)$ is a v-star biclique. Otherwise, let u be the vertex with maximum degree among all vertices in $N(v)$. Clearly, since G is $\{K_3, \text{false-twin}\}$-free, if there are two vertices of same maximum degree, then they have some different neighbors. Hence, $\{u\} \cup N(u)$ is a u-star biclique that contains the vertex v as desired. □

Based on the proof of last lemma, we obtain this immediate result.

Corollary 4.2. *Let G be a false-twin-free graph. Let v be a vertex such that $d(v) = \Delta(G)$ and v does not belong to a K_3. Then $\{v\} \cup N(v)$ is a biclique.*

The next result will help us to prove the main theorem of the section.

Lemma 4.3. *Let G be a $\{K_3, \text{false-twin}\}$-free graph. Let v be a vertex such that $d(v) = k$. Then v belongs to at least k different bicliques.*

Proof. Let $v_1, v_2, \ldots, v_k$ be the neighbors of v. Clearly they are an independent set. Let $x_1, x_2, \ldots, x_\ell$ be the vertices adjacent to $v_1, v_2, \ldots, v_k$ (not including v). Let G' be the subgraph induced by $\{v\} \cup \{v_1, v_2, \ldots, v_k\} \cup \{x_1, x_2, \ldots, x_\ell\}$. Clearly, $v_1, v_2, \ldots, v_k$ are not false-twins in G' and since G is K_3-free, v is not adjacent to any x_j, $1 \leq j \leq \ell$. Now, for each $1 \leq i \leq k$, let $S_{v_i} = \{N_{G'}(x_j) : v_i \in N_{G'}(x_j), 1 \leq j \leq \ell\} \cup \{N(v)\}$. Let $Cl(S_{v_i}) = \bigcap_{S \in S_{v_i}} S$. Observe first that $S_{v_i} \neq \emptyset$ for all i since $N(v)$ belongs to all of them. Observe then that $Cl(S_{v_i}) \cup (\{x_j : N_{G'}(x_j) \in S_{v_i}, 1 \leq j \leq \ell\} \cup \{v\})$ is a biclique in G' and therefore a biclique in G. We show now that $Cl(S_{v_i}) \neq Cl(S_{v_j})$ for all $i \neq j$, i.e., v belongs to k different bicliques in G. Suppose by contrary, that $Cl(S_{v_i}) = Cl(S_{v_j})$. Now, since $v_i \in Cl(S_{v_i})$, we have $v_i \in Cl(S_{v_j})$. Similarly, $v_j \in Cl(S_{v_i})$. So, for all $S \in S_{v_i}$, we have $v_j \in S$. Also, for all $S \in S_{v_j}$, we have $v_i \in S$. That is, $N(v_i) = N(v_j)$, a contradiction since G has no false-twin vertices. □

As a corollary, we obtain the following.

Corollary 4.4. *Let G be a $\{K_3, \text{false-twin}\}$-free graph. Suppose that there is a vertex v such that $G - \{v\}$ has k sets of false-twin vertices. Then $G - \{v\}$ has at least k bicliques less than G.*

Proof. Observe first that, since G has no false-twin vertices, every set of false-twin vertices in $G - \{v\}$ has size exactly two. Let $\{v_i, w_i\}$ be the k sets of false-twin vertices, such that v is adjacent to v_i, $i = 1, \ldots, k$. Observe now that, since v_i and w_i are false-twins in $G - \{v\}$, they belong to exactly the same bicliques but those bicliques containing the edge vv_i.

Consider now the subgraph induced by the vertices $\{v\} \cup \{v_1, v_2, \dots, v_k\} \cup N(v_1) \cup \dots \cup N(v_k)$. Call this graph G'. Clearly, $v_1, v_2, \dots, v_k$ are not false-twins in G'. Now, by Lemma 4.3, v belongs to k different bicliques in G'. These bicliques in G' are either bicliques or are contained in bigger bicliques in G, but they do not contain any of the vertices w_i. Therefore, after removing v, these k bicliques are lost in $G - \{v\}$ since any other biclique containing any v_i contains also w_i. □

Combining the last three results, it follows the main theorem of the section. It gives a tight lower bound on the number of bicliques for $\{K_3, \text{false-twin}\}$-free graphs.

Theorem 4.5. *Let G be a $\{K_3, \text{false-twin}\}$-free graph on $n \geq 4$ vertices. Then G has at least $\lceil \frac{n}{2} \rceil$ bicliques.*

Proof. The proof is by induction on n. For $n = 4$ the result trivially holds. Suppose $n \geq 5$. Now, by Lemma 4.1 there is a vertex v contained in a star biclique. Without loss of generality, we can suppose that v is the center. Consider the graph $G' = G - \{v\}$. We have the following two cases.

- G' is disconnected. Let $G_1, G_2, \dots, G_s$ be the connected components of G' on $n_1, n_2, \dots, n_s$ vertices respectively. Since G has no false-twin vertices, there can exist at most one G_i such that $n_i = 1$. Suppose that there are ℓ components, $\{G_{i_1}, G_{i_2}, \dots, G_{i_\ell}\} \subseteq \{G_1, G_2, \dots, G_s\}$ on $n_{i_1}, n_{i_2}, \dots, n_{i_\ell}$ vertices respectively, such that each G_{i_j} has k_{i_j} sets of false-twin vertices, $j = 1, \dots, \ell$. It is easy to see that each of these sets has exactly two vertices, otherwise G would have false-twin vertices. Now, by Corollary 4.4, G' has at least $k_{i_1} + k_{i_2} + \dots + k_{i_\ell}$ bicliques less than G. Also, since G' is disconnected, v along with at least one vertex of each of the s components is a biclique B in G isomorphic to $K_{1,r}$ that clearly, is not a biclique in G'. Now, consider for each G_{i_j} the graph $Tw(G_{i_j})$. Each of these graphs have $n_{i_j} - k_{i_j}$ vertices and no false-twin vertices. If $n_{i_j} - k_{i_j} = 2$ then $Tw(G_{i_j}) = K_2$, and therefore it has one biclique, i.e., at least $\left\lceil \frac{n_{i_j} - k_{i_j}}{2} \right\rceil$ bicliques. Note that $n_{i_j} - k_{i_j} \neq 3$ as $Tw(G_{i_j})$ is also K_3-free. If $n_{i_j} - k_{i_j} \geq 4$, by the inductive hypothesis, $Tw(G_{i_j})$ has also at least $\left\lceil \frac{n_{i_j} - k_{i_j}}{2} \right\rceil$ bicliques. Now, for all other G_i without false-twin vertices, if $n_i = 2$, G_i has, as before, $1 = \lceil \frac{n_i}{2} \rceil$ biclique, $n_i = 3$ is impossible as G_i is $\{K_3, \text{false-twin}\}$-free and for $n_i \geq 4$, by the inductive hypothesis, G_i has at least $\lceil \frac{n_i}{2} \rceil$ bicliques. If we sum up everything (and suppose the worst case, that is, there exists one G_i, say G_s, such that $n_i = 1$), then the number of bicliques of G is at least
$$\left(\sum_{j=1}^{\ell} \left\lceil \frac{n_{i_j} - k_{i_j}}{2} \right\rceil + k_{i_j} \right) + \left(\sum_{i=1, i \neq i_j}^{s-1} \left\lceil \frac{n_i}{2} \right\rceil \right) + 1 \geq$$
$$\left(\sum_{j=1}^{\ell} \left\lceil \frac{n_{i_j}}{2} \right\rceil \right) + \left(\sum_{i=1, i \neq i_j}^{s-1} \left\lceil \frac{n_i}{2} \right\rceil \right) + 1 \geq \left(\sum_{i=1}^{s-1} \left\lceil \frac{n_i}{2} \right\rceil \right) + 1 \geq \left\lceil \frac{n}{2} \right\rceil$$
as desired.
- G' is connected. Suppose first that in G' there are k sets of false-twin vertices. As before, each of these sets has two vertices. Then, by Corollary 4.4, G' has k bicliques less than G. Consider now the graph $Tw(G')$. This graph has $n - k - 1 \geq 4$ vertices (or just two vertices, i.e., one biclique. Remark that three vertices is not possible) and no false-twin vertices, therefore we can apply the inductive hypothesis. We conclude that $Tw(G')$ has at least $\left\lceil \frac{n-k-1}{2} \right\rceil$ bicliques. Then, G has at least $\left\lceil \frac{n-k-1}{2} \right\rceil + k \geq \lceil \frac{n}{2} \rceil$ bicliques as desired. Suppose last that G' has no false-twin vertices. By the inductive hypothesis, G' has at least $\lceil \frac{n-1}{2} \rceil$ bicliques. Finally, since the v-star biclique is not in G', we conclude that G has at least $\lceil \frac{n-1}{2} \rceil + 1 \geq \lceil \frac{n}{2} \rceil$ bicliques.

Since we covered all cases the proof is now complete. □

Clearly this bound is tight as the same family of graphs presented in Section 3 is $\{K_3, \text{false-twin}\}$-free.
As a consequence of Theorem 4.5, we have the following.

Corollary 4.6. *Let G be a false-twin-free bipartite graph on $n \geq 4$ vertices. Then G has at least $\lceil \frac{n}{2} \rceil$ bicliques.*

5. Discussion for bounds in false-twin-free graphs

In this section we address the following question: Is it true that every false-twin-free graph G has at least $\lceil \frac{n}{2} \rceil$ bicliques?
We answer this question showing a family of false-twin-free graphs with $k + 2^k - 1$ vertices and k^2 bicliques. Consider a graph G constructed as follows. Take a clique $K = \{v_1, v_2, \dots, v_k\}$ on k vertices and an independent set $I =$

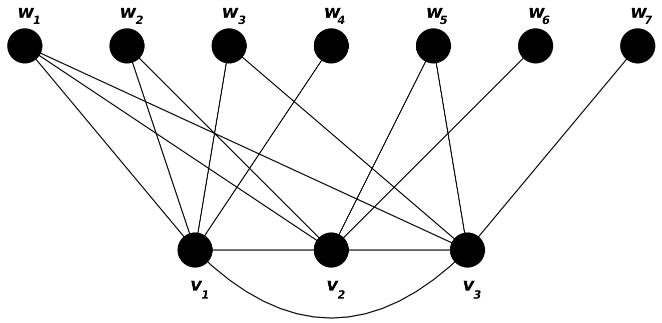

Fig. 4. Construction of graph G for $k = 3$.

$\{w_1, w_2, \ldots, w_{2^k-1}\}$. Consider the set of subsets $\mathcal{B} = \mathcal{P}(K) - \{\emptyset\}$, that is, the power set of K minus the subset containing the empty set. Clearly $|\mathcal{B}| = 2^k - 1$ and all its subsets are different. Finally set $N(w_i) = B_i$, for $B_i \in \mathcal{B}$, $i = 1, \ldots, 2^k - 1$. It is easy to see that the graph G constructed in this way has $k + 2^k - 1$ vertices and no false-twins. Moreover, it has no induced C_4 therefore all its bicliques are stars. See Fig. 4.

Now, $d(v_i) = 2^{k-1} + k - 1$ for all $i = 1, \ldots, k$ and $N(v_i) \neq N(v_j)$ for all $1 \leq i \neq j \leq k$. Then $\{v_i\} \cup N(v_i)$ is a v_i-star biclique. Also, as $N(v_i) - N(v_j) \neq \emptyset$ for all $1 \leq i \neq j \leq k$, we have that $\{v_i\} \cup \{v_j\} \cup (N(v_i) - N(v_j))$ is a v_i-star biclique. We can conclude that G has $k + k(k-1) = k^2$ bicliques. Finally if $n = k + 2^k - 1$, we can see that for $k = 6$ (i.e. $n = 69$), we have $35 = \lceil \frac{n}{2} \rceil < k^2 = 36$ but for $k = 7$ (i.e. $n = 134$), we have $67 = \lceil \frac{n}{2} \rceil > k^2 = 49$. In fact, for $k \in \mathbb{R}$, $k \geq 6.13$ (i.e. $n \geq 75$), we have that $\lceil \frac{n}{2} \rceil \geq k^2$. Following the idea of this example, we state the following conjecture.

Conjecture 5.1. *Let G be a false-twin-free graph on $n \geq 2$ vertices.*
If $k + 2^k - 1 \leq n < (k+1) + 2^{k+1} - 1$, for $k \in \mathbb{N}$, then:
If $n \leq 75$, then G has at least $\lceil \frac{n}{2} \rceil$ bicliques, otherwise it has at least k^2 bicliques.

Moreover, as it is possible to extend this idea to construct a graph G on n vertices, for each n, $k + 2^k - 1 \leq n \leq (k+1) + 2^{k+1} - 1$, and containing from k^2 to $(k+1)^2$ bicliques, we present this tighter conjecture.

Conjecture 5.2. *Let G be a false-twin-free graph on $n \geq 2$ vertices.*
If $k + 2^k - 1 \leq n < (k+1) + 2^{k+1} - 1$, for $k \in \mathbb{N}$, then:
If $n \leq 75$, then G has at least $\lceil \frac{n}{2} \rceil$ bicliques, otherwise it has at least $k^2 + \lfloor (2k+1) \frac{n-(k+2^k-1)}{2^k+1} \rfloor$ bicliques.

To finish this section, we present the following results about the structure of false-twin-free graphs.

Lemma 5.3. *Let G be a false-twin-free graph. If G has a K_3 as a subgraph then there is no vertex that belongs to all bicliques.*

Lemma 5.4. *Let G be a false-twin-free graph. There are at most two vertices v, w that belong to all bicliques and they must be adjacent. Moreover, for every other vertex u, u is adjacent to v if and only if u is not adjacent to w.*

Lemma 5.5. *Let G be a false-twin-free graph. For every biclique B there exists at most one vertex that belongs only to B.*

Lemma 5.6. *Let G be a false-twin-free graph. Let v a vertex such that $d(v) \geq 2$. Then v belongs to at least two different bicliques.*

From last lemma, we obtain this immediate result.

Corollary 5.7. *Let G be a false-twin-free graph. There are at least $\lceil \frac{n}{2} \rceil$ vertices that belong to at least two bicliques.*

Lemma 5.8. *Let G be a false-twin-free graph, $G \neq K_2$. If there are two vertices of degree one, they belong to different bicliques.*

6. Conclusions

In this paper we study bounds on the minimum number of bicliques in a graph. Since adding false-twin vertices to a graph does not change the number of bicliques, we restrict to false-twins-free graphs. We give a tight lower bound for a subclass of $\{C_4, \text{false-twin}\}$-free graphs and for the class of $\{K_3, \text{false-twin}\}$-free graphs. Also we discuss the problem for general false-twin-graphs showing that this bound does not hold. Finally, we present two conjectures for bounds in general false-twin-free graphs.

References

[1] G. Alexe, S. Alexe, Y. Crama, S. Foldes, P.L. Hammer, B. Simeone, Consensus algorithms for the generation of all maximal bicliques, Discrete Appl. Math. 145 (1) (2004) 11–21.
[2] J. Amilhastre, M.C. Vilarem, P. Janssen, Complexity of minimum biclique cover and minimum biclique decomposition for bipartite domino-free graphs, Discrete Appl. Math. 86 (2–3) (1998) 125–144.
[3] E.J. Baker, J.J. Jay, V.M. Philip, Y. Zhang, Z. Li, R. Kirova, M.A. Langston, E.J. Chesler, Ontological discovery environment: a system for integrating gene–phenotype associations, Genomics 94 (6) (2009) 377–387.
[4] Y. Cheng, G.M. Church, Biclustering of expression data, Proc. Int. Conf. Intell. Syst. Mol. Biol. 8 (2000) 93–103.
[5] E.J. Chesler, M.A. Langston, Combinatorial genetic regulatory network analysis tools for high throughput transcriptomic data, in: Proceedings of the 2005 Joint Annual Satellite Conference on Systems Biology and Regulatory Genomics, RECOMB'05, Springer-Verlag, Berlin, Heidelberg, 2007, pp. 150–165.
[6] P. Damaschke, Enumerating maximal bicliques in bipartite graphs with favorable degree sequences, Inform. Process. Lett. 114 (6) (2014) 317–321.
[7] V.M.F. Dias, C.M.H. de Figueiredo, J.L. Szwarcfiter, Generating bicliques of a graph in lexicographic order, Theoret. Comput. Sci. 337 (1–3) (2005) 240–248.
[8] D. Eppstein, Arboricity and bipartite subgraph listing algorithms, Inform. Process. Lett. 51 (4) (1994) 207–211.
[9] B. Ganter, R. Wille, Formal Concept Analysis, Springer-Verlag, Berlin, 1999.
[10] S. Gaspers, D. Kratsch, M. Liedloff, On independent sets and bicliques in graphs, in: Graph-Theoretic Concepts in Computer Science, in: Lecture Notes in Comput. Sci., vol. 5344, Springer, Berlin, Heidelberg, 2008, pp. 171–182.
[11] A. Gély, L. Nourine, B. Sadi, Enumeration aspects of maximal cliques and bicliques, Discrete Appl. Math. 157 (7) (2009) 1447–1459.
[12] M. Groshaus, L. Montero, Tight lower bounds on the number of bicliques in false-twin-free graphs, in: LAGOS'15—VIII Latin-American Algorithms, Graphs and Optimization Symposium, Electron. Notes Discrete Math. 50 (2015) 293–298.
[13] M. Groshaus, L. Montero, On the iterated biclique operator, J. Graph Theory 73 (2) (2013) 181–190.
[14] M. Habib, F. de Montgolfier, C. Paul, A simple linear-time modular decomposition algorithm for graphs, using order extension, in: Algorithm Theory—SWAT 2004, in: Lecture Notes in Comput. Sci., vol. 3111, Springer, Berlin, 2004, pp. 187–198.
[15] R. Kirova, M.A. Langston, X. Peng, A.D. Perkins, E.J. Chesler, A systems genetic analysis of chronic fatigue syndrome: combinatorial data integration from SNPs to differential diagnosis of disease, in: Proceedings, International Conference for the Critical Assessment of Microarray Data Analysis, CAMDA06, 2006.
[16] R. Kumar, P. Raghavan, S. Rajagopalan, A. Tomkins, Trawling the web for emerging cyber-communities, Comput. Netw. 31 (11–16) (1999) 1481–1493.
[17] S. Lehmann, M. Schwartz, L.K. Hansen, Biclique communities, Phys. Rev. E (3) 78 (1) (2008) 016108.
[18] J. Liu, W. Wang, Op-cluster: clustering by tendency in high dimensional space, in: Third IEEE International Conference on Data Mining, 2003, ICDM 2003, IEEE, 2003, pp. 187–194.
[19] K. Makino, T. Uno, New algorithms for enumerating all maximal cliques, in: Algorithm Theory—SWAT 2004, in: Lecture Notes in Comput. Sci., vol. 3111, Springer, Berlin, 2004, pp. 260–272.
[20] A.P. Mukherjee, S. Tirthapura, Enumerating maximal bicliques from a large graph using MapReduce, CoRR, http://arxiv.org/abs/1404.4910, 2014.
[21] R.A. Mushlin, A. Kershenbaum, S.T. Gallagher, T.R. Rebbeck, A graph-theoretical approach for pattern discovery in epidemiological research, IBM Syst. J. 46 (1) (2007) 135–149.
[22] E. Prisner, Bicliques in graphs I: bounds on their number, Combinatorica 20 (1) (2000) 109–117.
[23] M.J. Sanderson, A.C. Driskell, R.H. Ree, O. Eulenstein, S. Langley, Obtaining maximal concatenated phylogenetic data sets from large sequence databases, Mol. Biol. Evol. 20 (7) (2003) 1036–1042.
[24] M.R.D. Savio, A. Sankar, N.R. Vijayarajan, A novel enumeration strategy of maximal bicliques from 3-dimensional symmetric adjacency matrix, Internat. J. Artif. Intell. 12 (2) (2014).
[25] R. Schweiger, M. Linial, N. Linial, Generative probabilistic models for protein–protein interaction networks—the biclique perspective, Bioinformatics 27 (13) (2011) i142–i148.
[26] A. Tanay, R. Sharan, R. Shamir, Discovering statistically significant biclusters in gene expression data, Bioinformatics 18 (suppl 1) (2002) S136–S144.
[27] H. Wang, W. Wang, J. Yang, P.S. Yu, Clustering by pattern similarity in large data sets, in: Proceedings of the 2002 ACM SIGMOD International Conference on Management of Data, ACM, 2002, pp. 394–405.
[28] Y. Xiang, P. Payne, K. Huang, Transactional database transformation and its application in prioritizing human disease genes, IEEE/ACM Trans. Comput. Biol. Bioinform. 9 (1) (2012) 294–304.
[29] M.J. Zaki, M. Ogihara, Theoretical foundations of association rules, in: 3rd ACM SIGMOD Workshop on Research Issues in Data Mining and Knowledge Discovery, Citeseer, 1998, pp. 71–78.
[30] Y. Zhang, C.A. Phillips, G.L. Rogers, E.J. Baker, E.J. Chesler, M.A. Langston, On finding bicliques in bipartite graphs: a novel algorithm and its application to the integration of diverse biological data types, BMC Bioinformatics 15 (2014) 110.