Chromatic characterization of biclique covers

Denis Cornaz, Jean Fonlupt

Équipe Combinatoire, Case 189, UFR 921, Université Pierre et Marie Curie, 4 place Jussieu, 75252 Paris Cedex 05, France

Received 5 May 2003; received in revised form 23 January 2006; accepted 24 January 2006

Abstract

A biclique $B$ of a simple graph $G$ is the edge-set of a complete bipartite subgraph of $G$. A biclique cover of $G$ is a collection of bicliques covering the edge-set of $G$. Given a graph $G$, we will study the following problem: find the minimum number of bicliques which cover the edge-set of $G$. This problem will be called the minimum biclique cover problem (MBC). First, we will define the families of independent and dependent sets of the edge-set $E(G)$ of $G$: $F \subseteq E(G)$ will be called independent if there exists a biclique $B \subseteq E(G)$ such that $F \subseteq B$, and will be called dependent otherwise. From our study of minimal dependent sets we will derive a $0$–$1$ linear programming formulation of the following problem: find the maximum weighted biclique in a graph. This formulation may have an exponential number of constraints with respect to the number of nodes of $G$ but we will prove that the continuous relaxation of this integer program can be solved in polynomial time. Finally we will also study continuous relaxation methods for the problem (MBC). This research was motivated by an open problem of Fishburn and Hammer.

© 2006 Elsevier B.V. All rights reserved.

Keywords: Biclique; Bipartite; Chromatic number $\chi$; Clique number $\omega$

1. Introduction

Let $G$ be a simple undirected graph with node-set $V(G)$ and edge-set $E(G)$. $B \subseteq E(G)$ is a biclique if $B$ is the edge-set of a complete bipartite subgraph of $G$. A biclique cover $B$ of $G$ is a collection of bicliques covering $E(G)$ (every edge of $G$ belongs to at least one biclique of the collection). The biclique covering number $bc(G)$ of $G$ is the cardinality of a minimum biclique cover (MBC) of $G$; we use the notation of [14] (this parameter is called the bipartite dimension in [7]). The MBC problem is the problem of determining $bc(G)$ for any simple graph $G$.

The bipartite version of MBC when $G$ is a bipartite graph arises in many areas. Amilhastre et al. [2] give references concerning automata and language theories, graphs, partial orders, artificial intelligence and biology. In the general case Günlük [11] gives an application concerning the maximum multicommodity flow problem and the min-cut max-flow ratio for multicommodity flow. Fishburn and Hammer [7] show that $bc(G)$ equals the boolean interval dimension of the complementary graph of $G$. When $G$ is bipartite, $bc(G)$ is equal to the boolean rank of its adjacency matrix (see [14]). Monson et al. [14] provide a survey of known results and connections with factorizations of $0$–$1$ matrices.

Orlin [16] shows that the calculation of $bc(G)$ is $NP$-hard for general bipartite graphs and Müller [15] shows that it is $NP$-hard even for chordal bipartite graphs. However, it is possible to check whether or not $bc(G) = k$ for $k = 1, 2$
by verifying that $G$ does not contain some specific graphs as induced subgraphs (see [7]). Thus the decision problem $bc(G) = k$ is polynomial for $k \leq 2$. Amilhastre et al. [2] show that for bipartite domino-free graphs MBC is polynomial. The class of bipartite domino-free graphs includes the class of $C_4$-free bipartite graphs and bipartite distance-hereditary graphs. In [1] it is shown that the problem is also polynomial for the class of bipartite convex graphs.

Bermond [3] found a tight upper bound $b(n)$ ($n$ is the number of nodes of $G$). The result of Bermond is: $b(n) = n - \lfloor \log_2 n \rfloor$ for $n \leq 10$. Chung [4] proved the following result conjectured by Bermond: $b(n)/n \rightarrow 1$. Tuza [20] proved that $b(n) \leq n - \lfloor \log_2 n \rfloor + 1$ for all $n$. There are also other tight bounds: $[\log_2 \chi(G)] \leq bc(G) \leq \tau(G)$. In the left inequality, proved by Harary et al. [12], $\chi(G)$ is the chromatic number of $G$. This inequality is an equality if $G$ is a complete or a complete multipartite graph (see [3,13]). In the right inequality, $\tau(G)$ is equal to the minimum number of stars needed to cover the edge-set of $G$ (a star is a subset of edges incident to the same node; note that a star is a biclique). When $G$ has no cycle of length four, every complete bipartite subgraph is a star and then the right inequality becomes an equality. Let $G_E$ be the graph defined as follows: the node-set of $G_E$ is the edge-set of $G$ and two nodes are adjacent in $G_E$ if and only if the corresponding edges have distinct end-nodes and are not included in a cycle of length four. The size of the maximum clique of $G_E$ is a tight lower bound for $bc(G)$ (see [14]). Fishburn and Hammer [7] gave a better bound: $\chi(G_E) \leq bc(G)$. Moreover, they proved that this inequality is an equality in the case of triangle free graphs. This equality had already been noticed, but only for bipartite graphs (see [19]).

Our study of the biclique covering problem was motivated by its application to the multicommodity flow problem and its applications to telecommunication networks, and also by open questions and problems stated in [7].

This paper is organized as follows. In the next section we give further notation and definitions and present some preliminary results. $F \subseteq E(G)$ is independent if $F$ is contained in a biclique. A subset of $E(G)$ which is not independent will be called dependent. The maximal independent sets (with respect to inclusion), called also the complete sets, are the maximal bicliques. In Section 3, we give a complete characterization of the independent sets. In Section 4 we study the relation between the biclique covering problem and coloration problems of graphs and hypergraphs. We define the $r$-chromatic number of a graph noted $\chi_r(G)$ and we study this number. (The $r$-coloration of a graph is an extension of the usual notion of coloration for graphs.) The main result of this section is that for any $r > 0$, there exists a graph $G$ such that $\chi_r(G) < bc(G)$. This (negative) result means that there is no hope to reduce the biclique covering problem to a coloration problem. In Section 5 we study the minimal dependent sets. We give a complete description of these sets and as a consequence we establish necessary and sufficient conditions under which $\chi_r(G) = bc(G)$. This result provides an answer to an open problem of Fishburn and Hammer. We also prove that the problem of finding the minimal dependent set of maximum cardinality is NP-hard. In Section 6, we show how to find a minimal dependent set of minimum cost. In the last section, we formulate the maximum weighted biclique problem and the minimum biclique covering as 0–1 linear programs with an exponential number of constraints (with respect to the size of the node-set of this graph). However the continuous relaxation of those linear programs can be solved in polynomial time and this is one of the main result of this paper.

2. Notation, definitions and preliminary results

The node-set of $G$ will be designed by $V = V(G)$. The edge-set of $G$ is denoted by $E(G)$. The complement graph of $G$ is $\overline{G} = (V, E(\overline{G}))$. We will set $n = |V|$. $K_n = G \cup \overline{G}$ is the complete graph with node-set $V$. $E = E(G) \cup E(\overline{G})$ is the edge-set of $K_n$, so $K_n = (V, E)$.

We will assign to each edge $e$ of $K_n$ a valuation $p(e)$ defined as follows: $p(e) = 1$ if $e \in E(G)$ and $p(e) = 0$ if $e \in E(\overline{G})$.

2.1. Odd cycles

A walk is a sequence of nodes $P = v_1, \ldots, v_{k+1}$ of $K_n$ such that $k \geq 1$ and $e_i = (v_i, v_{i+1})$ is an edge of $K_n$ for $i = 1, \ldots, k$.

$e_1, \ldots, e_k$ is the edge-sequence of $P$; $k$ is the length of $P$. If all the nodes of $P$ are distinct, $P$ is a chain linking $v_1$ and $v_{k+1}$. An edge will be identified to a chain of length 1. If $v_1 = v_{k+1}$, $P$ is a closed walk. The node-set of $P$ is $V(P)$ and the edge-set of $P$ is $E(P)$. If $P$ is a closed walk, $k \geq 3$ and the nodes of the sequence $C = v_1, \ldots, v_k$ are all distinct, $C$ is a cycle. $e_1, e_2, \ldots, e_k$ is the edge-sequence of $C$. Two nodes (resp. two edges) of $C$ are consecutive on $C$ if they
are the endnodes of an edge of $C$ (resp. incident to a node of $C$). An edge linking two non-consecutive nodes of $C$ is a chord. A chordless cycle is a hole.

Given a cycle $C$ and two distinct nodes $u$ and $v$ of $C$, we say that a chain $P$ linking $u$ and $v$ is external to $C$ if no node of $P$ distinct from $u$ and $v$ belongs to $C$. Consider the two chains $P_1$ and $P_2$ linking $u$ and $v$ on $C$; there exists a cycle $C_1$ with node-set $V(P_1) \cup V(P)$ and edge-set $E(P_1) \cup E(P)$; similarly there exists a cycle $C_2$ with node-set $V(P_2) \cup V(P)$ and edge-set $E(P_2) \cup E(P)$. We say that $C_1$ and $C_2$ are the cycles induced by $C$ and $P$.

**Definition 1.** The valuation $p(P)$ of a walk $P$ is: $p(P) = \sum_{i=1}^{k} p(e_i)$. $P$ is an odd walk if $p(P)$ is odd and an even walk if $p(P)$ is even.

Parity properties refer to the function $p$ and not to the length of the sequence describing a walk or a cycle. A sequence is even (resp. odd) if its edge-sequence contains an even (resp. odd) number of edges of $E(G)$. Thus a sequence and the length of a sequence may be of distinct parity.

We will now prove two easy results:

**Lemma 2.** If $C_1$ and $C_2$ are two cycles induced by an odd cycle $C$ and an external chain $P$, $p(C_1)$ and $p(C_2)$ have distinct parity.

**Proof.** $p(C_1) + p(C_2) = p(C) + 2p(P)$; $p(C_1)$ and $p(C_2)$ have distinct parity since $p(C)$ is odd. □

**Lemma 3.** A subgraph $H$ of $K_n$ contains an odd closed walk if and only if it contains an odd hole.

**Proof.** Necessity follows from the definitions; to prove sufficiency, suppose that $H$ contains an odd closed walk. Consider the odd closed walk $P$ of minimum length. If $k = 2$, $P$ is even. So $k > 2$. If $P = v_1, \ldots, v_i, \ldots, v_{k+1}$ with $v_1 = v_i = v_{k+1}$ for some $1 < i < k + 1$, $P_1 = v_1, \ldots, v_i, v_1$ and $P_2 = v_i, \ldots, v_{k+1}, v_i$ are closed walks with length smaller than $k$. As $p(P) = p(P_1) + p(P_2)$, $P_1$ or $P_2$ is odd, which contradicts our choice of $P$.

So, $C = v_1, \ldots, v_i, \ldots, v_k$ is a cycle. If $C$ has a chord $e$, one of the two cycles induced by $C$ and $e$ is odd by Lemma 2 and the length of this cycle is smaller than $k$, a contradiction. □

2.2. Independent and minimal dependent sets

The property for a subset of $E(G)$ to be independent is hereditary: If $F$ is independent and $F' \subseteq F$, $F'$ is also independent. In particular $\emptyset$ is independent. $F \subseteq E(G)$ is a minimal dependent set or a circuit set if $F$ is dependent but any proper subset of $F$ is independent. We can assume that $E(G)$ is dependent (otherwise all the problems treated in this paper are trivial); hence the set of minimal dependent sets noted $\mathcal{C}(G)$ is nonempty.

2.3. Rooted graph

**Definition 4.** Let $F \subseteq E(G)$. The rooted graph of $F$ is the subgraph $H$ of $K_n = (V, E)$ defined as follows:

- a node $v \in V$ is in $V(H)$ if $v$ is incident to an edge of $F$,
- an edge $e \in E$ is in $E(H)$ if either $e \in F$, or $e \in E(G)$ and the endnodes of $e$ are in $V(H)$.

3. Independent sets

Let $H$ be a subgraph of $K_n$ and let us consider a set $\{+,-\}$ of two elements called labels. A labelling of $H$ is an application: $V(H) \rightarrow \{+,-\}$. We say that this labelling is a good labelling if it satisfies the following condition:

The labels of the two endnodes of any edge $e$ of $H$ are identical if $p(e) = 1$ and distinct if $p(e) = 0$.

If $H$ is a tree, a good labelling can be found by the following algorithm:

Assign the label $+$ to an initial node of $H$. While there exists an edge $e \in E(H)$ with one endnode $u$ labelled and the other $v$ not labelled, label $v$ with the same label as $u$ if $p(e) = 0$ and with the other label if $p(e) = 1$. 
The complexity of this procedure is linear in |V(H)|. We will now describe a simple algorithm to construct a good labelling on H or to decide that no good labelling exists. We can assume that H is connected (otherwise apply the following algorithm to each component of H).

**Labelling Algorithm.**

(a) If H is a tree, find a good labelling of H by the previous method.
(b) If H is not a tree, consider a spanning tree (V(H), T) of H and find a good labelling of this tree.
(c) If the good labelling of the tree is also a good labelling of (V(H), T ∪ {e}) for each e ∈ E(H) \ T, this labelling is a good labelling of H; otherwise, there exists no good labelling for H.

The justification of the validity of this procedure is immediate. The complexity of this algorithm is linear in |E(H)|.

A consequence of this algorithm is that a good labelling exists for H if and only if a good labelling exists for any cycle of H.

**Lemma 5.** H admits a good labelling if and only if H contains no odd cycle.

**Proof.** By the preceding remark we can assume that H itself is a cycle: the graph obtained by deleting an edge e = (u, v) of H is a chain P linking u and v with a good labelling. u and v have similar labels in this labelling if and only if P is an even chain. This labelling extends to C if and only if P is even and p(e) = 0 or P is odd and p(e) = 1. In both cases H is an even cycle.

We can now state the main result of this section:

**Theorem 6.** F ∈ E(G) is independent if and only if the rooted graph H of F contains no odd cycle.

**Proof.** Let us show first that the condition is necessary. If F is independent, F is contained in a biclique B ∈ E(G) which is the edge-set of a complete bipartite graph. Denote by {W₁, W₂} the partition of the node-set of this bipartite graph. Assign the label (+) to the nodes of W₁ and the label (−) to the nodes of W₂. If e ∈ E(H) and p(e) = 1, e ∈ F and the two end-nodes of e belong to the two distinct classes of the partition with distinct labels. If e ∈ E(H) and p(e) = 0, e /∈ B and the two end-nodes of e cannot be in the two distinct classes of the partition. Thus, these end-nodes have the same label. H has a good labelling and by Lemma 5 has no odd cycle.

We prove now that the condition is sufficient. Assume that H contains no odd cycle. By Lemma 5, a good labelling exists for H. Let W₁ (resp. W₂) be the set of nodes labelled with (+) (resp. (−)) in this labelling. If e ∈ F, p(e) = +1 and one endnode of e is in W₁ and the other in W₂. If B is the set of edges of Kₙ with one endnode in W₁ and the other in W₂, F ⊆ B and (W₁ ∪ W₂, B) is a complete bipartite graph of Kₙ. If e ∈ B e /∈ E(G); hence e ∈ E(G) and B is a biclique.

The Labelling Algorithm is polynomial which implies:

**Corollary 7.** The problem of deciding if a subset of edges of a graph is independent is polynomial. The problem of finding a biclique of a graph containing a given independent set is polynomial.

**4. Coloration of the edge-set of G**

In this section, we will be concerned with edge coloring problems. Given an edge coloration of G, the set of edges with the same color will be called a class of color.

**Definition 8.** Let r be an integer parameter ≥ 2.

- A coloration of G is a strong coloration if any dependent set belongs to at least two classes of colors.
- A coloration of G is a r-coloration if any dependent set F ⊆ E(G) such that |F| ≤ r belongs to at least two classes of colors.
In a strong coloring all the classes of colors are independent sets. Thus the minimum number of colors in a strong coloring is also the minimum number $bc(G)$ of bicliques which cover the set $E(G)$. In an $r$-coloration, every subset $F$ of a class of colors is independent if $|F| \leq r$. The $r$-chromatic number $\chi_r(G)$ of $G$ is the minimum number of colors in a $r$-coloration. Clearly a strong coloring is an $r$-coloration for all $r > 0$. Hence $\chi_r(G) \leq bc(G)$. Note also that if $r < s$, an $s$-coloration is also a $r$-coloration and $\chi_r(G) \leq \chi_s(G)$. With these (new) definitions in mind we can revisit previous results of Fishburn and Hammer (see [7]) for a graph $G$ without isolated nodes and with a nonempty edge-set. The main result of Fishburn and Hammer is that if $\omega(G) \leq 2$, then $\chi_2(G) = bc(G)$ (where $\omega(G)$ is the size of the largest clique of $G$). This result is obviously false if $\omega(G) > 2$ and a trivial counterexample is obtained when $G$ is the complete graph on three nodes $K_3$: $\chi_2(K_3) = 1$ but $bc(K_3) = 2$. So, if $\omega(G) > 2$, it may be possible that $\chi_2(G) = bc(G)$ provided that we consider 2-colorations with the following additional condition: no class of color contains a triangle $K_3$. But Fishburn and Hammer gave also an example (Fig. 1) showing that this statement is also false: the three colors associated to the edges of the graph are the three classes of colors in the optimal 2-coloration of $G$. The set of edges with color 3 is not an independent set; it is not hard to see that $bc(G) = 4$. However, no class of colors contains a triangle and the additional condition proposed by Fishburn and Hammer is satisfied. In their paper, they asked if there exists some positive integer $r$ for which $\chi_r(G) = bc(G)$ with eventually some other conditions on the $r$-coloration. We will now prove that this assertion is false; note that this result is negative since it implies that there is no way to reduce the minimum biclique covering problem to a coloring problem on hypergraphs if we fix the maximum size of an hyperedge.

Let $F$ be the edge-set of a cycle of size 5 of $K_5$. If we delete $F$ from the edge-set of $K_5$ we obtain another cycle $F_2$ of size 5. Delete one edge $f$ from $F$. Thus the rooted graph of $F$ is an odd cycle of size 5 and $F$ is not independent by Theorem 6 (see Fig. 2). But the rooted graph of $F' = F \setminus \{f\}$ is a path and therefore $F'$ is independent. Thus if we assign color 1 (resp. 2) to $F$ (resp. $F_2$) we obtain a 4-coloration of $K_5$ and $\chi_4(K_5) = 2$. But $bc(K_5) = 3$. Thus, $\chi_4(K_5) < bc(K_5)$. We can generalize this result as follows:

**Theorem 9.** For every integer $r$, there exits a graph $G$ such that $\chi_r(G) < bc(G)$.

**Proof.** Consider first the complete graph $K_{2r+1}$ with edge set $\Omega$ and let $F \subseteq \Omega$ be the edge-set of a cycle of $K_{2r+1}$ of size $2r + 1$. Let $G$ be the graph obtained from $K_{2r+1}$ by adding to each pair of nodes $u, v$ such that $e = (u, v) \in \Omega \setminus F$ a path $u, x, y, v$. $(x, y)$ will be called the copy of $e = (u, v)$ and $X$ will be the set of all copies ($|X| = |\Omega \setminus F|$). Finally we denote by $C_4(e)$ the graph of $G$ induced on $[u, x, y, v]$. Note that the edge-set of $C_4(e)$ is a biclique of $G$. Two distinct copies form a dependent set; a copy $(x, y)$ and an edge of $F$ also form a dependent set. As $F$ is dependent but any subset of $F$ is independent, we need two colors to color the set $F$ in a strong coloring but only one color to color $F$ in a 2r-coloration. From these observations we easily deduce that: $\chi_{2r}(G) = 1 + |X|$ and $bc(G) = 2 + |X|$; this proves our theorem. □
5. Minimal dependent sets

In this section we study rooted graphs of minimal dependent sets; these graphs will be called obstructions.

If $H$ is an obstruction, $H = (V_1, E_1)$ is the rooted graph of a dependent set $F$; therefore $H$ contains an odd cycle $C = v_1, \ldots, v_k$ by Lemma 6. We will set $F_1 = F \cap E(C)$ and $F_2 = F \setminus F_1$; $v \in V(C)$ will be called an exposed node of $C$ if no edge of $F_1$ is incident to $v$ (Fig. 3).

All the forthcoming proofs will be based on the following lemma:

Lemma 10. Let $v \in V(H) \setminus V(C)$ and let $f \in F$ be incident to $v$; the other endnode $u$ of $f$ is an exposed node of $C$ and $f$ is the unique edge of $F$ incident to $u$.

Proof. There exists always an edge $f \in F$ incident to $v$ (see the definition of a rooted graph). As $F$ is a minimal dependent set, $F \setminus f$ is independent and by Theorem 6, the rooted graph $H_1$ of $F \setminus f$ does not contain the odd cycle $C$; but $f$ is not an edge of $C$; therefore, there exists a node of $C$ which is not a node of $H_1$; but this node is one of the two end-nodes $u, v$ of $f$; as $v \notin V(C)$, this node is $u$; this implies that $f$ is the unique edge of $F$ incident to $u$. This proves the lemma. □

In the rest of this section we will make the following assumption (\(\alpha\)):

(\(\alpha\)) The odd cycle $C$ has no chord with valuation equal to 1 and among all the odd cycles which satisfy this property, $C$ is such that $p(C) = |F_1|$ is maximum.

Note that $C$ always exists since, by Lemma 3, $H$ contains an odd hole which has no chord at all. Note also that if $e$ is a chord of $C$, $p(e) = 0$ and $e \in E(\overline{G})$.

5.1. The set $F_2$

Proposition 11. (a) Any exposed node of $C$ is incident to (exactly) one edge of $F_2$.

(b) $F_2$ is a matching of the graph $H$.

Proof. If $u$ is an exposed node of $C$, $u$ is incident to an edge $f = (u, v) \in F_2$; Assumption (\(\alpha\)) implies that $f$ is not a chord of $C$ and $v \notin V(C)$; by Lemma 10, $f$ is the unique edge of $F$ incident to $u$; this proves part (a) of the proposition. Assume now that $v$ is incident to an other edge $(v, w) \in F$; again by Lemma 10, $w$ is an exposed node of $C$. The external chain to $C$ and the external chain $D = u, v, w$ to $C$ induce an even cycle $C_1$ and an odd cycle $C_2$ by Lemma 2; if $C_1$ contains a node $z$ distinct from $u$ and $w$, $z$ is incident to some $e = (z, z') \in F$ and by Lemma 10 (applied to $C_2$) $z'$ is an exposed node of $C_2$. Assumption (\(\alpha\)) ensures that $e$ is not a chord of $C$ and $z'$ is one of the three nodes $u, v, w$; but none of these three nodes is exposed on $C_2$; by Lemma 10, this is impossible: thus $u$ and $v$ are consecutive nodes of $C$. 

Assume that no other edge with valuation equal to 1. This proves the Proposition.

If there exists a third edge \((v, z) \in F, u, w, z\) are pairwise consecutive nodes of \(C, C = u, w, z\) and \(|F_1| = 0\) which is impossible. Thus \(C_1\) has no chord \(e\) with \(p(e) = 1\). But \(p(C_1) = p(C) + 2\) which contradicts assumption (z). \(f\) is the unique edge of \(F\) incident to \(v\) and \(F_2\) is a matching of \(H\). □

5.2. Chords of \(C\)

Let \(e = (u, v)\) be a chord of \(C\); we know that \(p(e) = 0\). Let \(C_1\) (resp. \(C_2\)) be the even (resp. odd) cycle induced by \(C\) and \(e\). We note by \(f_1 = (u, w)\) (resp. \(f_2 = (v, z)\)) the edge of \(C_1\) incident to \(u\) (resp. \(v\)).

We call \(e\) a short chord if the size of \(C_1\) is 3 (in this case, \(w = z\)) and a diagonal if the size of \(C_1\) is 4 (in this case, \((w, z)\) is an edge of \(C_1\) denoted \(f_3\)). The even cycle \(C_1\) will be called the fundamental cycle of \(e\).

**Proposition 12.** Either \(e\) is a short chord or a diagonal. \(p(f_1) = p(f_2) = 1\); \(f_1\) (resp. \(f_2\)) is the unique edge of \(F\) incident to \(u\) (resp. \(v\)). If \(e\) is a diagonal, \(p(w, z) = 0\) (Fig. 4).

**Proof.** Let \(w\) be a node of \(C_1\) distinct from \(u\) and \(v\); \(w\) is incident to an edge \(f = (w, w') \in F\). By Lemma 10, \(w'\) is an exposed node of \(C_2\). The only possible choices for \(w'\) are \(u\) or \(v\). So we can assume that \(w' = u\); by Lemma 10, \(f_1 = (u, w)\) is the unique edge of \(F\) incident to \(u\). As \(C_1\) is even, there exists an other edge of \(C\) with valuation equal to 1; this edge is incident to \(v\); therefore this edge is \(f = (v, z)\). \(C_1\) contains no other nodes than the four nodes \(u, v, z_1, z_2\) and no other edge with valuation equal to 1. This proves the Proposition. □

5.3. Wings

Finally, let \(e = (v, w) \in E(H)\) with \(p(e) = 0\) and \(v \in V(H) \setminus V(C)\) we know that there exists an edge \(f = (v, u) \in F_2\) where \(u\) is an exposed node of \(C\).

**Proposition 13.** The endnode \(w\) of \(e\) belongs to \(C\).

**Proof.** Assume that \(w \notin V(C)\); there exists \(g = (w, z) \in F_2\) with \(z \in F_2\). Let \(D\) be the chain \(u, v, w, z\) and let \(C_1\) (resp. \(C_2\)) be the even (resp. odd) cycle induced by \(C\) and \(D\). If \(C_1\) contains an edge \(e \in F_1\), this edge is incident to \(u\) or \(v\) by Lemma 10. But this is impossible since \(u\) and \(v\) are exposed nodes of \(C\); \(C_1\) contains no chord with valuation 1 since \(F_2\) is a matching of \(G\). But \(p(C_1) = p(C) + 2\) which contradicts Assumption (z). □

So we can assume from now on that \(e = (v, w)\) with \(w \in V(C)\). \(e\) will be called a wing of \(H\).

Let \(D\) be the chain \(u, v, w\) and \(C_1\) (resp. \(C_2\)) be the even (resp. odd) cycle induced by \(C\) and \(D\). Let \(f_1 = (w, z)\) be the edge of \(C_1\) incident to \(w\). \(C_1\) will be called the fundamental cycle of \(e\).

**Proposition 14.** \(C_1\) is a cycle of size 4; \(p(f_1) = 1\) and \(f_1\) is the unique edge of \(F\) incident to \(w\).

**Proof.** As \(p(f) = 1\), there exists an edge of \(C_1\) distinct from \(f\) with valuation equal to 1 and the size of \(C_1\) is at least 4.
Let \( z \) be a node of \( C_1 \) distinct from \( u, v, w, z \) is incident to an edge \( f_1 = (z, z') \in F \). By Lemma 10, \( z' \) is an exposed node of \( C_1 \). The only possible choices for \( z' \) are \( u, v \) or \( w \) but \( u \) and \( v \) are not exposed nodes of \( C_1 \). Hence \( z' = w \). This proves that \( C_1 \) is a cycle of size 4 and the result follows. \( \square \)

5.4. Characterization of obstructions

**Theorem 15.** A subgraph \( H \) of \( K_n \) is an obstruction if and only if it satisfies Propositions 11–14.

**Proof.** To prove sufficiency, we need to show by Lemma 5 and Theorem 6 that the rooted graph of \( F \backslash f \) has a good labelling for all \( f \in F \).

Let \( H_1 = (V_1, E_1) \) be the rooted graph of \( F \backslash f \) for some \( f \in F \). Let \( (V_1, T) \) be the subgraph of \( H_1 \) obtained by deleting the edges of \( H_1 \) which are also diagonals, short chords or wings of \( H \). The only possible cycle of \( (V_1, T) \) is \( C_1 \); clearly at least one edge of \( C \) does not belong to \( H_1 \) and \( T \) is a spanning tree of \( H_1 \). We can start from a good labelling on \( T \) and we will prove that this labelling remains a good labelling if we add to \( T \) an edge \( e = (u, v) \) of \( H_1 \) which is a diagonal of \( H \). (The cases of short chords or wings are similar and we will omit the proof.) So, assume that we add \( e \) to \( H_1 \); the fundamental cycle \( C_1 \) of \( e \) is: \( u, w, z, v \). If \((u, w)\) and \((z, v)\) are edges of \( H_1 \), \((u, w)\), \((z, v)\) and \((v, w)\) belong to \( T \), and the good labelling of \((V_1, T)\) is also a good labelling for \((V_1, T \cup \{e\})\) since \( C_1 \) is even. Assume now that one of the two edges \((u, w)\), \((z, v)\), for instance \((u, w)\), is not an edge of \( H_1 \). Since \((u, w)\) is the unique edge of \( F \) incident to \( w \) and \( u \) and \( w \) are not nodes of \( H_1 \); therefore, \( e \) is not an edge of \( H_1 \), a contradiction. This finishes the proof. \( \square \)

We now come back to the \( r \)-coloring problem of \( G \).

**Proposition 16.** Let \( F \) be a minimum dependent set and let \( W \) be the node-set of the obstruction \( H \) of \( F \). The induced subgraph \( G(W) \) contains a clique of size \(|F|\).

**Proof.** Assume first that every node of \( C \) is incident to one edge of \( F \); as \( F_2 \) is a matching of \( H, F \) is a perfect matching of \( H \). The size of \( C \) is: \(|C| = 2|F_1| + |F_2|\).

If \(|F_2|\) is even, \( C \) is an odd hole of even cardinality. Short chords cannot exist in this situation and the size of the fundamental cycles of wings and diagonals is 4. Hence \( H \) is bipartite and \( W \) can be partitioned into two stable sets \( W_1 \) and \( W_2 \) each of size \(|F|\). \( G(W_1) \) and \( G(W_2) \) are cliques of size \(|F|\) and the result follows.

If \(|F_2|\) is odd, \(|F| = |F_1| + |F_2|\) is even since \( p(C) = |F_1| \) is odd. We say that two edges \( f, f' \) of \( F \) are neighbouring edges if there exist an edge of \( C \) with valuation equal to 0 incident to \( f \) and incident to \( f' \). Each edge of \( F \) has two neighbouring edges; hence, we can assume that \( F = \{f_1, f_2, \ldots, f_{2r}\} \) with \( r \geq 1 \) and that the two neighbouring edges of \( f_i \) are \( f_{i-1} \) and \( f_{i+1} \) for \( 1 \leq i \leq 2r \) (with \( f_{2r+1} = f_1 \)). Moreover, by Theorem 15, if \( e \) is a diagonal or a wing, \( e \) is incident to two edges of \( F \) which are neighbouring edges. If \( W_1 \) (resp. \( W_2 \)) is the set of endnodes of \( F = \{f_1, \ldots, f_{2r-1}\} \) (resp. \( \{f_2, \ldots, f_{2r}\} \)), \( H(W_1) \) and \( H(W_2) \) are induced subgraphs of \( H \) with no edge \( e \) with \( p(e) = 0 \). Therefore, \( G(W_1) \) and \( G(W_2) \) are cliques of size \(|F|\); again the result follows.

Assume now that there is a unique node of \( C \), for instance \( v_1 \) incident to two edges of \( F \). The induced subgraph \( H(W \backslash v_2) \) is bipartite since \( C \) is not a cycle of this graph and \( F \backslash f_1 \) (where \( f_1 = (v_1, v_2) \)) is a perfect matching of this graph. As in the first part of the proof, \( W \backslash v_2 \) can be partitioned into two stable sets \( W_1 \) and \( W_2 \) each of size \(|F| - 1\). We can assume that \( v_1 \notin W_1 \). No edge with valuation 0 is incident to \( v_1 \), \( G(W_1 \cup v_1) \) is a clique of size \( F \).

Finally if there exists an other node \( v_i \) for some \( 2 \leq i \leq k - 1 \) of \( C \) incident to two edges of \( F \), we first delete the eventual short chords \((v_{i-1}, v_i), (v_{i+1}, v_i)\) and then we replace the two chains \( v_k, v_1, v_2 \) and \( v_{i-1}, v_i, v_{i+1} \) by two edges \( f_1 = (v_k, v_1) \) and \( f_2 = (v_{i-1}, v_{i+1}) \). Finally, we set \( p(f_1) = p(f_2) = 1 \). The graph \( H' \) obtained after this reduction satisfies all the properties of obstructions and \( C \) is reduced to an odd cycle \( C' \) with \( p(C') = p(C) - 2 \). By induction on \( p(C) \), there exists a set \( W_1 \subset V(H') \) which induces in \( G \) a clique of size \(|F'| = |F| - 2 \). As no edge with valuation equal to 0 is incident to \( v_1 \) or \( v_i \), \( G(W_1 \cup \{v_1, v_i\}) \) is a clique of size \(|F|\). \( \square \)

The following theorem generalizes the result: \( bc(G) = \chi_2(G) \) if \( \omega(G) \leq 2 \).

**Theorem 17.** \( bc(G) = \chi_2(G) \) if \( \omega(G) \leq r \).
Proof. Consider an optimal $r$-coloring of the edge-set of $G$. If a class of colors is not an independent set, it contains a minimal dependent set $F$ and by Proposition 16, $|F| \leq \omega(G)$. But in a $r$-coloring all the subset of a class of colors of cardinality $\leq r$ are independent. So each class of colors is independent. Thus, $bc(G) \leq \chi_r(G)$, as $bc(G) \geq \chi_r(G)$, the theorem is proved. □

Recall that we have assumed that $E$ is not a biclique, hence there always exists at least one nonempty dependent set of $G$.

Lemma 18. Let $F_{\text{max}}$ be a minimal dependent set of maximum cardinality and $\phi(G) = |F_{\text{max}}|$. Then:

- $\phi(G) = \omega(G) - 1$ or $\omega(G)$.
- If $\omega(G)$ is odd or $\phi(G)$ is even, $\phi(G) = \omega(G)$.

Proof. From Proposition 16, $\phi(G) \leq \omega(G)$. Let $F$ be a cycle whose node-set is $W$ where $W$ is an odd clique of $G$ of maximum size. $F$ is an odd cycle and the rooted graph of $F$ is $(W, F)$. Hence $F$ is a minimal dependent set and $|F| \leq \phi(G)$. But $|F| = |W| = \omega(G)$ if $\omega(G)$ is odd and $|F| = |W| = \omega(G) - 1$ if $\omega(G)$ is even. The result follows easily from these facts. □

Theorem 19. Finding the minimal dependent set of maximum size is NP-hard.

Proof. Let $G$ be a graph and consider the graph $G'$ obtained by adding a new node adjacent to all the nodes of $G$. If $\phi(G)$ is even, $\omega(G) = \phi(G)$; if $\phi(G')$ is even, $\omega(G) = \omega(G') - 1 = \phi(G') - 1$. Assume now that $\phi(G)$ and $\phi(G')$ are odd: if $\phi(G) = \omega(G) - 1$, then $\omega(G)$ is even, $\omega(G')$ is odd and $\phi(G') = \omega(G') = \omega(G) + 1 = \phi(G) + 2$. Hence, if $\phi(G) = \phi(G')$, $\omega(G) = \phi(G)$; if $\phi(G) = \phi(G') - 2$, $\omega(G) = \phi(G) + 1$. The algorithm which transforms $G$ into $G'$ is polynomial with respect to the size of the input $G$. Thus, we have reduced the maximum clique problem to the minimal dependent set of maximum size problem; this proves our statement. □

6. Minimal dependent set of minimum weight

Assume that a nonnegative weight $w(e)$ is assigned to each edge of $G$. The weight of $F \subseteq E(G)$ is $w(F) = \sum_{e \in F} w(e)$. We will set $w(e) = 0$ for all $e \in E(G)$.

We define now our first optimization problem:

Problem 1. Find a dependent set $F_{\text{min}} \in E(G)$ of minimum weight in the network $(K_n, w)$.

Assume now that there exists a nonnegative cost function $c$ defined on $E \times V$; we note by $c_1(e)$ the cost of $e \in E$ and $c_2(v)$ the cost of $v \in V$.

The cost of a closed walk $P = v_1, \ldots, v_k, v_1$ with edge-sequence $e_1, \ldots, e_k$ is: $c(P) = \sum_{i=1}^k c_1(e_i) + \sum_{i=1}^k d(v_i)$ where $d(v_i) = c_2(v_i)$ if the two edges $e_{i-1}, e_i$ are edges of $E(G)$ and $d(v_i) = 0$ otherwise, for $i = 1, \ldots, k$ (with $e_{-1} = e_k$). We can give the following interpretation for $c(P)$: we pay a cost $c_1(e)$ each time we visit edge $e$ on $P$ but we pay a cost $c_2(v)$ each time we visit node $v$ on $P$ provided that the edges preceding and succeeding $v$ in $P$ have a valuation equal to 0.

Our second optimization problem is

Problem 2. Find the odd closed walk $P_{\text{min}}$ with minimum cost in $(K_n, c)$.

Clearly this problem has a solution since the cost function $c$ is nonnegative. If we start from the edge-weights, we can define the node-costs as follows: $c_1(e) = w(e)$ and $c_2(v) = \min \{w(e)\}$ for all $e \in \delta(v)$ for all $v \in V$ (where $\delta(v)$ is the set of edges of $G$ incident to $v$).

The following proposition shows that Problem 1 and Problem 2 are equivalent:

Proposition 20. $w(F_{\text{min}}) = c(P_{\text{min}})$. 

Proof. Let $F_1$ be the set of edges of $G$ which are edges of the sequence $P_{\min} = v_1, \ldots, v_k$. Note that $\sum_{i=1}^{k} c_1(e_i) \geq w(F_1)$ (we may have a strict inequality since an edge of $F_1$ may appear more than once in the edge-sequence of $P_{\min}$). For each node $v_i$ of the sequence $v_1, \ldots, v_k$ for which we pay the cost $c_2(v_i)$, choose an edge $f_i$ of $G$ incident to $v_i$ and such that: $c_2(v_i) = w(f_i)$. Let $F_2$ be the set of edges obtained in this way. Note that $\sum_{i=1}^{k} d(v_i) = w(F_2)$. Hence, if $F = F_1 \cup F_2$, $c(P_{\min}) \geq w(F)$. But all the nodes of $P_{\min}$ are nodes of the rooted graph of $F$, thus $F$ is dependent and $c(P_{\min}) \geq w(F_{\min})$.

As the weight function $w$ is nonnegative, we can assume that $F_{\min}$ is a minimal dependent set. Let $H$ be the rooted graph of $F$. $H$ is an obstruction; $C = v_1, \ldots, v_k$ is the odd cycle of $H$, $F_1$ is the set of edges of $G$ which belong to $C$, and $F_2$ is the set of edges of $F$ incident to the exposed nodes of $C$. Let $P$ be the closed walk $P = v_1, \ldots, v_k, v_1$. $\sum_{i=1}^{k} c_1(e_i) = w(F_1)$. Moreover we pay a cost $c_2(v)$ for a node of $P$ if and only if this node is an exposed node of $C$; but if $f$ is the edge of $F_2$ incident to an exposed node $v$, $c(f) \geq c_2(v)$. Thus, $c(F_2) \geq \sum_{i=1}^{k} d(v_i)$ and $w(F_{\min}) \geq c(P) \geq c(P_{\min})$.

So, $w(F_{\min}) = c(P_{\min})$. □

We can now state the main result of this section.

Theorem 21. There is a polynomial algorithm for finding the minimal dependent set of minimum cost when the cost function is nonnegative.

Proof. In view of the preceding result, we just have to prove that Problem 2 can be solved in polynomial time. We will consider the following successive transformations: First we replace $K_n$ by the complete oriented graph $(V, A)$ on $n$ nodes; each edge $e \in E$ is replaced by two arcs $a_1$ and $a_2$ of opposite direction. We set $p(a_1) = p(a_2) = p(e)$ (recall that $p(e) \in \{0, 1\}$). Also we set $c'(a_1) = c'(a_2) = c_1(e)$, $c'(a)$ is the cost of $a$.

We now transform $(V, A)$ to obtain a new directed graph: $D = (W, A \cup B)$. We split each node $v \in V$ so that now all the arcs of $A$ have no common extremity, and we call $t(a)$ (resp. $h(a)$) the tail (resp. head) of $a$. So the node-set of $D$ is: $W = \bigcup_{a \in A} \{t(a), h(a)\}$. Note that $|W| = 2|A|$. After this first transformation we add a new set $B$ of arcs defined as follows: We create an arc $b$ from $h(a_1)$ to $t(a_2)$ if and only if there exists a node $v \in V$ which is the head of $a_1$ and the tail of $a_2$; the cost of $b$ is: $c'(b) = c_2(v)$ if $p(a_1) = p(a_2) = 0$, and $c'(b) = 0$, otherwise. We will say that the arc $b$ is associated to the node $v$. The parity of an edge in $B$ is even; i.e. we set $p(b) = 0$ for each $b \in B$. The network $(D, c')$ is now completely defined. (An example is depicted in Fig. 5.)

Note that the edge-sequence of a closed directed walk in $D$ alternatively uses arcs in $A$ and arcs in $B$. Let $\bar{P} = b_1, a_1, b_2, \ldots, b_k, a_k$ be a walk in $D$ with $a_i \in A$ and $b_i \in B$, for $i = 1, \ldots, k$. Since $b_i$ is associated to a node
of $V$, we can associate to $\tilde{P}$ a directed chain $P$ in $(V, A)$: $P = v_1, \ldots, v_k$ with edge-sequence $a_1, \ldots, a_k$. So $c'(\tilde{P}) = \sum_{i=1}^k c'(a_i) + \sum_{i=1}^k c'(b_i)$ and $c(P) = \sum_{i=1}^k c_1(a_i) + \sum_{i=1}^k d(v_i)$. But $c_1(a_i) = c'(a_i)$ and $d(v_i) = c'(b_i)$ for $i = 1, \ldots, k$; thus $c(P) = c'(\tilde{P})$ and this proves that the shortest closed directed walk problem in $D$ and the minimum cost closed directed walk problem in $(V, A)$ are equivalent. This proves our theorem since it is well-known how to find the odd directed walk of minimum length. (The procedure is quite simple and can be found in [8] for instance.) \[\square\]

Theorem 21 can also be proved by reducing Problem 1 to a matching problem (see [5]).

7. Maximum weighted biclique and minimum biclique cover

In this final section we will formulate the maximum weighted biclique problem and the MBC Problem as integer programs and we will study their continuous relaxation. It was shown in [17] that the maximum weighted biclique problem is NP-hard even when all the weights are 1.

7.1. The maximum weighted biclique problem

Let $\mathcal{C}(G)$ be the set of minimal dependent sets of $G$. Assign to each edge $e \in E(G)$ a variable $x_e$ and let $x \in \mathbb{R}^{E(G)}$ be the vector $x = (x_e; e \in E(G))$. If $d_e$ is the weight of $e$, the maximum weighted biclique problem is equivalent to the following integer program $(\mathcal{P}_1)$:

$$\begin{align*}
\forall e \in E(G) & \quad x_e \in \{0, 1\}, \\
\forall C \in \mathcal{C}(G) & \quad x(C) \leq |C| - 1, \\
\text{Maximise} & \quad \sum_{e \in E(G)} d_e x_e.
\end{align*}$$

If we replace constraints $(\mathcal{P}_1)$ by nonnegativity constraints:

$$0 \leq x_e \leq 1 \quad \text{for all } e \in E(G),$$

we obtain a linear program $(\mathcal{P})$ which is the continuous relaxation of $(\mathcal{P}_1)$. We state now the main result of this section:

Theorem 22. $(\mathcal{P})$ can be solved in polynomial time.

Proof. Recall the main fundamental result: the ellipsoid method [10] solves $(\mathcal{P})$ in polynomial time provided that the following separation problem is solvable in polynomial time: given $x \in \mathbb{R}^{E(G)}$, decide if $x$ satisfies the constraints of $(\mathcal{P})$ or find a constraint of $(\mathcal{P})$ which is violated by $x$. As the number of nonnegativity constraints is polynomial we only have to consider the minimal dependent sets constraints (whose number may be exponential). So assume that for a given $x$, there exists a minimal dependent set $C$ of $G$ such that:

$$x(C) = \sum_{e \in C} x_e > |C| - 1.$$

If we set $c(e) = 1 - x(e)$, the preceding inequality is equivalent to

$$\sum_{e \in C} (1 - x_e) = \sum_{e \in C} c(e) < 1.$$

So our problem reduces to the following problem: does there exist a minimal dependent set with weight strictly smaller than 1?

But to answer this question we need to find the minimal dependent set of minimum cost and this can be done in polynomial time by Theorem 21. \[\square\]

Consider the following auxiliary graph: $A(G) = (E(G), L)$ where $(e, f) \in L$ if and only if the set $\{e, f\}$ is a dependent set of $G$. If $K$ is a clique of the graph $A(G)$, any biclique contains at most one edge of $K$ in $G$ and thus we can add to the linear program $(\mathcal{P})$ the following family of constraints:

$$x(K) \leq 1 \quad \text{for all } K \in \mathcal{K}.$$

\[\square\]
\( \mathcal{K} \) is the set of cliques of \( A(G) \). But we have a negative result similar to the result for the weighted stable set problem where we cannot separate in polynomial time the clique constraints:

**Proposition 23.** The separation problem for constraints \((\gamma)\) is NP-complete.

**Proof.** Associate to a graph \( G \) the following bipartite graph \( B(G) \): each node \( v \) of \( G \) is replaced by two copies \( v' \in V' \) and \( v'' \in V'' \) of \( B(G) \). The two copies \( v' \) and \( v'' \) are linked by an edge \( e_v \) in \( B(G) \) called vertical edge. Moreover \( (v, w) \notin E(G) \) if and only if \( (v', w'') \) and \( (v'', w') \) are edges of \( B(G) \) (these two edges will be called transversal edges) and \( (v, w) \in E(G) \) if and only if \( (v', w'') \) and \( (v'', w') \) are not edges of \( B(G) \). Two vertical edges \( e_v, e_w \) induce a dependent set in \( B(G) \) if and only if \( (v, w) \in E \). Assume that in our original graph we have weights \( c_v \) assigned to the nodes \( v \in V \). Assign the weight \( c_v \) to the vertical edge \( e_v \), \( \forall v \in V \), and assign the weight 0 to all the transversal edges of \( B(G) \). Let \( \overline{K} \) be the solution of the following problem: find the subset of edges of \( B(G) \) such that any pair of elements of \( K \) is a dependent set and with maximum weight. The separation problem for constraints \((\gamma)\) is clearly equivalent to this maximization problem. But transversal edges have weight equal to 0 and we can assume that all the edges of \( \overline{K} \) are vertical edges. The set of nodes \( v \) of \( V \) such that \( e_v \in \overline{K} \) induces the maximum weighted clique of \( G \) and the maximum weighted clique problem is NP-hard. □

### 7.2. The minimum biclique cover problem

We define now a linear programming relaxation of the MBC problem similar to the formulation of the minimum coloration problem for the nodes of a graph. We can always assume that the maximum number of colors is known and equal to \( k \) (for instance, \( k = |E(G)| \)).

Let \( P(G) \) be the polytope defined by the nonnegativity constraints and the minimal dependent sets constraints. We assign to each color \( i \) a variable \( y^i \) (\( y^i = 1 \) (resp. \( y^i = 0 \)) means that color \( i \) is used (resp. not used) in the coloration of \( G \)).

Consider \( k \) vectors \( x^1, \ldots, x^k \) of dimension \( |E(G)| \). The constraints of our linear program are:

- \( x^i \in P(G) \) for \( i = 1, \ldots, k \).
- \( x^i_e \leq y^i \) for all \( e \in E \) and all \( i = 1, \ldots, k \).
- \( \sum_{i=1}^k x^i_e = 1 \) for all \( e \in E(G) \).

The objective function is: minimize \( \sum_{i=1}^k y^i \).

**Acknowledgments**

The authors wish to thank an anonymous referee and Julian Webster for their helpful suggestions.

**References**


Further reading