Chromatic characterization of biclique cover

Denis Cornaz

France Telecom R&D, 38-40 rue du Général Leclerc, Issy-les-Moulineaux 92794

Jean Fonlupt

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

Abstract

A biclique B of a simple graph G is the edge-set of a complete bipartite (not necessarily induced) 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 of G: $F \subseteq E$ will be called independent if there exists a biclique $B \subseteq E$ 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.

Key words: biclique, bipartite, chromatic number χ, clique number ω.

1 Introduction

Let $G = (V, E)$ be a simple graph. A subset B of E will be called a *biclique* if B is the edge-set of a complete bipartite (not necessarily induced) subgraph of G.

Email addresses: `denis.cornaz@francetelecom.com` (Denis Cornaz), `Jean.Fonlupt@ecp6.jussieu.fr` (Jean Fonlupt).

A *biclique cover* $\mathcal{B}$ of G is a collection of bicliques covering the edge-set of G (every edge of G belongs to at least one biclique of the collection). The *bipartite dimension* $d(G)$ of G is the cardinality of a minimum biclique cover of G. The minimum biclique cover problem (MBC) is the problem of determining $d(G)$ 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 O. Günlük [10] gives an application concerning the maximum multicommodity flow problem and the min-cut max-flow ratio for multicommodity flow. Fishburn and Hammer [6] show that $d(G)$ equals the boolean interval dimension of the complementary graph of G. When G is bipartite, $d(G)$ is equal to the boolean rank of its adjacency matrix (see [13]). Monson et al. [13] provide a survey of known results and connections with factorizations of $\{0, 1\}$ matrices.

J. Orlin [15] shows that the calculation of $d(G)$ is NP-hard for general bipartite graphs and H. Müller [14] shows that it is NP-hard even for chordal bipartite graphs. However, it is possible to check whether $d(G) = k$ or not for $k = 1, 2$ by verifying that G does not contain some graphs as induced subgraphs (see [6]). Thus the decision problem $d(G) = k$ 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 [2] we can find references asserting that the problem is also polynomial for the class of bipartite convex graphs.

J.-C. 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 $n \leq 10$. F.R.K Chung [4] proved the following result conjectured by Bermond: $b(n)/n \to 1$. Z. Tuza [18] proved that $b(n) \leq n - \lfloor \log_2 n \rfloor + 1$ for all n. There are also other tight bounds: $\lceil \log_2 \chi(G) \rceil \leq d(G) \leq \tau(G)$. In the left inequality, proved by Harary et al. [11], $\chi(G)$ is the chromatic number of G. This inequality is an equality if G is a complete or a complete multipartite graph (see [3,12]). In the right inequality, $\tau(G)$ is equal to the minimum number of stars which 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 of the maximum clique of G_E is a tight lower bound for $d(G)$ (see [13]). Fishburn and Hammer [6] gave a better bound: $\chi(G_E) \leq d(G)$. 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 [17]).

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 [6].

This paper is organized as follows: in the next section we give more notation and definitions and present some preliminary results. A subset $F \subseteq E$ is independent if F is contained in a biclique. A subset of E which is not independent will be called dependent. In Section 3 we give a complete characterization of independent sets. In Section 4 we study the relation between the biclique covering problem and coloration problems of graphs and hypergraphs. We introduce the notion of the r-coloration problem of the edge-set of a graph, for any positive integer r; 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 $\chi_r(G) < d(G)$. 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 $\chi_r(G) = d(G)$. 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

2.1 Basic definitions

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

b) We note by $\overline{E}$ the edge-set of the complement graph of G: $(v, w) \in E$ if and only if $(v, w) \notin \overline{E}$. E will be also called the set of red edges and $\overline{E}$ will be the set of black edges. In the rest of this paper we will study subgraphs $H = (W, F \cup \overline{F})$ of the complete graph $(V, E \cup \overline{E})$ where $F \subseteq E$ is a set of red edges and $\overline{F} \subseteq \overline{E}$ is a set of black edges. Such a subgraph will be called a painted graph.

Definition 2.1 *Let $F \subseteq E$. The rooted graph of F is the graph $H = (W, F \cup \overline{F})$ defined as follows:*

(1) W is the set of nodes incident to at least one edge of F,
(2) $\overline{F}$ is the set of edges of the complement graph of G having their two endnodes in W.

Given a painted graph H we denote by D the oriented graph obtained by replacing each edge $e = (v, w)$ of H by two arcs (oriented edges) $a_1 = (v, w)$ and $a_2 = (w, v)$. a_1 and a_2 are the images of e. They are red (resp. black) arcs if e is a red (resp. black) edge. If $a = (v, w)$ is an arc, v is the head of a and w is the tail of a. In an oriented graph, w (resp. v) is a successor of v (resp. predecessor) of v (resp. w).

2.2 Red-odd cycles

In a graph (resp. a directed graph) a walk is a sequence of nodes $P = v_1, v_2 ... v_k, v_{k+1}$ such that $k \geq 1$ and $e_i = (v_i, v_{i+1})$ is an edge (resp. an arc) of the graph for $i = 1, 2, ..., k$.
$e_1, e_2 ... e_k$ is the edge-sequence of P; k is the length of P. P is odd (resp. even) if k is odd (resp. even). Note that if P is a walk in H, P is also a walk in D and vice versa. If all the nodes of P are distinct, P is a path linking v_1 and v_{k+1} if the graph is non oriented, and linking v_1 to v_{k+1} in the case of directed graphs. If $v_1 = v_{k+1}$, P is a closed walk. If P is a closed walk and the nodes of the sequence $C = v_1, v_2 ... v_k$ are all distinct, C is a directed cycle in the case of directed graphs. In the case of graphs, C is a cycle provided that $k \geq 3$. $e_1, e_2 ... e_k$ is also the edge-sequence of the (directed) cycle C and $\{e_1, e_2 ... e_k\}$ is the edge-set 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.

Definition 2.2 *A closed walk or a cycle, (resp. a directed cycle) of a painted graph is red-odd (resp. red-even) if its edge-sequence contains an odd (resp. even) number of red edges.*

We will establish now the following easy preliminary lemma:

Lemma 2.1 *A painted graph contains a red-odd hole if and only if it contains a red-odd closed walk.*

Proof. If $v_1, v_2 ... v_k$ is a red-odd hole, $v_1, v_2 ... v_k, v_1$ is a red-odd closed walk. Suppose now that H contains a red-odd closed walk. In the graph D obtained by orienting H, the edge-sequence of this closed walk can be partitioned into edge-sequences of directed cycles. So, at least one of this directed cycle $C = v_1, v_2 ... v_k$ is red-odd. If $k = 2$, the two oriented edges of C are the images of the same edge of H and they are both red or both black, which is impossible if C is red-odd. Thus $k \geq 3$ and C is a cycle in H. Now, we can assume that C is the smallest red-odd cycle of H. If f is a chord of C, there exists a partition (E_1, E_2) of the edge-set of C such that $E_1 \cup f$ and $E_2 \cup f$ are the edge-sets of two cycles of smaller size C; but one of these two cycles is red-odd contradicting our assumption on the size of C. $\square$

Let H' be the graph obtained by replacing each black edge (v, w) of H by a chain of length 2: v, z, w. (the two edges of this chain will be also black edges). We can now state the following lemma:

Lemma 2.2 *H contains no red-odd closed walk if and only if H' is bipartite.*

Proof. Replace in a hole C of H each pair v, w of nodes linked by a black edge by a chain v, z, w. The sequence obtained after this transformation is a cycle C' of H'; the converse is also true: If C' is a cycle of H', there exists a cycle C such that C' is obtained from C by this transformation. The number of black edges of C' is even; hence C is red-odd if and only if C' is odd; by the preceding lemma and the definition of bipartite graphs, our lemma is proved. $\square$

3 Independent sets

Recall that $F \subseteq E$ is called independent if F is contained in a biclique B of G. Otherwise F is dependent. Let H be the rooted graph of a subset of red edges F and H' be the graph obtained by replacing each black edge of H by a chain of length 2.

Consider now one of the following rules for labelling the node-set of a graph with two labels $(+)$ and $(-)$.

* **Rule A:** The labels of the two endnodes of an edge of the graph are distinct.

* **Rule B:** The labels of the two endnodes of an edge of the graph are distinct if this edge is red and identical if this edge is black.

A graph is well labelled by **rule A** (resp. **rule B**) if we can assign one of the two labels to each node of the graph so that all the pair of adjacent nodes of the graph satisfy this rule. We can now state the main result of this section:

Theorem 3.1 *The following statements are equivalent,*

(a) F is independent,

(b) H contains no red-odd closed walk,

(c) H' is bipartite,

(d) H' is well labelled by rule A,

(e) H is well labelled by rule B.

Proof. (b)$\Leftrightarrow$ (c) by Lemma (2.2).
(c)$\Leftrightarrow$ (d) is a well known fact.
(d)$\Rightarrow$ (e): If H' is well labelled by rule A, two endnodes of a red edge of H have distinct labels and satisfy rule B; if (v, w) is a black edge of H, when we replace (v, w) by a chain v, z, w in H', rule A applied to this chain in H' implies that v and w have identical labels. So the pair v, w satisfies rule B in H.
(e)$\Rightarrow$ (d): The proof is similar to the previous case.
(a)$\Rightarrow$ (e): If F is independent, F is included in a biclique B and there exists a complete bipartite subgraph $(W_1 \cup W_2, B)$ of G. Assign label $(+)$ to the nodes of W_1 and label $(-)$ to the nodes of W_2. If (v, w) is a red edge of H, v and w belong to distinct sides of the bipartition and have distinct labels. If (v, w) is a black edge of H, $(v, w) \notin E$. Hence, v and w belong to the same side of the bipartition and have identical labels. This proves (e).
(e)$\Rightarrow$ (a): Consider a labelling of H which satisfies rule B and let W_1 (resp. W_2) be the set of nodes labelled with $(+)$ (resp. $(-)$) in this labelling. Let B be the set of edges of G with one endnode in W_1 and one endnode in W_2. $F \subseteq B$; rule B implies that no red edge of H links two nodes which belong both to W_1 (resp. W_2); hence the graph $(W_1 \cup W_2, B)$ is bipartite. Again by rule B if $v \in W_1$ and $w \in W_2$, $(v, w)$is not a black edge and $(v, w) \in B$. Thus B is a biclique. □

From an algorithmic point of view, we can implement the following procedure: Assign first the label $(+)$ to some node in each connected component of H; then label successively the other nodes of H in such a way that rule B is always satisfied by the previously labelled nodes. Stop when all the nodes of H are labelled in a unique way or when a node receives two distinct labels. This algorithm is implementable in $O(m)$ elementary operations where m is the number of edges of H. As $m \leq |V|^2$ we can state the following result:

Corollary 3.1 *The problem of deciding if a subset of edges of a graph is independent is polynomial.*

Assume now that $F \subseteq E$ is independent and let H be the rooted graph of F. As H is well labelled by rule B, the set of edges linking two nodes of W with distinct labels is red and is a biclique containing F. So we have:

Corollary 3.2 *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 interested by 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 4.1 *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 $F \subseteq E$ such that $|F| \leq r$ belongs to at least two classes of colors.*

In a strong coloration all the classes of colors are independent sets. Thus the minimum number of colors in a strong coloration is also the minimum number $d(G)$ of bicliques which cover the set E. In a r-coloration, every subset F of a class of color with $|F| \leq r$ is independent. The r-chromatic number $\chi_r(G)$ of G is the minimum number of colors in a r-coloration. Clearly a strong coloration is a r-coloration $\forall r > 0$. Hence $\chi_r(G) \leq d(G)$. Note also that if $r < s$, a 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 [6]) for a graph G without isolated nodes and with nonempty edge-set. The main result of Fishburn and Hammer is that if $\omega(G) \leq 2$, then $\chi_2(G) = d(G)$ ($\omega(G)$ is the size of the largest clique of G). This result is obviously false if $\omega(G) > 2$ and a trivial counterexample to this statement is obtained when G is the complete graph on three nodes K_3: $\chi_2(K_3) = 1$ but $d(K_3) = 2$. So, if $\omega(G) > 2$, it may be possible that $\chi_2(G) = d(G)$ provided that we consider 2-colorations with the additional condition: no class of color contains

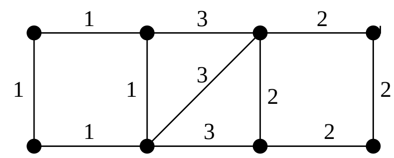

Fig. 1.

a triangle K_3. But Fishburn and Hammer gave also an example depicted in Figure 1 showing that this statement is also false: the three colors associated to the edges of the graph of the picture 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 $d(G) = 4$; however no class of color contains a triangle and the additional condition proposed by Fishburn and Hammer is satisfied. In their paper, these two authors asked if there exists some positive integer r for which $\chi_r(G) = d(G)$ with eventually some other conditions on the r-coloration. We will now prove that this assertion is false; note that this is result is negative since it implies that there is no way to reduce the minimum biclique covering problem to a coloration problem on graphs or hypergraphs. Consider the following example: Here, G is the graph K_5. Let F be the edge-

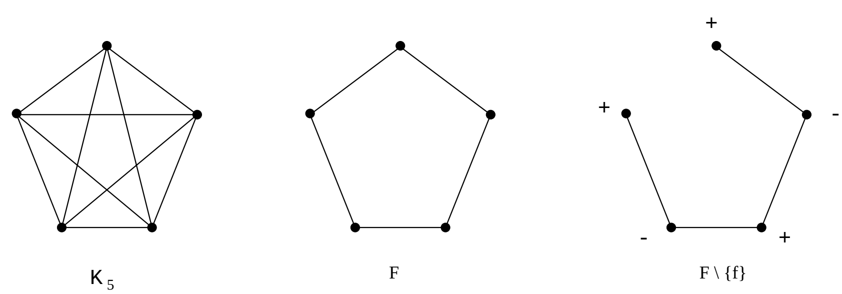

Fig. 2.

set of a cycle of size 5 of K_5. If we delete F from the edge-set of K_5 we obtain an other cycle F' of size 5. Delete one edge f from F. Note that here the set of black edges is empty. Thus the rooted graph H of F is an odd cycle of size 5 and F is not independent by Theorem (3.1) (see Figure 2). But the rooted graph H_1 of $F \setminus \{f\}$ is a path and therefore $F \setminus \{f\}$ is independent. Thus if we assign color 1 (resp. 2) to F (resp. F') we obtain a 4-coloration of K_5 and

$\chi_4(K_5) = 2$. But $d(K_5) = 3$. Thus we can make the following remark:

Remark 1 $\chi_4(K_5) < d(K_5)$.

Theorem 4.1 *For every integer r, there exits a graph G such that $\chi_r(G) < d(G)$.*

Proof. Consider first the complete graph $K_{2r+1} = (V, E)$ and let $F \subset E$ 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 $e = (u, v) \in E \setminus F$ a path u, x, y, v. (x, y) will be called the copy of $e = (u, v)$ and $\overline{E}$ will be the set of all the copies ($|\overline{E}| = |E \setminus F|$). Finally we note 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 $f \in \overline{E}$ and an edge of F form also a dependent set. As F is dependent but any subset of F is independent we need two colors for coloring the set F in a strong coloration but only one color for coloring F in a $2r$-coloration. From these observations we easily deduce that: $\chi_{2r}(G) = 1 + |\overline{E}|$ and $d(G) = 2 + |\overline{E}|$; this proves our theorem. □

5 Minimal Dependent Sets

Observe that since a single edge is an independent set, every dependent set with cardinality 2 is minimal. We will now characterize the minimal dependent sets but first we need more definitions.

5.1 Obstructions

Let $V_1 \subseteq V$. For each $v \in V_1$, choose a red edge f_v incident to v. We will make the convention (for avoiding heavy notations in our presentation), that if $f_v = f_w$ for two nodes v, w of V_1 (which may occur if v and w are adjacent and $f_v = f_w = (v, w)$), we will create a double edge ${f_v}^1$ and ${f_w}^2$ with endnodes v and w so that all the edges chosen will be distinct. We will set $\delta(V_1) = \{f_v \quad |v \in V_1\}$; $\delta(V_1)$ will be called a rooting set for V_1.
Note that without our convention, $\delta(V_1)$ should be defined as a family rather than a set.

Lemma 5.1 *Let $\delta(V_1)$ be a rooting set for V_1.*

(a) $|\delta(V_1)| = |V_1|$.
(b) If $V_2 \subset V_1$, there exists a rooting set $\delta(V_2)$ for V_2 with $\delta(V_2) \subset \delta(V_1)$.

Proof. (a) is a consequence of our definition of rooting sets. To prove (b), take for rooting set of V_2 the set: $\delta(V_2) = \{f_v | f_v \in \delta(V_1); v \in V_2\}$. $\square$

Lemma 5.2 *Let F be a set of red edges and let W be the set of nodes incident to at least one edge of F. If $V_2 \subset W$, there exists a rooting set $\delta(V_2)$ with the property that each element of $\delta(V_2)$ is an edge of F or is a copy of an edge of F.*

Proof. Take $v \in V_2$ and choose an edge $f_v = (v, w) \in F$. Let $F_1 = F \setminus \{f_v\}$. If $w \notin V_2$ let $V_3 = V_2 \setminus \{v\}$; if $w \in V_2$ let $V_3 = V_2 \setminus \{v, w\}$. By induction on the size of V_2, there exists a rooting set $\delta(V_3) \subset F_1$. $\delta(V_2)$ is now obtained by adding f_v to $\delta(V_3)$, if $w \notin V_2$, and two copies ${f_v}^1$ and ${f_w}^2$ of f_v if $w \in V_2$. $\square$

A node v of a cycle C of $(V, E \cup \overline{E})$ will be called exposed if the two edges of C incident to v are black.

Definition 5.1 *We say that a set of red edges F_1 and a rooting set $\delta(V_2)$ for some $V_2 \in V$ induce an obstruction $\mathcal{O}$ if there exists a red-odd cycle C of the complete graph $(V, E \cup \overline{E})$ such that:*

- *F_1 is the set of red edges of C,*
- *V_2 is the set of exposed nodes of C.*

We will write $\mathcal{O} = (F_1, \delta(V_2))$. The size of the obstruction $\mathcal{O}$ is $|F_1| + |\delta(V_2)|$; by Lemma (5.1) the size of $\mathcal{O}$ is also $|F_1| + |V_2|$.
We note by $\mathcal{O}(G)$ the set of obstructions of G. Note that the cycle C that we will call the cycle of the obstruction is well and uniquely defined; note also that if there is no duplicated edge in $\delta(V_2)$, $F_2 = \delta(V_2)$ is a subset of red edges and in this case we will write $\mathcal{O} = (F_1, F_2)$. Such a situation will occur in the case of strong obstructions that we now define:

Definition 5.2 *An obstruction is a strong obstruction if no chord of the cycle of the obstruction belongs to $\delta(V_2)$. An obstruction which is not a strong obstruction is a weak obstruction.*

An obstruction $\mathcal{O} = (F_1, \delta(V_2))$ strictly contains an obstruction $\mathcal{O}' = (F_1', \delta(V_2'))$ if: $F_1' \cup \delta(V_2') \subset F_1 \cup \delta(V_2)$. We have now the following proposition:

Proposition 5.1 *A weak obstruction always strictly contains a strong obstruction.*

Proof. If this statement is false, there exists a weak obstruction $\mathcal{O} = (F_1, \delta(V_2))$ containing no other obstruction. Let $C = v_1, v_2 ... v_k$ be the cycle of $\mathcal{O}$. By defi-

nition of a weak obstruction there exists a chord $e = (v, w)$ of C which belongs to $\delta(V_2)$. We can assume that $v = v_1$ and $w = v_r$ with $2 < r < k$. We can also assume that the cycle $C' = v_1, v_2 ... v_r$ is a red-odd cycle and the cycle $C'' = v_1, v_r ... v_k$ is a red-even cycle. Let F_1' the set of red edges of C' and V_2' be the set of exposed nodes of C'. F_1' is obtained by adding first e to F and then by subtracting the red edges of C which belong to C''. But at least one such edge exists since C'' is red-even. Thus $|F'| \leq |F|$.
$V_2' \subseteq V_2$ but $(v, w) \in \delta(V_2)$ and at least one of the two nodes v,w is exposed in C and is not exposed in C'. So, $V_2' \subset V_2$ and by Lemma (5.1) there exists a rooting set $\delta(V_2') \subset \delta(V_2)$. Thus the obstruction $\mathcal{O}' = (F_1', \delta(V_2'))$ is included in $\mathcal{O}$, and as the size of $\mathcal{O}'$ is strictly smaller than the size of $\mathcal{O}$, the inclusion is strict, a contradiction. □

5.2 Obstructions and Minimal Dependent Sets

Proposition 5.2 *The two following implications are true:*

(a) If $\mathcal{O} = (F_1, F_2)$ is a strong obstruction, $F = F_1 \cup F_2$ is a dependent set.
(b) If F is a minimal dependent set, there exists a partition F_1, F_2 of F such that $\mathcal{O} = (F_1, F_2)$ is a strong obstruction.

Proof. (a) Let C be the cycle of $\mathcal{O}$. C is a red-odd cycle contained in the rooted graph H of F. The result follows by Theorem (3.1).
(b) As F is a dependent set, the rooted graph H of F contains a red-odd cycle C by Theorem (3.1). Let F_1 be the set of red edges of C and V_2 be the set of exposed nodes of C. By Lemma (5.2) we can construct a rooting set $\delta(V_2)$ and each element of this set is an element of F or a copy of an element of F (in case of duplication). The obstruction $\mathcal{O} = (F_1, \delta(V_2))$ contains a strong obstruction; by the minimality of F and part (a) of this theorem, $\mathcal{O}$ is a strong obstruction and $F = F_1 \cup F_2$ where $F_2 = \delta(V_2)$. □

Remark 2 *Let $\mathcal{O} = (F_1, F_2)$ be a strong obstruction and C be is its cycle. Let H be the rooted graph of $F = F_1 \cup F_2$. The nodes of H which are not nodes of C will be called external nodes of C. The following observations are immediate consequences of the definition of strong obstructions:*

(1) No chord of C is a red edge of H.
(2) If v is a non-exposed node of C, v is incident to one or two red edges of H and these edges belong to the edge-set of C.
(3) If v is an exposed node of C, v is incident to exactly one red edge (v, w) of H, where w is an external node of C.
(4) No red edge of H links two external nodes.

The rest of this section is devoted to the complete characterization of minimal dependent sets. A minimal dependent set is a strong obstruction with some additional properties that we are going to prove. But, first it is important to remark that there is not a unique way to associate a strong obstruction $\mathcal{O} = (F_1, F_2)$ to a minimal dependent set. In the example depicted in Figure 3, there are two ways to associate a strong obstruction to $F = \{e_1, e_2, e_3\}$ (where $e_1 = (v_1, v_2), e_2 = (v_3, v_5), e_3 = (v_4, v_6)$).

(1) $F_1 = \{e_1\}$; $C = v_1, v_2, v_3, v_4$; $F_2 = \{e_2, e_3\}$ or
(2) $F_1 = \{e_1, e_2, e_3\}$; $C = v_1, v_2, v_3, v_5, v_6, v_4$; $F_2 = \emptyset$.

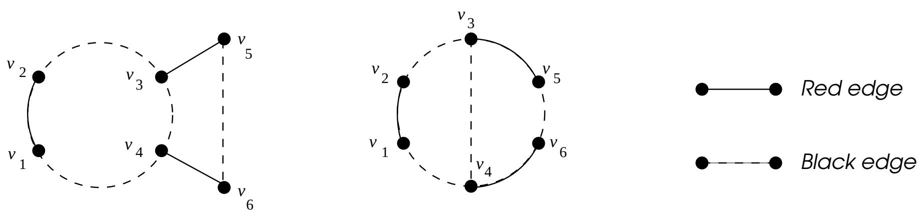

Fig. 3.

Main assumption: We will choose always a decomposition (F_1, F_2) of F for which $|F_1|$ is as large as possible.

For the rest of this section, $C = v_1, ..., v_i, ..., v_j, ..., v_k$ will be the cycle of the obstruction (F_1, F_2). We will set $v_{k+i} = v_i$ and $v_{-i} = v_{k-i}$ for $i = 0, 1, 2$. Recall that H is the rooted graph of the obstruction.
Let P be a path in H with endnodes v_i, v_j with $(1 \leq i < j \leq k)$ such that all the nodes of P except its two endnodes are distinct external nodes of C. Let P_1 be the subsequence $v_i, ..., v_j$ of C. The set of nodes which belong to $P \cup P_1$ (resp. $P \cup (C \setminus P_1)$) induces a cycle C_1 (resp. C_2). One of the two cycle (for instance C_1) is red-even and the other C_2 is red-odd; this implies also that the parity of the number of red edges of P and P_1 is the same.

First, we will prove the following preliminary lemma:

Lemma 5.3 *Assume that $F = (F_1, F_2)$ is a minimal dependent set.*

a) If f is a red edge of P_1, f is incident to v_i or v_j.
b) Let $v \in P_1$ be an exposed node and $g = (v, w)$ be the unique red edge of the obstruction incident to v; $w \in P$.

c) v_i (resp. v_j) is incident to exactly one red edge.

Proof. If a) is false, the red-odd cycle C_2 belongs to the rooted graph of $F \setminus \{f\}$ which is not an independent set by Theorem (3.1) contradicting the definition of a minimal dependent set.
b) By definition of a strong obstruction, $w \notin P_2$; if $w \notin P$, C_2 belongs to the rooted graph of $F \setminus \{g\}$ and the same conclusion as in (a) holds.
c) If c) is false, there exists a red edge f incident to v_i and distinct from (v_{i-1}, v_i). Again C_2 belongs to the rooted graph of $F \setminus \{f\}$ and the same conclusion as in (a) holds. □

We need now to introduce some definitions illustrated on Figure 4:

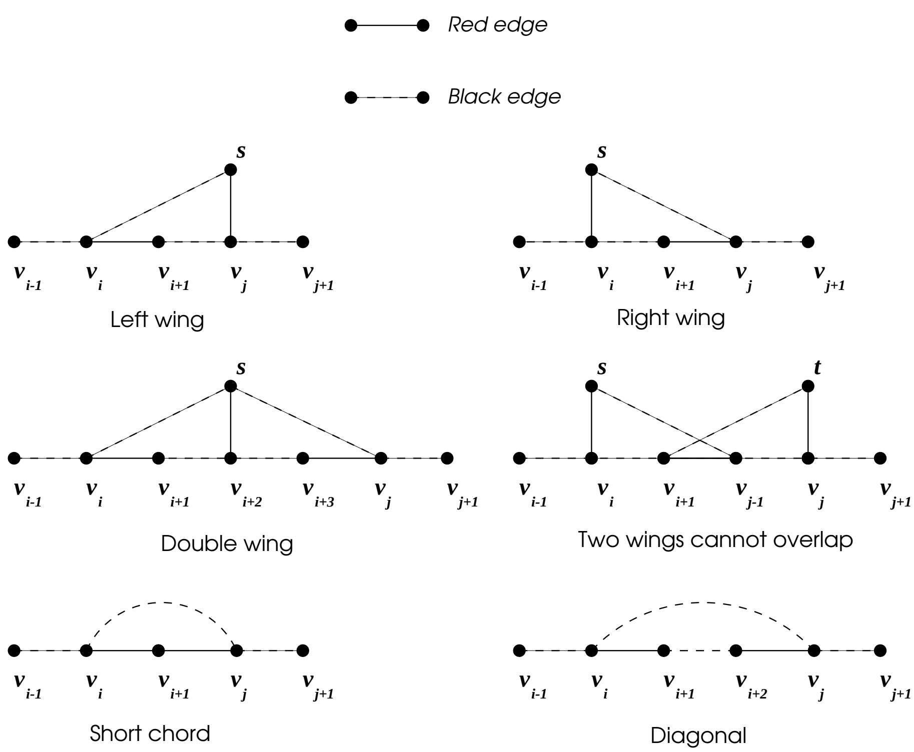

Fig. 4.

Definition 5.3 *Let (v_i, v_j) be a chord of the cycle C of the obstruction.*

(1) (v_i, v_j) will be called a short chord if $j = i + 2$, (v_i, v_{i+1}), (v_{j-1}, v_j) are red edges and (v_{i-1}, v_i), (v_j, v_{j+1}) are black edges.

(2) (v_i, v_j) *will be called a diagonal if* $j = i + 3$, (v_i, v_{i+1}), (v_{j-1}, v_j) *are red edges and* (v_{i-1}, v_i), (v_{i+1}, v_{i+2}), (v_j, v_{j+1}) *are black edges.*

Definition 5.4 *A left (resp. a right) wing is a subgraph of the rooted graph of* F *induced on four nodes* $W = \{v_i, v_{i+1}, v_{i+2}, s\}$ *such that:*

(1) s *is an external node of the obstruction and* v_i, v_{i+1}, v_{i+2} *is a chordless path of* C *with a red (resp. a black) edge* (v_i, v_{i+1}) *and a black (resp. a red) edge* (v_{i+1}, v_{i+2}).
(2) s *is adjacent to the exposed node* v_{i+2} *(resp.* v_i*) by a red edge and to* v_i *(resp.* v_{i+2}*) by a black edge.*
(3) (v_{i-1}, v_i) *and* (v_{i+2}, v_{i+3}) *are black edges.*

Note that s may belong to a left and a right wing. In this case we say that the two wings form a double wing.

Definition 5.5 *A pair of wings overlaps if there exists an edge of the cycle* C *which belongs to these two wings.*

We can now state the following proposition.

Proposition 5.3 *Assume that* $F = (F_1, F_2)$ *is a minimal dependent set.*

(1) If e *is a chord of* C, e *is a short chord or a diagonal.*
(2) If s *is an external node incident to the red edge.* (s, v_i) *where* $v_i \in C$, *the two other possible neighbors of* s *are* v_{i-2} *and* v_{i+2}. *If* v_{i-2} *(resp.*v_{i+2}*) is adjacent to* s, s *belongs to a left (resp. right) wing.*
(3) No pair of wings overlaps.

Proof. 1) If $e = (v_i, v_j)$ is a chord, we can apply Lemma (5.3) to the path $P = v_i, v_j$. As P has no red edge, P_1 has an even number of red edges. By Lemma (5.3.b), P_1 contains no exposed node; therefore P_1 contains at least one red edge. Hence Lemma (5.3.a) implies that the two red edges of P_1 are (v_i, v_{i+1}) and (v_{j-1}, v_j). By Lemma (5.3.c), (v_{i-1}, v_i) and (v_j, v_{j+1}) are black edges. Finally, as P_1 has no exposed node, either $i = j - 2$ and e is a short chord or $i = j - 3$ and e is a diagonal.
2) Assume first that two external nodes s, t are adjacent. By definition of a strong obstruction, there exists two red edges (s, v_i) and (t, v_j) where v_i and v_j are distinct exposed nodes of C. let P be the path $P = v_i, s, t, v_j$. Lemma (5.3.a) implies that P_1 has no red edge. But the red-odd cycle C_2 has more red edges than C which is impossible by our main assumption.
Assume now that there exist two exposed nodes v_i and v_j of C such that (s, v_i) and (s, v_j) are red edges, let P be the path $P = v_i, s, v_j$. The proof is now similar to the proof of the previous case and again this is impossible.

So we can assume that (s, v_i) is the unique red edge incident to s. If there exists a black edge (s, w) in H, $w \in C$ by the preceding result and we can suppose that $w = v_j$ for some $j \neq i$. Assume first that $j > i$ and let P be the path $P = v_i, s, v_j$. P contains one red edge and therefore P_1 contains an odd number of red edges. By Lemma (5.3.a), (v_{j-1}, v_j) is the unique red edge of P_1. By Lemma (5.3.c), (v_j, v_{j+1}) is a black edge. Moreover v_i is the unique exposed node of P_1 by Lemma (5.3.b). Thus $j = i + 2$ and the set $\{v_i, v_{i+1}, v_{i+2}, s\}$ induces a right wing. The case $j < i$ is similar: $j = i - 2$ and $\{v_{i-2}, v_{i-1}, v_i, s\}$ induces a left wing. Finally the only possible neighbors of s besides v_i are v_{i-2} and v_{i+2}; if s is adjacent to these three nodes, s belongs to a double wing.
3) Assume by contradiction that there exists two external nodes s, t and two wings induced on the sets $\{v_i, v_{i+1}, v_{i+2}, s\}$ and $\{v_{j-2}, v_{j-1}, v_j, t\}$ which overlap. The only possibility to overlap (see Figure 4) is that the first wing is a left wing, the second a right wing and $i + 1 = j - 2$. The path $Q = v_i, s, v_{i+2}, v_{i+1}, t, v_j$ has three red edges and the path $Q' = v_i v_{i+1} v_{i+2} v_{i+3}$ has one red edge. The cycle obtained by replacing Q' by Q in the sequence describing C is a red odd cycle and has a larger number of red edges than C which is impossible by our main assumption. □

The following theorem states that the converse of this proposition is also true:

Theorem 5.1 *A set F is minimal dependent if and only if it induces an obstruction which satisfies properties (1), (2), (3) of proposition 5.3.*

Proof. Let $f \in F$ and H_1 be the rooted graph of $F \setminus \{f\}$. If f is an edge of C, C cannot be a subgraph of H_1. If f links an external node to an exposed node v of C, v is not a node of H_1 and again C is not a subgraph of H_1.
Assume by contradiction that $F \setminus \{f\}$ is not independent. By Lemma (3.1) H_1 contains a red-odd closed walk D. Assume first that D contains a diagonal $e = (v_i, v_{i+3})$; f cannot be (v_i, v_{i+1}) or (v_{i+2}, v_{i+3}) since v_i and v_{i+2} have not been deleted from H. If we replace in D the subsequence $..., v_i, v_{i+3}, ...$ by $..., v_i, v_{i+1}, v_{i+2}, v_{i+3}, ...$ we get a new closed path which does not contain e and which is still red-odd. We can eliminate reiterating this process all the diagonals and similarly all the short chords. If D contains a black edge (v_{i-2}, s) linking a node v_{i-2} of the cycle to an external node s, s and v_{i-2} belong to a wing, for instance a left wing $\{v_{i-2}, v_{i-1}, v_i, s\}$ of H and these four nodes belong also to H_1. Replace in D the subsequence $..., v_{i-2}, s, ...$ by $..., v_{i-2}, v_{i-1}, v_i, s, ...$ Again the new closed walk is red-odd and we can eliminate all the black edges which are not edges of C. If an external node s belongs to D after all these elimination steps, we have a subsequence of D which is : $..., v_i, s, v_i, ...$ where v_i is the exposed node adjacent to s. But we can delete from the preceding sequence the subsequence s, v_i and reiterating this process, we can eliminate

all the external nodes from our red-odd closed walk. Finally we can assume that D contains only edges of the cycle C. By Lemma (2.1), D contains a red-odd cycle included in C and strictly included in C since C is not a subgraph of H_1 which is impossible since any proper subgraph of C is bipartite. □

In the next figure we give three examples of graphs $H = (W, F \cup \overline{F})$, where F is a minimal dependent set.

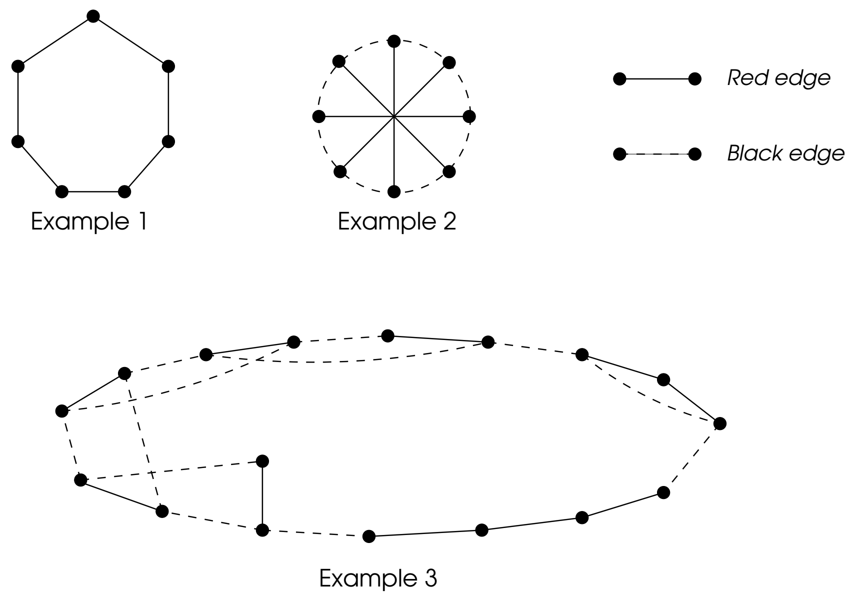

Fig. 5.

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

Proposition 5.4 *Let W be the set of nodes of the rooted graph of F where F is a minimum dependent set. The induced subgraph $G(W)$ contains a clique of size $|F|$.*

Proof. If v_i is an exposed node of C, we will denote by s_i the (unique) external node of H adjacent to v_i.
Consider the subgraph $(W, \overline{F})$ of H. In this graph each node is incident to at most two edges. Each connected component of this graph is an isolated node, a path or a cycle. If $f = (v_i, v_j) \in \overline{F}$ (resp. $f = (v_i, s_j) \in \overline{F}$) orient f from v_i to v_j (resp. s_j) if $i < j$ and in the opposite direction if $i > j$; the properties

of minimal dependent sets imply that in the oriented graph, there exists at most one edge leaving and at most one edge entering each node; each cycle of $(W, \overline{F})$ is transformed into a directed cycle in this orientation. Note now that for each couple of indices $i, i+1$ for $1 \leq i \leq k$, (we can choose $i = 1$), there exists at most two black edges with initial end v_j or s_j with $j \leq 1$ and terminal end v_l or s_l with $l \geq 2$. We say that such an edge crosses the pair $1, 2$. If no edge crosses the pair $1, 2$ our orientation has no directed cycle and $(W, \overline{F})$ has no cycle. If one edge crosses the pair $1, 2$, and if there exists a cycle D in the orientation of $(W, \overline{F})$, consider all the nodes of D which are also nodes of C. If v and w are two such nodes consecutive on C, the subchain of C with endnodes v and w has always an even number of red edges of C (in fact 0 or 2). As the edge-set of C is partitioned into subchains of this type, C has an even number of red edges which is impossible. So, if a cycle D exists in $(W, \overline{F})$, there exists two black edges crossing the pair $i, i+1$ for $1 \leq i \leq k$. But this means that each node of W is incident to two black edges, and that D contains all the nodes of W. But our graph H is now cubic and has an even number of nodes. So, D is even. Therefore $(W, \overline{F})$ is a bipartite graph and there exists in this graph a stable set K with $|K| \geq \frac{|W|}{2} \geq |F|$. But K induces a clique in G, which proves our result. □

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

Theorem 5.2 *$d(G) = \chi_r(G)$ if $\omega(G) \leq r$.*

Proof. Consider an optimal r-coloring of the edge-set of G: If a class of color is not an independent set, it contains a minimal dependent set F and by the preceding proposition, $|F| \leq \omega(G)$. But in a r-coloring all the subset of a class of colour of cardinality $\leq r$ are independent. So this is impossible and each class of color is independent. Thus, $d(G) \leq \chi_r(G)$ and the theorem is proved. □

Recall for the rest of this section that we have assumed in the begining of the paper that E is not a biclique, hence there always exists at least one nonempty dependent set of G.

Lemma 5.4 *Let F_{max} be a minimal dependent set of maximum cardinality and let us set: $\phi(G) = |F_{max}|$.*

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

Proof. From the previous proposition, $\phi(G) \leq \omega(G)$. 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 5.3 *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 $\phi(G) = \omega(G) - 1$ $\omega(G)$ 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 cost

Assume that in the complete graph $(V, E \cup \overline{E})$ a non negative cost $c(e)$ is assigned to red edge e $\forall e \in E$. Define the cost $d(\mathcal{O})$ where $\mathcal{O} = (F_1, \delta(V_2))$ by:

$$d(\mathcal{O}) = \sum_{e \in F_1 \cup \delta(V_2)} c(e).$$

Note that if an edge f is duplicated in the rooting set $\delta(V_2)$, the cost of each duplicated edge will be equal to $c(f)$ and $c(f)$ will be counted twice in the sum.

Problem 1 *Find the obstruction of minimum cost (P_1).*

Proposition 6.1 *The minimal dependent set of minimum cost is a solution of P_1.*

Proof. Let $\mathcal{O} = (F_1, \delta(V_2))$ be a solution of P_1. If $\mathcal{O}$ is not a strong obstruction, there exists by Theorem (5.1) a strong obstruction F included in $\mathcal{O}$ and since the costs are non negative, $d(F) \leq d(\mathcal{O})$ and F is a solution of P_1. If F is not a minimal dependent set, F includes an obstruction F' which is a minimal dependent set and $d(F') \leq d(F) \leq d(\mathcal{O})$; this proves the result. □

Note that by our previous results, we can find in polynomial time a minimal dependent set included in an obstruction if this obstruction is given; so we will concentrate now on problem P_1.

We will define now the following optimization problem P_2: As in the previous situation a non negative cost is assigned to each red edge; a non negative cost $c(v)$ is also assigned to each node v of G. Let C be a red-odd cycle; if F_1 is the set of red edges of C and V_2 is the set of exposed nodes of C, the cost $d(C)$ of C is:

$$d(C) = \sum_{e \in F_1} c(e) + \sum_{v \in V_2} c(v).$$

So this cost can be interpreted in the following way: when we describe the cycle C we pay a cost when we use a red edge but we pay also a cost when we pass through an exposed node.

Problem 2 *Find the red-odd cycle of minimum cost (P_2).*

Proposition 6.2 *The minimal dependent set of minimum cost is solution of P_2.*

Proof. Let f_v be the red edge incident to v of minimum cost and set $c(v) = c(f_v)$ $\forall v \in V$. (By eventually deleting nodes of G we can assume that such an edge always exists $\forall v \in V$). For any subset W of V we will set: $\delta(W) = \{f_v | v \in W\}$. Let C be the red odd-cycle of minimum cost with this system of cost: If F_1 is the set of red edges of C and V_2 is the set of exposed nodes of C, consider the (not necessarily strong) obstruction $\mathcal{O} = (F_1, \delta(V_2))$:

$$d(\mathcal{O}) = \sum_{e \in F_1} c(e) + \sum_{f_v \in \delta(V_2)} c(f_v).$$

Thus,

$$d(\mathcal{O}) = \sum_{e \in F_1} c(e) + \sum_{v \in (V_2)} c(v) = d(C).$$

If $\mathcal{O}' = (F_1', \delta(V_2'))$ is an other obstruction, let C' be the cycle of this obstruction:

$$\sum_{e \in \delta(V_2')} c(e) \geq \sum_{f_v \in \delta(V_2')} c(f_v).$$

Hence $d(\mathcal{O}') \geq d(C') \geq d(C) = d(\mathcal{O})$. Thus, the obstruction of minimum cost is a solution of P_2 and the result follows by the preceding proposition. $\square$

Finally we will introduce a third optimization problem: Let $C = v_1, v_2, ..., v_k, v_{k+1}$ (with $v_{k+1} = v_1$) a closed red-odd closed walk. Describing C from v_1 to v_{k+1} we pay a cost $c(e)$ if $e = (v_i, v_{i+1})$ is a red edge and a cost $c(v)$ if (v_{i-1}, v_i) and $(v_i, v_{i+1}))$ are both black edges (we take the convention that $v_{k+2} = v_1$). The total cost paid is the cost $d(C)$ of C.

Problem 3 *Find the red-odd closed walk of minimum cost (P_3).*

Note that when C is a red-odd cycle the cost of C is the same for problems P_2 and P_3. A second elementary observation is that the cost of C does not depend on the initial node of C: If $C' = v_2, ..., v_k, v_{k+1}, v_1, v_2$, $d(C') = d(C)$. We can now prove the following result:

Proposition 6.3 *The solution of P_2 is solution of P_3.*

Proof. Assume this theorem false and let $C = v_1, v_2, ..., v_k, v_{k+1}$ be the red-odd closed walk of minimum cost with also the minimum number of nodes among all the solutions. As $v_1, v_2, ..., v_k$ is not a cycle, there exists i and j such that $i < j$ and $v_i = v_j$. Moreover as the choice of the first node of the closed walk is indifferent, we can assume that $1 < i < j < k+1$ and also that the subchain $C' = v_i, ...v_j$ of C is red-even. Let us delete $v_{i+1}, ...v_j$ from C and let us set $D = v_1, ...v_{i-1}, v_i, v_{j+1}, .., v_k, v_{k+1}$. D is a red-odd chain shorter than C. Let $f_1 = (v_i, v_{i+1})$, α (resp. β) the cost paid when we pass through v_i (resp. v_j) in C and γ when we pass through v_i in the second chain D. As all the cost are non negative, $d(D) \leq d(C) + \gamma - \alpha - \beta - c(f_1)$. If $\gamma = 0$ $d(D) \leq d(C)$ which is a contradiction with respect to the assumption on the length of C. If $\gamma > 0$ and $\alpha > 0$ $\gamma = \alpha = c(v)$ and again $d(D) \leq d(C)$. The same result holds if $\beta > 0$. Finally if $\gamma = c(v)$ and $\alpha = \beta = 0$ the two edges (v_{i-1}, v_i) and (v_j, v_{j+1}) are black edges but $f_1 = (v_i, v_{i+1})$ is a red edge and $c(f_1) \geq c(v)$. Again $d(D) \leq d(C)$ and we get a contradiction. □

Note that this result is similar (but in a more complicated context) to the well known result for shortest path algorithms: if all the costs of the edges in a network are non negative, the shortest path is the solution of the shortest walk problem and can be found in polynomial time by Dijkstra algorithm.

To finish this section we will prove that P_3 can be reduced to the shortest path problem in a network $\mathcal{N}$ with non negative costs and we can use Dijkstra algorithm. As an alternative proof we will also show that P_3 can be reduced to the shortest path problem in an undirected network $\mathcal{G}$ with (eventually) negative costs but with no absorbing cycle. This problem can be solved by the use of weighted matching algorithms (involving T-joins; see for instance [7]). So our main result will be:

Theorem 6.1 *There is a polynomial algorithm for finding the minimal dependent set of minimum cost when the costs of the red edges are non negative.*

Proof A. Let us describe $\mathcal{N}$; first we consider the directed graph D orientation of G. Each node q of $\mathcal{N}$ will be represented by a signature $\sigma(q)$ and a symbol $s(q)$. Two nodes q and q' will be identical if they have the same symbol and the same signature and we will write $q = (s(q); \sigma(q))$. The signature of each node will be $-$ or $+$. The nodes with signature $-$ (resp. $+$) belong to the first (resp. second) level. For each arc $a = (v, w)$ of D create two symbols $I(a)$ and $T(a)$. We will say that $I(a)$ refer to v and $T(a)$ refer to w. If $|V| = n$ the total number of nodes of the complete oriented graph D is $n(n-1)$. The total number of symbols is $2n(n-1)$ and as we have two possible signatures, the total number of nodes of the network is $4n(n-1)$. Let us describe now the set of arcs of the network with their costs:

* Red arcs: Associate to any red arc $a = (v, w)$ of D two red arcs called the copies of a : $f_1 = (p_1, q_1)$ and $f_2 = (p_2, q_2)$. $p_1 = (I(a); -)$, $q_1 = (T(a); +)$, $p_2 = (I(a); +)$, $q_2 = (T(a); -)$. Each red edge e of G has four copies in the network and the cost of each copy will be $c(e)$.
* Black arcs: Associate to any black arc $a = (v, w)$ of D two black arcs called the copies of a: $f_1 = (p_1, q_1)$ and $f_2 = (p_2, q_2)$. $p_1 = (I(a); -)$, $q_1 = (T(a); -)$, $p_2 = (I(a); +)$, $q_2 = (T(a); +)$. Each red edge e of G has four copies in the network and the cost of each copy will be 0.
* Transition arcs: To any pair of consecutive arcs $a = (v, w)$ and $b = (w, z)$ of D associate two arcs: $f_1 = (p_1, q_1)$ and $f_2 = (p_2, q_2)$. $p_1 = (T(a); -)$, $q_1 = (I(b); -)$, $p_2 = (T(a); +)$, $q_2 = (I(b); +)$. The costs of f_1 and f_2 are $c(w)$ if a and b are both black arcs and 0 otherwise.

Note that the red and black arcs go from nodes with symbol $I(a)$ to nodes with symbol $T(a)$ for some arc of D; on the other hand a transition arc goes from a node with symbol $T(a)$ to a node with symbol $I(b)$ where $T(a)$ and $I(b)$ refer to the same node of G. The black edges and the transition edges stay on the same level; only the red edges start make transitions between the two levels. If we forget for a moment the signatures of the nodes our transformation of D into a network is similar to the relation between the task-and-node representation and the task-on-arc representation of a simple scheduling problem.
Let $C_1 = q_1, q_2, ..., q_{2k}, q_{2k+1}$ be a walk in the network. Assume also that the symbol of q_1 is $I(a_1)$ for some arc a_1 of the graph D; hence, there exists in D a sequence of edges $a_1, ..., a_k$ in D such that:
The symbol of q_{2i+1} is $I(a_i)$ for $0 \leq i \leq k$ and the symbol of q_{2i} is $T(a_i)$ for $1 \leq i \leq k$. The arc $f_{2i-1} = (q_{2i-1}, q_{2i})$ is a copy of a_i for $1 \leq i \leq k$ and the arc $f_{2i} = (q_{2i}, q_{2i+1})$ is a transitive arc for $1 \leq i \leq k$. The symbol $I(a_{i+1})$ of q_{2i+1} refer to a node v_{i+1} for $0 \leq i \leq k$. We have also $a_i = (v_i, v_{i+1})$ for $1 \leq i \leq k$. $C = v_1, v_2, ..., v_{k+1}$ is the walk of D associated to C_1. If the symbols of q_1 and q_{2k+1} are identical C is a closed walk. If q_1 and q_{2k+1} have identical symbols but distinct signatures (and we will always assume that the signature of q_1 is $-$) C is a red-odd walk since if the initial and final levels are not the same we need to use an odd number of edges going from a level to the other level; only red edges do this.

In the network, the cost $d_1(C_1)$ is:

$$d_1(C_1) = \sum_{i=1}^{2k} c(f_i).$$

If we use an arc which is the copy of a red edge we pay the cost of this edge; if we use a black arc we pay nothing since the cost of this arc is 0. If we use a transition arc (p, q), p and q are related to the same node v of G and we pay the cost c_v if and only if the two arcs of C incident to v are both black. This is exactly our definition of the cost $d(C)$ of the chain C in problem P_3 and $d(C) = d_1(C_1)$. Therefore the solution of P_3 is obtained by solving the shortest path problem in the network (by Dijkstra algorithm) between two nodes with the same symbol $I(a)$ (where a is an arc of D) and distinct signature. (We need to apply Dijkstra algorithm $n(n-1)$ times). This gives a polynomial algorithm to solve P_3 and this proves our theorem. $\square$

Proof B. We will show now that P_3 is equivalent to the minimum weighted path problem in an undirected graph $\mathcal{G}$ with no absorbing cycle. Each node has a level (+) or (-) and have a subscript r (for red) or b (for black); $\mathcal{G}$ has $4n$ nodes. Let us now describe completely $\mathcal{G}$:

* Nodes: Associate to each node $v \in V$ four copies of v: v_r^+, v_b^+, v_r^-, and v_b^-; v is the image of these four nodes.
* Red edges: Associate to any red edge $e = \{u, v\} \in E$ two red edges of $\mathcal{G}$: $\{u_r^+, v_r^-\}$ and $\{u_r^-, v_r^+\}$ with cost $c(e)$. These two edges are copies of e and e is the image of these two edges.
* Black edges: Associate to each black edge $e = \{v, w\} \in \overline{E}$ two copies black edges $\{v_b^+, w_b^+\}$ and $\{v_b^-, w_b^-\}$ with cost $(c(v) + c(w))/2$. These two edges are copies of e and e is the image of these two edges.
* Transition edges: Associate to any node $v \in V$ two transition edges $\{v_r^+, v_b^+\}$ and $\{v_r^-, v_b^-\}$ with negative cost $-c(v)/2$. The image of these two edges will be a loop (v, v) that we will add to our original graph G.

Let $P = u_1, u_2, ..., u_{l+1}$ be a walk in $\mathcal{G}$ linking v_r^+ and v_r^-; if w_i is the image of u_i in G for $i = 1, 2, ..., k$, let D be the chain $w_1, w_2, ..., w_{l+1}$ in G; finally let $C = v_1, v_2, ..., v_{k+1}$ obtained from D by deleting all the loops. We will also assume that the first edge (u_1, u_2) of P is a red edge. Since the images of u_1 and u_{l+1} are v, $w_1 = v_1 = w_{l+1} = w_{k+1} = v$ and D and C are closed walks; moreover the first node of P has a signature $+$ and the last node of P has a signature $-$; as in the previous proof, C is a red-odd walk in G. Note that in $\mathcal{G}$ a red edge cannot be incident to a black edge, and a transition edge cannot be incident to an other transition edge. Thus the edge-set of P can be partitioned into subchains $P' = u_r, ..., u_s$ with $1 \leq r < s \leq l+1$ which are of one of the following type:

(I) All the edges of P' are red edges. The cost $d(P')$ of P' is the sum of the costs of the red edges of the chain C which are images of the edges of P'.

(II) The first and the last edge of P' are transition edges and the other edges are black edges. So $s-r \geq 4$; if $s-r>4$, the images $v_{r+2},...,v_{s-2}$ of the nodes $u_{r+2},...,u_{s-2}$ are nodes of C incident to two black edges of C. The cost of this subchain is:

$$d(P') = -c(v_r)/2 + \sum_{r+1 \leq i \leq s-2} \left\{ \frac{c(v_i)+c(v_{i+1})}{2} \right\} - c(v_{s-1})/2 = \sum_{r+2 \leq i \leq s-2} c(v_i).$$

Any red edge of C belongs a subchain of type I and a node of C incident to two black edges belongs to a subchain of type II. Hence the cost $c(P)$ of P is equal to the cost $d(C)$ of problem P_3. As $c(P) \geq 0$, $\mathcal{G}$ has no absorbing cycle and we can solve in polynomial time the minimum weighted path problem in the network using matching algorithms. The solution of P_3 is obtained by solving the minimum weighted path problem in $\mathcal{G}$ beetween v_r^+ and v_r^- for every $v \in V$. □

7 Maximum weighted biclique and minimum biclique cover

In this final section we will formulate the maximum weighted biclique problem and the minimum biclique cover problem as integer programs and we will study their continuous relaxation.

7.1 The maximum weighted biclique problem

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

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

If we replace constraints (α) by nonnegativity constraints:

$$0 \leq x_e \leq 1 \quad \forall e \in E,$$

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

Theorem 7.1 *$(\mathcal{P})$ can be solved in polynomial time.*

Proof. Recall the main fundamental result: the *ellipsoid method* [9] solves $(\mathcal{P})$ in polynomial time provided that the following separation problem: given $x \in \mathbb{R}^E$, decide if x satisfies the constraints of $(\mathcal{P})$ or find a constraint of $(\mathcal{P})$ which is violated by x is solvable in polynomial time. As the number of nonnegativity constraints is polynomial we just 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 less 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 (6.1). □

Remark
We cannot expect this relaxation to be a strong relaxation for the maximum weighted biclique problem since the point x whose components are all equal to $\frac{1}{2}$ is always a feasible solution of $(\mathcal{P})$. We have to remind that minimal dependent sets constraints are never strong constraints in Combinatorial Optimization. For instance, if we consider the Stable Set polytope of a graph $G = (V, E)$, the minimal dependent sets are the pair of nodes linked by an edge: so our constraints are somehow related to the edge constraints: $x_v + x_w \leq 1$, $\forall e = (v, w) \in E$. But in our problem the number of minimal dependent sets is exponential and our approach seems to be the first tractable attack of this problem using linear programming. From the minimal dependent sets inequalities we can of course deduce new arithmetic (or Gomory) cuts using the same technique as in the case of the Stable Set polytope where Clique constraints and Odd Holes constraints are obtained from the edge constraints by arithmetic rounding.
Consider the following auxiliary graph: $A(G) = (F, U)$ where $(f, g) \in U$ if and only if the set $\{f, g\}$ 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 :

$$x(K) \leq 1 \quad \forall K \in \mathcal{K} \qquad (\gamma)$$

($\mathcal{K}$ 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 7.1 *The separation problem for constraints (γ) is NP-complete.*

Proof. Associate to a graph $G=(V,E)$ the following bipartite graph $B(G)$: each node v of G is replaced by two copies $v' \in V'$ and $v" \in V"$ in $B(G)$. The two copies v' and $v"$ are linked by an edge e_v in $B(G)$ called vertical edge; moreover $(v,w) \notin E$ if and only if $(v',w")$ and $(v",w')$ are edges of $B(G)$ (These two edges will be called transversal edges) and $(v,w) \in E$ 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 $e_v \quad \forall v \in V$ 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 (γ) 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 a NP-hard problem. $\square$

Finally we conjecture that the minimal dependent sets inequalities are facet inducing inequalities.

7.2 The minimum biclique cover problem

We define now a linear programming relaxation of the minimum biclique cover 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 $k=|E|$).
Let $K(G)$ 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 $x^1,...,x^k$ of dimension $|E|$. The constraints of our linear program are:

- $x^i \in K(G)$ for $i=1,...,k$
- $x^i_e \leq y^i$ for every $e \in E$ and $i=1,...,k$

- $\sum_{i=1}^{k} x_e^i = 1$ for every $e \in E$.

The objective function is: Minimize $\sum_{i=1}^{k} y^i$.

References

[1] G. Alexe, S. Alexe, S. Foldes, P.L. Hammer, and B. Simeone, *Consensus algorithms for the generation of all maximal bicliques*, DIMACS Tech. Rep. 2000-14 (May 2000).

[2] J. Amilhastre, P. Janssen, and M.-C. Vilarem, *Complexity of minimum biclique cover and minimum biclique decomposition for bipartite domino-free graphs*, Disc. App. Math. 86 (1998) 125-144.

[3] J.-C. Bermond, *Couverture des arêtes d'un graphes par des graphes bipartis complets*, preprint, Univ. of Paris-Sud, Centre d'Orsay, Rapport de Recherche No. 10. (1978).

[4] F.R.K. Chung, *On the coverings of graphs*, Discrete Math. 30 (1980) 89-93.

[5] P. Duchet, *Hypergraphs*, Handbook of Combinatorics (eds. R. Graham, M. Grötschel, and L. Lovász), Elsevier Science B.V. (1995) 381-432.

[6] P.C. Fishburn and P.L. Hammer, *Bipartite dimensions and bipartite degrees of graphs*, Discrete Math. 160 (1996) 127-148.

[7] A. Frank, *Connectivity and network flows*, Handbook of Combinatorics (eds. R. Graham, M. Grötschel, and L. Lovász), Elsevier Science B.V. (1995) 111-177.

[8] M.R. Garey and D.S. Johnson, *Computers and intractability : A guide to the theory of NP-completness*, W.H. Freeman, San Fransisco (1979).

[9] M. Grötschel, L. Lovàsz and A. Schrijver, *The ellipsoid method and its consequences in combinatorial optimization*, Combinatorica 1(2) (1981) 169-197.

[10] O. Günlük, *A new min-cut max-flow ratio for multicommodity flows*, in process (2001).

[11] F. Harary, D. Hsu, and Z. Miller, *The biparticity of a graph*, Journal of graph theory 1 (1977) 131-133.

[12] K.F. Jones, J.R. Lundgren, N.J. Pullman, and R. Rees, *A note on the covering numbers of* $K_n \setminus K_m$ *and complete t-partite graphs*, Congr. Num. 66 (1988) 181-184.

[13] S.D. Monson, N.J. Pullman, and R. Rees, *A survey of clique and biclique coverings and factorizations of (0,1)-matrices*, Bulletin of the ICA 14 (1995) 17-86.

[14] H. Müller, *On edge perfectness and classes of bipartite graphs*, Discrete Math. 149 (1996) 159-187.

[15] J. Orlin, *Contentment in Graph Theory*, Indag. Math. 39 (1977) 758-762.

[16] D.J. Siewert, *Biclique covers and partitions of bipartite graphs and digraphs and related matrix ranks of* $\{0,1\}$*-matrices*, PhD Thesis of the University of Colorado at Denver (2000).

[17] H.U. Simon, *On approximable solution for combinatorial optimization problems*, SIAM J. Disc. Math. 3 (2) (1990) 294-310.

[18] Z. Tuza, *Covering of graphs by complete bipartite subgraphs; complexity of 0-1 matrices*, Combinatorica 4 (1) (1984) 111-116.