Available online at www.sciencedirect.com
ScienceDirect
Theoretical Computer Science 362 (2006) 140–161
Theoretical Computer Science
www.elsevier.com/locate/tcs

Minimal invariant sets in a vertex-weighted graph

Francesco M. Malvestuto*, Mauro Mezzini, Marina Moscarini

Dipartimento di Informatica, Università "La Sapienza", Via Salaria 113, 00198 Roma, Italy

Received 13 June 2005; received in revised form 28 May 2006; accepted 1 June 2006

Communicated by Ding-Zhu Du

Abstract

A weighting of vertices of a graph is admissible if there exists an edge weighting such that the weight of each vertex equals the sum of weights of its incident edges. Given an admissible vertex weighting of a graph, an *invariant set* is an edge set such that the sum of the weights of its edges is the same for every edge weighting, and a nonempty invariant set is *minimal* if none of its nonempty proper subsets is an invariant set. It is easily seen that every nonempty invariant set is a disjoint union of minimal invariant sets. A graphical characterisation of minimal invariant sets in a bipartite graph is known both in the case the vertex weights are reals and in the case the vertex weights are nonnegative reals. We shall state a graphical characterisation of minimal invariant sets in an arbitrary vertex-weighted graph. Moreover, we give a linear algorithm to test an invariant set for minimality. Finally, we state a complete axiomatisation of invariant sets and give a polynomial algorithm to find a set of minimal invariant sets that completely characterise the set of all invariant sets.

Keywords: Algebraic set; Invariant set

1. Introduction

A graph G is a pair $(V(G), E(G))$, where $V(G)$ is a nonempty finite set called the set of its *vertices*, and $E(G)$ is a set of unordered pairs of its vertices called the set of its *edges*. Let $e = (u, v)$ be an edge; the vertices u and v are the *endpoints* of the edge e, which is a *loop* if $u = v$. By $h(e)$ we denote the set of endpoints of e; thus, the cardinality of $h(e)$ is 1 or 2 depending on whether e is or is not a loop. Let $\boldsymbol{G} = (g_{ve})_{v\in V(G), e\in E(G)}$ be the binary matrix with $g_{ve} = 1$ if and only if $v \in h(e)$. We call $\boldsymbol{G}$ the *matrix* of G. Note that, according to the terminology used in hypergraph theory [2], $\boldsymbol{G}$ is the incidence matrix of the hypergraph with vertex set $V(G)$ and hyperedge set $\{h(e) : e \in E(G)\}$.

Henceforth, R and R_+ will denote the set of real numbers and the set of nonnegative real numbers, respectively. A *vertex weighting* of G is a $|V(G)|$-dimensional vector of real numbers. A vertex weighting $\boldsymbol{a} = (a_v)_{v\in V(G)}$ of G is *admissible over* R (or R_+) if there exists at least one real-valued (nonnegative real-valued, respectively) solution of the following system of linear equations:

$$\boldsymbol{G}\boldsymbol{x} = \boldsymbol{a}, \tag{1}$$

* Corresponding author. Tel.: +390649918310; fax: +39068541842.
E-mail address: malvestuto@di.uniroma1.it (F.M. Malvestuto).

0304-3975/$ - see front matter
doi:10.1016/j.tcs.2006.06.005

where $\boldsymbol{G}$ is the matrix of G, and $\boldsymbol{x} = (x_e)_{e \in E(G)}$ denotes an $|E(G)|$-dimensional variable. It is well-known [9,13] that a vertex weighting $\boldsymbol{a}$ of a connected graph G is not admissible over R if and only if G is bipartite and $\sum_{v \in U} a_v \neq \sum_{v \in W} a_v$ where (U, W) is a bipartition of G, and that $\boldsymbol{a}$ is admissible over R_+ if and only if each $a_v \in \mathsf{R}_+$, $v \in V(G)$, and $\boldsymbol{a}$ is admissible over R [11,20]. If $\boldsymbol{a}$ is a vertex weighting of G that is admissible over R (or R_+), then we call the pair $(G, \boldsymbol{a})$ a *vertex-weighted graph over* R (R_+, respectively), and a real-valued (nonnegative real-valued, respectively) solution of system (1) a *feasible edge weighting* of $(G, \boldsymbol{a})$. In the theory of magic graphs [9,13], a vertex weighting is also called an "indexing vector'' or a "vertex labelling'', and a feasible edge weighting is also called a "labelling induced'' by (or a "labelling compatible'' with) the given vertex weighting.

Let $(G, \boldsymbol{a})$ be a vertex-weighted graph over R or over R_+, and let S be a subset of $E(G)$; the sum-expression $\sum_{e \in S} x_e$ is a *sum invariant* of $(G, \boldsymbol{a})$ [14–18] if either $S = \emptyset$ or, for every two feasible edge weightings $\boldsymbol{x}' = (x'_e)_{e \in E(G)}$ and $\boldsymbol{x}'' = (x''e)_{e \in E(G)}$ of $(G, \boldsymbol{a})$, one has $\sum_{e \in S} x'_e = \sum_{e \in S} x''_e$. Let $\sum_{e \in S} x_e$ be a sum invariant of $(G, \boldsymbol{a})$; its *value* is taken to be 0 if $S = \emptyset$; otherwise, it is taken to be the sum $\sum_{e \in S} x_e$, where $\boldsymbol{x} = (x_e)_{e \in E(G)}$ is any feasible edge weighting of $(G, \boldsymbol{a})$. A *zero invariant* of $(G, \boldsymbol{a})$ is a sum invariant with value zero.

Let $\sum_{e \in S} x_e$ be a sum invariant of $(G, \boldsymbol{a})$ with value α, then we call S an *invariant set* (also called the "effective area'' of the sum-invariant in [14–18]) of $(G, \boldsymbol{a})$ with value α. If the singleton $\{e\}$, for some edge e of G, is an invariant set of $(G, \boldsymbol{a})$, then we call e an *invariant edge* of $(G, \boldsymbol{a})$. A *zero invariant edge* of $(G, \boldsymbol{a})$ is an invariant edge with value zero. An invariant set of $(G, \boldsymbol{a})$ is *minimal* if it is not empty and none of its nonempty proper subsets is an invariant set of $(G, \boldsymbol{a})$. It is easy to see that the family of invariant sets of $(G, \boldsymbol{a})$ is always closed under disjoint union and proper difference. Note that, by the closure under proper difference and disjoint union, a nonempty edge set is an invariant set of $(G, \boldsymbol{a})$ if and only if it is a disjoint union of one or more minimal invariant sets.

In this paper, we address the problem of finding a graphical characterisation of invariant sets and of minimal invariant sets in vertex-weighted graphs over R and R_+. The interest in (minimal) invariant sets was first motivated by the security issues connected with the publication of a two-dimensional statistical table [8,11,14–19] and with the interrogation of a statistical database [19,20,22–24]. Recently, invariant sets have found applications in sum-query processing [10,12,21]. Finally, minimal invariant sets can be used to find a compact representation of the information content of a statistical summary [14,22].

At the present, the only known results on the subject are the graphical characterisations of minimal invariant sets in bipartite vertex-weighted graphs [17,19]. In this paper we state a graphical characterisation of minimal invariant sets in an arbitrary vertex-weighted graph; moreover, we also give a linear test for recognising minimal invariant sets and a complete axiomatisation of invariant sets.

The organisation of the paper is as follows. Section 2 contains some basic graph-theoretic definitions. In Section 3, after recalling existing analytic characterisations of invariant sets and of minimal invariant sets, we introduce the notion of an "algebraic set" of a graph (which proves to be equivalent to that of an invariant set in a vertex-weighted graph over R), and we show that the problem of finding a graphical characterisation of (minimal) invariant sets reduces to the problem of finding a graphical characterisation of (minimal) algebraic sets. Section 4 contains properties of the so-called "signatures'' of a graph induced by an algebraic set. In Section 5 a graphical characterisation of minimal algebraic sets is given. In Section 6 it is proven that recognising a minimal algebraic set is a linear problem. In Section 7 we state a complete axiomatisation of algebraic sets and give a polynomial algorithm to find a set of minimal algebraic sets that completely characterise the set of all algebraic sets. Directions for future research are suggested in Section 8.

2. Definitions

In this section, we recall further more-or-less standard definitions on graphs [3], which will be used in the sequel.

A graph G is *empty* if $E(G) = \emptyset$. The *star* of a (possibly empty) subset U of $V(G)$, denoted by $star(U)$ (by $star(v)$ if $U = \{v\}$), is the set of edges having at least one endpoint in U; moreover, the star of U is called a *starset* [19] if, for every two distinct vertices u and v in U, $star(u) \cap star(v) = \emptyset$.

A *subgraph* of graph G is a graph H with $V(H) \subseteq V(G)$ and $E(H) \subseteq E(G)$; if, in addition, one has that $E(H) = \{e \in E(G) : h(e) \subseteq V(H)\}$, then H is called the *subgraph* of G *induced* by $V(H)$. The subgraph of G induced by $V(G) - \{u\}$ is denoted by $G - u$. If S is a subset of $E(G)$, then by $G - S$ we denote the graph with vertex set $V(G)$ and edge set $E(G) - S$. If u and v are vertices of G and (u, v) is not an edge of G, by $G + e$ we denote the graph with vertex set $V(G)$ and edge set $E(G) \cup \{e\}$.

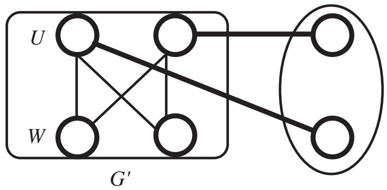

Fig. 1. A bipartite bond.

A *path* of length k, $k \geqslant 0$, is a sequence $(v_0, v_1, \ldots, v_k)$ of distinct vertices such that, if $k > 0$, then $(v_{i-1}, v_i) \in E(G)$, $1 \leqslant i \leqslant k$; the vertices v_0 and v_k are said to be *joined* by the path. A graph is *connected* if every two vertices are joined by a path. A *component* of a graph is a maximal connected induced subgraph.

A *cycle* of length k, $k \geqslant 1$, is a sequence $(v_0, v_1, \ldots, v_k)$ of vertices such that $(v_0, v_1, \ldots, v_{k-1})$ and $(v_1, \ldots, v_k)$ are paths and $v_k = v_0$. A cycle is *odd* or *even* depending on whether its length is odd or even. Accordingly, a loop is an odd cycle of length 1. A graph G is *bipartite* if it contains no odd cycles, that is, if and only if either G is an empty graph or there are two nonempty disjoint subsets U and W of $V(G)$ such that $U \cup W = V(G)$ and each edge of G has one endpoint in U and the other endpoint in W; in the latter case, we call (U, W) a *bipartition* of G. Note that, if a nonempty bipartite graph G is connected, then there is exactly one bipartition of G.

Let G be a connected graph. A *spanning tree* of G is a connected subgraph of G with vertex set $V(G)$ and with no cycles. If G is not bipartite, a *spanning L-tree* [7] of G is a spanning tree of G with the addition of one edge of G that creates an odd cycle.

Given a subset Z of $V(G)$, by $[Z]_1$ and $[Z]_2$ we denote the following two subsets of $E(G)$:

$$[Z]_1 = \{e \in E(G) : |h(e) \cap Z| = 1\}, \quad [Z]_2 = \{e \in E(G) : |h(e) \cap Z| = 2\}.$$

Note that a loop belongs to $[Z]_1$ if and only if its endpoint belongs to Z, and that $[Z]_2$ is loopless. A nonempty subset S of $E(G)$ is an *edge cut* [3] if there is a nonempty proper subset Z of $V(G)$ such that $S = [Z]_1$ and S is loopless. A minimal edge cut is called a *bond* [3] (a "basic set'' in [14–16]). A bond S is *bipartite* [14–16] ("simple'' in [19]) if there is a bipartite component G' of $G - S$ that either G' is an empty graph or the bipartition (U, W) of G' is such that either $S \cap star(U)$ or $S \cap star(W)$ is empty (see Fig. 1).

3. Invariant sets and algebraic sets

Let G be a graph and S a subset of $E(G)$. The *characteristic vector* of S (in G) is the binary vector $\boldsymbol{s} = (s_e)_{e \in E(G)}$ with $s_e = 1$ if and only if $e \in S$. The following two results characterise invariant sets in terms of their characteristic vectors.

Proposition 1 (*Chin and Ozsoyoglu [6]*). *Let $(G, \boldsymbol{a})$ be a vertex-weighted graph over* R. *A subset of $E(G)$ is an invariant set of $(G, \boldsymbol{a})$ if and only if its characteristic vector can be expressed as a linear combination of rows of the matrix of G.*

Proposition 2 (*Malvestuto and Moscarini [24]*). *Let $(G, \boldsymbol{a})$ be a vertex-weighted graph over* R_+, *and let Z be the set of its zero invariant edges. A subset S of $E(G)$ is an invariant set of $(G, \boldsymbol{a})$ if and only if the characteristic vector of $S - Z$ (in $G - Z$) can be expressed as a linear combination of rows of the matrix of $G - Z$.*

By Proposition 1, given a vertex-weighted graph $(G, \boldsymbol{a})$ over R, invariance is independent of the weighting $\boldsymbol{a}$ of G; moreover, if a subset S of $E(G)$ is an invariant set of $(G, \boldsymbol{a})$ and the characteristic vector of S can be expressed as a linear combination of rows of the matrix of G with coefficients c_v, $v \in V(G)$, then the value of S is $\sum_{v \in V(G)} c_v a_v$.

By Proposition 2, given a vertex-weighted graph $(G, \boldsymbol{a})$ over R_+, invariance depends on the weighting $\boldsymbol{a}$ of G only through the set Z of the zero invariant edges of $(G, \boldsymbol{a})$; moreover, if a subset S of $E(G)$ is an invariant set of $(G, \boldsymbol{a})$ and the characteristic vector of $S - Z$ (in $G - Z$) can be expressed as a linear combination of rows of the matrix of $G - Z$ with coefficients c_v, $v \in V(G)$, then the value of S is $\sum_{v \in V(G)} c_v a_v$.

We shall re-state Propositions 1 and 2 using the notion of an "algebraic set'' [19] of a graph G, which is now recalled.

A (possibly empty) subset of $E(G)$ is an *algebraic set* of G if its characteristic vector can be expressed as a linear combination of rows of the matrix of G. If the singleton $\{e\}$ is algebraic, then e is an *algebraic edge* of G. Thus, a subset S of $E(G)$ is an algebraic set of G if and only if there exists at least one solution of the following equation system:

$$\boldsymbol{G}^{\top}\boldsymbol{y} = \boldsymbol{s}, \tag{2}$$

where $\boldsymbol{G}^{\top}$ denotes the transpose of the matrix of G. Note that $\boldsymbol{c}$ is a solution of system (2) if and only if

$$c_w = s_{(w,w)} \quad \text{for every loop } (w, w) \text{ of } G, \text{ and}$$
$$c_u + c_v = s_{(u,v)} \quad \text{for every edge } (u, v) \text{ of } G,\ u \neq v.$$

Trivial examples of an algebraic set of G are the empty subset of $E(G)$ and starsets. Nontrivial examples are a bipartite bond (with reference to Fig. 1, $c_v = 1$ for $v \in U$, $c_v = -1$ for $v \in W$ and $c_v = 0$ for $v \notin U \cup W$) and the edge set of G whenever G is loopless ($c_v = \frac{1}{2}$ for all $v \in V(G)$).

A subset of $E(G)$ is a *minimal algebraic set* of G if it is a nonempty algebraic set of G and none of its nonempty proper subsets is an algebraic set of G. Like the family of invariant sets of a vertex-weighted graph, also the family of algebraic sets of a graph is closed under disjoint union and proper difference [4,5], so that every nonempty algebraic set is a disjoint union of one or more minimal algebraic sets.

Using the notion of an algebraic set, Propositions 1 and 2 can be re-phrased as follows:

Proposition 1′. *Let $(G, \boldsymbol{a})$ be a vertex-weighted graph over* R. *A subset of $E(G)$ is an invariant set of $(G, \boldsymbol{a})$ if and only if it is an algebraic set of G.*

Proposition 2′. *Let $(G, \boldsymbol{a})$ be a vertex-weighted graph over* R_+, *and let Z be the set of its zero invariant edges. A subset S of $E(G)$ is an invariant set of $(G, \boldsymbol{a})$ if and only if $S - Z$ is an algebraic set of $G - Z$.*

By Propositions 1′ and 2′, minimal invariant sets can be characterised as follows.

Theorem 1. *Let $(G, \boldsymbol{a})$ be a vertex-weighted graph over* R. *A subset of $E(G)$ is a minimal invariant set of $(G, \boldsymbol{a})$ if and only if it is a minimal algebraic set of G.*

Theorem 2. *Let $(G, \boldsymbol{a})$ be a vertex-weighted graph over* R_+, *and let Z be the set of its zero invariant edges. A subset S of $E(G)$ is a minimal invariant set of $(G, \boldsymbol{a})$ if and only if either $S = \{e\}$ for some edge e in Z, or $S \cap Z = \emptyset$ and S is a minimal algebraic set of $G - Z$.*

A graphical characterisation of the set of zero invariant edges of a vertex-weighted graph over R_+ was given by Gusfield [11] in the bipartite case, and by two of the authors [20] in the general case. Therefore, by Theorems 1 and 2, the problem of finding a graphical characterisation of minimal invariant sets reduces to the problem of finding a graphical characterisation of minimal algebraic sets, which will be given in next section and is based on some properties of solutions of system (2), which are now stated.

First of all, observe that, by the closure under disjoint union, a subset S of $E(G)$ is an algebraic set of G if and only if, for each component G' of G, the intersection of S with $E(G')$ is an algebraic set of G'. Moreover, a subset S of $E(G)$ is a minimal algebraic set of G if and only if there is a component G' of G such that S is a minimal algebraic set of G'. Therefore, without loss of generality, henceforth we assume that G is connected.

The following results concern the number and the type of solutions of system (2), where $\boldsymbol{s}$ is the characteristic vector of an algebraic set of a connected graph G. We shall see that they are determined by the rank of the matrix of G. In the special case that G is loopless, the following result is known.

Proposition 3 (*Conforti and Rao [7], Van Nuffelen [25]*). *Let G be a loopless and connected graph, and let* $\boldsymbol{G}$ *be the matrix of G. If G is bipartite, then the rank of* $\boldsymbol{G}$ *is $|V(G)| - 1$ and the columns of* $\boldsymbol{G}$ *indexed by the edges of a spanning tree of G form a maximal linearly independent set. If G is not bipartite, then the rank of* $\boldsymbol{G}$ *is $|V(G)|$ and the columns of* $\boldsymbol{G}$ *indexed by the edges of a spanning L-tree of G form a maximal linearly independent set.*

Fig. 2. The path joining the endpoints of e'' and e'.

Proposition 3 can be easily generalised to an arbitrary connected graph.

Lemma 1. *Let G be a connected graph. The rank of the matrix of G is $|V(G)| - 1$ or $|V(G)|$ depending on whether G is or is not bipartite.*

Proof. Let $\boldsymbol{G}$ be the matrix of G. If G is loopless, then the statement holds by Proposition 3. Consider now the case that G contains loops. Let E' be the set of loops of G. Distinguish two cases depending on whether $G - E'$ is or is not bipartite.

Case 1: $G - E'$ is bipartite. Let T be the edge set of a spanning tree of $G - E'$, and let $\boldsymbol{B} = \{\boldsymbol{g}_e : e \in T\}$ be the set of columns of $\boldsymbol{G}$ indexed by T. By Proposition 3, $\boldsymbol{B}$ is a maximal linearly independent set of columns of the matrix of $G - E'$ and $|\boldsymbol{B}| = |V(G)| - 1$. Let e' be a loop of G, let $T' = T \cup \{e'\}$ and let $\boldsymbol{B}' = \{\boldsymbol{g}_e : e \in T'\}$ be the set of columns of $\boldsymbol{G}$ indexed by T'. It is easy to see that the column $\boldsymbol{g}_{e'}$ of $\boldsymbol{G}$ indexed by e' cannot be expressed as a linear combination of columns in $\boldsymbol{B}$; therefore, $\boldsymbol{B}'$ is a linearly independent set of columns of $\boldsymbol{G}$. Let e'' be any element (if any) of $E' - \{e'\}$. Let $(v_0, v_1, \ldots, v_k)$, $k \geqslant 1$, be the path in $T' \cup \{e''\}$ that joins the endpoints of e'' and e' (see Fig. 2).

It is easy to see that the column $\boldsymbol{g}_{e''}$ of $\boldsymbol{G}$ indexed by e'' can be expressed as a linear combination of columns in $\boldsymbol{B}'$ as follows:

$$\boldsymbol{g}_{e''} = \boldsymbol{g}_{e_1} - \boldsymbol{g}_{e_2} + \cdots + (-1)^{k-1}\boldsymbol{g}_{e_k} + (-1)^k \boldsymbol{g}_{e'}.$$

Therefore, $\boldsymbol{B}' \cup \{\boldsymbol{g}_{e''}\}$ is a linearly dependent set. Since this holds for every loop in $E' - \{e'\}$, $\boldsymbol{B}'$ is a maximal linearly independent set of columns of $\boldsymbol{G}$, so that the rank of $\boldsymbol{G}$ is

$$|\boldsymbol{B}'| = |\boldsymbol{B}| + 1 = (|V(G)| - 1) + 1 = |V(G)|.$$

Case 2: $G - E'$ is not bipartite. Let T be the edge set of a spanning L-tree of $G - E'$, and let $\boldsymbol{B} = \{\boldsymbol{g}_e : e \in T\}$ be the set of columns of $\boldsymbol{G}$ indexed by T. By Proposition 3, $\boldsymbol{B}$ is a maximal linearly independent set of columns of the matrix of $G - E'$ and $|\boldsymbol{B}| = |V(G)|$. Let e' be any element of E'. The addition of e' to T creates a couple of odd cycles joined by a path, say $(v_0, v_1, \ldots, v_k)$, $k \geqslant 0$, such that, if $k > 0$, then no edge (v_{i-1}, v_i) is in either odd cycle (see Fig. 3).

Then, it is easy to see that the column $\boldsymbol{g}_{e'}$ of $\boldsymbol{G}$ indexed by e' can be expressed as a linear combination of columns in $\boldsymbol{B}$ as follows: if $k = 0$ then

$$\boldsymbol{g}_{e'} = \tfrac{1}{2}(\boldsymbol{g}_{e_1} - \boldsymbol{g}_{e_2} + \cdots + \boldsymbol{g}_{e_{2m+1}})$$

otherwise

$$\boldsymbol{g}_{e'} = \boldsymbol{g}_{e_1} - \boldsymbol{g}_{e_2} + \cdots + (-1)^{k-1}\boldsymbol{g}_{e_k} + \frac{(-1)^k}{2}(\boldsymbol{g}_{e_{k+1}} - \boldsymbol{g}_{e_{k+2}} + \cdots + \boldsymbol{g}_{e_{k+2m+1}}).$$

Since this holds for every loop in E', $\boldsymbol{B}$ is a maximal linearly independent set of columns of $\boldsymbol{G}$, so that the rank of $\boldsymbol{G}$ is $|\boldsymbol{B}| = |V(G)|$. $\square$

Since the rank of a matrix is equal to the rank of its transpose, by Lemma 1, the rank of the coefficient matrix of system (2) is $|V(G)| - 1$ or $|V(G)|$ depending on whether G is or is not bipartite, so that the following holds.

Corollary 1. *Let G be a connected graph, and* **s** *the characteristic vector of an algebraic set of G. System* (2) *has either ∞^1 solutions or exactly one solution depending on whether G is or is not bipartite.*

Corollary 1 has the following important algorithmic implication.

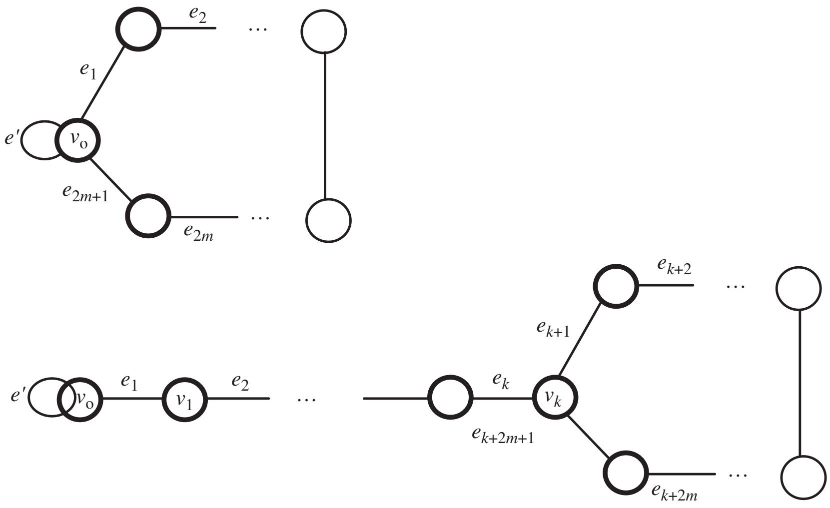

Fig. 3. The addition of loop e' to T.

Fact 1 (*Malvestuto and Moscarini [24]*). *Given a connected graph G and a subset S of* $E(G)$, *one can decide whether or not S is an algebraic set of G in time linear in the size of G and*, *in the affirmative case*, *one can find a solution of system* (2) *in time linear in the size of G*.

Given a vertex-weighted graph $(G, \boldsymbol{a})$ over R, by Theorem 1 and Fact 1, invariance can be tested in time linear in the size of G.

Given a vertex-weighted graph $(G, \boldsymbol{a})$ over R_+ and the set Z of its zero invariant edges, by Theorem 2 and Fact 1, invariance can be tested in time linear in the size of G. In the case that G is bipartite, Gusfield [11] proved that Z can be found in time linear when a feasible edge weighting of $(G, \boldsymbol{a})$ is known, and that a feasible edge weighting of $(G, \boldsymbol{a})$ can be obtained in cubic time [1] by solving a maximum-flow problem on a flow network that can be constructed from $(G, \boldsymbol{a})$ in linear time. Two of the authors [20] extended Gusfield's approach to the nonbipartite case, and proved that a feasible edge weighting of $(G, \boldsymbol{a})$ can be still obtained in cubic time and that Z can be found in time linear from a feasible edge weighting of $(G, \boldsymbol{a})$.

Corollary 1 concerned the number of solutions of system (2). The next result (see Lemma 3) concerns the type of solutions of system (2).

Lemma 2. *Let G be a connected graph*, $\boldsymbol{s}$ *the characteristic vector of an algebraic set of G and* $\boldsymbol{c}$ *a solution of system* (2). *For every two distinct vertices u and v of G*, *one has*

$$c_v = p + (-1)^k c_u,$$

where p is an integer and k is the length of a path joining u and v.

Proof. Let $(v_0, v_1, \ldots, v_k)$ be a path with $v_0 = u$ and $v_k = v$. Since $\boldsymbol{c}$ is a solution of system (2), one has

$$c_{v_i} = s_{(v_{i-1}, v_i)} - c_{v_{i-1}} \quad (i = 1, \ldots, k)$$

so that one has

$$c_v = p + (-1)^k c_u,$$

where

$$p = s_{(v_{k-1}, v_k)} - s_{(v_{k-2}, v_{k-1})} + \cdots + (-1)^{k-1} s_{(v_0, v_1)}.$$

Since $\boldsymbol{s}$ is a binary vector, p is an integer, which proves the statement. $\square$

The following is a straight consequence of Lemma 2.

Corollary 2. *Let S be an algebraic set of a connected graph G, and $\boldsymbol{s}$ the characteristic vector of S. For every solution $\boldsymbol{c}$ of system* (2), *either each component of $\boldsymbol{c}$ is an integer or no component of $\boldsymbol{c}$ is an integer.*

Lemma 3. *Let G be a connected graph, and $\boldsymbol{s}$ the characteristic vector of an algebraic set of G.*
(A) *If G is bipartite, then there always exists an integral solution of system* (2).
(B) *If G contains a loop, then there exists exactly one solution of system* (2) *and it is integral.*
(C) *If G is not bipartite and loopless, then there exists exactly one solution of system* (2) *and it is either integral or half-integral* (*that is, each component of the solution is of the type $q/2$ where q is an odd integer*).

Proof. (A) By Corollary 1, one component, say c_u, of a solution $\boldsymbol{c}$ of system (2) can be assigned an arbitrary value. By taking $c_u = 0$, we obtain from Corollary 2 that all the components of $\boldsymbol{c}$ are integers.

(B) By Corollary 1, there is exactly one solution of system (2), say $\boldsymbol{c}$. If (w, w) is a loop of G, then c_w is equal to either 1 or 0 depending on whether the loop is or is not in S. By Corollary 2, all the components of $\boldsymbol{c}$ are integers.

(C) Let $(v_0, v_1, \ldots, v_k)$ be an odd cycle of G. Since v_0 and $v_k - 1$ are joined by a path of even length, by Lemma 2 one has $c_{v_{k-1}} = p + c_{v_0}$ for some integer p. It follows that

$$c_{v_k} = s_{(v_{k-1},v_k)} - c_{v_{k-1}} = q - c_{v_0},$$

where $q = s_{(v_{k-1},v_k)-p}$. Since $v_0 = v_k$, one has $c_{v_0} = q - c_{v_0}$ and, hence, $c_{v_0} = q/2$. If q is an even integer, then c_{v_0} is an integer and, by Corollary 2, all the components of $\boldsymbol{c}$ are integers. If q is an odd integer, then c_{v_0} is a half-integer and, by Lemma 2, for every vertex $v \neq v_0$, c_v is the sum of an integer with a half-integer, which proves that the components of $\boldsymbol{c}$ are all half-integers. □

Corollary 3. *A nonempty set S of loops in a connected graph G is an algebraic set of G if and only if $G - S$ is either an empty graph or a nonempty bipartite graph with bipartition (U, W) such that either $S \cap star(U)$ or $S \cap star(W)$ is empty.*

Proof. Let $\boldsymbol{s}$ be the characteristic vector of S.

(*only if*): Assume that S is algebraic and let (u, u) be a loop in S. By Lemma 3, the solution $\boldsymbol{c}$ of system (2) is integral and $c_u = 1$. By Lemma 2, for every vertex $v \neq u$, one has either $c_v = +1$ or $c_v = -1$ depending on whether u and v are joined by a path of even or odd length. Since c_v is never zero, there is no loop in $E(G) - S$. Moreover, if (v, v) is another loop in S, the length of every path joining u and v must be even. Finally, there is no vertex v such that there are two paths joining u and v one of even length and the other of odd length. To sum up, $G - S$ is bipartite.

(*if*): Assume that $G - S$ is bipartite. The statement is trivial if $G - S$ is empty; otherwise, let (U, W) be the bipartition of $G - S$. Without loss of generality, we can assume that $S \cap star(W) = \emptyset$. Then, we can obtain a solution $\boldsymbol{c}$ of system (2) by setting $c_u = 1$ for all $u \in U$, and setting $c_u = -1$ for all $v \in W$, which proves that S is an algebraic set. □

4. The signature induced by an algebraic set

Let S be an algebraic set of a connected graph G, and let $\boldsymbol{s}$ be the characteristic vector of S. The *signed support* of a solution $\boldsymbol{c}$ of system (2) is the ordered pair (P, N), where $P = \{v \in V(G) : c_v > 0\}$ and $N = \{v \in V(G) : c_v < 0\}$; in what follows, (P, N) will be referred to as a *signature* of G *induced* by S. Note that, by Corollary 1, if G is connected and is not bipartite, then there is exactly one signature of G induced by an algebraic set of G. Finally, by Fact 1, a signature of G induced by an algebraic set can be found in linear time.

Example 1. Consider the algebraic set $S = \{(u, x), (v, y)\}$ of the bipartite connected graph G shown in Fig. 4. The general solution of system (2), where $\boldsymbol{s}$ is the characteristic vector of S, is ($c_u = \lambda$, $c_v = \lambda - 1$, $c_x = -\lambda + 1$, $c_y = -\lambda + 2$), where λ is a real parameter. Therefore, there are seven possible signatures of G induced by S, of which five are for $\lambda > 0$ (see Fig. 5a), one for $\lambda = 0$ (see Fig. 5b) and one for $\lambda < 0$ (see Fig. 5c).

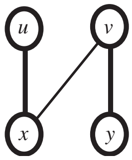

Fig. 4. An algebraic set of a bipartite graph.

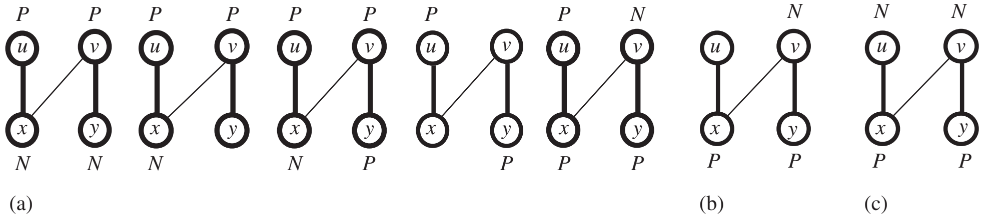

Fig. 5. Signatures of a bipartite graph induced by an algebraic set.

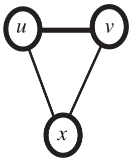

Fig. 6. An algebraic set of a nonbipartite graph.

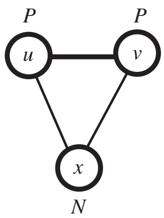

Fig. 7. The signature of a nonbipartite graph induced by an algebraic set.

Example 2. Consider the algebraic set $S = \{(u, v)\}$ of the nonbipartite connected graph G shown in Fig. 6. System (2), where $\boldsymbol{s}$ is the characteristic vector of S, has exactly one solution $\boldsymbol{c}$ with $c_u = c_v = \frac{1}{2}$ and $c_x = -\frac{1}{2}$. Therefore, there is only one signature of G induced by S (see Fig. 7). The next two lemmas state some graphical properties of signatures.

Lemma 4. *Let S be a nonempty algebraic set of a connected graph G. For every signature (P, N) of G induced by S, one has that*

(A) $P \neq \emptyset$ *and* $P \cap N = \emptyset$,
(B) *star*(P) *is the disjoint union of star* (N), $[P \cup N]_1$ *and* $[P]_2$,
(C) *star*$(P)-$*star*(N) *is a (possibly empty) subset of* S,
(D) *if star* $(P) \neq$ *star*(N), *then either* $[P \cup N]_1$ *or* $[P]_2$ *is empty*,
(E) *if* $V(G) = P \cup N$, *then* $G - S$ *is bipartite and either star*(P)–*star* (N) *is a set of loops or star* $(P)-$*star*$(N) = [P]_2$,
(F) *if* $V(G) \neq P \cup N$, *then*$[P]_2 = \emptyset$.

Proof. Let $\boldsymbol{s}$ be the characteristic vector of S, $\boldsymbol{c}$ any solution of system (2) and (P, N) the signed support of $\boldsymbol{c}$.

(A) $P \neq \emptyset$ since S is not empty, and P and N are trivially disjoint.

(B) For every edge (u, v) of G, since $c_u + c_v = s_{(u,v)} \in \{0, 1\}$, if $c_u < 0$ then $c_v > 0$, which proves that $star(N) \subseteq star(P)$. Moreover, if (u, v) is an edge in $star(P) - star(N)$ and u is in P, then either $v \in P$ or $v \notin P \cup N$. In the former case, $(u, v) \in [P \cup N]_1$ if $u = v$ and $(u, v) \in [P]_2$ otherwise; in the latter case, $(u, v) \in [P \cup N]_1$.

(C) The statement is trivially true if $star(P) = star(N)$. Let us assume that $star(P) \neq star(N)$, and let (u, v) be an edge in $star(P) - star(N)$. Without loss of generality, we can assume that u is in P so that $c_u > 0$ and $c_v \geqslant 0$. Therefore, $c_u + c_v > 0$ and, since $c_u + c_v = s_{(u,v)} \in \{0, 1\}$, one has $c_u + c_v = s_{(u,v)} = 1$, which proves that (u, v) is in S.

(D) Let us distinguish two cases depending on whether $star(P) - star(N)$ is or is not loopless.

Case 1: $star(P) - star(N)$ is loopless. Let (u, v) be any edge in $star(P) - star(N)$. Without loss of generality, we can assume that $u \in P$ so that $v \notin N$. By (C), $star(P) - star(N)$ is a subset of S so that $c_u + c_v = s_{(u,v)} = 1$. Since, by Corollary 2, either both c_u and c_v are integers or both c_u and c_v are half-integers, one has that either $c_u = 1$ and $c_v = 0$ or $c_u = c_v = \frac{1}{2}$. In the former case, $(u, v) \in [P \cup N]_1$; in the latter case, $(u, v) \in [P]_2$. It follows that, if $\boldsymbol{c}$ is integral, then $[P]_2$ is empty; otherwise, $[P \cup N]_1$ is empty.

Case 2: $star(P) - star(N)$ contains a loop, say (w, w). Then, $c_w = 1$ so that, by Lemma 3, $\boldsymbol{c}$ is integral. Therefore, for every two distinct vertices u and v in P, one has $c_u + c_v \geqslant 2$, which excludes the existence of an edge with endpoints u and v, that is, $[P]_2 = \emptyset$.

(E) If $V(G) = P \cup N$ then, by (C) of Lemma 4, $G - S$ is a subgraph of the graph $G - (star(P) - star(N))$ which is bipartite since its edge set is $star(N)$. Therefore, $G - S$ is bipartite. Moreover, the endpoints of every edge of G are in $P \cup N$. Therefore, $[P \cup N]_1$ is either a set of loops or an empty set. In the former case, $star(P) \neq star(N)$ and, by (D), $[P]_2 = \emptyset$ so that, by (B), $star(P) - star(N) = [P \cup N]_1$. In the latter case, by (B), $star(P) - star(N) = [P]_2$.

(F) If $V(G) \neq P \cup N$, then $c_w = 0$ for each vertex w in $V(G) - (P \cup N)$. By Corollary 2, $\boldsymbol{c}$ is integral. Therefore, if u and v are two distinct vertices in P, then both c_u and c_v are positive integers so that $c_u + c_v \geqslant 2$, which excludes the existence in G of the edge (u, v), which proves that $[P]_2 = \emptyset$. □

Lemma 5. *Let S be a nonempty algebraic set of a connected graph G.*

(A) *If there is a signature (P, N) of G induced by S such that $star(P) = star(N)$, then G is bipartite.*

(B) *If there is a signature (P, N) of G induced by S such that $[P]_2 \neq \emptyset$, then G is loopless, $G - S$ is bipartite and (P, N) is a bipartition of $G - S$.*

Proof. (A) Let $\boldsymbol{s}$ the characteristic vector of S, $\boldsymbol{c}$ a solution of system (2) such that $star(P) = star(N)$, where (P, N) is the signed support of $\boldsymbol{c}$. Then, for each edge (u, v) of G with $c_u > 0$ one has $c_v < 0$, so that there is no vertex w with $c_w = 0$ (for, otherwise, G would not be connected). Therefore, G is bipartite.

(B) Let $\boldsymbol{s}$ be the characteristic vector of S, $\boldsymbol{c}$ a solution of system (2) such that there is an edge (u, v) of G with $u \neq v$, $c_u > 0$ and $c_v > 0$. Since $c_u + c_v \in \{0, 1\}$, neither c_u nor c_v can be an integer so that $\boldsymbol{c}$ is not integral, which, by (B) of Lemma 3, excludes the existence of loops and of vertices w with $c_w = 0$. Therefore, $V(G) = P \cup N$ so that, by (E) of Lemma 4, $G - S$ is bipartite and (P, N) is a bipartition of $G - S$. □

5. Minimal algebraic sets

As said in Section 3, in light of Theorems 1 and 2, the problem of finding a graphical characterisation of minimal invariant sets reduces to the problem of finding a graphical characterisation of minimal algebraic sets. At the present, there exist a graphical characterisation of minimal algebraic sets of a bipartite graph (see Theorem 3), and a graphical characterisation of algebraic edges of an arbitrary graph (see Theorem 4).

Theorem 3 (*Kao [14], Malvestuto [19]*). *The minimal algebraic sets of a bipartite graph are exactly its bipartite bonds.*

Theorem 4 (*Malvestuto and Mezzini [20]*). *An edge of a graph is algebraic if and only if its removal creates one more bipartite component.*

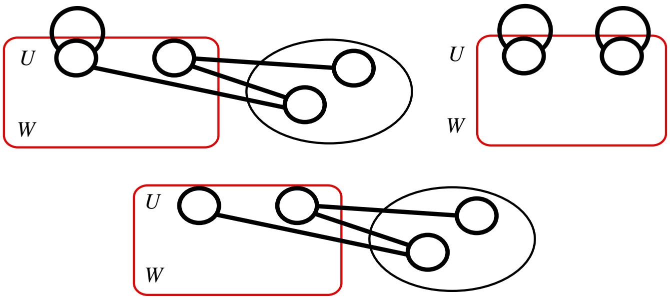

Fig. 8. Bip sets with $[U]_2 = \emptyset$.

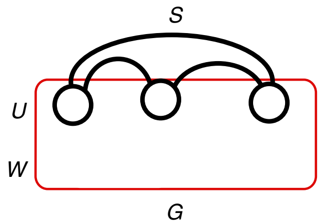

Fig. 9. A bip set with $[U \cup W]_1 = \emptyset$.

We shall state a graphical characterisation of minimal algebraic sets of a (connected) graph, which subsumes Theorems 3 and 4. Our characterisation is based on the notion of a "bip set", which is now introduced. (Note that our definition of a "bip set" is different from that used in [7].)

A nonempty subset S of $E(G)$ is a *bip set* of G if $S = star(U) - star(W)$, where U and W are subset of $V(G)$ such that:

(i) $U \neq \emptyset$ and $U \cap W = \emptyset$,
(ii) $star(U) \supseteq star(W)$, and
(iii) either $[U \cup W]_1 = \emptyset$ or $[U]_2 = \emptyset$.

Fact 2. *Let $S = star(U) - star(W)$ be a bip set. One has that*
(A) *$star(W)$ is a starset*;
(B) *if $[U]_2 = \emptyset$, then $S = [U \cup W]_1$ (see Fig.* 8);
(C) *if $[U \cup W]_1 = \emptyset$ then $S = [U]_2$ (see Fig.* 9).

A bip set of G is *minimal* if none of its proper subsets is a bip set of G. We shall show (see Theorem 5) that every minimal algebraic set is a minimal bip set and vice versa.

Lemma 6. *Let $S = star(U) - star(W)$ be a bip set of a connected graph G. Then, S is an algebraic set and (U, W) is a signature of G induced by S.*

Proof. If $[U]_2 = \emptyset$ (see Fig. 8) then, by (B) of Fact 2, the characteristic vector of S can be written as

$$\sum_{v \in U} \boldsymbol{g}_v - \sum_{v \in W} \boldsymbol{g}_v$$

and, hence, S is algebraic and (U, W) is a signature of G induced by S. If $[U \cup W]_1 = \emptyset$ (see Fig. 9) then, since G is connected, $V(G) = U \cup W$ and, by (C) of Fact 2, $S = [U]_2$. Therefore, the characteristic vector of S can be written as

$$\frac{1}{2} \sum_{v \in U} \boldsymbol{g}_v - \frac{1}{2} \sum_{v \in W} \boldsymbol{g}_v$$

and, hence, S is algebraic and (U, W) is a signature of G induced by S. $\square$

The next lemma shows that every nonempty algebraic set contains a bip set.

Lemma 7. *Let S be a nonempty algebraic set of a connected graph G. There exists a signature (P, N) of G induced by S such that star(P)–star(N) is a bip set and is contained in S.*

Proof. Let $\boldsymbol{s}$ be the characteristic vector of S. If G is not bipartite, then system (2) has exactly one solution, say $\boldsymbol{c}$. Otherwise, arbitrarily chosen an edge (u^*, v^*) in S, let $\boldsymbol{c}$ be the solution of system (2) with $c_{u*} = \frac{1}{2}$. (Note that the existence of such a solution is ensured by Corollary 1, and that $c_{v*} = \frac{1}{2}$). Let (P, N) be the signed support of $\boldsymbol{c}$. In order to prove that $star(P) - star(N)$ is a bip set contained in S, by (A–D) of Lemma 4, it is sufficient to show that $star(P) \neq star(N)$. If G is bipartite, then both u^* and v^* are in P so that $(u^*, v^*) \in star(P) - star(N)$ and, hence, $star(P) \neq star(N)$. If G is not bipartite then, by (A) of Lemma 5, $star(P) \neq star(N)$ for, otherwise, G would be bipartite. □

Corollary 4. *Let S be a minimal algebraic set of a connected graph G. There is a signature (P, N) of G induced by S such that S equals the bip set $star(P) - star(N)$.*

Proof. By Lemma 7, there is a signature (P, N) of G induced by S such that S contains the bip set $star(P) - star(N)$. Then, S must coincide with this bip set for, otherwise, since $star(P) - star(N)$ is an algebraic set by Lemma 6, S would not be a minimal algebraic set. □

Finally, we are in a position to state the following characterisation of minimal algebraic sets.

Theorem 5. *A minimal algebraic set is a minimal bip set and vice versa.*

Proof. Let S be a minimal algebraic set of graph G. We now prove that S is a minimal bip set of G. Since S is a minimal algebraic set of G, S is contained in the edge set of a component of G, say G'. By Corollary 4, there is a signature (P, N) of G' induced by S such that S equals the bip set $star(P) - star(N)$ of G'. Let S' be a minimal bip set contained in the bip set $star(P) - star(N)$. Since S' is an algebraic set by Lemma 6, S must coincide with S' for, otherwise, S would not be a minimal algebraic set, which proves that S is a minimal bip set.

Let S be a minimal bip set of graph G. We now prove that S is a minimal algebraic set. By Lemma 6, S is an algebraic set of G. Let S' be a minimal algebraic set contained in S. Since S' is a minimal algebraic set of G, S' is contained in the edge set of a component of G, say G'. By Corollary 4, there is a signature (P, N) of G' induced by S' such that S' equals the bip set $star(P) - star(N)$. Then, S must coincide with S' for, otherwise, S would not be a minimal bip set, which proves that S is a minimal algebraic set. □

6. Testing an algebraic set for minimality

In this section we provide a linear algorithm to recognise minimal algebraic sets of a connected graph G. In the next two subsections, we separately discuss two cases depending on whether or not G is bipartite.

6.1. The bipartite case

Recall that, by Theorem 3, minimal algebraic sets are bipartite bonds and vice versa.

Theorem 6. *Let G be a bipartite and connected graph. A nonempty algebraic set S of G is a minimal algebraic set if and only if $G - S$ has exactly two components.*

Proof. (*only if*). Let S be a minimal algebraic set of G. By Theorem 3, S is a bipartite bond and the statement holds.

(*if*): Let S be a nonempty algebraic set of G, and assume that $G - S$ has exactly two components. Since S is a nonempty algebraic set, there exists a minimal algebraic set S' contained in S. Suppose, by contradiction, that $S \neq S'$. Since, by Theorem 3, S' is a (bipartite) bond, $G - S'$ has exactly two components, say G' and G''. On the other hand, the set $S - S'$ is a subset of the edge set of $G - S'$ and, since $S - S'$ is a nonempty algebraic set, by Theorem 3 there is

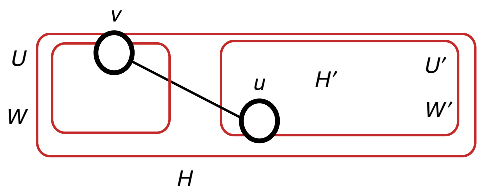

Fig. 10.

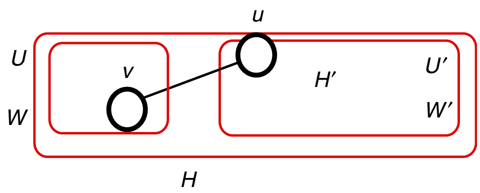

Fig. 11.

a bipartite bond of $G - S'$ contained in $S - S'$. Such a bond of $G - S'$ should be a bond of either G' or G''. But, then $G - S$ should have at least three components (contradiction). □

Theorem 6 provides a test for minimality in the bipartite case, which only requires counting the components of $G - S$. It is then easy to see that this minimality test runs in linear time.

6.2. The nonbipartite case

We need the following technical lemma.

Lemma 8. *Let $S = star(U) - star(W)$ be a bip set, H the bipartite subgraph of $G - S$ with $V(H) = U \cup W$ and $E(H) = star(W)$, H' a component of H, $U' = U \cap V(H')$ and $W' = W \cap V(H')$. If $[U']_2 = \emptyset$, then $S \cap star(V(H'))$ is a bip set.*

Proof. If H is connected, then $S = S \cap star(V(H))$ so that the statement is trivially true. Let us assume that H is not connected. Let H' be a component of H, $U' = U \cap V(H')$ and $W' = W \cap V(H')$. Since $star(W) \subseteq star(U)$ and $W' \subseteq W$, $star(W') \subseteq star(U)$. On the other hand, for each edge (u, v) with $u \in W'$, v cannot belong to $U - U'$ for, otherwise, $(u, v) \notin S$ and H' would not be a component of H (see Fig. 10). Therefore, $star(W') \subseteq star(U')$. Moreover, $star(U') \neq star(W')$ for, otherwise, $H' = H$ and H would be connected. So, since $[U']_2 = \emptyset$ by hypothesis, $star(U') - star(W')$ is a bip set. What remains is to prove that $S \cap star(V(H')) = star(U') - star(W')$. Of course, $S \cap star(V(H'))$ is a subset of $star(U') - star(W')$. Suppose, by contradiction, that there is an edge (u, v) in $star(U') - star(W')$ but not in S. Then, $u \in U'$ and $v \in W - W'$ and, hence, H' would not be a component of H (see Fig. 11). It follows that $S \cap star(V(H'))$ coincides with $star(U') - star(W')$ and, hence, is a bip set. □

The following two conditions (a) and (b) will be used to characterise minimal algebraic sets (see Theorem 7). Recall that there is exactly one signature of a nonbipartite connected graph induced by an algebraic set. Let S be a nonempty algebraic set of a nonbipartite and connected graph G, and let (P, N) be the signature of G induced by S.

Condition (a)
$V(G) = P \cup N$ and

(a1) *either S is a set of loops, or*
(a2) *$S = [P]_2$ and no component of $G - S$ is an induced subgraph of G.*

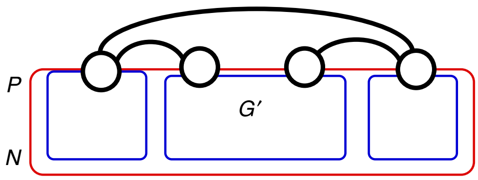

Fig. 12.

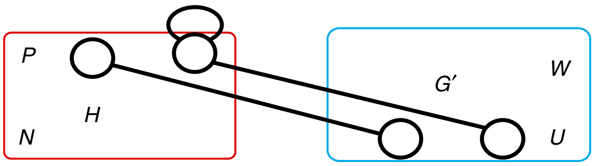

Fig. 13.

Condition (b)
$V(G) \neq P \cup N$ and
(b1) $S = [P \cup N]_1$,
(b2) *the graph $H = (P \cup N,\ star(N))$ is connected,*
(b3) *for each bipartite component $G' \neq H$ of $G - S$, G' is not an empty graph and the bipartition (U, W) of G' is such that neither $S \cap star(U)$ nor $S \cap star(W)$ is empty,*
(b4) *if S is loopless, then $G - S$ is not bipartite.*

Lemma 9. *Let G be a nonbipartite and connected graph, S a minimal algebraic set of G, and (P, N) the signature of G induced by S. Then, either condition* (a) *or condition* (b) *holds.*

Proof. First of all, observe that, by Corollary 4, S coincides with the bip set $star(P) - star(N)$. Distinguish two cases depending on whether or not $V(G) = P \cup N$.

Case 1: $V(G) = P \cup N$. By (E) of Lemma 4, either S is a set of loops or $S = [P]_2$. If S is a set of loops, then condition (a1) holds. Let us assume that $S = [P]_2$. By (B) of Lemma 5, G is loopless and $G - S$ is bipartite with bipartition (P, N). Suppose, by contradiction, that there is a component G' of $G - S$ that is an induced subgraph of G (see Fig. 12). By Lemma 8, the set $S' = S \cap star(V(G'))$ is a bip set and, by Lemma 6, is an algebraic set. Moreover, since G is not bipartite, S' must be properly contained in S. Therefore, S' is a nonempty algebraic set that is properly contained in S, which contradicts the minimality of S. To sum up, condition (a) holds.

Case 2: $V(G) \neq P \cup N$. We show that condition (b) holds by proving (b1)–(b4).

(b1) By (F) of Lemma 4, $[P]_2 = \emptyset$ so that, by (B) of Lemma 4, one has that $S = star(P) - star(N) = [P \cup N]_1$.

(b2) Suppose, by contradiction, that H is not connected and let H' be a component of H. Let $S' = S \cap star(V(H'))$. By Lemma 8, S' is a bip set and, by Lemma 6, is an algebraic set. Moreover, since H is not connected, S' must be properly contained in S. Therefore, S' is a nonempty algebraic set that is properly contained in S, which contradicts the minimality of S.

(b3) Suppose, by contradiction, that there is a bipartite component $G' \neq H$ of $G - S$ such that either G' is empty or the bipartition (U, W) of G' is such that either $S \cap star(U)$ or $S \cap star(W)$ is empty. In the latter case, without loss of generality, we assume that $S \cap star(W) = \emptyset$ (see Fig. 13). Let $S' = S \cap star(V(G'))$. Note that, since G is connected, $S' \neq \emptyset$. Moreover, if G' is an empty graph, then S' is a star and, hence, a bip set; otherwise, since $S \cap star(W) = \emptyset$, one has that $S' = star(U) - star(W)$ and that, since $S' = [U \cup W]_1$ and $S' \subseteq S$, $[U]_2 = \emptyset$, which makes S' a bip set. In both cases, by Lemma 6, S' is an algebraic set. Moreover, since G is not bipartite, S' must be properly contained in S. To sum up, S' is a nonempty algebraic set that is properly contained in S, which contradicts the minimality of S.

(b4) Suppose, by contradiction, that S is loopless and $G - S$ is bipartite. Let (U, W) be a bipartition of $G - S$ such that $P \subseteq U$ and, hence, $N \subseteq W$ (see Fig. 14). Let $S' = star(U) - star(W)$. Note that, since G is not bipartite, S' is

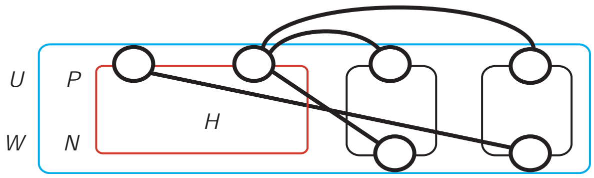

Fig. 14.

not empty. Moreover, since $P \neq \emptyset$, U is not empty. Observe that, for each edge (u, v) in S, if $u \in W$ then $v \in P$; moreover, since $G - S$ is bipartite, for each edge (u, v) of $G - S$, if $u \in W$ then $v \in U$. Therefore, for each edge (u, v) of G, if $u \in W$ then $v \in U$, and, hence, $star(W) \subseteq star(U)$. Indeed, S' is a bip set owing to the fact that, since G is not bipartite, $[U]_2 \neq \emptyset$ and, since $V(G) = U \cup W$ and S is loopless, $[U \cup W]_1 = \emptyset$. By Lemma 6, S' is a nonempty algebraic set. Finally, every edge in $S'(= [U]_2)$ has one endpoint in P and the other in $V(G) - (P \cup N)$, so that, since $S = [P \cup N]_1$, every edge in S' belongs to S. To sum up, S' is a nonempty algebraic set that is properly contained in S, which contradicts the minimality of S. $\square$

We now show that the converse of Lemma 9 also holds.

Lemma 10. *Let G be a nonbipartite and connected graph, S a nonempty algebraic set of G, and (P, N) the signature of G induced by S. Then, S is a minimal algebraic set if $S = star(P) - star(N)$ and either condition* (a) *or condition* (b) *holds.*

Proof. Let us assume that $S = star(P) - star(N)$.

We first prove that condition (a) is sufficient for S to be a minimal algebraic set. Let us assume that $V(G) = P \cup N$ and that (a1) or (a2) holds.

(a1) Assume that S is a set of loops. By Corollary 3, no proper nonempty subset of S is algebraic and, hence, S is a minimal algebraic set.

(a2) Assume that $S = [P]_2$ and no component of $G - S$ is an induced subgraph of G. Let S' be a minimal algebraic set contained in S. We shall prove that $S' = S$. First of all, note that, by Theorem 5, S' is a (minimal) bip set, say $star(U) - star(W)$. Then, by Lemma 6, (U, W) is the signature of G induced by S'. We first prove that (i) $[U \cup W]_1 = \emptyset$ and, then, (ii) $S' = S$.

Proof of (i). Since S is not empty and $S = [P]_2$, by (B) of Lemma 5, G is loopless and, hence $[U \cup W]_1$ is loopless. Suppose by contradiction that $[U \cup W]_1 \neq \emptyset$. Let H' be the bipartite graph with $V(H') = U \cup W$ and $E(H') = star(W)$. Note that, since $[U \cup W]_1 \neq \emptyset$ (and, hence, $[U]_2 = \emptyset$) and $[U \cup W]_1$ is loopless, H' is an induced subgraph of G and, since S' is a minimal algebraic set, H' is connected by Lemma 9. Since S' is a subset of S, there exists a component G' of $G - S$ that is a subgraph of H'. Moreover, since no component of $G - S$ is an induced subgraph of G, there is an edge e in S whose endpoints are both in $V(G')$. Therefore, G' is a nonempty bipartite graph with bipartition (P', N'), where $P' = P \cap V(G')$ and $N' = N \cap V(G')$, and the endpoints of e are both in P' so that the graph $G' + e$ is not bipartite (see Fig. 15). Finally, since G' is a subgraph of H' and since H' is an induced subgraph of G, $G' + e$ is a subgraph of H' so that H' is not bipartite (contradiction). Therefore, $[U \cup W]_1$ must be empty.

Proof of (ii). Since S' is a bip set with $[U \cup W]_1 = \emptyset$, one has that $S' = [U]_2$ and $G - S'$ equals the bipartite graph $H' = (U \cup W, star(W))$. Therefore, since each component of $G - S$ is a subgraph of some component of $G - S'$, each edge in S having both endpoints in the same component of $G - S$ must belong to S' (for, otherwise, $G - S'$ would not be bipartite). We now prove that also each edge in S whose endpoints are in distinct components of $G - S$ must belong to S', which completes the proof that $S' = S$. Suppose, by contradiction, that there is an edge (u, v) in $S - S'$ such that u and v are in two distinct components of $G - S$, say G_1 and G_2, respectively. Since no component of $G - S$ is an induced subgraph of G, there is an edge (u_1, v_1) in S having both endpoints in G_1 and there is an edge (u_2, v_2) in S having both endpoints in G_2 (see Fig. 16). Since (u, v) does not belong to S' and since S' is a subset of S, there is a component G' of the bipartite graph $G - S'$ containing both G_1 and G_2. Note that G' has the bipartition (U', W'),

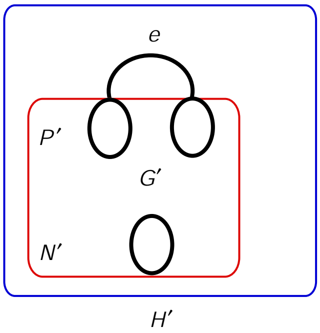

Fig. 15.

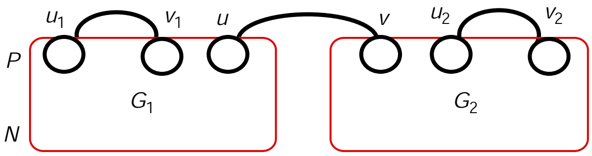

Fig. 16.

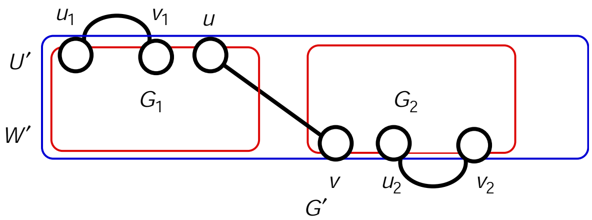

Fig. 17.

where $U' = U \cap V(G')$ and $W' = W \cap V(G')$, so that, since $S' = [U]_2$, one has $S' \cap star(W') = \emptyset$. Since (u_1, v_1) and (u_2, v_2) are both in S' (see above), the vertices u_1 and u_2 are joined in G' by a path of odd length (see Fig. 17), which implies that either (u_1, v_1) or (u_2, v_2) is in $S' \cap star(W')$, which contradicts the fact that $S' \cap star(W') = \emptyset$.

We now prove that condition (b) is sufficient for S to be a minimal algebraic set. Let us assume that $V(G) \neq P \cup N$ and that (b1)–(b4) hold. Let S' be a minimal algebraic set that is contained in S. We shall prove that $S' = S$. First of all, observe that, by Theorem 5, S' is a minimal bip set, say $S' = star(U) - star(W)$. Then, by Lemma 6, (U, W) is the signature of G induced byS'. Let us distinguish two cases depending on whether S is or is not loopless.

Case 1: S is loopless. By hypothesis (b4), $G - S$ is not bipartite and, since S' is a subset of S, $G - S'$ is not bipartite. Therefore, $[U]_2 = \emptyset$ (for, otherwise, $G - S'$ would be bipartite by (B) of Lemma 5) and, hence, $S' = [U \cup W]_1$. Let H' be the bipartite graph with $V(H') = U \cup W$ and $E(H') = star(W)$. In order to prove that $S' = S$, it is sufficient to show that $H' = H$, where H is the connected graph involved in condition (b2), that is, $H = (P \cup N, star(N))$. First of all, note that, since S' is a minimal algebraic set, H' is a component of $G - S'$ by Lemma 9, and, since S (and, hence, S') is loopless, H' is the subgraph of G induced by $U \cup W$. Moreover, since S' is a subset of S, H' contains at least one component of $G - S$. Finally, since H' is bipartite, H' can contain only bipartite components of $G - S$. In order to prove that $H' = H$, suppose, by contradiction, that H' contains a bipartite component $G' \neq H$ of $G - S$. By condition (b3), G' is not an empty graph. Let (U', W') be the bipartition of G'. Note that, since G' is contained in H', either U' or W' is contained in W so that, since $S' = star(U) - star(W) = [U \cup W]_1$, if $U' \subseteq W$ then no edge in $star(U')$ is in S'; otherwise (that is, if $W' \subseteq W$), no edge in $star(W')$ is in S'. Again by condition (b3), neither $S \cap star(U')$ nor $S \cap star(W')$ is empty (see Fig. 18). But, then, H' should also contain H. Moreover, since H' is an induced subgraph of G, H' should also contain all the edges in $S \cap star(V(G'))$ (see Fig. 19), which would make H' a nonbipartite graph

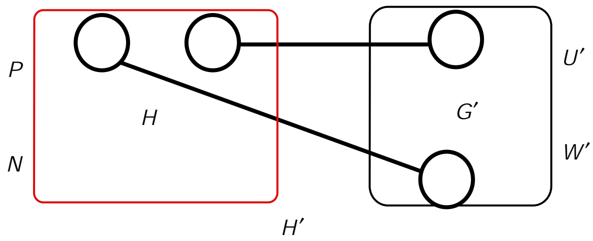

Fig. 18.

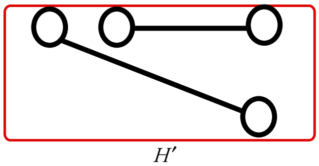

Fig. 19.

(contradiction). Therefore, H' contains no bipartite component $G' \neq H$ of $G - S$ and, since H' contains at least one bipartite component of $G - S$, one has $H' = H$ and, hence, $S' = S$.

Case 2: S contains loops. In this case, $[U]_2 = \emptyset$ (for, otherwise, G would be loopless by (B) of Lemma 5) and, hence, $S' = [U \cup W]_1$. Moreover, S' cannot be a set of loops for, otherwise, $G - S'$would not be bipartite by hypothesis (b3) and, by Corollary 3, S' would not be an algebraic set. On the other hand, using the same argument as in Case 1, the bipartite graph $H' = (U \cup W,\ star(W))$ is an induced subgraph of G and contains H, so that S' contains all the loops in S for, otherwise, H' would not be bipartite. To sum up, S' must contain all the loops in S and some further edges in S. Using the same argument as in Case 1, one can prove that $H' = H$ which implies that $S' = S$. □

Combining Lemmas 9 and 10 we obtain the following.

Theorem 7. *Let G be a nonbipartite and connected graph, S a nonempty algebraic set of G, and (P, N) the signature of G induced by S. Then, S is a minimal algebraic set if and only if $S = star(P) - star(N)$ and either condition* (a) *or condition* (b) *holds.*

Proof. (*only if*). By Lemma 9. (*if*) By Lemma 10. □

Finally, we prove that one can decide in linear time whether a given set of edges of a connected graph is or is not a minimal algebraic set.

Theorem 8. *Given a connected graph G and a subset S of $E(G)$, one can decide whether S is or is not a minimal algebraic set of G in time linear in the size of G.*

Proof. Let $n = |V(G)|$ and $m = |E(G)|$. By Fact 1, one can decide whether S is or is not an algebraic set in $O(m+n)$ time. Let us assume that S is an algebraic set. If G is bipartite then, by Theorem 6, one can decide in $O(m+n)$ time whether or not S is a minimal algebraic set. Let us assume that G is not bipartite. In $O(1)$ time one can check that $S \neq \emptyset$. Moreover, by Fact 1, one can find the signature (P, N) of G induced by S in $O(m+n)$ time. At this point, we can use Theorem 7 and proceed as follows. Initially, each vertex in P is assigned the sign "+" and each vertex in N the sign "−"; moreover, we "mark" the endpoints of each edge in S. Next, we test condition $V(G) = P \cup N$, which requires $O(n)$ time. We now prove that

(i) Condition (a) can be tested in $O(m+n)$ time.
(ii) Condition (b) can be tested in $O(m+n)$ time.

Proof of (i). In O(m) time one can check whether or not S is a set of loops, as prescribed by (a1). As to (a2), again in O(m) time one can check that, for each edge e in S, the endpoints of e are both signed by "+", that is, $S = [P]_2$. Moreover, the components of $G - S$ can be constructed in O($m + n$) time. Let k be their number; then, we order them using a progressive number from 1 to k and, for each component G' of $G - S$, we label each vertex v of G' by the order number of G'. Let $K := \{1, \ldots, k\}$. After finding the components of $G - S$, we scan S. When an edge (u, v) in S is examined, if the labels of u and v are the same, then we set $K := K - \{h\}$, where h is the label of u. Note that each component of $G - S$ contains the endpoints of at least one edge in S if and only if the final value of K is the empty set. It is then easy to see that checking (a2), takes O($m + n$) time. To sum up, condition (a) can be tested in O($m + n$) time.

Proof of (ii). In O(m) time one can check that, for each edge e in S, either e is a loop or exactly one endpoint of e is a signed vertex (by "+"). Next, using a traversal of $G - S$ with start point at a signed vertex, in O($m + n$) time one can test condition (b2), find the bipartite components of $G - S$, and test condition (b4). As to (b3), note that

- a bipartite component G' of $G - S$ is an empty graph if and only if $V(G')$ is a singleton, and
- for a nonempty bipartite graph G' with bipartition, one has $S \cap star(U) \neq \emptyset$ (or $S \cap star(W) \neq \emptyset$) if and only if there is a marked vertex in U (in W, respectively).

Therefore, condition (b3) can be tested in O($m + n$) time. □

Before closing this section, we discuss the algorithmic implications of Theorem 8 for the problem of recognising minimal invariant sets.

Let $(G, \boldsymbol{a})$ be a vertex-weighted graph over R. By Theorem 1, every a minimal invariant set of $(G, \boldsymbol{a})$ is a minimal algebraic set and vice versa. Therefore, given a subset S of $E(G)$, by Theorem 8 one can decide in linear time whether or not S is a minimal invariant set of $(G, \boldsymbol{a})$.

Let $(G, \boldsymbol{a})$ be a vertex-weighted graph over R_+, and let Z be the set of its zero invariant edges. By Theorem 2, a subset S of $E(G)$ is a minimal invariant set of $(G, \boldsymbol{a})$ if and only if either $S = \{e\}$ for some e in Z, or $S \cap Z = \emptyset$ and S is a minimal algebraic set of $G - Z$. As we recalled in Section 3, the set Z can be found (once and for all) in cubic time. So, given Z, by Theorem 8 one can decide in linear time whether or not S is a minimal invariant set of $(G, \boldsymbol{a})$.

7. A complete axiomatisation

In the previous sections, we gave a graphical characterisation of invariant sets and of minimal invariant sets of a vertex-weighted graph. In this section we first present an axiomatisation and, then, address a problem raised by Kao [14,18] which consists in finding a collection of minimal invariant sets that completely characterise the whole family of invariant sets. For the sake of simplicity, we only consider vertex-weighted graphs over R. (It does not need much effort [21] to adjust our results to vertex-weighted graphs over R_+). Let $(G, \boldsymbol{a})$ be a vertex-weighted graph over R. By Proposition 1′ (and Theorem 1), a subset of $E(G)$ is an invariant set (or a minimal invariant set) of $(G, \boldsymbol{a})$ if and only if it is an algebraic set (a minimal algebraic set, respectively) of G. The following result provides an insight into the structure of the family of algebraic sets of G, which without loss of generality we assume to be connected.

Theorem 9. *Let G be a connected graph. If G is bipartite or contains loops, then the family of algebraic sets of G is the smallest family containing stars, and is closed under disjoint union and proper difference. If G is not bipartite and is loopless, then the family of algebraic sets of G is the smallest family containing stars, and is closed under disjoint union, proper difference and complementation.*

Proof. If G is bipartite then, by Theorem 3, the minimal algebraic sets are exactly the bipartite bonds of G. Therefore, every nonempty algebraic set is a disjoint union of bipartite bonds. The statement then follows from the fact every bipartite bond is either a star or can be obtained as a proper difference of two distinct starsets.

If G contains a loop then, by Theorem 5, every minimal algebraic set is a minimal bip set, say $star(P) - star(N)$, with $[P]_2 = \emptyset$ and, since $star(P)$ is a starset, $star(P) - star(N)$ is the proper difference of two starsets. Therefore, every nonempty algebraic set is a disjoint union of proper differences of starsets.

If G is not bipartite and is loopless then, by Theorem 5, every minimal algebraic set is a minimal bip set, say $star(P) - star(N)$. If $[P]_2 = \emptyset$, then $star(P) - star(N)$ is the proper difference of two starsets (see above); otherwise (that is, if $[P \cup N]_1 = \emptyset$), one has that $V(G) = P \cup N$ and $E(G) = star(P)$ so that $star(P) - star(N)$ equals the complement of the starset $star(N)$. Therefore, every nonempty algebraic set is a disjoint union either of proper differences of starsets or of complements of starsets. □

From Theorem 9 we get the following complete axiomatisation of invariant sets of vertex-weighted graph $(G, \boldsymbol{a})$ over R, where G is either bipartite or contains loops.

(I) The star of a vertex is an invariant set.
(II) If S and S' are two invariant sets and $S' \subseteq S$, then $S - S'$ is an invariant set.
(III) If S and S' are two disjoint invariant sets, then $S \cup S'$ is an invariant set.

If G is not bipartite and contains no loops, then axioms (I)–(III) with the addition of the following axiom provide a complete axiomatisation of invariant sets.

(IV) If S is an invariant set, then the complement of S (i.e., $E(G) - S$) is an invariant set.

Consider now the family of minimal invariant sets. Since there may be an exponential number of minimal invariant sets [14,23], Kao [14,18] raised the problem of finding a collection of minimal invariant sets that completely characterise the whole family of invariant sets. By Theorem 9, a family of minimal invariant sets that completely characterise the whole family of invariant sets can obtained by collecting, for each vertex v of G, the minimal algebraic sets contained in $star(v)$. As proved by Kao [14,18], if G is bipartite then such a family of minimal algebraic sets can be obtained in time linear in the size of G. We shall solve the same problem in the case that G is not bipartite by giving a polynomial procedure (see the "decomposition algorithm'') for finding a partition of $star(u)$ into minimal algebraic sets. It is based on the following result.

Corollary 5. *Let G be a nonbipartite and connected graph, and let u be a vertex of G. The star of u is a minimal algebraic set of G if and only if*

(b′) *for each bipartite component G' of $G - u$, one has that*
  (b′.1) *G' is not an empty graph, and*
  (b′.2) *if (U, W) is the bipartition of G', then neither $star(u) \cap star(U)$ nor $star(u) \cap star(W)$ is empty*;
(b″) *if $star(u)$ is loopless, then $G - star(u)$ is not bipartite.*

Proof. Of course, $star(u)$ is an algebraic set of G. Moreover, the signature (P, N) of G induced by $star(u)$ has $P = \{u\}$ and $N = \emptyset$. Then, the statement follows from condition (b) of Theorem 7. □

In our decomposition algorithm for $star(u)$, we first test condition (b′) and, then, condition (b″). Moreover, for each bipartite component G' of $G - u$ for which condition (b′.1) or condition (b′.2) does not hold, we construct a nonempty algebraic subset of $star(u)$. Finally, when we test condition (b″), if it does not hold, then we construct two disjoint algebraic subsets of $star(u)$. In Theorem 10, we shall prove that the algebraic subsets of $star(u)$ found by the decomposition algorithm form a partition of $star(u)$ into minimal algebraic sets.

Decomposition Algorithm.
(1) $\boldsymbol{D} := \emptyset$, $S := \emptyset$.
(2) Mark all the vertices v such that (u, v) is an edge of G.
(3) Find the components of $G - u$. For each bipartite component G' of $G - u$,
  if G' is an empty graph and $V(G') = \{v\}$, then add $star(v)$ to $\boldsymbol{D}$ and set $S := S \cup star(v)$;
  otherwise, let U' be a set in the bipartition of G' that contains marked vertices; if $V(G') - U'$ contains no marked vertices, then
  add $star(u) \cap star(U')$ to $\boldsymbol{D}$, and set $S := S \cup star(u) \cap star(U')$
(4) Set $S := star(u) - S$. If S contains a loop or $G - S$ is not bipartite, then add S to $\boldsymbol{D}$; otherwise, find a bipartition (U, W) of $G - S$ such that U contains u, and add to $\boldsymbol{D}$ the two sets $[U]_2$ and $S - [U]_2$ of edges of G.

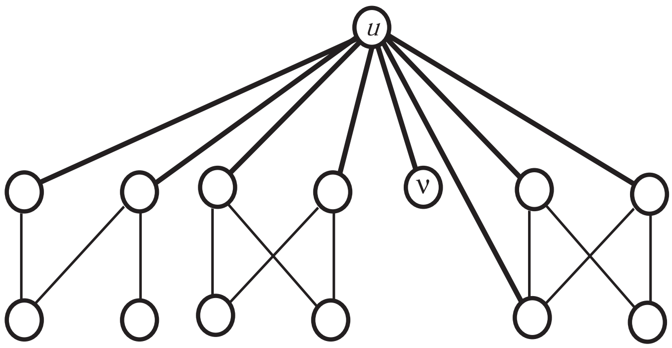

Fig. 20. A loopless nonbipartite graph G.

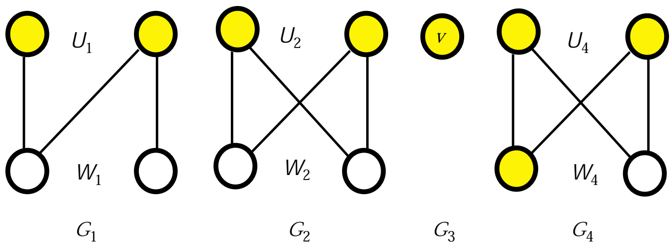

Fig. 21. The (bipartite) components of $G - u$.

Example 3. Consider the nonbipartite and loopless graph shown in Fig. 20. We apply the decomposition algorithm to find a partition $\boldsymbol{D}$ of the star of the vertex u into minimal algebraic sets.

Step 1: $\boldsymbol{D} := \emptyset$, $S := \emptyset$.
Step 2: The vertices adjacent to v are marked.
Step 3: $G - u$ has four components, each of which is bipartite (see Fig. 21).
G_1 is a nonempty graph with bipartition (U_1, W_1). Since U_1 contains marked vertices and W_1 contains no marked vertices, we add $star(u) \cap star(U_1)$ to $\boldsymbol{D}$ which becomes

$$\boldsymbol{D} = \{star(u) \cap star(U_1)\};$$

moreover, $S = star(u) \cap star(U_1)$.
G_2 is a nonempty graph with bipartition (U_2, W_2). Since U_2 contains marked vertices and W_2 contains no marked vertices, we add $star(u) \cap star(U_2)$ to D which becomes

$$\boldsymbol{D} = \{star(u) \cap star(U_1), star(u) \cap star(U_2)\};$$

moreover, $S = star(u) \cap star(U_1) \cup star(u) \cap star(U_2)$.
G_3 is an empty graph. We add $star(v)$ to D which becomes

$$\boldsymbol{D} = \{star(u) \cap star(U_1), star(u) \cap star(U_2), star(v)\};$$

moreover, $S = star(u) \cap star(U_1) \cup star(u) \cap star(U_2) \cup star(v)$.
G_4 is a nonempty graph with bipartition (U_4, W_4). Since both U_4 and W_4 contains marked vertices, we do not modify $\boldsymbol{D}$ and S.
Step 4: $S := star(u) - S$ (see Fig. 22).
Since S contains no loops and $G - S$ is bipartite, then we find a bipartition (U, W) of $G - S$ such that $u \in U$. We can take (see Fig. 21) either

$$U = \{u\} \cup W_1 \cup W_2 \cup W_4 \quad \text{and} \quad W = \{v\} \cup U_1 \cup U_2 \cup U_4,$$

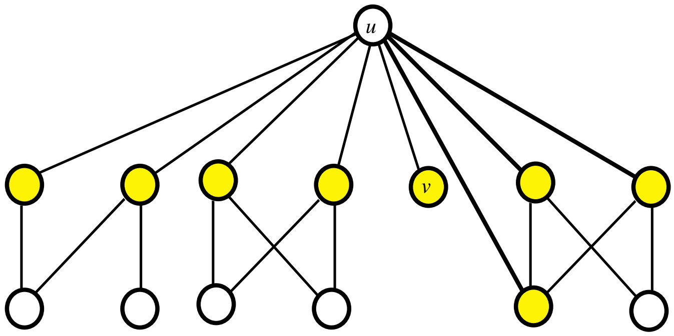

Fig. 22. The subset S of $star(u)$.

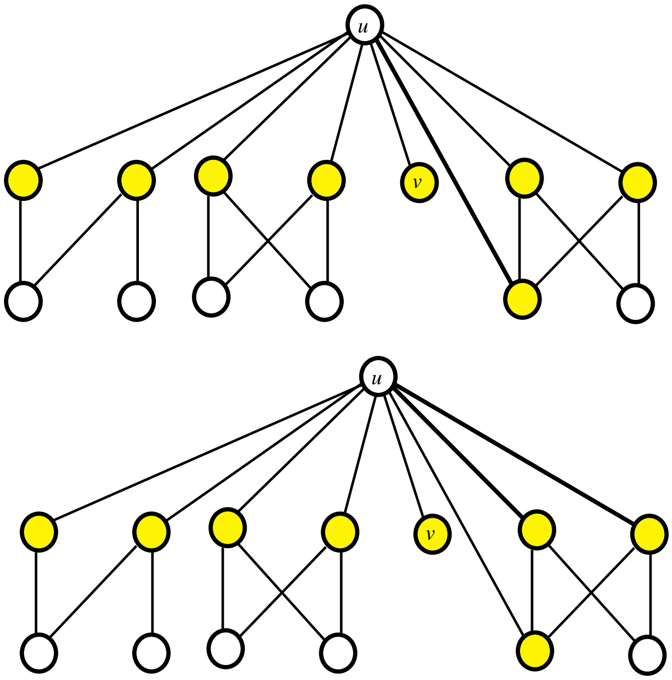

Fig. 23. A decomposition of S.

or

$$U = \{u\} \cup W_1 \cup W_2 \cup U_4 \quad \text{and} \quad W = \{v\} \cup U_1 \cup U_2 \cup W_4.$$

In both cases, we add to $\boldsymbol{D}$ both the two subsets of S shown in Fig. 23.

Theorem 10. *The decomposition algorithm correctly finds a partition of a star into minimal algebraic sets.*

Proof. Let u be a vertex of G. If $star(u)$ is a minimal algebraic set then, by Theorem 7, after performing step 3 we find $\boldsymbol{D} = \emptyset$ and $S = \emptyset$, and after performing step 4 we find $\boldsymbol{D} = \{S\}$ where $S = star(u)$, which proves the correctness of the result of the algorithm. Assume that $star(u)$ is not a minimal algebraic set. By Corollary 5, condition (b′) or condition (b″) does not hold. Assume that condition (b′) does not hold. Then, there is a bipartite component G' of $G - u$ for which either (b′.1) or (b′.2) does not hold. Let us distinguish the two cases.

Case 1: Let G' be an empty bipartite component of $G - u$ and let $V(G') = \{v\}$. Then, $star(v) = \{(u, v)\}$ and, by Theorem 4, (u, v) is an algebraic edge since its removal creates one bipartite component (namely, G'). Therefore, $star(v)$ is a minimal algebraic set which is contained in $star(u)$. So, $star(v)$ is correctly added to $\boldsymbol{D}$.

Case 2: Let G' be a nonempty bipartite component with bipartition (U', W') such that W' contains no marked vertices, that is, $star(u) \cap star(U') \neq \emptyset$ and $star(u) \cap star(W') = \emptyset$. Let $S' = star(u) \cap star(U')$. Of course, S' is properly contained in $star(u)$. Moreover, since $star(u) \cap star(W') = \emptyset$, S' coincides with the bip set $star(U') - star(W')$. Therefore, S' is an algebraic set of G and (U', W') is the signature of G induced by S'. Using condition (b) of Theorem 7, we now prove that S' is a minimal algebraic set of G. First of all, $V(G) \neq V(G')$ and $S' = [U' \cup W']_1$ so that condition (b1) holds. Moreover, G' is connected so that condition (b2) holds. Finally, $G - S'$ has exactly two components, G' and the subgraph G'' of G induced by $V(G) - V(G')$, and G'' is not bipartite (for, otherwise, G would be bipartite) so that conditions (b3) and (b4) hold. Therefore, condition (b) of Theorem 7 holds and, hence, S' is a minimal algebraic set. So, S' is correctly added to $\boldsymbol{D}$.

Let S be the subset of $star(u)$ obtained by deleting all the algebraic subsets of $star(u)$ found at step 3. Note that S is not empty for, otherwise, G would be bipartite. Moreover, since the algebraic subsets of $star(u)$ found at step 3 are pairwise disjoint, S is an algebraic set. Let (P, N) be the signature of G induced by S. Then, it is easy to see that $V(G) \neq P \cup N$ and, by construction, parts (b1)–(b3) of condition (b) in Theorem 7 hold. Therefore, if S contains a loop or $G - S$ is not bipartite, then condition (b) of Theorem 7 holds and, hence, S is a minimal algebraic set, which proves that S is correctly added to $\boldsymbol{D}$. Assume that S contains no loops and $G - S$ is bipartite. By Theorem 7, S is not a minimal algebraic set. Then, for each bipartite component G' of $G - S$ that does not contains u, one has that G' is not an empty graph and $S \cap star(U') \neq \emptyset$ and $S \cap star(W') \neq \emptyset$, where (U', W') is the bipartition of G'. Let (U, W) be a bipartition of $G - S$ such that $u \in U$, and let $S' = [U]_2$ and $S'' = S - S' = \{(u, v) : v \in W\}$. Then, S' is the bip set $star(U) - star(W)$ and, hence, is algebraic; moreover, (U, W) is the signature of G induced by S'. Since S' contains no loops and $G - S'$ is connected, by (a) of Theorem 7, S' is a minimal algebraic set. As to S'', it is an algebraic set for it is the proper difference of two algebraic sets. Let $U^* = W \cup \{u\}$ and $W^* = V(G) - U^*$. Then, it is easy to see that (U^*, W^*) is the signature of G induced by S''. Since S'' contains no loops and $G - S''$ is connected, again by (a) of Theorem 7, S'' is a minimal algebraic set. So, S' and S'' are correctly added to $\boldsymbol{D}$. $\square$

By Theorem 10, we can find a partition of a star into minimal algebraic sets in time polynomial in the size of the underlying graph.

8. Conclusion

Invariant sets of a vertex-weighted graph (and, more in general, of a vertex-weighted hypergraph) represent the amount of information released when the weights of its vertices are made public. They do coincide with algebraic sets in the case that weights are reals, and are closely related to algebraic sets in the case that weights are nonnegative reals. We stated a graphical characterisation of minimal invariant sets, and a complete axiomatisation of invariant sets. Future research is required to find analogous results for the case that weights are nonnegative integers, or to consider invariant sets of a vertex-weighted hypergraph.

References

[1] R.K. Ahuja, T.L. Magnanti, J.B. Orlin, Network Flows, Prentice-Hall, Englewood Cliffs, NJ, 1993.

[2] C. Berge, U.S.R. Murty, Hypergraphs, North-Holland, Amsterdam, 1989.

[3] J.A. Bondy, U.S.R. Murty, Graph Theory With Applications, North-Holland, New York, 1976.

[4] M.C. Chen, L. McNamee, A model of summary data and its applications to statistical databases, in: G. Goos, J. Hartmanis (Eds.), Proc. Fourth Internat. Conf. on "Statistical & Scientific Database Management" 1988, Lecture Notes in Computer Sciences, Vol. 339, Springer, Berlin, 1989, pp. 354–372.

[5] M.C. Chen, L. McNamee, M. Melkanoff, On the data model and access method of summary data management, IEEE Trans. Knowledge Data Eng. 1 (1989) 519–529.

[6] F.Y. Chin, G. Ozsoyoglu, Auditing and inference control in statistical databases, IEEE Trans. Software Eng. 8 (1982) 574–582.

[7] M. Conforti, M.R. Rao, Some new matroids on graphs: cut sets and the max cut problem, Math. Oper. Res. 12 (1987) 193–204.

[8] L.H. Cox, L.V. Zayatz, An agenda for research on statistical disclosure limitation, J. Official Statist. 11 (1995) 205–220.

[9] M. Doob, Generalization of magic graphs, J. Combin. Theory B 17 (1974) 205–217.

[10] S. Grumbach, L. Tininini, On the content of materialized aggregate views, J. Comput. System Sci. 66 (2003) 133–168 (A preliminary version appeared in Proc. XIX ACM Symp. on “Principles of Database Systems”, 2000.).

[11] D. Gusfield, A graph-theoretic approach to statistical data security, SIAM J. Comput. 17 (1988) 552–571.

[12] A.Y. Halevy, Answering queries using views: a survey, VLDB J. 10 (2001) 270–294.

[13] R.H. Jeurissen, Magic graphs, a characterization, European J. Combin. 9 (1988) 363–368.

[14] M.-Y. Kao, Minimal linear invariants, in: K. Kanchanasut, J.J. Levy (Eds.), “Algorithms, Concurrency and Knowledge”, Proc. of the Asian Computing Science Conference, Lecture Notes in Computer Science, Vol. 1023, Springer, Berlin, 1995, pp. 23–33.

[15] M.-Y. Kao, Data security equals graph connectivity, SIAM J. Discrete Math. 9 (1996) 87–110.

[16] M.-Y. Kao, Efficient detection and protection of information in cross-tabulated tables II: minimal linear invariant sets, J. Combin. Optim. 1 (1997) 187–202.

[17] M.-Y. Kao, Total protection of analytic-invariant set information in cross-tabulated tables, SIAM J. Comput. 26 (1997) 231–242.

[18] M.-Y. Kao, D. Gusfield, Efficient detection and protection of information in cross tabulated tables I: linear invariant test, SIAM J. Discrete Math. 6 (1993) 460–476.

[19] F.M. Malvestuto, A universal-scheme approach to statistical databases containing homogeneous summary tables, ACM Trans. Database Systems 18 (1993) 678–708.

[20] F.M. Malvestuto, M. Mezzini, A linear algorithm for finding the invariant edges of an edge-weighted graph, SIAM J. Comput. 31 (2002) 1438–1455.

[21] F.M. Malvestuto, M. Mezzini, M. Moscarini, Answering statistical sum-queries using materialised sum-queries: an analytic approach, unpublished manuscript, 2005.

[22] F.M. Malvestuto, M. Mezzini, M. Moscarini, Auditing sum-queries to make a statistical database secure, ACM Trans. Inform. System Security 9 (2006) 31–60.

[23] F.M. Malvestuto, M. Moscarini, Computational issues connected with the protection of sensitive statistics by auditing sum-queries, in: M. Rafanelli, M. Jarke (Eds.), Proc. Tenth Internat. Conf. on “Scientific & Statistical Database Management”, IEEE Computer Society, Los Alamitos, CA, 1998, pp. 134–144.

[24] F.M. Malvestuto, M. Moscarini, An audit expert for large statistical databases, in: Statistical Data Protection, EUROSTAT, 1999, pp. 29–43.

[25] C. Van Nuffelen, On the incidence matrix of a graph, IEEE Trans. Circuits Systems 23 (1976) 572.