



## The spanning galaxy problem

Daniel Gonçalves, Frédéric Havet, Alexandre Pinlou, Stéphan Thomassé

► **To cite this version:**

Daniel Gonçalves, Frédéric Havet, Alexandre Pinlou, Stéphan Thomassé. The spanning galaxy problem. *Discrete Applied Mathematics*, Elsevier, 2012, 160 (6), pp.744-754. <hal-00749191>

**HAL Id: hal-00749191**

**<https://hal.inria.fr/hal-00749191>**

Submitted on 23 Oct 2016

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# On Spanning Galaxies in Digraphs

Daniel Gonçalves<sup>a</sup>, Frédéric Havet<sup>b</sup>, Alexandre Pinlou<sup>a</sup>, Stéphan Thomassé<sup>a</sup>

<sup>a</sup>*Université Montpellier 2 - CNRS, LIRMM, 161 rue Ada, 34392 Montpellier Cedex 5, France.*

<sup>b</sup>*Projet Mascotte I3S (CNRS & UNSA) and INRIA, INRIA Sophia-Antipolis, 2004 route des Lucioles, BP 93, 06902 Sophia-Antipolis Cedex, France.*

---

## Abstract

In a directed graph, a *star* is an arborescence with at least one arc, in which the root dominates all the other vertices. A *galaxy* is a vertex-disjoint union of stars. In this paper, we consider the SPANNING GALAXY problem of deciding whether a digraph  $D$  has a spanning galaxy or not. We show that although this problem is NP-complete (even when restricted to acyclic digraphs), it becomes polynomial-time solvable when restricted to strong digraphs. In fact, we prove that restricted to this class, the SPANNING GALAXY problem is equivalent to the problem of deciding if a strong digraph has a strong digraph with an even number of vertices. We then show a polynomial-time algorithm to solve this problem. We also consider some parameterized version of the SPANNING GALAXY problem. Finally, we improve some results concerning the notion of directed star arboricity of a digraph  $D$ , which is the minimum number of galaxies needed to cover all the arcs of  $D$ . We show in particular that  $dst(D) \leq \Delta(D) + 1$  for every digraph  $D$  and that  $dst(D) \leq \Delta(D)$  for every acyclic digraph  $D$ .

*Keywords:* spanning galaxy; even strong subdigraph; directed star arboricity; algorithms; fixed parameter tractable.

---

## 1. Introduction

All digraphs considered here are finite and loopless. We rely on [4] for classical notations and concepts. The out-neighbourhood (resp. in-neighbourhood) of a vertex  $u$  is denoted by  $N^+(u)$  (resp.  $N^-(u)$ ). A digraph is said to be *even* if it has an even number of vertices.

A *directed path* in a digraph is a subdigraph  $P$  with vertex set  $\{v_1, \dots, v_k\}$  and arc set  $\{v_1v_2, \dots, v_{k-1}v_k\}$ . It is said to go from  $v_1$  to  $v_k$ . A digraph  $D = (V, A)$  is *strongly connected* or *strong* if for every pair  $(u, v) \in V^2$  there is a directed path from  $u$  to  $v$ .

---

*Email addresses:* Daniel.Goncalves@lirmm.fr (Daniel Gonçalves),  
Frederic.Havet@sophia.inria.fr (Frédéric Havet), Alexandre.Pinlou@lirmm.fr  
(Alexandre Pinlou), thomasse@lirmm.fr (Stéphan Thomassé)

A *circuit* in a digraph is a subdigraph  $C$  with vertex set  $\{v_1, \dots, v_k\}$  and arc set  $\{v_1v_2, \dots, v_{k-1}v_k, v_kv_1\}$ . A digraph is *acyclic* if it contains no circuit.

An *arborescence* is a connected digraph in which every vertex has in-degree 1 except one, called the *root*, which has in-degree 0. A *diforest* is a vertex-disjoint union of arborescences. A *star* is an arborescence with at least one arc, in which the root dominates all the other vertices. A *galaxy* is a diforest of stars. A galaxy  $S$  in a digraph  $D$  is *spanning* if  $V(S) = V(D)$ .

In this paper, we mainly study the the following decision problem:

SPANNING GALAXY problem

Instance: A digraph  $D$ .

Question: Does  $D$  have a spanning galaxy?

We prove the following complexity results.

1. *The SPANNING GALAXY problem is linear-time solvable for arborescences* (Proposition 1). We also explore the relations between spanning galaxies and winning diforests for the parity game.
2. *The SPANNING GALAXY problem is NP-complete even for digraphs that are acyclic, planar, bipartite, subcubic, with arbitrarily large girth and with maximum out-degree 2* (Theorem 7).
3. *The SPANNING GALAXY problem is polynomial-time solvable on strong digraphs* (Corollary 12). In order to prove this, we show in Theorem 9 that a strong digraph has a spanning galaxy if and only if it contains an even strong subdigraph. We then describe (Theorem 11) a polynomial-time algorithm for deciding if a strong digraph has an even strong subdigraph.
4. *The problems of testing if a digraph contains a spanning galaxy that uses or avoids, respectively, a prespecified arc is NP-complete, even if the given digraph is strong* (Theorem 14). In contrast, we show in Theorem 13 that the problems of testing if a strong digraph contains an even strong subdigraph that uses or avoids, respectively, a prespecified arc is polynomial-time solvable.
5. *The SPANNING GALAXY problem is  $W[2]$ -hard when parameterized by the number of stars of the galaxy and Fixed Parameter Tractable when parameterized by the number  $k$  of vertices of the digraph that are spanned* (does a digraph contain a galaxy on at least  $k$  vertices?). See Section 5.

In Section 6, using the notion of spanning galaxy, we improve some results regarding directed star arboricity. In fact, our interest to spanning galaxy originated in this concept. The *directed star arboricity* of a digraph  $D$ , denoted by  $dst(D)$ , is the minimum number of galaxies needed to cover  $A(D)$ . Amini et

al. [3] conjectured that every digraph  $D$  with maximum degree  $\Delta \geq 3$  satisfies  $dst(D) \leq \Delta$ . We prove this conjecture for acyclic digraphs and show that  $dst(D) \leq \Delta(D) + 1$  for every digraph  $D$ .

Finally, we conclude with some related open questions.

## 2. Spanning Galaxy and Winning Diforest

**Proposition 1.** *The SPANNING GALAXY problem can be solved in linear time for arborescences.*

**Proof .** If an arborescence  $T$  has no vertices, it vacuously admits a spanning galaxy. If  $T$  is restricted to its root, it has none. Now if  $T$  has at least two vertices, we consider a furthest leaf  $v$  from  $r$  and we denote by  $u$  the in-neighbour of  $v$ . By definition of  $v$ , all the out-neighbours of  $u$  are leaves. Thus, if  $T$  admits a spanning galaxy, this galaxy contains the star with root  $u$  whose leaves are the out-neighbours of  $u$ . Hence,  $T$  admits a spanning galaxy if and only if  $T - (\{u\} \cup N^+(u))$  does. This gives a simple linear-time algorithm for arborescences.  $\square$

The proof of the above proposition also implies the following lemma.

**Lemma 2.** *Every arborescence  $T$  contains a galaxy spanning every vertex except possibly the root.*

The *parity game* is a widely studied game. Its restriction to arborescences is played on an arborescence  $T$  (with root  $r$ ) by two players, Player 1 and Player 2, as follows. At the beginning of a play, a token is placed on the root  $r$  and is then moved over  $V(T)$  following the transitional relation: if the token is placed on a vertex  $v$ , then the next position of the token is one of the out-neighbours of  $v$ . The players move the token alternatively (starting with Player 1) until the token reaches a leaf. A player wins if its opponent cannot move anymore. Since our arborescences are finite, one of the two players has a winning strategy. If Player 1 has a winning strategy, we say that  $T$  is *winning*; otherwise,  $T$  is *losing*. By convention, an arborescence  $T$  with zero vertices is winning.

**Lemma 3.** *An arborescence  $T$  admits a spanning galaxy if and only if  $T$  is winning.*

**Proof .** This directly follows from:

- An arborescence  $T$  with one vertex is losing.
- Given an arborescence  $T$  with at least two vertices, where  $v$  is any furthest leaf from the root, and where  $u$  is the in-neighbour of  $v$ ;  $T$  is winning if and only if  $T - (\{u\} \cup N^+(u))$  is winning.

$\square$

A diforest is *winning* if all its arborescences are winning, otherwise it is *losing*. Since stars are winning arborescences Lemma 3 implies the following:

**Lemma 4.** *A digraph  $D$  admits a spanning galaxy if and only if  $D$  contains a winning spanning diforest.*

The directed path of length  $l$ ,  $P_l = (r, v_1, v_2, \dots, v_l)$ , admits a spanning galaxy if and only if  $l$  is odd (recall that the length of a path is its number of arcs). Given two arborescences  $T$  and  $T'$  and a vertex  $v$  of  $T$ , we denote by  $T \vee_v T'$  the arborescence obtained by identifying  $v$  in  $T$  with the root of  $T'$ . When  $v$  is the root of  $T$ , we simply write  $T \vee T'$ . Observe that  $T \vee T'$  is winning if and only if  $T$  or  $T'$  is winning. Similarly, if  $T'$  is losing, then  $T \vee_v T'$  is winning if and only if  $T$  is winning.

Thus, we have the following two lemmas, which we will use in Section 6.

**Lemma 5.** *For every arborescence  $T$  and every odd integer  $l$ , the arborescence  $T \vee P_l$  is winning.*

**Lemma 6.** *For every arborescence  $T$ , every vertex  $v$  of  $T$ , and every even integer  $l$ , the arborescence  $T \vee_v P_l$  is winning if and only if  $T$  is winning.*

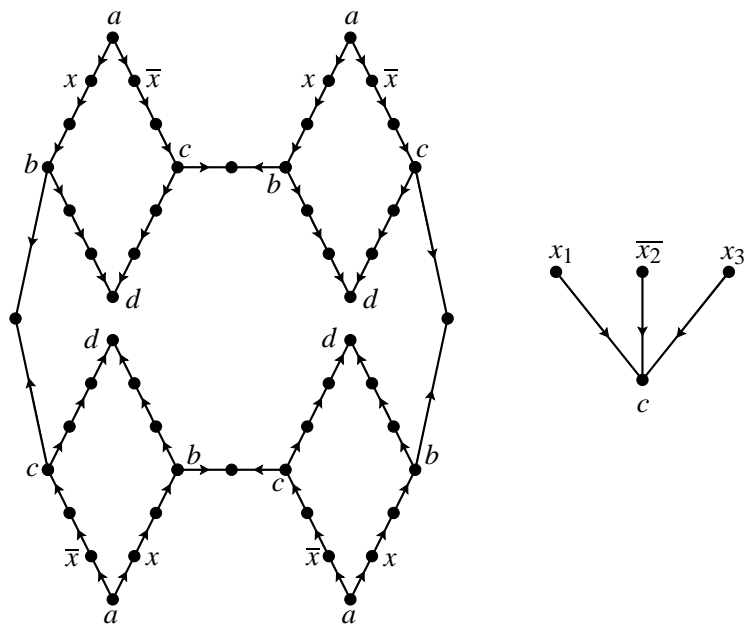
### 3. Spanning Galaxy in Acyclic Digraphs

**Theorem 7.** *The SPANNING GALAXY problem is NP-complete, even when restricted to digraphs which are acyclic, planar, bipartite, subcubic, with arbitrarily large girth, and with maximum out-degree 2.*

**Proof .** This problem is clearly in NP and we prove now that it is NP-hard for this restricted family of digraphs. Kratochvíl proved that PLANAR  $(3, \leq 4)$ -SAT is NP-complete [8]. In this restricted version of SAT, the variable-clause incidence graph of the input formula is planar, every clause is a disjunction of three literals, and every variable occurs in at most four clauses. We reduce PLANAR  $(3, \leq 4)$ -SAT to the SPANNING GALAXY problem. Given an instance  $I$  of PLANAR  $(3, \leq 4)$ -SAT, we shall construct a planar digraph  $D_I$  such that  $I$  is a satisfiable instance of PLANAR  $(3, \leq 4)$ -SAT if and only if  $D_I$  has a spanning galaxy. For this, we take one copy of the graph depicted in Figure 1(a) per variable of  $I$ , and one copy of the graph depicted in Figure 1(b) per clause of  $I$ . Whenever the literal  $x$  (resp.  $\bar{x}$ ) appears in a clause  $c$  in  $I$ , we identify one vertex labelled  $x$  (resp.  $\bar{x}$ ) of the variable gadget of  $x$  with a source of the clause gadget of  $c$ .

Let us observe that the digraph  $D_I$  is acyclic, planar, bipartite, subcubic, with maximum in-degree 3 and with maximum out-degree 2.

The variable gadget of  $x$  in the graph  $D_I$  is connected to the rest of the graph by the vertices labelled by  $x$  or  $\bar{x}$ . The vertices which are not labelled by  $x$  or  $\bar{x}$  are called *internal vertices* of the variable gadget of  $x$ . One can observe that there are only two possible galaxies that span all the internal vertices of a variable gadget. Actually, these two galaxies span all the vertices of the variable gadget. Moreover, in the first galaxy, every vertex  $x$  is the root of a star and every vertex  $\bar{x}$  is a leaf of a star; in the second one, every vertex  $\bar{x}$  is a root of a star and every vertex  $x$  is a leaf of a star.



(a) The gadget for the variable  $x$ .

(b) The gadget for the clause  $c = (x_1 \vee \bar{x}_2 \vee x_3)$ .

Figure 1: The gadgets for Theorem 7.

In addition, one can observe that the previous remark, stating that the vertices  $x$  are roots of stars whenever the vertices  $\bar{x}$  are leaves, holds for any odd paths linking  $a$  and  $b$  (resp.  $a$  and  $c$ ,  $b$  and  $d$ ,  $c$  and  $d$ ). Therefore, the girth of the graph  $D_I$  can be made arbitrarily large.

Let  $I$  be an instance of PLANAR  $(3, \leq 4)$ -SAT.

Suppose first that  $I$  is satisfiable by some truth assignment  $\phi$ . Let us exhibit a spanning galaxy of  $D_I$ . For every variable  $x$ , we span its gadget with a galaxy in such a way that the vertices labelled  $x$  are roots of stars if and only if  $\phi(x) = \text{TRUE}$ . In this way, we can span the internal vertices  $c$  of the clause gadgets. Indeed, since  $c$  is satisfied by  $\phi$ , the vertex  $c$  has an in-neighbour  $x_1$  which is the root of a star. We then add the arc  $x_1c$  to our galaxy to span  $c$ . Suppose now that  $D_I$  has a spanning galaxy  $T$ . Let  $\phi$  be the truth assignment  $\phi$  defined by  $\phi(x) = \text{TRUE}$  if and only if the vertices labelled  $x$  are roots of stars of  $T$ . Then  $\phi$  satisfies  $I$  since every clause vertex  $c$  needs one of its in-neighbours to be the root of some star.  $\square$

#### 4. Spanning Galaxy and even strong subdigraph

Let  $D$  be a strong digraph. A *handle*  $h$  of  $D$  is a directed path  $(s, v_1, \dots, v_\ell, t)$  from  $s$  to  $t$  (where  $s$  and  $t$  may be identical, or the handle possibly restricted to the arc  $st$ ) such that:

- for all  $1 \leq i \leq \ell$ ,  $d^-(v_i) = d^+(v_i) = 1$ , and
- the digraph  $D \setminus h$  obtained from  $D$  by *suppressing*  $h$ , that is removing the arcs and the internal vertices of  $h$ , is strong.

The vertices  $s$  and  $t$  are the *endvertices* of  $h$  while the vertices  $v_i$  are its *inner vertices*. The vertex  $s$  is the *tail* of  $h$  and  $t$  its *head*. The *length* of a handle is the number of its arcs, here  $\ell + 1$ . A handle of length one is said to be *trivial*. For any  $1 \leq i, j \leq \ell$ , we say that  $v_i$  *precedes* (resp. *strictly precedes*)  $v_j$  on the handle  $h$  if  $i \leq j$  (resp.  $i < j$ ).

Given a strong digraph  $D$ , a *handle decomposition* of  $D$  starting at  $v \in V(D)$  is a triplet  $(v, (h_i)_{1 \leq i \leq p}, (D_i)_{0 \leq i \leq p})$ , where  $(D_i)_{0 \leq i \leq p}$  is a sequence of strong digraphs and  $(h_i)_{1 \leq i \leq p}$  is a sequence of handles such that:

- $V(D_0) = \{v\}$ ,
- for  $1 \leq i \leq p$ ,  $h_i$  is a handle of  $D_i$  and  $D_i$  is the (edge-disjoint) union of  $D_{i-1}$  and  $h_i$ , and
- $D = D_p$ .

A handle decomposition is uniquely determined by  $v$  and either  $(h_i)_{1 \leq i \leq p}$ , or  $(D_i)_{0 \leq i \leq p}$ . The number of handles  $p$  in any handle decomposition of  $D$  is exactly  $|A(D)| - |V(D)| + 1$ . The value  $p$  is also called the *cyclomatic number* of  $D$ . Observe that  $p = 0$  when  $D$  is a singleton and  $p = 1$  when  $D$  is a circuit. A digraph  $D$  with cyclomatic number two is called a *theta*.

The following lemma is straightforward.

**Lemma 8.** *For every strong digraph  $D'$  of some strong digraph  $D$ , there is a handle decomposition  $(v, (h_i)_{1 \leq i \leq p}, (D_i)_{0 \leq i \leq p})$  of  $D$  such that  $D' = D_i$  for some  $i$ .*

A handle is *even* if its length is even. A handle decomposition is *even* if one of its handles is even. A strong digraph is *even* if it has an even number of vertices. Handles, handle decompositions and strong digraphs are *odd* when they are not even. Given a digraph  $D$ ,  $\overline{D}$  is obtained from  $D$  by reversing every arc.

**Theorem 9.** *Given a strong digraph  $D$ , the following are equivalent:*

- (1)  $D$  has a spanning galaxy.
- (2)  $\overline{D}$  has a spanning galaxy.
- (3)  $D$  contains a winning spanning arborescence.
- (4)  $D$  has an even handle decomposition.
- (5)  $D$  contains an even circuit or an even theta.
- (6)  $D$  contains an even strong subdigraph.

**Proof .**

**(3)  $\Rightarrow$  (1).** Consider a digraph  $D$  containing a winning spanning arborescence  $T$ . Lemma 3 implies that  $T$  contains a spanning galaxy, which also spans  $D$ .

**(4)  $\Rightarrow$  (3).** Let  $(v, (h_i)_{1 \leq i \leq p}, (D_i)_{0 \leq i \leq p})$  be an even handle decomposition of  $D$ . Let  $q$  be the largest integer such that  $h_q$  is an even handle. Since  $D_{q-1}$  is strong, it contains a spanning arborescence  $T_{q-1}$  rooted at  $s_q$ , the first vertex of  $h_q$ . Now for every  $q \leq r \leq p$ , we define a spanning arborescence  $T_r$  of  $D_r$  as follows. For every  $h_r = (s_r, v_1, \dots, v_\ell, t_r)$ , we let  $T_r = T_{r-1} \vee_{s_r} P_r$  where  $P_r$  is the path  $(s_r, v_1, \dots, v_\ell)$ , i.e. the handle  $h_r$  minus its last arc. By Lemma 5, the arborescence  $T_q$  is winning since  $T_{q-1} \vee_{s_q} P_q$  is  $T_{q-1} \vee P_q$ . Therefore, by Lemma 6,  $T_r$  is winning, for every  $q \leq r \leq p$ . Thus  $T_p$  is a winning spanning arborescence of  $D$ .

**(1)  $\Rightarrow$  (4).** By way of contradiction, suppose that there exists a strong digraph  $D$  with no even handle decomposition admitting a spanning galaxy. Observe that in particular,  $D$  has no even circuit. Choose such a  $D$  with minimum number of arcs. Let  $F$  be a spanning galaxy of  $D$ . Observe that

- (o) every trivial handle  $st$  of  $D$  belongs to  $E(F)$ ,

otherwise deleting the arc  $st$  from  $D$  leaves a strong digraph spanned by  $F$  and with no even handle decomposition, against the minimality of  $D$ .

Consider a handle decomposition  $(v, (h_i)_{1 \leq i \leq p}, (D_i)_{0 \leq i \leq p})$  of  $D$  which minimizes the number of trivial handles. Let  $q$  be the largest integer such that



$h_q = (v_0, \dots, v_{\ell+1})$  is non trivial (here we adopt the notation  $s := v_0$  and  $t := v_{\ell+1}$ ). Hence, every handle  $h_i$  is trivial for every  $q < i \leq p$ . Moreover, since  $h_q$  is odd and non trivial, we have  $\ell \geq 2$ . Since the number of trivial handles in this decomposition is minimum, we have the following straightforward properties.

- (i) there is no arc  $v_i v_j$  with  $j \geq i + 2$ , except possibly  $v_0 v_{\ell+1}$ ;
- (ii) for  $2 \leq i \leq \ell$ , the vertex  $v_i$  has no in-neighbour in  $D_{q-1}$ ;
- (iii) for  $1 \leq i \leq \ell - 1$ , the vertex  $v_i$  has no out-neighbour in  $D_{q-1}$ .

In addition, Observation (iii) implies that:

- (iv)  $v_1$  has no in-neighbours in  $D_{q-1} - v_0$ .

Indeed if  $u$  is such an in-neighbour, both arcs  $uv_1$  and  $v_0 v_1$  would be trivial handles of  $D$ . Hence, according to the previous observation, they both are in  $E(F)$  which is impossible.

Furthermore,

- (v) there is no arc  $v_j v_i$  with  $0 \leq i < j \leq \ell + 1$ .

We prove (v) as follows. An arc  $v_j v_i$  is *short* if there is no distinct arc  $v_{j'} v_{i'}$  for which  $i \leq i' < j' \leq j$ . By way of contradiction, consider a short arc  $v_j v_i$  which minimizes  $i$ . By (i) and since there is no even circuit, the vertices  $v_i, v_{i+1}, \dots, v_j$  induce an odd circuit. Moreover, since deleting the arc  $v_j v_i$  leaves  $D$  strongly connected, we have  $v_j v_i \in F$  (by (o)). Hence there is at least one vertex in  $X = \{v_{i+1}, \dots, v_{j-1}\}$  which has a neighbour in  $F - X$ . Let  $i'$  be the smallest index of such a vertex. By (i), (ii), (iii), (iv) and the choice of  $v_j v_i$ , we have either  $j = i + 1$  and an arc  $v_l x$  with  $x \in V(D_{q-1})$ , or there is an arc  $v_{j'} v_{i'}$  such that  $i < i' < j < j'$ . In the first case,  $v_l v_{l+1}$  is a trivial handle and so by (o), it is in  $E(F)$ . This is a contradiction to  $v_{l+1} v_i$  in  $E(F)$ . Hence we may assume that we are in the second case. If  $i' - i$  is odd then  $(v_j, v_i, v_{i+1}, \dots, v_{i'})$  is an even handle on the circuit  $(v_{i'}, v_{i'+1}, \dots, v_{j'})$ , contradicting the fact that  $D$  has no even handle decomposition. If  $i' - i$  is even then  $X' = \{v_{i+1}, \dots, v_{i'-1}\}$  has odd cardinality, and both arcs  $v_{j'} v_{i'}$  and  $v_j v_i$  are in  $E(F)$ . Hence there must be a vertex in  $X'$  which has a neighbour in  $F - X'$ , contradicting the definition of  $i'$ . This proves (v).

The above properties imply that the only arc entering  $S = \{v_1, \dots, v_\ell\}$  is  $v_0 v_1$  and the only arcs leaving  $S$  are those leaving  $v_\ell$ . Moreover  $(v_0, v_1, \dots, v_\ell)$  is an induced path. If  $\{v_1 v_2, v_3 v_4, \dots, v_{\ell-1} v_\ell\} \subseteq E(F)$  then the digraph  $D_{q-1}$  would also be a counterexample, contradicting the minimality of  $D$ . Thus  $E(F)$  contains the arcs  $v_0 v_1, v_2 v_3, \dots, v_{\ell-2} v_{\ell-1}$  and all the arcs leaving  $v_\ell$  (by (iv)). Hence  $v_1$  has only  $v_0$  as in-neighbour (by (o)). Thus the digraph obtained from  $D$  by contracting  $v_0 v_1$  and  $v_1 v_2$  has a spanning galaxy and no even handle decomposition. This contradicts the minimality of  $D$ .

(4)  $\Rightarrow$  (5). By way of contradiction, suppose that there are strong digraphs with an even handle decomposition containing no even circuits nor even thetas. Consider such a digraph  $D$  with an even handle decomposition  $(v, (h_i)_{1 \leq i \leq p}, (D_i)_{0 \leq i \leq p})$  minimizing  $p$ . It is clear, by minimality of  $p$  that the only even handle of this decomposition is  $h_p$ . Otherwise  $D_{p-1}$  would contradict the minimality of  $p$ .

In the remainder, we denote by  $s$  and  $t$  the tail and the head respectively of the handle  $h_p$ .

**Claim 1.**  $p > 2$ .

If  $p = 1$ , then  $D$  would be an even circuit. If  $p = 2$ , then  $h_1$  has odd length and thus  $D$  would either be an even theta or contain an even circuit. This proves Claim 1.

By Lemma 8, there is a handle decomposition  $(s, (h'_i)_{1 \leq i \leq p}, (D'_i)_{0 \leq i \leq p})$  of  $D$  starting at  $s$  and such that  $h'_p = h_p$ . For every  $1 < i < p$ , let us denote by  $s_i$  the tail of  $h'_i$  and by  $t_i$  its head.

**Claim 2.** For every  $1 < i < p$ , the endvertices of  $h'_i$  are inner vertices of  $h'_{i-1}$ .

Suppose for a contradiction that the claim does not hold. Let  $q$  be the largest integer such that one of the two endvertices of  $h'_q$  is not an inner vertex of  $h'_{q-1}$ . One of the endvertices of  $h'_q$  is an inner vertex of  $h'_{q-1}$ . Otherwise  $h'_{q-1}$  would be a handle of  $D$  and the digraph obtained from  $D$  by suppressing  $h'_{q-1}$  would contradict the minimality of  $p$ . By directional duality, we may assume that  $s_q$  is an inner vertex of  $h'_{q-1}$  and  $t_q$  is not. Let us divide  $h'_{q-1}$  into two paths, the path  $R$  with tail  $s_{q-1}$  and head  $s_q$  and the path  $S$  with tail  $s_q$  and head  $t_{q-1}$ . Then  $S$  is a handle of  $D$  and the digraph obtained from  $D$  by suppressing  $S$  contradicts the minimality of  $p$  (the handles  $h'_{q-1}$  and  $h'_q$  are replaced by the concatenation of  $R$  and  $h_q$ ). This proves Claim 2.

**Claim 3.** For every  $1 < i < p$ , the vertex  $t_i$  precedes  $s_i$  on  $h'_{i-1}$ .

Suppose not. Then  $s_i$  strictly precedes  $t_i$  on  $h'_{i-1}$ . Let  $R$  be the subpath of  $h'_{i-1}$  with tail  $s_i$  and head  $t_i$ . Then  $R$  is a handle of  $D$  and the digraph obtained from  $D$  by suppressing  $R$  contradicts the minimality of  $p$  (the handles  $h'_{i-1}$  and  $h'_i$  are replaced by a single handle going from  $s_{i-1}$  to  $t_{i-1}$  containing  $h'_i$ ). This proves Claim 3.

The circuit  $h'_1$  can be divided into two paths:  $P_1$  with tail  $s_2$  and head  $t_2$  and  $P_2$  with tail  $t_2$  and head  $s_2$ . If  $s_2$  and  $t_2$  are identical, we assume that  $P_2$  has no arc. If  $s \in V(P_2)$ , then according to Claims 2 and 3,  $P_1$  is a handle of  $D$  which suppression leaves a digraph with an even handle decomposition and no even circuit or theta. This contradicts the minimality of  $D$ . If  $s \notin V(P_2)$ , then it is an internal vertex of  $P_1$ . Let  $P_3$  be the subpath of  $P_1$  with tail  $s$  and head  $t_2$ . Then by Claims 2 and 3,  $P_3$  is a handle of  $D$  which suppression leaves a digraph with an even handle decomposition and no even circuit or theta. This contradicts the minimality of  $D$ .

(5)  $\Rightarrow$  (6). Trivial since even circuits and thetas are strong digraphs with an even number of vertices.

(6)  $\Rightarrow$  (4). By Lemma 8 consider a handle decomposition  $(v, (h_i)_{1 \leq i \leq p}, (D_i)_{0 \leq i \leq p})$  of  $D$  such that some digraphs  $D_i$  have an even number of vertices. Let  $q$  be the smallest integer such that  $D_q$  has an even number of vertices. Then the handle  $h_q$  has an odd number of inner vertices, thus has even length.

(4)  $\Leftrightarrow$  (2). It is clear that a strong digraph  $D$  has an even handle decomposition if and only if  $\overline{D}$  does. Thus, since (4)  $\Leftrightarrow$  (1), a strong digraph  $D$  has an even handle decomposition if and only if  $\overline{D}$  has a spanning galaxy.  $\square$

Since every vertex  $v$  of every strong digraph  $D$  is the root of an arborescence  $T$ , by Lemma 2, for every strong digraph  $D$  and every vertex  $v$  of  $D$ ,  $D$  has a galaxy spanning every vertex except possibly  $v$ . This can be improved as follows.

**Theorem 10.** *A strong digraph  $D$  has either a spanning galaxy or for every  $v \in V(D)$ ,  $D$  has a matching spanning every vertex except  $v$ .*

This comes from the characterization of factor critical graphs by Lovász [9]. A non-oriented graph  $G$  is *factor critical*, if for every  $v \in V(G)$  the graph  $G - v$  has a perfect matching and Lovász showed that a graph  $G$  is factor critical if and only if it has an odd handle decomposition.

**Theorem 11.** *Deciding if a strong digraph contains an even strong subdigraph is polynomial-time solvable.*

**Proof .** Let us describe a polynomial-time algorithm to decide whether a strong digraph  $D$  contains an even strong subdigraph (ESS for short). The algorithm performs as follows. We first find a handle decomposition  $(v, (h_i)_{1 \leq i \leq p}, (D_i)_{0 \leq i \leq p})$  where  $h_q = (x_0, x_1, \dots, x_\ell)$  is the last non-trivial handle. If there exists an arc with tail  $y$  in  $V(D_{q-1})$  and head  $x_i$ ,  $2 \leq i \leq \ell-1$ , then replacing the two handles  $h_q$  and  $(y, x_i)$  by  $(y, x_i, \dots, x_\ell)$  and  $(x_0, \dots, x_i)$  we obtain a new decomposition with less trivial handles. If there exists an arc with tail  $x_i$ ,  $1 \leq i \leq \ell-2$ , and head  $y$  in  $V(D_{q-1})$ , then replacing the two handles  $h_q$  and  $(x_i, y)$  by  $(x_0, \dots, x_i, y)$  and  $(x_i, \dots, x_\ell)$  we obtain a new decomposition with less trivial handles. If there exists a handle  $x_i x_j \neq x_0 x_\ell$ , then replacing the two handles  $h_q$  and  $(x_i x_j)$  by  $(x_0, \dots, x_i, x_j, \dots, x_\ell)$  and  $(x_i, \dots, x_j)$  we obtain a new decomposition with less trivial handles. The three above operations are done in constant time and since the initial number of trivial handles is polynomial, one can compute in polynomial time a handle decomposition where there are no such arcs. If the decomposition has an even handle then return “YES” thanks to Theorem 9. We can then suppose in the remainder that  $\ell$  is odd. Let  $D'$  be the digraph obtained from  $D_{q-1}$  by adding all the arcs between  $N_{\overline{D}}(x_1)$  and  $N_D^+(x_{\ell-1})$ . Let  $S = \{x_1, x_2, \dots, x_{\ell-1}\}$  be the set of inner vertices of  $h_q$ .

**Claim 4.**  $D$  has an ESS if and only if  $D[S]$  has an ESS or  $D'$  has an ESS.

Since every subdigraph of  $D[S]$  is a subdigraph of  $D$ , if  $D[S]$  has an ESS, then  $D$  has an ESS. Therefore, we may suppose that  $D[S]$  does not have an ESS. For every ESS  $E$  of  $D$ , the digraph  $E' = D'[V(E) \setminus S]$  is an ESS of  $D'$ . Indeed, if  $|V(E)| \neq |V(E')|$ , the handle  $h_q$  is a subdigraph of  $E$ ; however  $h_q$  is odd so  $|V(h_q)|$  is even, and thus  $|V(E')| = |V(E)| - |V(h_q)| + 2$  is also even. Furthermore, since the paths of  $E$  from  $N_D^-(x_1)$  to  $N_D^+(x_{\ell-1})$  are replaced by single arcs in  $E'$ ,  $E'$  is strong. Finally, it is also clear that given any ESS  $F'$  of  $D'$  one of the graphs  $D[V(F')]$  or  $D[V(F') \cup S]$  is an ESS of  $D$  – according to whether or not there is an arc  $uv$  in  $A(F') \cap (N_D^-(x_1) \times N_D^+(x_{\ell-1}))$ . This proves Claim 4.

Checking if  $D[S]$  has an ESS can be done in polynomial time. We first check if there exists a *backward* arc (i.e. an arc  $x_b x_a$  such that  $a < b$ ) such that  $a$  and  $b$  have distinct parity. If there is such an arc, the graph  $D[\{x_a, x_{a+1}, \dots, x_b\}]$  is an ESS.

If there exists no such arc, we distinguish two types of backward arcs  $x_b x_a$  of  $D[S]$ : the arcs where  $a$  and  $b$  are both even, called *e-arcs*, and those where  $a$  and  $b$  are both odd, called *o-arcs*. Observe that the vertex set  $F$  of an ESS of  $D[S]$  is of the form  $\{x_i, x_{i+1}, \dots, x_j\}$ . Indeed, since there is no arc  $x_a x_b$  with  $a+1 < b$  in  $D[S]$  and since there is a path from the vertex with smaller index in  $V(F)$  (here  $x_i$ ) to the one with higher index (here  $x_j$ ), all the vertices between  $x_i$  and  $x_j$  are in  $F$ .

Furthermore since  $F$  is even,  $i$  and  $j$  have distinct parity. Consider a set  $A$  of backward arcs such that the union of the directed path  $\{x_i, x_{i+1}, \dots, x_j\}$  and  $A$  is strong, and such that in addition  $A$  is minimum with respect to inclusion. The arcs of  $A$ , when ordered increasingly according to the index of their tail, are such that two consecutive arcs  $x_c x_a$  and  $x_d x_b$  satisfy  $a < b < c < d$ . Note that since  $i$  and  $j$  have distinct parity, there exists two consecutive backward arcs of distinct types (one is an e-arc and the other one is an o-arc), say  $x_c x_a$  and  $x_d x_b$ . Then  $D[x_a, x_{a+1}, \dots, x_d]$  is an ESS.

Hence to summarize  $D[S]$  contains an ESS if and only if it contains an ESS with at most two backward arcs. So this can be checked in polynomial time.

In the case of  $D'$ , we check whether it contains an ESS or not by applying the algorithm recursively.  $\square$

Theorems 9 and 11 imply the following.

**Corollary 12.** *The SPANNING GALAXY problem is polynomial-time solvable for strong digraphs.*

#### 4.1. Prescribing an arc

Thomassen [12] showed that it is NP-complete to decide whether a digraph  $D$  has an even circuit containing a given arc. On the other hand, there is a polynomial-time algorithm to decide whether a digraph  $D$  has an even circuit

avoiding a given arc  $uv$ , it suffices to test whether  $D \setminus \{uv\}$  has an even circuit or not (by [10, 11]). Similarly, one can decide in polynomial time whether a digraph  $D$  has an even circuit or an even theta avoiding a given arc  $uv$ , by testing the strong components of  $D \setminus \{uv\}$ . We now show that one can also decide in polynomial time if a strong digraph has an even strong digraph containing a given arc.

**Theorem 13.** *It is polynomial-time decidable whether a strong digraph has an even strong digraph (ESS) containing (resp. avoiding) a given arc.*

**Proof .** The algorithm to decide if a strong digraph  $D$  has an ESS avoiding a given arc  $uv$  is very simple. It just has to check if some strong connected component of  $D \setminus \{uv\}$  has an ESS.

The algorithm to decide if a strong digraph  $D$  has an even strong subdigraph containing a given arc  $uv$  is very similar to the algorithm described in the proof of Theorem 11. Indeed, consider a handle decomposition  $(u, (h_i)_{1 \leq i \leq p}, (D_i)_{0 \leq i \leq p})$  such that  $uv \in h_1$  and such that the number of trivial handles is minimized.

If  $h_1 = (x_0 = u, x_1 = v, x_2, \dots, x_{l-1}, x_l = u)$  is the last (and only) non-trivial handle, then the other arcs of  $D$  are arcs  $x_i x_j$  with  $0 < j < i \leq l$ . Thus  $D$  has only one path from  $v$  to  $u$ , namely  $(v, x_2, \dots, x_{l-1}, u)$ . So, every strong digraph containing  $uv$  spans  $D$  and  $D$  has an ESS containing  $uv$  if and only if  $D$  is even, which is easily checked.

Now suppose that the last non-trivial handle is  $h_q = (x_0, x_1, \dots, x_l)$  with  $q > 1$ . Then there is no arc from  $D_{q-1}$  to  $\{x_2, \dots, x_{l-1}\}$ , from  $\{x_1, \dots, x_{l-2}\}$  to  $D_{q-1}$ , nor arc  $x_i x_j \neq x_0 x_l$  with  $i + 1 < j$ . If  $h_q$  is even then one of  $D_{q-1}$  or  $D_q$  is an even strong subdigraph containing  $uv$ . Thus assume  $h_q$  is odd, and consider the digraph  $D'$  obtained from  $D_{q-1}$  by adding all the arcs between  $N_D^-(x_1)$  and  $N_D^+(x_{l-1})$ . One can verify that  $D$  has an ESS containing  $uv$  if and only if  $D'$  has an ESS containing  $uv$ . Thus the algorithm just has to consider the smaller strong digraph  $D'$ .  $\square$

We now prove that the similar variants of the SPANNING GALAXY problem are NP-complete.

**Theorem 14.** *It is NP-complete to decide, given a strong digraph and one of its arc, whether there exists a spanning galaxy containing (resp. avoiding) this arc.*

**Proof .** The reduction from the SPANNING GALAXY problem in the acyclic case is straightforward. Given an acyclic digraph  $D$ , we construct  $D'$  from  $D$  by adding a disjoint directed path  $(a_1, a_2, a_3, a_4)$ , all possible arcs from  $a_4$  to sources of  $D$ , and all possible arcs from sinks of  $D$  to  $a_1$ . Note that  $D'$  is strong. Observe that  $D'$  has a spanning galaxy  $F$  containing the arc  $a_1 a_2$  (resp. avoiding  $a_2 a_3$ ) if and only if  $D$  has a spanning galaxy.  $\square$

## 5. Parameterizations of Galaxy Problems

The SPANNING GALAXY problem being hard in the general case, it is natural to ask if some parameterized version is tractable. A first attempt could be to ask for a fixed parameter tractable algorithm on parameter  $k$  (i.e. admitting an algorithm in time  $O(f(k)n^c)$  for some constant  $c$ ) deciding if a digraph admits a spanning galaxy with at most  $k$  stars. Unfortunately, the problem  $k$ -DOMINATION (which is  $W[2]$ -complete [6]) admits a straightforward reduction to this problem. Indeed, every minimal dominating set  $A$  of a graph  $G$  (with no isolated vertex) corresponds to the set of roots of a spanning galaxy of the digraph  $D$  obtained from  $G$  by replacing each edge  $ab \in E(G)$  by the arcs  $ab$  and  $ba$ . Hence the SPANNING GALAXY problem is at least as hard as  $k$ -DOMINATION, thus it is  $W[2]$ -hard.

However, the following problem is easier to handle:

$k$ -GALAXY problem

Instance: A digraph  $D$  and an integer  $k$ .

Parameter:  $k$ .

Question: Does  $D$  have a galaxy spanning at least  $k$  vertices?

This problem is very easily fixed parameter tractable, but we will show a much stronger result. Indeed, there is a polynomial-time algorithm (in size of  $D$ ) which transforms every instance  $(D, k)$  of  $k$ -GALAXY into an instance  $(D', k')$  which is equivalent to  $(D, k)$  and such that  $k' \leq k$  and  $D'$  has at most  $2k - 2$  vertices. This algorithm is called a  $(2k - 2)$ -kernelization algorithm, and the output  $D'$  is called a *kernel*. Observe that applying a brute force algorithm on  $D'$  to check if it admits a galaxy spanning at least  $k'$  vertices takes  $O(f(k))$  time. Hence the existence of the kernelization algorithm gives an FPT algorithm for  $k$ -GALAXY running in  $O(f(k) + n^c)$  time.

The general idea of the proof is the following. Given an input  $(D, k)$ , first grow a galaxy by some local procedures in order to obtain a galaxy that is *locally maximal*. Then we compute some matchings in order to get a larger galaxy. When no more improvement is obtained, we stop the process and we check if the largest obtained galaxy has at least  $k$  vertices. If so,  $(D, k)$  is a "Yes" instance of the  $k$ -GALAXY problem and we return the small "Yes" instance  $(S_k, k)$  with  $S_k$  the star of order  $k$ . If not, we can find a kernel.

A galaxy  $F$  of  $D$  is *locally maximal* if it satisfies the following conditions:

- (a) The vertices of  $V(D) \setminus V(F)$  form a stable set.
- (b) If  $uv \in A(F)$  and  $uw \in A(D)$ , then we have  $w \in V(F)$ .
- (c) If  $u \in V(F)$  and  $uv, uw \in A(D)$ , at least one of  $v$  and  $w$  belong to  $V(F)$ .
- (d) If  $uv, uw \in A(F)$  and  $wx \in A(D)$ , then  $x \in V(F)$ .

Given a galaxy  $G$  one can compute a locally maximal galaxy  $\text{lm}(G)$  spanning at least as many vertices as  $G$  with the following polynomial-time algorithm.

**Algorithm 1** ( $\text{lm}(G)$ ).

Step 1:  $F := G$ .

Step 2: If  $uv \in D \setminus F$  then  $V(F) := V \cup \{u, v\}$ ;  $E(F) := E(F) \cup \{uv\}$ ; Go to Step 2.

Step 3: If  $uv \in A(F)$ ,  $uw \in A(D)$  and  $w \notin V(F)$  then  $V(F) := V \cup \{w\}$ ;  $E(F) := E(F) \cup \{uw\}$ ; Go to Step 2.

Step 4: If  $tu \in E(F)$ ,  $uv, uw \in A(D)$  and  $v, w \notin V(F)$ , then  $V(F) := V \cup \{v, w\}$ ;  $E(F) := E(F) \setminus \{tu\} \cup \{uv, uw\}$ ; If  $d_F^+(t) = 0$ ,  $V(F) := V(F) \setminus \{t\}$ ; Go to Step 2.

Step 5: If  $uv, uw \in E(F)$ ,  $wx \in A(D)$  and  $x \notin V(F)$ , then  $V(F) := V(F) \cup \{x\}$ ;  $E(F) := E(F) \setminus \{uw\} \cup \{wx\}$ ; Go to Step 2.

Step 6: Return  $F$ .

This procedure being defined, one can now describe the kernelization algorithm.

**Algorithm 2** ( $\text{Ker}(D, k)$ ).

Step 1:  $G := (\emptyset; \emptyset)$ ;

Step 2:  $G := \text{lm}(G)$ ;

Step 3:  $N_G^+ = \{v \in V(D) \setminus V(G) \mid \exists u \in V(G), uv \in A(D)\}$ ;  $N_G^- = V(D) \setminus (V(G) \cup N_G^+)$ ;

Step 4: Compute a maximum matching  $M$  in the bipartite graph induced by the arcs from  $N^-(G)$  to  $V(G)$ ;

Step 5: If  $|V(M)| > |V(G)|$  then  $G := M$  and go to Step 2;

Step 6: If  $|V(G)| \geq k$ , return  $(S_k, k)$ .

Step 7: Else, return  $(D[V(G) \cup N^+(G) \cup V(M)], k)$ .

**Theorem 15.** *Algorithm 2 is a  $(2k - 2)$ -kernelization of the  $k$ -GALAXY problem.*

**Proof .** As one can compute  $\text{lm}(G)$  and the maximum matching in a graph in polynomial time, Algorithm 2 runs in polynomial time. Indeed it goes back at most  $|V(D)|$  times to Step 2 (from Step 5) since the order of the galaxy strictly increases each time.

Let us now show that  $(D, k)$  is a "Yes" instance of the  $k$ -GALAXY problem if and only if  $Ker(D, k)$  is. This is trivially true when Algorithm 2 stops at Step 6. Hence we may assume that it stops at Step 7.

The galaxy  $G$  (at the end of the running) is a locally maximal galaxy because of Step 2. Condition (a) implies that  $N_G^-$  is a stable set of  $D$ . Note that  $N_G^+$  may contain in-neighbours of  $V(G)$  while  $N_G^-$  contains no out-neighbours of  $V(G)$ .

$D' = D \setminus (V(G) \cup N^+(G) \cup V(M))$  is a subdigraph of  $D$ . Hence if  $Ker(D, k) = (D', k)$  is a "Yes" instance, so is  $(D, k)$ . Reciprocally, we shall prove that if  $(D, k)$  is a "Yes" instance so is  $(D', k)$ . Suppose not. Then there is a galaxy  $G^*$  of  $D$  such that  $D'$  does not contain a galaxy spanning  $|V(G^*)|$  vertices. Among the possible choices of  $G^*$ , select one which minimizes its number of vertices in  $N_G^- \setminus M$ , and then which minimizes its number of arcs between  $N_G^- \setminus M$  and  $V(G)$ . Since  $G^* \not\subseteq D'$ ,  $G^*$  has a vertex  $u \in N_G^- \setminus M$ , and thus  $G^*$  has an arc  $uv_1$  with  $v_1 \in V(G)$ . Since  $uv_1 \notin M$  there is an arc  $u_1v_1$  in  $M$ . We inductively define the vertices  $u_i$  and  $v_i$ , for  $i \geq 2$ , as follows. If  $u_{i-1}$  does not belong to  $G^*$  then  $u_j$  and  $v_j$  are not defined for  $j \geq i$ . Otherwise, let  $v_i$  be any vertex such that  $u_{i-1}v_i$  is an arc of  $G^*$ . Note that  $v_i \neq v_j$  with  $j < i$  (otherwise it would have two incoming arcs in  $G^*$ ) and that  $v_i \in V(M)$ , otherwise the path  $(v_i, u_{i-1}, v_{i-1}, \dots, u_1, v_1, u)$  would be an augmenting path with respect to  $M$ , contradicting the maximality of  $M$ . Thus let  $u_i$  be the vertex such that  $u_iv_i \in M$ . Let  $t$  be the greater index for which the vertices  $u_t$  and  $v_t$  are defined. Then  $u_t \notin V(G^*)$  and we can replace the arcs  $uv_1$  and  $u_{i-1}v_i$ , for  $2 \leq i \leq t$ , by the arcs  $u_iv_i$ , for  $1 \leq i \leq t$ . Note that since  $u_t$  was not previously spanned, the obtained galaxy spans at least as many vertices as  $G^*$  and covers more arcs of  $M$ , a contradiction to the choice of  $G^*$ .

$G$  has at most  $k - 1$  vertices because of Step 6. Since  $G$  is locally maximal, we have  $|N^+(G)| \leq |V(G)|/2$  because of the conditions (b), (c) and (d). Finally,  $|V(M)| \leq |V(G)|$  because of Step 5 and  $|V(M) \setminus V(G)| = |V(M)|/2$  by definition of  $M$ . Hence  $|V(D')| \leq 2|V(G)| \leq 2k - 2$ .  $\square$

## 6. Directed Star Arboricity

Recall that the *directed star arboricity* of a digraph  $D$ , denoted by  $dst(D)$ , is the minimum number of galaxies needed to cover  $A(D)$ . This notion has been introduced in [7] and is an analog of the *star arboricity* defined in [1].

Let us denote the maximum multiplicity of an edge in a multigraph  $G$  by  $\mu(G)$ . By Vizing's theorem [13], one can colour the edges of a multigraph with  $\Delta(G) + \mu(G)$  colours so that two edges have different colours if they are incident. Since the multigraph underlying a digraph has maximum multiplicity at most two, for any digraph  $D$ ,  $dst(D) \leq \Delta + 2$ . Amini et al. [3] conjecture the following:

**Conjecture 16 (Amini et al. [3]).** *Every digraph  $D$  with maximum degree  $\Delta \geq 3$  satisfies  $dst(D) \leq \Delta$ .*



The condition  $\Delta \geq 3$  in the above conjecture is necessary since the odd circuits have maximum degree 2 and directed star arboricity 3. This conjecture would be tight since every digraph with a vertex with in-degree  $\Delta$  (and out-degree 0) has directed star arboricity at least  $\Delta$ . In [3], Amini et al. proved that Conjecture 16 holds when  $\Delta = 3$ .

A *nice galaxy* in a digraph  $G$  is a galaxy spanning all the vertices of maximum degree. To prove Conjecture 16 by induction on the maximum degree, it suffices to show that every digraph with maximum degree  $\Delta \geq 4$  has a nice galaxy.

**Conjecture 17 (Amini et al. [3]).** *Every digraph with maximum degree  $\Delta \geq 4$  has a nice galaxy.*

Amini et al. [3] showed the conjecture for 2-diregular digraphs. In this section, we prove Conjecture 17 for acyclic digraphs, which implies Conjecture 16 for acyclic digraphs. We also prove that every digraph has a galaxy spanning the vertices with in-degree at least two and derive that  $dst(D) \leq \Delta(D) + 1$  for every digraph  $D$ .

### 6.1. Acyclic digraphs

In this subsection, we settle Conjecture 17 for acyclic digraphs and derive that Conjecture 16 holds for acyclic digraphs. To do so, we need the following lemma on odd-cycles+matching graphs. An *odd-cycles+matching graph* is the disjoint union of odd cycles and a matching.

**Lemma 18.** *Every graph with at least one edge has an odd-cycles+matching subgraph spanning all the vertices of maximum degree.*

**Proof .** Let  $G$  be a graph of maximum degree  $\Delta$  and  $V_\Delta$  be the set of vertices of degree  $\Delta$ . The result holds trivially if  $\Delta = 1$  so we may assume that  $\Delta \geq 2$ . Let  $H$  be an odd-cycles+matching subgraph that spans the maximum number of vertices of  $V_\Delta$ . Let  $C_1, \dots, C_p$  be the odd cycles of  $H$  and  $M$  its matching. Suppose by way of contradiction that there is a vertex  $v$  in  $V_\Delta \setminus V(H)$ . An *alternating  $v$ -path* is a path starting at  $v$  such that every even edge is in  $M$  (and so every odd edge is not in  $M$ ). Let  $A_0$  (resp.  $A_1$ ) be the set of vertices  $u$  such that there exists a  $v$ -alternating path of even (resp. odd) length ending at  $u$ . Note that  $v \in A_0$  as  $(v)$  is an alternating  $v$ -path of length 0.

**Claim 5.**  $A_0 \subset V_\Delta$ .

Suppose that  $A_0 \not\subset V_\Delta$ . Then there is a vertex  $x \in A_0 \setminus V_\Delta$ . Let  $P$  be the even alternating  $v$ -path ending at  $x$ . Then the odd-cycles+matching subgraph obtained from  $H$  by replacing the matching  $M$  by  $M' = M \Delta P$  spans one more vertex of  $V_\Delta$ , namely  $v$ , than  $H$ . This is a contradiction.

**Claim 6.**  $A_1 \subseteq V(H)$ .

Suppose by way of contradiction that a vertex  $x \in A_1$  is in  $V(G) \setminus V(H)$ . Let  $P$  be an odd alternating  $v$ -path ending at  $x$ . Then the odd-cycles+matching subgraph obtained from  $H$  by replacing the matching  $M$  by  $M' = M \Delta P$  spans one more vertex of  $V_\Delta$ , namely  $v$ , than  $H$ . This is a contradiction.

**Claim 7.**  $A_1 \subseteq V(M)$ .

Suppose by way of contradiction that a vertex  $x \in A_1$  is in  $\bigcup_{i=1}^p C_i$ , say in  $C_p$ . Then  $C_p - x$  has a matching  $M_1$ . Let  $P$  be an odd alternating  $v$ -path ending at  $x$ . This path of odd length has a perfect matching  $M_2 = P \setminus M$ . Thus the disjoint union of  $C_1, \dots, C_{p-1}$  and  $(M \setminus P) \cup M_1 \cup M_2$  is an odd-cycles+matching subgraph spanning one more vertex of  $V_\Delta$ , namely  $v$ , than  $M$ . This is a contradiction.

**Claim 8.**  $|A_0| = |A_1| + 1$ .

Indeed,  $M$  matches every vertex of  $A_0$ , except  $v$ , with a vertex of  $A_1$ , and vice versa.

**Claim 9.**  $A_0$  is a stable set.

Suppose to the contrary that there exist two adjacent vertices  $x$  and  $y$  in  $A_0$ . Let  $P_x$  and  $P_y$  be two even alternating  $v$ -paths ending at  $x$  and  $y$ , respectively. We choose  $x, y, P_x$  and  $P_y$  in such a way that  $|V(P_x) \cup V(P_y)|$  is minimum. Note that  $P_x$  and  $P_y$  may share common vertices and arcs at the beginning. If  $xy \in M$ , then  $x$  is the predecessor of  $y$  in  $P_y$  and vice-versa. In this case let  $Q_y = P_x - y$  and  $Q_x = P_y - x$ . Otherwise let  $Q_x = P_x$  and  $Q_y = P_y$ . In both cases,  $Q_x$  and  $Q_y$  are alternating  $v$ -paths of same parity. Note that by minimality of  $|V(P_x) \cup V(P_y)|$  there exists only one vertex  $z \in V(Q_x) \cap V(Q_y)$  (possibly  $z = v$ ) and three paths  $Q_{v-z}, Q_{z-x}$  and  $Q_{z-y}$ , going respectively from  $v$  to  $z$ , from  $z$  to  $x$  and from  $z$  to  $y$  such that  $Q_x = Q_{v-z} \cup Q_{z-x}$ ,  $Q_y = Q_{v-z} \cup Q_{z-y}$ , and  $V(Q_{z-x}) \cap V(Q_{z-y}) = \{z\}$ . Note that we necessarily have  $z \in A_0$  since every odd vertex in  $Q_x$  and  $Q_y$  is followed by its neighbour in  $M$ . Let  $C_{p+1}$  be the odd cycle formed by the paths  $Q_{z-x}$  and  $Q_{z-y}$ , and by the edge  $xy$ . Then the odd-cycles+matching subgraph obtained from  $H$  by replacing the matching  $M$  by  $M' = M \Delta Q_{v-z}$  and adding the odd cycle  $C_{p+1}$  spans one more vertex of  $V_\Delta$  than  $H$ . This is a contradiction.

By Claim 9, all the edges with an end in  $A_0$  have the other end in  $A_1$  and thus, by Claims 5 and 8, there are  $|A_0| \times \Delta = (|A_1| + 1) \times \Delta$  edges between  $A_0$  and  $A_1$ . This is impossible because the vertices in  $A_1$  have degree at most  $\Delta$ . This contradicts the existence of  $v$  and thus proves the Lemma.  $\square$

**Theorem 19.** *Every acyclic digraph has a nice galaxy.*

**Proof .** Let  $D$  be an acyclic digraph and  $G$  its underlying undirected graph  $D$ . By Lemma 18,  $G$  has an odd-cycles+matching subgraph  $H$  spanning all the

vertices of maximum degree. The subdigraph  $D'$  of  $D$  which is an orientation of  $H$  is the union of oriented odd circuits and a matching. Each oriented circuit is not directed because  $D$  is acyclic, and thus has a spanning galaxy. Hence  $D'$  has a spanning galaxy, which is a nice galaxy of  $D$ .  $\square$

**Corollary 20.** *If  $D$  is an acyclic digraph then  $dst(D) \leq \Delta(D)$ .*

**Proof .** We prove the result by induction on  $\Delta(D)$ , the result holding trivially when  $\Delta(D) = 1$ . Suppose now  $\Delta(D) = k > 1$ . By Theorem 19,  $D$  has a nice galaxy  $F_k$ . Hence  $D' = D \setminus E(F_k)$  has maximum degree at most  $k - 1$ . By induction,  $D'$  has an arc-partition into  $k - 1$  galaxies  $F_1, \dots, F_{k-1}$ . Thus  $(F_1, \dots, F_k)$  is an arc-partition of  $D$  into  $k$  galaxies.  $\square$

### 6.2. Galaxy spanning the vertices with in-degree at least two

Let  $D$  be a digraph. The *out-section* of a vertex  $x$  is the set  $S^+(x)$  of vertices  $y$  to which there exists a directed path from  $x$ . An *out-generator* of  $D$  is a vertex  $x \in V(D)$  such that  $S^+(x) = V(D)$ . Note that if  $D$  is strong, every vertex is an out-generator. Every out-generator is the root of a spanning arborescence, so by Lemma 2 we get the following:

**Corollary 21.** *Let  $v$  be an out-generator of a digraph  $D$ . Then  $D$  contains a galaxy  $F$  spanning all the vertices of  $D - v$ .*

**Theorem 22.** *Every digraph  $D$  has a galaxy spanning all the vertices with in-degree at least 2.*

In order to prove this theorem, we need the following folklore proposition. We give its proof for sake of completeness.

**Proposition 23.** *Let  $D$  be a strong digraph with minimum in-degree 2. Then there is a vertex  $x$  such that  $D - x$  is strong.*

**Proof .** Consider a handle decomposition minimizing the number of trivial handles. Let  $(x_0, \dots, x_l)$  be the last non-trivial handle. The vertex  $x_{l-1}$  has in-degree at least two, hence the other arcs entering  $x_{l-1}$  are trivial handles. If  $l$  is greater than 2, any of these trivial handles, together with  $x_0, \dots, x_l$  would result in two non-trivial handles - which is impossible by assumption. Thus  $l = 2$ , and then the vertex  $x_1$  can be deleted.  $\square$

**Proof of Theorem 22.** We prove the result by induction on the number of arcs.

If  $D$  has an arc  $a$  entering a vertex of in-degree either 1 or more than 2, then by induction  $D \setminus a$  has a galaxy  $G$  spanning all the vertices with in-degree at least 2 in  $D \setminus a$ . But the vertices with in-degree at least 2 in  $D$  have also in-degree at least 2 in  $D \setminus a$ . Hence  $G$  spans all the vertices with in-degree at least 2 in  $D \setminus a$ .

Henceforth, we assume that every vertex of  $D$  has in-degree 2 or 0. Suppose first that  $D$  contains a vertex  $v$  of in-degree 0. Set  $D^+ = D[S^+(v)]$  and  $D' = D - D^+$ . By definition of out-section, there are no arcs leaving  $D^+$ . So the vertices of  $D'$  have the same in-degree in  $D'$  and  $D$ . By the induction hypothesis, there is a galaxy  $F'$  spanning all the vertices of  $D'$  with in-degree 2 and by Corollary 21, there is a galaxy  $F^+$  spanning all the vertices of  $D^+$  with in-degree 2. The union of  $F'$  and  $F^+$  is the desired galaxy.

Suppose now that all the vertices of  $D$  have in-degree 2. Consider an initial strong component  $C$  of  $D$ , that is a strong component  $C$  such that all the arcs with head in  $C$  have their tail in  $C$ . By Proposition 23, there exists a vertex  $v$  of  $C$  such that  $C - v$  is strong. Let  $S^+$  be the out-section of  $v$  in  $D - (C \setminus \{v\})$  and  $T = S_D^+(v) \setminus S^+$  and  $D' = D - S_D^+(v)$ . Note that  $v$  is an out-generator of  $D[S^+]$  and  $D_1 = D[T \cup \{v\}]$ . Moreover since  $C - v$  is strong, every vertex of  $C - v$  is an out-generator of  $D_2 = D[T]$ .

By the induction hypothesis, there is a galaxy  $F'$  spanning all the vertices of  $D'$  with in-degree 2. By Corollary 21, there is a galaxy  $F^+$  of  $D[S^+]$  spanning all the vertices of  $S^+ \setminus \{v\}$  in which  $v$  is either not spanned or a root. If  $v$  is a root of  $F^+$  then, by Corollary 21, there is a galaxy  $F_1$  of  $D_1$  spanning all the vertices of  $T$  in which  $v$  is either not spanned or a root. The union of  $F'$ ,  $F^+$  and  $F_1$  is a spanning galaxy of  $D$ . If  $v$  is not a root of  $F^+$ , let  $u$  be an in-neighbour of  $v$ . By Corollary 21, there is a galaxy  $F_2$  of  $D_2$  spanning all the vertices of  $T \setminus \{u\}$  in which  $u$  is either not spanned or a root. The union of  $F'$ ,  $F^+$ ,  $F_2$  and the arc  $uv$  is a spanning galaxy of  $D$ .  $\square$

Note that Theorem 22 implies the result of Amini et al. [3] that a 2-diregular digraph has a spanning galaxy.

**Theorem 24.** *Let  $D$  be a digraph with maximum degree  $\Delta \geq 2$ . Then  $dst(D) \leq \Delta + 1$ .*

**Proof .** Set  $D_0 = D$ . For every  $i$  from 1 to  $\Delta - 2$ , let  $F_i$  be a galaxy spanning all the vertices of in-degree at least 2 in  $D_{i-1}$  and  $D_i = D_{i-1} \setminus E(F_i)$ . Observe that a vertex of  $D' = D_{\Delta-2}$  has either in-degree at most one or in-degree 2 and out-degree 0. Now we just have to prove that  $dst(D') \leq 3$ . For this, choose one arc entering each vertex with in-degree two and denote the set of these arcs by  $F$ . In the graph  $D' \setminus F$  every vertex has in-degree exactly 1. Consider a 3-colouring of the arcs of  $D' \setminus F$  such that each colour class induces a galaxy. This is possible because  $D' \setminus F$  is a disjoint union of functional graphs. Then for every arc  $xy$  of  $F$ , at most two colours are forbidden, one by the other arc entering  $x$ , and another by the arc entering  $y$ . Indeed,  $x$  has no out-neighbour except  $y$  and  $y$  has in-degree at most one (since  $d^+(y) > 0$ ).  $\square$

## 7. Open questions

Deciding if a strong digraph has an even strong subdigraph is polynomial-time solvable. By Theorem 9, this is equivalent to deciding if strong digraph

contains an even circuit or an even theta. Deciding whether a strong digraph contains an even circuit can be solved in polynomial time [10, 11]. Hence a natural question is the following.

**Problem 25.** Can we decide in polynomial time whether a strong digraph contains an even theta?

One can also search for even thetas or circuits using or avoiding a prespecified arc. Thomassen [12] showed that it is NP-complete to decide whether a digraph  $D$  has an even circuit containing a given arc. Theorem 13 states that it is polynomial-time decidable if a digraph has an even strong subdigraph containing a given arc. Hence it is natural to ask the following two questions.

**Problem 26.** Can we decide in polynomial time whether a digraph has an even theta containing a given arc?

**Problem 27.** Can we decide in polynomial time whether a digraph has an even circuit or an even theta containing a given arc ?

Observe that Theorems 13 and 9 do not imply an affirmative answer to Problem 27. Indeed, there are even strong subdigraphs with some arcs in no even circuit nor even theta. For example, consider the even digraph with vertex set  $\{a_1, \dots, a_8\}$  and edges  $\{a_i a_{i+1} \mid 1 \leq i \leq 6\} \cup \{a_4 a_1, a_5 a_8, a_8 a_4, a_7 a_5\}$ . It is easy to check that the arcs  $a_5 a_6$ ,  $a_6 a_7$  and  $a_7 a_5$  are in no even circuit and in no even theta.

It is easy to find a polynomial-time algorithm to decide whether a digraph  $D$  has an even circuit avoiding a given arc  $uv$ : it suffices to test whether  $D \setminus \{uv\}$  has an even circuit or not. Similarly, one can decide in polynomial time whether a digraph  $D$  has an even circuit or an even theta avoiding a given arc  $uv$ , by testing if one of the strong components of  $D \setminus \{uv\}$  contains an even strong subdigraph. But we do not know about the complexity of testing if a digraph has an even theta avoiding a prespecified arc.

**Problem 28.** Can we decide in polynomial time whether a digraph has an even theta avoiding a given arc ?

## References

- [1] I. Algor and N. Alon, The star arboricity of graphs, *Discrete Mathematics* **75** (1989), 11–22.
- [2] N. Alon, C. McDiarmid, B. Reed, Star arboricity, *Combinatorica* **12** (1992), 375–380.
- [3] O. Amini, F. Havet, F. Huc and S. Thomassé, WDM and directed star arboricity, *Combinatorics, Probability and Computing* **19** (2010), 161–182.
- [4] J. Bang-Jensen and G. Gutin, *Digraphs: theory and applications*, 2nd edition xxii+ 816 pp, January 2009.

- [5] R. Brandt, *Multicasting using WDM in Multifiber Optical Star Networks*, Thesis, UCSB, September 2003.
- [6] R. Downey and M. Fellows, Fixed parameter tractability and completeness I: basic results, *SIAM J. Comput.*, **24** (1995), 873–921.
- [7] B. Guiduli, On incidence coloring and star arboricity of graphs, *Discrete Mathematics* **163** (1997), 275–278.
- [8] J. Kratochvíl, A special planar satisfiability problem and a consequence of its NP-completeness, *Discrete Applied Mathematics* **52(3)** (1994), 233–252.
- [9] L. Lovász, A note on factor-critical graphs, *Studia Sci. Math. Hungar.*, **7** (1972), 279–280.
- [10] W. McCuaig, Pólya’s permanent problem, *Electron. J. Combin.* **11(1)** (2004), Research Paper 79, 83 pp. (electronic).
- [11] N. Robertson, P. D. Seymour and R. Thomas, Permanents, Pfaffian orientations, and even directed circuits, *Ann. of Math. (2)* **150(3)** (1999), 929–975.
- [12] C. Thomassen, Even cycles in directed graphs, *European J. Combin.*, **6(1)** (1985), 85–89.
- [13] V. G. Vizing. On an estimate of the chromatic class of a p-graph. *Metody Diskret. Analiz.* 3:25–30, 1964.