# REPORTS IN INFORMATICS

ISSN 0333-3590

## On the Complexity of Reconstructing $H$-free Graphs from their Star Systems

Fedor V. Fomin Jan Kratochvíl
Daniel Lokshtanov Federico Mancini
Jan Arne Telle

REPORT NO 360 August 2007

*Department of Informatics*
UNIVERSITY OF BERGEN
*Bergen, Norway*

This report has URL http://www.ii.uib.no/publikasjoner/texrap/ps/2007-360.ps

Reports in Informatics from Department of Informatics, University of Bergen, Norway, is available at http://www.ii.uib.no/publikasjoner/texrap/.

Requests for paper copies of this report can be sent to:

Department of Informatics, University of Bergen, Høyteknologisenteret,
P.O. Box 7800, N-5020 Bergen, Norway

# On the Complexity of Reconstructing $H$-free Graphs from their Star Systems

Fedor V. Fomin* Jan Kratochvíl† Daniel Lokshtanov* Federico Mancini*
Jan Arne Telle*

**Abstract**

In the Star System problem we are given a set system and asked whether it is realizable by the multi-set of closed neighborhoods of some graph, i.e. given subsets $S_1, S_2, ..., S_n$ of an $n$-element set $V$ does there exist a graph $G = (V, E)$ with $\{N[v] : v \in V\} = \{S_1, S_2, ..., S_n\}$? For a fixed graph $H$ the $H$-free Star System problem is a variant of the Star System problem where it is asked whether a given set system is realizable by closed neighborhoods of a graph containing no $H$ as an induced subgraph. We study the computational complexity of the $H$-free Star System problem. We prove that when $H$ is a path or a cycle on at most 4 vertices the problem is polynomial time solvable. In complement to this result, we show that if $H$ belongs to a certain large class of graphs the $H$-free Star System problem remains NP-complete. In particular, the problem is NP-complete when $H$ is either a cycle or a path on at least 5 vertices. This yields a complete dichotomy for paths and cycles.

*Department of Informatics, University of Bergen, N-5020 Bergen, Norway. Email: {fomin, daniello, federico, telle}@ii.uib.no.

†Dept. of Applied Mathematics and Institute for Theoretical Computer Science, Charles University, Praha, Czech Republic. Supported by Czech research grant 1M0545.

# 1 Introduction

The closed neighborhood of a vertex in a graph is sometimes called the "star" of the vertex. The "star system" of a graph is then the multi-set of closed neighborhoods of all the vertices of the graph and the Star System problem is the problem of deciding whether a given system of sets is a star system of some graph. The Star System problem is a natural combinatorial problem that fits into a broader class of *realizability* problems. In a realizability problem we are given a list $P$ of invariants or properties (like a sequence of vertex degrees, set of cliques, number of colorings, etc) and the question is whether the given list is *graphical*, i.e. corresponds to the list of parameters of some graph. One of the well studied problems of realizability is the case when $P$ is a degree sequence. This can be seen as a modification of the Star System problem where, instead of stars, the list $P$ contains only the sizes of the stars. In this case, graphic sequences can be characterized by the Erdős-Gallai Theorem [6].

The Star System problem (also known as the Closed Neighborhood Realization problem) is equivalent to a number of other interesting problems. For example, it is equivalent to the question of whether a given $0-1$ matrix $A$ is symmetrizable, i.e. whether by permuting rows (or columns) $A$ can be turned into a symmetric matrix with all diagonal entries equal to 1. We refer to the recent survey of Boros et al. [5] for further equivalent problems related to the Matrix Symmetrization and Star System problems.

The question of the computational complexity of the Star System problem was first posed by Gert Sabidussi and Vera Sós at a conference in the mid-70s [7] (and since this appears to be the oldest reference to the problem, we choose to use the Star System terminology). At the same conference Babai observed that the Star System problem was at least as hard as the Graph Isomorphism problem. There are strong similarities with Graph Isomorphism, e.g., the Star System problem is equivalent to deciding if a given bipartite graph allows an automorphism of order 2 such that each vertex is adjacent to its image. In view of these connections to Graph Isomorphism the NP-hardness of the Star System problem came as a surprise. The proof of this fact was achieved in two steps. First, a related effort of Lubiw [9] showed that deciding whether an arbitrary graph has an automorphism of order 2 is NP-complete. Then Lalonde [8] showed that the Star System problem was NP-complete by a reduction from Lubiw's problem. This reduction came as a small surprise considering that, after Lubiw's proof, Babai had written that between Lubiw's problem and the Star System problem he "did not believe there was a deeper relationship" [3].

The result of Lalonde was rediscovered by Aigner and Triesch [1, 2] who proved it in a stronger form, and indeed discovered a subproblem which is equivalent to Graph Isomorphism. It is easy to see that the problems of reconstructing graphs from their closed neighborhood hypergraphs and from their open neighborhood hypergraphs are polynomially equivalent. It is more convenient however, to describe the results of Aigner and Triesch in the language of open neighborhoods. They proved that deciding if a set system is the open neighborhood hypergraph of a *bipartite* graph is Graph Isomorphism-complete, while deciding if the open neighborhood hypergraph of a bipartite graph can be realized by a nonisomorphic (and non-bipartite) graph becomes again NP-complete.

Since bipartite graphs (and their complements) are hereditary classes of graphs, it is natural to pay closer attention to restriction of the Star System problem to classes of graphs defined by forbidden induced subgraphs. The problem we investigate in this paper is the following variation of the Star System problem, for a fixed graph $H$:

**$H$-free Star System Problem**
Input: A set system $\mathcal{S}$ over a ground set $V$
Question: Does there exist an $H$-free graph $G = (V, E)$ such that $\mathcal{S}$ is the star system of $G$?

Our main result is a complete dichotomy in the case when $H$ is either a cycle $C_k$, or a path $P_k$ on $k$ vertices. We prove that the $H$-free Star System problem for $H \in \{C_k, P_k\}$ is polynomial time solvable when $k \leq 4$ (Section 3) and NP-complete when $k > 4$ (Section 4). Our NP-completeness result for paths and cycles follows from a more general result, which shows that there exists a much larger family of graphs for which if $H$ belongs to it then the $H$-free Star System problem is NP-complete.

## 2 Preliminaries

We use standard graph notation with $G = (V, E)$ being a simple loopless undirected graph with vertex set $V$ and edge set $E$. We denote by $N[v]$ and $N(v)$ the closed and open neighborhoods of a vertex $v$, respectively, and by $\overline{G}$ the complement of a graph $G$ having an edge $uv$ iff $u \neq v$ and $uv \notin E(G)$. We also call $N[v]$ the star of $v$ and say that $v$ is the center of $N[v]$. An *automorphism* of a graph $G = (V, E)$ is an isomorphism $f : V \to V$ of the graph to itself, and it has order 2 if for every vertex $x$ we have $f(f(x)) = x$, i.e. the image of its image is itself. For a graph $G = (V, E)$ we define the $|V|$-element multi-set $Stars(G) = \{N[v] : v \in V\}$. For a fixed graph $H$ we say that a graph $G$ is $H$-free if $G$ does not contain an induced subgraph isomorphic to $H$.

## 3 Forbidding short paths and cycles

### 3.1 Forbidding short paths

In this section we show that $P_k$-free Star System Problem is solvable in polynomial time for $k \leq 4$. For $k \leq 2$ the $P_k$-free System Problem is trivially polynomial time solvable. For $k = 3$ the realizable graph is a disjoint union of cliques, and in this case the problem is again trivial. The proof that $P_4$-free Star System can be solved in polynomial time occupies the remaining part of this subsection.

The graphs without induced $P_4$ are called *cographs*. We exploit the following characterization of cographs.

**Proposition 1 ([4]).** *A graph $G$ is a cograph if and only if every non-trivial induced subgraph of $G$ contains at least one pair of vertices $x$ and $y$, such that either $N[x] = N[y]$ or $N(x) = N(y)$.*

For a set system $\mathcal{S}$ over a ground set $V$, we say that $(x, y) \in V^2$ is a *closed pair* (of $\mathcal{S}$) if for every $S \in \mathcal{S}$ we have that $x \in S$ if and only if $y \in S$. Also, we define $(x, y)$ to be an *open pair* (of $\mathcal{S}$) if there is exactly one set $S_x \in \mathcal{S}$ containing $x$ and not $y$, exactly one set $S_y \in \mathcal{S}$ containing $y$ and not $x$, and $S_x \setminus \{x\} = S_y \setminus \{y\}$.

We will show that a given set system $\mathcal{S}$ is the star system of a cograph $G$ if and only if it can be reduced to one set on a single element by sequentially contracting closed and open pairs. We start by showing that, if $\mathcal{S}$ is the star system of a cograph, then $\mathcal{S}$ contains a closed or an open pair.

**Lemma 2.** *If $\mathcal{S} = Stars(G)$ for a nontrivial cograph $G = (V, E)$ then $\mathcal{S}$ has either a closed or an open pair.*

*Proof.* First of all recall that, since $\mathcal{S} = Stars(G)$, for every $v \in V$ we have that $N[v] = S_v \in \mathcal{S}$. By Proposition 1, $G$ has a pair of vertices $x$ and $y$ such that either $N[x] = N[y]$ or $N(x) = N(y)$. In the first case $(x, y)$ is a closed pair in $\mathcal{S}$, because for every $z \in V$ we have that $x \in N[z]$ if and only if $y \in N[z]$. Hence every element of $\mathcal{S}$ contains either both $x$ and $y$, or none of them. In the latter case, $N[x]$ contains $x$ and not $y$, $N[y]$ contains $y$ and not $x$, and $N[x] \setminus \{x\} = N(x) = N(y) = N[y] \setminus \{y\}$. Thus for every $z \in V \setminus \{x, y\}$ it follows that $x \in N[z]$ if and only if $y \in N[z]$. This means that $(x, y)$ is an open pair because every element in $S$ contains either both $x$ and $y$, or none of them, and there are exactly two elements, $S_x$ and $S_y$, such that $S_x \setminus \{x\} = S_y \setminus \{y\}$. □

In the rest of the subsection, we will show how to contract closed and open pairs respectively, such that the resulting star system represents a cograph if and only if it did before the contraction. We start with the closed pairs.

**Lemma 3.** *Given a set system $\mathcal{S}$, let $R$ be an inclusion maximal subset of $V$ containing no closed pairs, and let $\mathcal{S}_R$ be the set $\{Z \mid \exists S \in \mathcal{S}$ such that $S \cap R = Z\}$. Then there is a cograph $G = (V, E)$ with $Stars(G) = \mathcal{S}$ if and only if there is a cograph $G'$ with $Stars(G') = \mathcal{S}_R$.*

*Proof.* Let us assume that there is a cograph $G$ with $Stars(G) = \mathcal{S}$ and let $G[R] = G'$. Then we claim that $Stars(G') = \mathcal{S}_R$. First, note that for every $v \in R$, $N_{G'}[v]$ is in $\mathcal{S}_R$. Additionally, consider $S \in \mathcal{S}_R$. Let $x$ be a vertex of $G$ such that $S = N[x] \cap R$. By maximality of $R$ there is $x' \in R$ such that $N[x'] = N[x]$, thus $N_{G'}[x'] = N[x] \cap R = S$. Hence, for every $S \in \mathcal{S}_R$ there is a $v$ in $R$ such that $S = N_{G'}[v]$. Finally, observe that by construction $\mathcal{S}_R$ has no duplicate elements, and that $Stars(G')$ also has no duplicate elements because a duplicate element would imply that $R$ contains a closed pair, contradicting that $R$ has none. Together, this implies that $Stars(G') = \mathcal{S}_R$.

On the other hand, suppose that there is a cograph $G'$ with $Stars(G') = \mathcal{S}_R$. For every element $v$ in $V$ there is a unique $x$ in $R$ such that $(v, x)$ is a closed pair. Let $f : V \to R$ be the mapping such that for every vertex $v \in V$, the pair $(v, f(v))$ is a closed pair. Also, let $f^{-1}$ be the inverse image of $f$. That is, for a vertex $x \in R$, we have that $f^{-1}(x) = \{v \in V \mid f(v) = x\}$. Finally, let $G = (V, \{(u, v) : (f(u), f(v)) \in E(G')\})$. We claim that $Stars(G) = \mathcal{S}$ and that $G$ is a cograph.

For each $v \in V$, $N[v] = \bigcup_{x \in N_{G'}[f(v)]} f^{-1}(x)$ by the definition of $G$ and $N_{G'}[f(v)] \in \mathcal{S}_R$. Furthermore, for every set $S \in \mathcal{S}_R$, we have that $\bigcup_{x \in S} f^{-1}(x) \in \mathcal{S}$ by definition of $\mathcal{S}_R$. Thus, since for every $v \in V$ there exists $S \in \mathcal{S}_R$ such that $S = N_{G'}[f(v)]$, we can conclude that $N[v]$ is in $\mathcal{S}$. On the other hand, for every $S \in \mathcal{S}$, there exists $S'$ in $\mathcal{S}_R$ such that $S = \bigcup_{x \in S'} f^{-1}(x)$. Let $u$ be the element of $R$ such that $N_{G'}[u] = S'$. Then $S = \bigcup_{x \in N_{G'}[u]} f^{-1}(x) = N[u]$. This means that $S \in Stars(G)$. Finally, we know that for each vertex $u \in R$, there are $|f^{-1}(u)|$ copies of the star $\bigcup_{x \in N_{G'}[u]} f^{-1}(x) = N[u]$ in $S$. Also, we know that in $V$ there are $|f^{-1}(u)|$ vertices with the same closed neighborhood $N[u]$ for each $u \in R$. This proves that $Stars(G) = \mathcal{S}$.

We will now prove that $G$ is also a cograph, by showing that it does not contain an induced $P_4$. Observe first that $G[R] = G'$ is a cograph by definition. For the sake of contradiction, let us assume that there is a set $P \subseteq V$ of 4 vertices that induces a $P_4$ in $G$. If there is a pair $u$ and $v$ of distinct vertices in $P$ such that $f(u) = f(v)$ then $N[u] = N[v]$, and specifically $u$ and $v$ have the same neighbourhood in $P$ which is impossible because no pair of distinct vertices of a $P_4$ have the same neighbourhood. If no such pair exists, then $P' = \{f(x) : x \in P\}$ must induce a $P_4$ in $G'$ as well. This leads to a contradiction, concluding the proof of the lemma. ☐

Before we show how to contract the open pairs, we give a lemma to resolve some ambiguity about open pairs and cographs. Notice, in fact, that given a cograph $G$ and the corresponding star system, there might be an open pair $(x, y)$ such that $N_G(x) \neq N_G(y)$. However, we will prove that given any open pair, we can always find a cograph $G'$ isomorphic to $G$, such that $N'_G(x) = N'_G(y)$.

**Lemma 4.** *If $Stars(G) = \mathcal{S}$ for a cograph $G = (V, E)$ and $(x, y)$ is an open pair of $\mathcal{S}$, there is a cograph $G' = (V', E')$ such that $Stars(G') = \mathcal{S}$, $xy \notin E'$ and $N_{G'}(x) = N_{G'}(y)$.*

*Proof.* If $xy \notin E$, then $N[x]$ contains $x$ and not $y$, $N[y]$ contains $y$ and not $x$, so by uniqueness of $S_x$ and $S_y$ we have that $N(x) = N(y)$. By letting $G' = G$ we are done. Now, let us assume $xy \in E$. Then there are vertices $x'$ and $y'$ such that $N[x']$ contains $x$ and not $y$ and $N[y']$ contains $y$ and not $x$. Clearly, $x$, $y$, $x'$ and $y'$ must be distinct vertices. Following this, if $x'y' \notin E$, then $\{x', x, y, y'\}$ induces a $P_4$ in $G$. Thus, as $G$ is a cograph, $x'y' \in E$. This means that $C = \{x', x, y, y'\}$ induces a $C_4$ in $G$.

We now proceed to show that $C$ is a module of $G$, that is, for any $z \in V \setminus C$, either $N[z] \cap C = C$ or $N[z] \cap C = \emptyset$. Observe that as $x$ and $y$ are an open pair, $N[x'] \setminus \{x\} = N[y'] \setminus \{y\}$ so $x' \in N[z]$ if and only if $y' \in N[z]$. As $x'$ is the only vertex such that $N[x']$ contains $x$ but not $y$ and $y'$ is the only vertex such that $N[y']$ contains $y$ and not $x$ it follows that $x \in N[z]$ if and only if $y \in N[z]$. For the sake of contradiction, let us suppose that $x \in N[z]$ and $y' \notin N[z]$. Then, by the discussion above $x' \notin N[z]$ so $\{z, x, x', y'\}$ induces $P_4$ in $G$, contradicting that $G$ is a cograph. Let us assume now that

$x \notin N[z]$ and $y' \in N[z]$. Similarly to the previous case, $x' \notin N[z]$ so $\{z, y', x', x\}$ induces a $P_4$ in $G$, again giving a contradiction. From this it follows that $x \in N[z]$ if and only if $y' \in N[z]$. Together with the equivalences above this proves that each of the vertices $z, y', x'$ and $x$, is in $N[z]$ for a given $z$, if and only if the other three are as well. This means that $C$ is a module of $G$.

We build $G'$ from $G$ by simply switching the labels of $y$ and $y'$. We prove that $G'$ meets the requirements of the statement of the Lemma. Clearly $(x, y) \notin E(G')$ and $N_{G'}(x) = N_{G'}(y)$. Furthermore for any $z \in V \setminus C$. $N_{G'}[z] = N[z]$. It remains to show that $\{N[x], N[y], N[x'], N[y']\} = \{N_{G'}[x], N_{G'}[y], N_{G'}[x'], N_{G'}[y']\}$. But $N[x] = N_{G'}[x']$, $N[x'] = N_{G'}[x]$, $N[y] = N_{G'}[y']$ and $N[y'] = N_{G'}[y]$. Thus $Stars(G') = Stars(G) = \mathcal{S}$ which concludes the proof. □

We are now ready to give the contraction rule for open pairs.

**Lemma 5.** *Let $(x, y)$ be an open pair of $\mathcal{S}$, and let $\mathcal{S}'$ be the set system obtained by deleting the unique star of $S$ containing $x$ but not $y$, and removing $x$ from all the other elements of $\mathcal{S}$. There is a cograph $G$ with $Stars(G) = \mathcal{S}$ if and only if there is a cograph $H$ with $Stars(H) = \mathcal{S}'$.*

*Proof.* Suppose there is a cograph $G$ with $Stars(G) = \mathcal{S}$. By Lemma 4, there exists a cograph $G'$ for which $Stars(G') = \mathcal{S}$ and the only star in $\mathcal{S}$ containg $x$ but not $y$, is exactly $N_{G'}[x]$. This means that removing the set representing $N_{G'}[x]$ from $\mathcal{S}$ and $x$ from all the other elements of $\mathcal{S}$, we get exactly the star system of $H = G' \setminus \{x\}$ which clearly is a cograph.

Suppose now that there is a cograph $H$ with $\mathcal{S}' = Stars(H)$. We build $G$ from $H$ by adding the vertex $x$ and making $x$ adjacent to the open neighbourhood of $y$. Clearly $Stars(G) = \mathcal{S}$. We prove that $G$ is a cograph by obtaining a contradiction. Observe that both $G \setminus \{x\}$ and $G \setminus \{y\}$ are isomorphic to $H$, meaning that they cannot contain any induced $P_4$. Thus, if there is a $P_4$ in $G$, it contains both $x$ and $y$. However $x$ and $y$ have the same open neighbourhood, which leads to a contradiction because no pair of distinct vertices of a $P_4$ has the same open neighbourhood. □

We are now in the position to prove the main result of this subsection.

**Theorem 6.** *The $P_4$-free Star System Problem is solvable in $O(n^4)$ time.*

*Proof.* To decide whether a given set system $\mathcal{S}$ is a star system of a $P_4$-free graph, we use the following algorithm. If $\mathcal{S}$ has an open pair, apply Lemma 5 to create a new and smaller set system $\mathcal{S}'$ that is a star system of a cograph if and only if $\mathcal{S}$ is. Apply the algorithm recursively on $S'$. If $\mathcal{S}$ has a closed pair, apply Lemma 3 to create a new and smaller set system $\mathcal{S}_R$ that is a star system of a cograph if and only if $\mathcal{S}$ is. Apply the algorithm recursively on $\mathcal{S}_R$. If $\mathcal{S}$ contains a single set on a single element, answer yes. If neither of the above cases apply, answer no. Correctness follows directly from Lemma 2. Let us argue for the runtime. We store our set system so that we can insert, delete and check membership in a set in constant time. At every step of the algorithm, if we do not answer "no", we reduce the set system by at least one element by applying Lemma 5 or Lemma 3. Hence the algorithm can have at most $n$ main steps. To check whether a given pair is an open pair or a closed pair takes $O(n)$ time, therefore finding all closed, or all open pairs takes $O(n^3)$ time. When we apply Lemma 5 to reduce the graph we need to remove a set and an element from all other sets. This can be done in $O(n)$ time. When Lemma 3 is applied, we need to delete at most n elements from all sets, and then remove all duplicate sets. Deleting the elements takes $O(n^2)$ time, while finding and deleting all duplicates takes $O(n^3)$ time. Thus we can conclude that the algorithm terminates in $O(n^4)$ time. □

## 3.2 Forbidding $C_3$ and $C_4$

In this subsection we show that $C_3$-free and $C_4$-free Star System Problems are solvable in polynomial time.

**Theorem 7.** *The $C_3$-free Star System problem is solvable in $O(n^3)$ time.*

*Proof.* Let $\mathcal{S}$ be a set system on a ground set $V$. The crucial observation is that if $\mathcal{S}$ is a star system of a $C_3$-free graph $G = (V, E)$, then for every edge $uv \in E$ there are exactly two sets containing $u$ and $v$. In fact, since $uv \in E$, we have that $u$ and $v$ should be in at least two stars, one of which is centered in $u$ and one centered in $v$. Let $S_u$ and $S_v$ be these stars. If there is a third star $S$ containing $u$ and $v$, then the center of this star, $x \neq u, v$ is adjacent to $u$ and $v$, and thus $xuv$ forms a $C_3$ in $G$, which is a contradiction.

Let us assume that the system $\mathcal{S}$ is connected, i.e. for every two elements $u$ and $v$ there is a sequence of elements $u = u_1, u_2, \ldots, u_k = v$ such that for every $i \in \{1, \ldots, k-1\}$ there is a set $S \in \mathcal{S}$ containing $u_i$ and $u_{i+1}$. (If $\mathcal{S}$ is not connected, then we apply our arguments for each connected component of $\mathcal{S}$.)

Assume that we have correctly guessed the star $S_v \in \mathcal{S}$ of a vertex $v$ in some $C_3$-free graph $G$ with $Stars(G) = \mathcal{S}$. Then each $x \in S_v$, $x \neq v$, is adjacent to $v$ in $G$. Thus there is a unique star $S_x \neq S_v$ containing both $v$ and $x$, and vertex $x$ should be the center of $S_x$. Now every vertex $y$ from $S_x$ should have a unique star containing $x$ and $y$, and so on. Since $\mathcal{S}$ is connected, we thus have that after guessing the star for the first vertex $v$ we can uniquely assign stars to the remaining vertices. There are at most $n$ guesses to be made for the first vertex and we can in $O(n^2)$ time check the correctness of the guess, i.e. check if the star system of the constructed graph corresponds to $\mathcal{S}$, to prove the theorem. □

**Theorem 8.** *The $C_4$-free Star System Problem is solvable in $O(n^3)$ time.*

*Proof.* The proof is based on the following observation. Let $G = (V, E)$ be a $C_4$-free graph and let $x, y \in V$. Let $S_1, S_2, \ldots, S_t$ be the set of stars of $G$ containing both $x$ and $y$. Then

$$|\bigcap_{i=1}^{t} S_i| \leq t \text{ if } xy \in E \tag{1}$$

$$|\bigcap_{i=1}^{t} S_i| \geq t + 2 \text{ if } xy \notin E \tag{2}$$

In fact, if $xy \in E$, then $x$ and $y$ have $t - 2$ common neighbors. Every vertex $v \in \cap_{i=1}^{t} S_i \setminus \{x, y\}$ is adjacent to $x$ and $y$, thus $v$ is the center of the star $S_i$ for some $i \in \{1, \ldots, t\}$ and (1) follows.

If $xy \notin E$, then $x$ and $y$ have $t$ neighbors in common. Moreover, because $G$ is $C_4$-free, these neighbors form a clique in $G$. Thus $\cap_{i=1}^{t} S_i$ contains all these $t$ vertices plus the vertices $x$ and $y$ which yields (2).

Given a set system $\mathcal{S}$ on ground set $V$, the algorithm checking if $\mathcal{S}$ is a star system of some $C_4$-free graph is simple. First, if there is a pair $x, y \in V$ with $S_1, ..., S_t$ being the set of stars containing both vertices and such that

$$|\bigcap_{i=1}^{t} S_i| = t + 1$$

we immediately conclude that the answer is no.

If there is no such pair, then we construct a graph $G = (V, E)$ with $xy \in E$ if and only if the sets of $\mathcal{S}$ containing both $x$ and $y$ satisfy (1). Finally, we check if $Stars(G) = \mathcal{S}$. If this is the case then the answer is yes, otherwise the answer is no.

□

# 4 Forbidding long paths and cycles

In this section we show that there exists an infinite family of graphs $H$ for which the $H$-free Star System problem is NP-complete. In particular both $P_k$ and $C_k$, with $k > 4$, belong to it.

**Definition 9.** For an arbitrary graph $H$, we define $B(H)$ to be its bipartite neighborhood graph, i.e., the bipartite graph with both color classes having $|V(H)|$ vertices labelled by $V(H)$ and having an edge between a vertex labelled $u$ in one color class and a vertex labelled $v$ in the other color class iff $uv \in E(H)$.

For example, for the cycle on 5 vertices $C_5$, we have $\overline{C_5} = C_5$ and $B(\overline{C_5}) = C_{10}$. Our main NP-completeness result is that the $H$-free Star System problem is NP-complete whenever $B(\overline{H})$ has a cycle or two vertices of degree larger than two in the same connected component. For a bipartite graph $G = (V, E)$ with color classes $V_1, V_2$ we say that an automorphism $f : V \to V$ is side-switching if $f(V_1) = V_2$ and $f(V_2) = V_1$. Consider the following two problems.

**AUT-BIP-2SS**
Input: A bipartite graph $G$
Question: Does $G$ have an automorphism of order 2 that is side-switching?

**AUT-BIP-2SS-NA**
Input: A bipartite graph $G$
Question: Does $G$ have an automorphism of order 2 that is side-switching where every vertex and its image are non-adjacent?

Lalonde [8] has shown that the AUT-BIP-2SS problem is NP-complete. Together with Sabidussi he also reduced AUT-BIP-2SS to AUT-BIP-2SS-NA. The proof of our main NP-completeness result is a (nontrivial) refinement of the reduction of Lalonde-Sabidussi, which will ensure that AUT-BIP-2SS-NA remains NP-complete for various restricted classes of bipartite graphs.

To relate NP completeness of AUT-BIP-2SS-NA to the Star System Problem, we use the following lemma.

**Lemma 10.** *If AUT-BIP-2SS-NA is NP-complete on bipartite $B(\overline{H})$-free graphs, then the $H$-free Star System Problem is NP-complete.*

*Proof.* We reduce the first problem, which takes as input a bipartite $B(\overline{H})$-free graph $F$, to the second, which takes as input a set system $S$. We may assume the two partition sides of $F$ are of equal size, since otherwise an automorphism switching the two sides cannot exist. Let the vertices of one color class of $F$ be $\{v_1, v_2, ..., v_n\}$ and of the other $\{w_1, w_2, ..., w_n\}$. The set system we construct will be $S = \{S_1, S_2, ..., S_n\}$ where $S_i = \{w_j : v_i w_j \notin E(F)\}$, i.e. the non-neighbors of $v_i$ on the other side.

As already noted by Babai [3], it is not hard to see that $F$ is a Yes-instance of AUT-BIP-2SS-NA iff there exists a graph $G$ with $Stars(G) = S$. Let us give the argument. The equivalence of those two problems is most naturally proved by noting that they are both equivalent to the question if the bipartite complement $C_F$ of $F$, with $V(C_F) = V(F)$ and $E(C_F) = \{v_i w_j : v_i w_j \notin E(F)\}$, has an automorphism of order 2 such that every vertex *is adjacent* to its image, and thus also side-switching.

It remains to show that if $Stars(G) = S$ then $G$ must be $H$-free. First note that if $Stars(G) = S$, then its bipartite closed neighborhood graph $C(G)$ - constructed by adding to its bipartite neighborhood graph $B(G)$ all $|V(H)|$ edges between pairs of vertices having the same label - is isomorphic to $C_F$. We therefore have that $B(\overline{G}) = F$, in other words, the bipartite neighborhood graph of the complement of $G$ is isomorphic to $F$. Moreover, if $H$ is an induced subgraph of $G$ then clearly $B(\overline{H})$ is an induced subgraph of $B(\overline{G}) = F$ and thus since $F$ is $B(\overline{H})$-free we must have $G$ being $H$-free. □

**Definition 11.** Let $D_p$ be the class of bipartite graphs of girth larger than $p$ where any two vertices of degree three or more have distance at least $p$.

**Theorem 12.** *For any integer $p$ the problem AUT-BIP-2SS-NA is NP-complete even when restricted to graphs in $D_p$.*

*Proof.* We reduce from the NP-complete AUT-BIP-2SS problem and adapt the construction given by Lalonde and Sabidussi [8] for our purposes.

Given a bipartite graph $G = (V, E)$ with color classes $A$ and $B$ we describe how to construct $H \in D_p$ with the property that $G$ is a yes-instance of AUT-BIP-2SS iff $H$ is a yes-instance of AUT-BIP-2SS-NA. Note firstly that we can assume $G$ has no vertex $v$ of degree 1 since if we remove each such $v$ (simultaneously) and add a cycle of length $2k$, where $k$ is greater than the maximum cycle length in $G$, attached to the unique neighbor of $v$, then $G$ has a side-switching automorphism of order 2 if and only if the new graph has one.

Let $p'$ be the smallest even integer at least as large as $p$. Let $H$ be the graph obtained by replacing each edge of $G$ by two paths of length $p' + 1$. Note that the inner vertices of these paths are then the only vertices of degree 2 in $H$. Moreover, we have $H \in D_p$ and the two color classes of $H$ respect $A$ and $B$.

If $f : V(G) \to V(G)$ is an order-two side-switching automorphism of $G$, then define $g : V(H) \to V(H)$ as follows:

- $g(v) = f(v)$ for every $v \in A \cup B$,
- for the newly added vertices of degree 2, let $u, uv_1^1, uv_2^1, \ldots, uv_{p'}^1, v$ and $u, uv_1^2, uv_2^2, \ldots, uv_{p'}^2, v$ be the two paths joining $u$ and $v$, and let $x, xy_1^1, xy_2^1, \ldots, xy_{p'}^1, y$ and $x, xy_1^2, xy_2^2, \ldots, xy_{p'}^2, y$ be the two paths joining $x = f(v)$ and $y = f(u)$. Then set $g(uv_j^i) = xy_{p'+1-j}^{3-i}$ for $i = 1, 2$ and $j = 1, 2, \ldots, p'$.

It is straightforward to see that $g$ is an order-two side-switching automorphism. The only place where $ug(u)$ might be an edge would be in the middle of a path $u, uv_1^1, uv_2^1, \ldots, uv_{p'}^1, v$ when $f(u) = v$, but note that then vertices of one path are mapped onto vertices of the other one and $xg(x) \notin E(H)$ is fulfilled.

On the other hand, suppose $g : V(H) \to V(H)$ be an order-two side-switching automorphism of $H$. Since the original vertices of $G$ have degrees greater than 2 in $H$, the restriction of $g$ to $V(G)$ is a correctly defined mapping $g : V(G) \to V(G)$. Since the paths of length $p' + 1$ uniquely correspond to edges of $G$, this restriction of $g$ is an automorphism of $G$. It is obviously of order 2, and since the sides of $H$ respect the sides of $G$, it is side-switching. (Note that we even did not need to assume that $ug(u) \notin E(H)$ for this implication.) □

**Definition 13.** Let $H$ be a graph. We define a function $f(H)$ from graphs to integers and infinity. If $B(\overline{H})$ is acyclic with no connected component having two vertices of degree larger than two then let $f(H) = \infty$. Otherwise, let $f(H)$ be the smallest of i) the length of the smallest induced cycle of $B(\overline{H})$, and ii) the length of the shortest path between any two vertices of degree larger than two in $B(\overline{H})$.

For example, for the cycle on 5 vertices $C_5$, we have $B(\overline{C_5}) = C_{10}$ and thus $f(C_5) = 10$. Note that if $f(H) \neq \infty$ then $D_{f(H)}$ is contained in the class of bipartite $B(\overline{H})$-free graphs. We therefore have the following Corollary of Lemma 10 and Theorem 12.

**Corollary 14.** *The $H$-free Star System Problem is NP-complete whenever $f(H) \neq \infty$. Moreover, if $\mathcal{F}$ is a set of graphs for which there exists an integer $p$ such that for any $H \in \mathcal{F}$ we have $f(H) \leq p$, then the $\mathcal{F}$-free Star System Problem (i.e., deciding on an input $S$ if there is a graph having no induced subgraph isomorphic to* any *graph in $\mathcal{F}$) is NP-complete.*

Since $B(\overline{C_k})$ contains a cycle for any $k \geq 5$ we have the corollary.

**Corollary 15.** *For any $k \geq 5$, the $C_k$-free Star System problem is NP-complete.*

Similarly, $B(\overline{P_k})$ is connected and contains at least 2 vertices of degree greater or equal to 3 for any $k \geq 5$. Hence we also have the following corollary.

**Corollary 16.** *For any $k \geq 5$, the $P_k$-free Star System problem is NP-complete.*

## 5 Closing remarks

In this paper we obtained a complete dichotomy for the $H$-free Star System problem when the forbidden graph $H$ is either a path or a cycle. Moreover, our NP-completeness result holds for $H$ taken from a much larger family of graphs, so that the remaining cases in which the problem might not be NP-complete are very restricted. It is tempting to ask if the $H$-free Star System problem has P vs NP-completeness dichotomy in general, i.e. whether for any graph $H$ the $H$-free Star System problem is either polynomial-time solvable or NP-complete (and presumably not Graph Isomorphism-complete.)

A closely related question is on the complexity of the Star System problem restricted to graph classes defined by several forbidden induced subgraphs as in Corollary 14. By the result of Aigner and Triesch [1, 2] (see also [5]) we do then not have dichotomy in general, as there are classes of graphs defined by an infinite set of forbidden induced subgraphs (like forbidding the complements of odd cycles) such that the Star System problem is Graph Isomorphism complete on these classes. However, we do not know whether there is a graph class characterized by a *finite* set of forbidden induced subgraphs such that the Star System problem on this class is Graph Isomorphism complete, or if instead dichotomy may hold in this case.

# References


[1] M. Aigner and E. Triesch, *Reconstructing a graph from its neighborhood lists*, Combin. Probab. Comput., 2 (1993), pp. 103–113.

[2] M. Aigner and E. Triesch, *Realizability and uniqueness in graphs*, Discrete Math., 136 (1994), pp. 3–20.

[3] L. Babai, *Isomorphism testing and symmetry of graphs*, Ann. Discrete Math., 8 (1980), pp. 101–109.

[4] D. G. Corneil, H. Lerchs, and L. S. Burlingham, *Complement reducible graphs*, Discrete Appl. Math., 3 (1981), pp. 163–174.

[5] V. G. E. Boros and I. Zverovich, *Neighborhood hypergraphs of bipartite graphs*, tech. rep., RUTCOR, 2006.

[6] P. Erdős and T. Gallai, *Graphs with prescribed degrees of vertices (in hungarian)*, Matematikai Lapok, 11 (1960), pp. 264–274.

[7] A. Hajnal and V. Sós, eds., *Combinatorics. Vol. II*, vol. 18 of Colloquia Mathematica Societatis János Bolyai, North-Holland, 1978.

[8] F. Lalonde, *Le problème d'étoiles pour graphes est NP-complet*, Discrete Math., 33 (1981), pp. 271–280.

[9] A. Lubiw, *Some NP-complete problems similar to graph isomorphism*, SIAM J. Comput., 10 (1981), pp. 11–21.