On the Complexity of the G-Reconstruction Problem

Zdeněk Dvořák*
Vít Jelínek†

Abstract

Let G be a fixed unoriented graph. A G-structure of a graph F is a hypergraph H with the same set of vertices as F and with the property that a set h is a hyperedge of H iff the subgraph of F induced on h is isomorphic to G. We consider a G-reconstruction problem – given a hypergraph H, decide whether there exists a graph F such that H is a G-structure of F. It has been proved that this problem is polynomial if $|V(G)| \leq 4$ ([H96], [HHR02]). We investigate the complexity of the problem for larger graphs G and prove that there are cases where this problem is NP-complete – in fact we prove that it is NP-complete for almost all graphs G.

Introduction and Basic Definitions

All the graphs considered in this paper are simple unoriented graphs without multiple edges or loops. If $G = (V, E)$ is a graph and h a subset of V, we denote by $G[h]$ the subgraph of G induced by h, i.e. $G[h] = \left(h, E \cap \binom{h}{2}\right)$. We denote the vertex connectivity of G by $\kappa(G)$.

A *hypergraph* $H = (V, E)$ consists of a vertex set V and a set of hyperedges $E \subseteq 2^V$. The hypergraph H is called *k-uniform* if the size of each of its hyperedges is exactly k. A k-uniform hypergraph H is called *the G-structure of graph F* if $k = |V(G)|$, $V(H) = V(F)$ and for every $h \subseteq V(F)$ such that $|h| = k$, $h \in E(H)$ if and only if $F[h]$ is isomorphic to G.

*rakdver@kam.ms.mff.cuni.cz
†jelinek@kam.mff.cuni.cz

The G*-reconstruction problem* parametrized by a fixed graph G is defined as follows:

Input: a k-uniform hypergraph H, where $k = |V(G)|$.

Question: does there exist a graph F such that H is the G-structure of F?

This question naturally arises when we consider the amount of information needed to reconstruct a graph from information about its smaller parts (see [U60] and [BH77] for examples of this kind of problems). The other motivation for this problem is the *Semi-Strong Perfect Graph Theorem* that says that if two graphs have the same P_4-structure, then one is perfect iff the other one is ([R87]). Therefore to recognize perfect graphs, it suffices to be able to recognize the P_4-structures corresponding to perfect graphs. This observation motivated Chvátal to ask whether it is possible to recognize P_4-structures of arbitrary (i.e. not necessarily perfect) graphs in a polynomial time. This motivated a lot of research in the area, see for example [CH85], [H85] or [C87]. The question was finally answered by Hayward et al. ([HHR02]), who proved that indeed there exists a polynomial time algorithm for this problem.

It follows that the G-reconstruction problem is polynomial for all graphs on at most 4 vertices. This led Hayward et al. ([HHR02]) to ask whether there is a graph G such that the G-reconstruction problem is NP-complete.

In this paper, we prove that the G-reconstruction problem is NP-complete for many graphs G (in fact it is NP-complete for a sufficiently large random graph G with high probability). The smallest graph for which we are able to prove the NP-completeness is $K_2 + 3K_1$ (an edge plus three isolated vertices). With respect to the original question of Chvátal and the results of [HHR02] it is noteworthy that we are also able to prove that the P_n-reconstruction problem is NP-complete for all $n \geq 6$.

We first prepare some technical tools (section 1). In section 2 we prove the key theorem of this paper that shows that the G-reconstruction problem is NP-complete for all graphs G that satisfy some simple properties. In the following section 3 we prove those properties for several interesting classes of graphs. Since the membership in NP is obvious, we are only concerned with NP-hardness of the problem.

1 Preliminaries

Consider a G-reconstruction problem for a fixed G, $k = |V(G)|$. In the rest of the paper, we always suppose that G is neither edgeless nor complete graph (in both of these cases, the G-reconstruction problem is trivially polynomial). Let $P(x_1, \ldots, x_m)$ be some predicate on boolean variables. If F is a graph, $U = (u_1, \ldots, u_m)$, $u_i \in \binom{V(H)}{2}$ an ordered m-tuple, we say that U satisfies P if the assignment x_i is true $\Leftrightarrow u_i \in E(F)$ satisfies P.

An ordered pair $T = (H, U)$ is called a *gadget for* P, if the following conditions hold:

- H is a k-uniform hypergraph.
- $U = (u_1, \ldots, u_m)$ is an ordered m-tuple, where $u_i \in \binom{V(H)}{2}$ are mutually different, but not necessarily disjoint pairs of vertices of H.
- For any assignment to variables x_i, the predicate $P(x_1, \ldots, x_m)$ is satisfied iff there exists a graph F with G-structure H such that $u_i \in E(F) \Leftrightarrow x_i$ is true.

We call the pairs u_i as well as the vertices that belong to them *special*.

Let $T = (H_g, U_g)$ be a gadget for some P, $f = \{u_f, v_f\} \in U_g$. Let H be a k-uniform hypergraph and $e = \{u_e, v_e\}$ a pair of vertices of $V(H)$. Let $h : V(H_g) \to V(H_g) \cup V(H)$ be defined as identity for vertices other than u_f and v_f and by $h(u_f) = u_e$ and $h(v_f) = v_e$. The hypergraph H' is *obtained from* H *by connecting the gadget* T *through* f *to it on* e if $V(H') = V(H) \cup (V(H_g) \setminus \{u_f, v_f\})$ and $E(H') = E(H) \cup \{h[l] | l \in E(H_g)\}$ (i.e. we identify the vertices u_f and v_f with u_e and v_e, respectively).

Lemma 1. *Suppose that* $\kappa(G) > 2$. *Let* $T^1 = (H^1, U^1)$, $T^2 = (H^2, U^2)$ *be gadgets for some* $P^1(x_1, \ldots, x_{m^1})$ *and* $P^2(x_1, \ldots, x_{m^2})$, *respectively. Let* H *be obtained from* H^1 *by connecting* T^2 *through* u_1^2 *to it on* u_1^1. *Then* $T = (H, (u_1^1, u_2^1, \ldots, u_{m^1}^1, u_2^2, \ldots, u_{m^2}^2))$ *is a gadget for*

$$P(x_1, \ldots, x_{m^1+m^2-1}) = P^1(x_1, \ldots, x_{m^1}) \wedge P^2(x_1, x_{m^1+1}, x_{m^1+2}, \ldots, x_{m^1+m^2-1}).$$

Proof. Let F be a graph with structure H. Then since the size of G is at least 3 (due to the connectivity restriction), no hyperedge of H is fully contained in u_1^1. Therefore structures restricted to $V(H^1)$ and $V(H^2) \setminus u_1^2 \cup u_1^1$ are isomorphic to H^1 and H^2. Due to the definition of gadgets the corresponding edges satisfy the predicates P^1 and P^2, therefore F satisfies the predicate P.

For the second implication suppose that $x_1, \ldots, x_{m^1+m^2-1}$ satisfy the predicate P. Therefore when restricted to $V(H^1)$ and $V(H^2) \setminus u_1^2 \cup u_1^1$, we may construct graphs F^1 and F^2 with structures H^1 and H^2 and edges agreeing with x_i. These two graphs overlap only on u_1^1 and there they have prescribed edge or non-edge by x_1, so they do not conflict and we may join them into graph F on $V(H)$. It has of course all the special pairs agreeing with the assignment x_i, so we just need to prove that H is structure of F. The only problem might arise from the k-tuples that intersect both $V(H^2) \setminus u_1^2$ and $V(H^1) \setminus u_1^1$. No such k-tuple belongs to H, so we need to prove that G is not isomorphic to any graph induced by such k-tuple. This however follows straightforwardly from the condition on connectivity of G, since if G were isomorphic to a graph induced by some such k-tuple K, then $K \cap u_1^1$ would be a cut of size at most 2 in G. $\square$

Next we are going to state several lemmata about the existence of gadgets for various properties. The gadget that forces the presence of an edge at the pair of special vertices (i.e. the one for the property $P(x) = x$) is denoted by O.

Lemma 2. *Suppose that $\kappa(G) > 2$. Then the gadget O exists if and only if there is a gadget Z for the property $P(x) = \overline{x}$ (that enforces the absence of an edge).*

Proof. We construct Z as follows: first, we take k vertices and put this k-tuple K into Z. Then we embed a fixed copy of G on these vertices, and connect a copy of the gadget O to every pair that forms an edge of this copy. We choose $U = \{u\}$, where u is any non-edge of the prescribed copy of G. By Lemma 1 this is a gadget that enforces that the subgraph of any reconstruction of Z induced by K is exactly the prescribed copy of G and therefore that u is a non-edge.

By an analogical construction we may obtain O from Z. $\square$

Lemma 3. *Suppose that $\kappa(G) > 2$ and that the gadget O exists. Then there also exists a gadget N for the property $P(x, y) = (x \neq y)$.*

Proof. We take k vertices and put this k-tuple K into N. Let us denote this hypergraph by N_0. Since G is neither complete nor edgeless, there are at least two ways how to reconstruct the N_0. We construct the sequence $N_0, N_1, \ldots, N_o$ of hypergraphs satisfying the following properties:

- $N_i \subset N_{i+1}$

- Let n_i be the number of different subgraphs of reconstructions of N_i induced by K. Then $n_i \leq 2n_{i+1}$.
- $n_i \geq 2$ for $i \neq o$, $n_o = 1$.

If we succeed, then $n_{o-1} = 2$. The graphs F_1 and F_2 obtainable as subgraphs of reconstructions of N_{o-1} induced by K must have the same number of edges (since they are isomorphic to G) and therefore there exist two pairs of vertices u_1 and u_2 of K such that u_i is the edge in F_i and the nonedge in F_{3-i}, therefore $(N_{o-1}, (u_1, u_2))$ is a gadget for P.

We construct the sequence as follows: Let $e_1, e_2, \dots$ be a sequence of pairs of vertices of K, in any order. Let N_i^O and N_i^Z be hypergraphs obtained from N_{i-1} by connecting O and Z on e_i, respectively, n_i^O and n_i^Z numbers of subgraphs in reconstructions of these hypergraphs induced by K. Then $n_i^O + n_i^Z = n_{i-1}$, let n_i^X be the greater of them (any of them if they are equal) and put $N_i = N_i^X$. If $n_i^X = 1$, we let $o = i$, otherwise we continue with the construction. □

Lemma 4. *Suppose that $\kappa(G) > 2$ and that the gadget N exists. Then there also exists a gadget E for the property $P(x, y) = (x = y)$.*

Proof. We take two copies of the gadget N and connect them using Lemma 1. □

Lemma 5. *Suppose that $\kappa(G) > 2$ and that the gadget O exists. Then there also exists at least one of the following gadgets:*

- *The T_{SAT} gadget for the property $P(x, y, z) = x \vee y \vee z$.*
- *The T_{NAE} gadget for the property $P(x, y, z) = (x \neq y) \vee (x \neq z) \vee (y \neq z)$.*
- *The $T_{1\ in\ 3}$ gadget for the property $P(x, y, z) = (x \wedge \overline{y} \wedge \overline{z}) \vee (\overline{x} \wedge y \wedge \overline{z}) \vee (\overline{x} \wedge \overline{y} \wedge z)$.*

Proof. We take a k-tuple of vertices K and add K to the constructed hypergraph. Let v be an arbitrary vertex in K and let as $e_1, \dots, e_{k-1}$ be the pairs of vertices of K containing v.

Due to the connectivity condition and the fact that G is not complete we have $\delta(G) \geq 3$ and $k \geq 5$. The following cases may occur:

- There are vertices in G that have degrees $\delta(G)+1$ and $\delta(G)+2$. Then we attach copies of O to $e_1, \ldots, e_{\delta(G)-1}$ and copies of Z to $e_{\delta(G)}, \ldots, e_{k-4}$. The pairs e_{k-3}, e_{k-2} and e_{k-1} are the special ones. This forms a T_{SAT} gadget, as in any reconstruction at least one of the special pairs may and must be chosen to be an edge.

- There is a vertex in G that has degree $\delta(G)+1$, no vertex of degree $\delta(G)+2$ and $\delta(G) \leq k-2$. Then the same construction applies and the resulting gadget is T_{NAE}.

- There is no vertex in G with degree $\delta(G)+1$, then we attach copies of O to $e_1, \ldots, e_{\delta(G)-2}$ and copies of Z to $e_{\delta(G)-1}, \ldots, e_{k-4}$. Furthermore, we attach copies of N to e_{k-3}, e_{k-2} and e_{k-1} and take the other special pairs of these copies as special for the resulting gadget. We thus obtain $T_{1\ in\ 3}$.

- $\delta(G) = k-2$, $\Delta(G) = k-1$, then G is a complement of a matching. If there are exactly two vertices of degree $k-2$, we attach copies of O to all pairs of O except of a single triangle. Then we choose negations (through N) of the pairs of the triangle as special and the resulting gadget is $T_{1\ in\ 3}$. Otherwise we force edges and nonedges in K using gadgets O and Z as given by a fixed copy of G except for pairs on four vertices on which two nonedges are in this fixed copy. Of these four we take any three and by taking negated pairs in this triangle as special we again obtain $T_{1\ in\ 3}$.

□

Lemma 6. *Suppose that the gadget E exists. If $\kappa(G) > 2$ and there exists a gadget $T = (H, U)$ for a predicate P, then there also exists a gadget $T' = (H', U')$ for the same predicate such that the members of U' are mutually disjoint.*

Proof. It is enough to construct gadget E' derived from E with this property, since if we have it, we may attach copies of E' using Lemma 1 to the pairs in U and thus obtain mutually disjoint pairs U' which are copies of U.

If the special edges of E are not disjoint, we mark them as $\{x, y\}$ and $\{y, z\}$. Let E_1 and E_2 be two copies of E, let us denote the special pairs of the latter by $\{x', y'\}$ and $\{y', z'\}$. Join E_1 and E_2 using Lemma 1 so that z is identified with y' and y with z'. Then the resulting hypergraph with special edges $\{x, y\}$ and $\{x', z\}$ satisfies the property. □

We may consider all of the gadgets we have constructed to be of this type. We use this in the following theorems especially for N and E gadgets, where it makes the reasoning a bit easier.

Lemma 7. *Suppose that the gadget E exists, $\kappa(G) > 2$ and that a gadget $T = (H, U)$ for a predicate $P(x_1, x_2, \ldots, x_m)$ exists. Then there also exists a gadget $T' = (H', U')$ for the predicate $P'(x_2, \ldots, x_m) = P(x_2, x_2, x_3, \ldots, x_m)$.*

Proof. This lemma is similar to Lemma 1, but we must be a bit more careful since we are working only on one graph. Let u_1 and u_2 be the vertex pairs that correspond to variables x_1 and x_2 of P. We copy the pairs u_1 and u_2 to free pairs u_1' and u_2' by attaching a copy of E to each of them (using Lemma 1). We add a chain consisting of $2k+1$ copies of gadget E and identify the special pair at its start with u_1' and the one at its end with u_2'. Let H' be the hypergraph obtained in this way. We now prove that this hypergraph together with the pairs $(u_2, u_3, \ldots, u_m)$ is the sought gadget.

Let F be any reconstruction of H'. When we restrict it to $V(H)$, we obtain a reconstruction of H (since we copied u_1 and u_2 to free edges, no new hyperedges could be added entirely on $V(H)$), thus we know that $P(x_1, x_2, \ldots, x_n)$ is satisfied. When restricted to the chain of E's, it is again a reconstruction, therefore $x_1 = x_2$.

On the other hand, suppose that we are given $x_2, \ldots, x_m$ such that the predicate $P(x_2, x_2, x_3, \ldots, x_m)$ is satisfied. Then there exists a reconstruction F_1 of H with edges on U given by this assignment and reconstructions F_2^i of all the copies of E with assignment x_2, x_2. These reconstructions only meet at edges corresponding to x_2, therefore they do not conflict. We want to prove that F' obtained as the union of these reconstructions is a reconstruction of H'. There is no problem with k-tuples that are entirely contained in one of the gadgets (where we consider H with the two copies of E attached to be one gadget, which is possible due to Lemma 1). Take one of the other k-tuples K that is not contained in any single gadget. If K is entirely contained in the chain, there is again no problem since the chain itself forms a gadget due to Lemma 1. Otherwise, since every vertex is contained in at most two E gadgets forming the chain, there must be a gadget E_0 in the chain that does not contain any vertex of this k-tuple. Let E_1 be the E gadget from the chain that contains a vertex of K not in u_1' and u_2'. WLOG suppose that E_1 separates E_0 from u_1' in the chain. Then u_1' is a cut of size 2 in the graph induced by K – one of the components it separates is the intersection of K and the part of chain between E_0 and u_1' (nonempty because of the presence of E_1), the other one is the rest. It is indeed a cut, as all edges of F' are covered by some gadget and they can pass neither through E_0 nor through u'_1. Therefore the graph induced by K is not isomorphic to G, which is right as K is not a hyperedge of H'. $\square$

Lemma 8. *Suppose that $\kappa(G) > 2$, G is self-complementary (i.e. G is isomorphic to $\overline{G}$) and that the gadget E exists. Then also the gadget T_{NAE} exists.*

Proof. We take a k-tuple of vertices K and add K to the constructed hypergraph. Let v be a vertex in K and let $e_1, \ldots, e_{k-1}$ be the pairs of vertices of K containing v. Due to connectivity and the fact that G is self-complementary we see that $3 \leq \delta(G) \leq \Delta(G) \leq k-4$, since $\delta(G) + \Delta(G) = k-1$.

We add the E constraints between edges e_i so that e_1, e_2 and e_3 form the special edges of the sought gadget – we may do this using Lemma 7. We add the E constraint between e_3 and e_i for i between 4 and $\Delta(G)+1$ – this means that in any reconstruction these $\Delta(G)-1$ pairs are either all edges, or all non-edges. Suppose they are edges, then there is at most one edge left and at least one of e_1 and e_2 must be non-edge. Symmetrically we see that not all of e_1, e_2 and e_3 may be non-edges.

We just need to observe that indeed all of the cases allowed by NAE condition can be realized. But this is easy – if e_3 should be in the reconstruction, we put the vertex of highest degree in G to v, $\Delta(G)-1$ of its neighbors to the other vertices of $e_3, e_4, \ldots, e_{\Delta(G)+1}$, the last neighbor either to e_1, e_2 or e_{k-1} depending on the assignment we want to simulate and the other vertices arbitrarily. The case when e_3 is not in the reconstruction is symmetrical due to the self-complementarity of G. $\square$

2 NP-Completeness of the Graph Reconstruction Problem

We are now ready to prove the NP-completeness of the graph reconstruction problem modulo a few conditions on G:

Theorem 9. *Let G be such that $\kappa(G) > 2$ and that the gadget O exists. Then the G-reconstruction problem is NP-complete.*

Proof. Using the previous lemmata, T_{SAT}, T_{NAE} or $T_{1\ in\ 3}$ exist. We proceed by reduction from the well-known NP-complete problems $3-SAT$, $3-NAE-SAT$ or 3-exact set cover ([GJ79]), respectively. Each of them can be interpreted as a problem having on its input a CNF-formula with clauses of size exactly 3 and looking for assignment of truth values to variables such that at least 1 (2 or 3, exactly 1) variable in each clause is satisfied. In the second and third case we even do not need negations to be present in the clauses, but allowing them makes the problems only harder. Let T_X be the gadget for the appropriate problem and a formula F its instance.

We want to create an equivalent instance of G-reconstruction problem. For each clause of F we add a copy of T_X to it. For each variable that occurs in n clauses, we add a star with n rays from copies of E joined by single special vertex pair. Whether the variable is true or false will be determined by whether the edge (and therefore also its copies) is chosen to the reconstruction or not. If the occurrence of variable is negated, we attach N gadget to the appropriate ray. For each clause C of F, we then copy the special edges of gadget T_X added for C to the ends of the rays that correspond to the appropriate occurrences of variables of C, using Lemma 1 and Lemma 7. It follows that the created instance is an F-gadget (where the special edges are the centers of the stars corresponding to variables, and clauses of F are interpreted according to X), and therefore it is reconstructible if and only if F is satisfiable.

Since G is fixed, all the used gadgets have constant size, therefore the reduction is polynomial. □

Note also that the gadget O does not exist if G is self-complementary. However this doesn't prevent the existence of the gadget N for this class of graphs, which justifies the following theorem:

Theorem 10. *Let G be self-complementary, $\kappa(G) > 2$ and such that the gadget N exists. Then the G-reconstruction problem is NP-complete.*

Proof. The proof is analogical to the proof of Theorem 9 – we may construct gadget E using just gadget N and T_{NAE} is obtained using Lemma 8. □

We obtain NP-completeness for some more graphs using the following trivial theorem:

Theorem 11. *Assume that the G-reconstruction problem is NP-complete. Then the $\overline{G}$-reconstruction problem is NP-complete as well.*

Proof. Follows trivially from the fact that F is a G-reconstruction of a hypergraph H if and only if $\overline{F}$ is a $\overline{G}$-reconstruction of H. □

3 Classes of Graphs for Which the Problem is Hard

We show that there are several important classes of graphs that satisfy the conditions of Theorem 9.

To satisfy the condition on the connectivity of G, we need to assume that either G or its complement (due to Theorem 11 and Lemma 2) are 3-connected. This is true for most of the interesting classes of graphs, with the notable exception of stars.

The existence of the O or N gadgets is the interesting part. First we show that in fact most of the graphs satisfy this condition:

Theorem 12. *Suppose that G is a graph on k vertices which satisfies the following conditions:*

(i) $\forall x \in V(G)$ $G - x$ *has no nontrivial automorphism*

(ii) $\forall x, y \in V(G)$ *if* $x \neq y$ *then* $G - x$ *is not isomorphic to* $G - y$

Then the following holds:

1. *Let H be a G-structure of a graph F which contains two hyperedges h_1 and h_2 that intersect in $k-1$ vertices, i. e. we have $h_1 = \{a, v_1, v_2, \ldots, v_{k-1}\}$ and $h_2 = \{b, v_1, v_2, \ldots, v_{k-1}\}$. Then $\{a, v_i\} \in E(F)$ if and only if $\{b, v_i\} \in E(F)$.*

2. *If G is not self-complementary, then the gadgets O and Z exist. On the other hand, if G is self-complementary, then the gadget N exists.*

Proof. First we prove part 1 of the theorem. Let f_i be the isomorphism that maps G onto $F[h_i]$ for $i \in \{1, 2\}$. For contradiction, assume that F does not have the required property, WLOG assume that $\{a, v_1\} \in E(F)$ and $\{b, v_1\} \notin E(F)$. Two cases will be distinguished:

1. There is a vertex $x \in V(G)$ such that $f_1(x) = a$ and $f_2(x) = b$. Since $\{a, v_1\} \in E(F)$ and $\{b, v_1\} \notin E(F)$, we know that the preimage of v_1 under f_1 is different from the preimage of v_1 under f_2. Hence if we compose f_1 with the inverse f_2 and restrict this mapping to the graph $G - x$, we obtain a nontrivial automorphism, contrary to assumption (i) of the theorem.

2. There are two distinct vertices $x, y \in V(G)$ such that $f_1(x) = a$ and $f_2(y) = b$. But then the graphs $G-x$ and $G-y$ are both isomorphic to $F[\{v_1, \ldots, v_{k-1}\}]$, which contradicts assumption (ii) of the theorem.

To prove the second part of the theorem, we first introduce the following notation: let $\{x, y\} \in E(G)$ be an arbitrary edge of G and $\{\widehat{x}, \widehat{y}\} \in \binom{V(G)}{2} \setminus E(G)$ be an arbitrary non-edge of G. We construct a graph K by the following procedure:

1. Let V_0 be an arbitrary set of size k, let $G_0 = (V_0, E_0)$ be a fixed isomorphic embedding of G on the set V_0.

2. With each element $x_0 \in V_0$ we associate a set $S(x_0)$ of $k-2$ new vertices denoted $\{w(x_0); w \in V(G) \setminus \{x, y\}\}$. For each vertex $y_0 \in V_0$ such that $\{x_0, y_0\}$ is an edge of E_0 we add new edges induced by the set $\{x_0, y_0\} \cup S(x_0)$ so that this set contains an isomorphic copy of G defined by the isomorphism that maps x to x_0, y to y_0 and w to $w(x_0)$ for each $w \in V(G) \setminus \{x, y\}$.

3. With each element $x_0 \in V_0$ we associate a set $\widehat{S}(x_0)$ of $k-2$ new vertices denoted $\{\widehat{w}(x_0); w \in V(G) \setminus \{\widehat{x}, \widehat{y}\}\}$. For each vertex $y_0 \in V_0$ such that $\{x_0, y_0\}$ is not an edge of G_0 we add new edges induced by the set $\{x_0, y_0\} \cup \widehat{S}(x_0)$ so that this set contains an isomorphic copy of G defined by the isomorphism that maps $\widehat{x}$ to x_0, $\widehat{y}$ to y_0 and w to $\widehat{w}(x_0)$ for each $w \in V(G) \setminus \{\widehat{x}, \widehat{y}\}$.

We denote by K the graph on $k \cdot (2k-3)$ vertices obtained by the previous three steps. Let H be the G-structure of K. Let K' be any graph whose G-structure is H, let $G'_0 = K'[V_0]$. Obviously, G'_0 and G_0 are both isomorphic to G, because their vertex set V_0 is a hyperedge of H. However, we can establish a stronger property:

- If G is self-complementary then $G'_0 = G_0$ or $G'_0 = \overline{G_0}$, where $\overline{G_0}$ is the complement of G_0.

- If G is not self-complementary then $G'_0 = G_0$.

This property immediately implies the second part of the theorem. To prove the property, we introduce a binary relation on $\binom{V_0}{2}$ called *friendship* defined as follows: we say that the two pairs $\{x, y_1\}$ and $\{x, y_2\}$ are friends in G_0, if they share a common vertex and if either both of them or neither of them belong to E_0. From the first part of the theorem we obtain that if two pairs of vertices $\{x, y_1\}$ and $\{x, y_2\}$ are friends in G_0, then they are also friends in G_0', because H contains the two hyperedges $h_i = \{x, y_i\} \cup S'(x)$ for $i \in \{1, 2\}$, where $S'(x)$ is defined as $S(x)$ if the two friends are edges of G_0, and $S'(x) = \widehat{S}(x)$ otherwise. Let $\sim$ denote the relation obtained as the transitive closure of the friendship relation in G_0. Note that $\sim$ is an equivalence on $\binom{V_0}{2}$ whose blocks are the edge sets of the connected components of G_0 and of the connected components of $\overline{G_0}$. Since friendship in G_0 implies friendship in G_0', we have that the blocks of $\sim$ are subsets of the blocks of $\sim'$, where $\sim'$ is the closure of the friendship relation in G_0'. But the number of blocks of $\sim'$ is the number of connected components of G_0' and its complement, which is equal to the number of blocks of $\sim$. Hence the two equivalence relations $\sim$ and $\sim'$ are equal and the two corresponding friendship relations are equal as well. Since every disconnected graph has a connected complement, we know that at least one block B of the relation $\sim$ is spanning in V_0, i. e. for each $x \in V_0$ at least one edge incident to x belongs to B. If the members of B are edges of G_0' then all the other blocks must only contain non-edges (and vice versa), because any edge outside of B would necessarily become a friend of some edge in B by the spanning property of B. This implies that the edge set of G_0' is either equal to B or equal to the complement of B. Clearly if G is not self-complementary, then only one of these two options is available.

This concludes the proof of the theorem. □

Theorem 13. *Let G be a random graph. Then with probability $1 - o(1)$ the G-reconstruction problem is NP-complete.*

Proof. It suffices to show that a random graph is almost always at least 3-connected, and that it almost always satisfies the assumptions of Theorem 12. Since the probability that in a random graph there exist two vertices u, v such that $|N(u) \cap N(v)| \leq n/5$ is $o(1)$, the connectivity condition is almost always satisfied. We thus only need to prove that the prerequisites of the Theorem 12 hold for almost all graphs. This is just a simple extension of the well-known Theorem stating that almost no random graph has a nontrivial automorphism (see for example [BES80]).

Let us take any of $k^2(k-1)!$ bijections f between $G-x$ and $G-y$, where x and y are two (not necessarily different) vertices of G. Take a directed graph F with edges $(v, f(v))$. F consists of oriented cycles and one path starting in y and ending in x. Let f be an isomorphism of $G-x$ and $G-y$. Then not all edges of G are independent – there exist a set S of pairs of vertices G such that the information about which of these pairs are edges uniquely determines G – therefore there are at most $2^{|S|}$ graphs for which f is the "bad" isomorphism. We try to find a set as small as possible having this property.

Let l_1, l_2, $\ldots$, l_m be the lengths of cycles of f in nondecreasing order and let l be the length (number of vertices) of the path between vertices x and y. Then the size of a minimal set S is at most

$$\left(\sum_{i=1}^{m}\left\lfloor\frac{l_i}{2}\right\rfloor\right) + (l-1) + \left(\sum_{i=1}^{m-1}(m-i)l_i\right) + \left(\sum_{i=1}^{m} l_i\right).$$

The first term corresponds to the edges inside single cycle – there are $\lfloor\frac{l_i}{2}\rfloor$ different distances between vertices. The second term is for edges inside the path. The third term counts the number of independent edges between the pairs of cycles – at worst we must join a single vertex of the longer cycle to all the vertices of the short one. The last term is for edges between cycles and the path – all the edges from the start of the path are independent.

Suppose that f has less then $k-x$ fixpoints. In this case $m \le k-\frac{x}{2}$. The third term is always at most $\frac{km}{2}$ (to see this, add $\sum_{i=1}^{m-1} il_i$ to it and observe that the original third term is smaller than the added sum), therefore it sums to at most $\frac{k^2}{2}-\frac{kx}{4}$. The remaining terms sum to at most $2k$. If $x \ge 8\log_2 k$, we see that the fraction of graphs for which any such f may be the "bad" isomorphism is at most

$$O\left(\frac{k^2(k-1)!}{2^{2k\log_2 k}}\right) \le O\left(\frac{k^k}{2^{2k\log_2 k}}\right) \le O\left(\frac{1}{2^{k\log_2 k}}\right) = o(1),$$

so we do not have to care about them.

Now consider the case when f fixes all but $8\log_2 k$ vertices. The number of such bijections is at most $O(2^{\log^4 k})$. Unless f is the identity, there is a such that $f(a) = b \ne a$. But then edges incident to b are determined by the edges incident to a, therefore at least $k-2$ edges do not belong to S. The fraction of graphs for that some such f is the "bad" isomorphism is at most

$$O\left(\frac{2^{\log^4 k}}{2^{k-2}}\right) = o(1).$$

Therefore indeed only $o(1)$ fraction of graphs have the "bad" isomorphism. $\square$

Many natural classes of graphs however do not have the properties required by Theorem 12. We separately investigate the existence of O gadget for some special classes:

- For paths, because of the motivation for the problem described in the introduction.
- For graphs containing just a single edge, since they make the smallest examples for that the connectivity condition holds.

Theorem 14. *If G is a path on k vertices with $k \geq 5$, then a gadget O and a gadget Z for G exist.*

Proof. Let C_{k+2} be the cycle on $k{+}2$ vertices, with vertex set $V = \{x_0, \ldots, x_{k+1}\}$ and edge set $E = \{\{x_i, x_{i+1}\};\ i = 0, \ldots, k+1\}$, where the indices of x_i in the previous expression, as well as in the rest of this proof, are evaluated modulo $k+2$. If a pair of vertices $x, y \in V$ is connected by an edge of C_{k+2}, we say that the two vertices are close to each other and we call {x,y} a close pair. Let H be the P_k-structure of C_{k+2}. Clearly a set $h \subset V$ is a hyperedge of H if and only if the two vertices in $V \setminus h$ are close. We claim that C_{k+2} is the only graph with P_k-structure H. This claim implies that H can be turned either into a gadget O or into a gadget Z, by choosing appropriate special pairs of vertices. To prove the claim, we denote by F an arbitrary graph whose P_k-structure is H. We proceed by a sequence of observations:

1. Every vertex $x \in V$ has the degree at least two in F. Indeed if the degree of x were less than two, we could choose a close pair $\{y_1, y_2\}$ not containing x, in a way that x is an isolated vertex in the subgraph of F induced by the complement of $\{y_1, y_2\}$. This is impossible, because every complement of a close pair must induce a subgraph of F isomorphic to P_k.
2. Every vertex of F has the degree at most two. This is obvious if $k \geq 7$, because then for every vertex x of degree at least three we might choose a close pair $\{y_1, y_2\}$ which avoids x as well as at least three vertices connected to x by an edge of F. This is again impossible, because the complement of a close pair must induce a path. In the two remaining cases ($k = 5$ and $k = 6$), this simple argument merely shows that the maximum degree is less than 4. We treat the two cases separately:

$k=5$: For contradiction, let us assume that some vertex of F (WLOG the vertex x_0) has the degree equal to 3. Each close pair must intersect at least one of the three edges incident to x_0. For this to be possible, x_0 can not be connected by an edge to both x_1 and x_6. Let us assume that WLOG x_0 is not connected to x_1. This implies that the close pair $\{x_0, x_1\}$ intersects at least five edges of F. On the other hand, F has at least one vertex of degree two (the number of vertices of degree three must be even) and every close pair containing this vertex intersects at most five edges of F. Since every close pair must intersect the same number of edges, we conclude that every close pair intersects exactly five edges. It follows that F has nine edges, which implies that F has four vertices of degree three and three vertices of degree two. There must then be a close pair $\{x_i, x_{i+1}\}$ containing two vertices of degree three. Since the close pair intersects five edges, the two vertices of the pair must be connected by an edge. It follows that the neighbors of x_i are exactly the vertices x_{i+1}, x_{i+3} and x_{i+5}, because any other choice yields a close pair avoiding x_i and all its neighbors. Similarly, x_{i+1} is connected with x_i, x_{i+3} and x_{i+5}. Then x_i, x_{i+1} and x_{i+3} induce a triangle, which shows that the subgraph induced by the complement of the close pair $\{x_{i+4}, x_{i+5}\}$ is not a path, which is a contradiction.

$k=6$: Let us assume that the vertex x_i has degree three. Then its neighbors must be the vertices x_{i+2}, x_{i+4} and x_{i+6}. Again we see that every close pair intersects at least five edges, which implies that F has at least four vertices of degree three. We may WLOG assume that at least two of these vertices have even indices. In this case the two vertices together with any other even-numbered vertex induce a triangle which is avoided by some close pair of vertices, which yields a contradiction.

3. Now that we know that F is a 2-regular graph, we easily conclude that $F = C_{k+2}$: F has exactly $k+2$ edges, and the complement of any close pair induces $k-1$ edges, so every close pair intersects three edges, which implies that every close pair is itself an edge of F.

The statement of the theorem follows, as noted above. □

Note that the complement of P_n satisfies $\kappa(\overline{P_n}) > 2$ whenever $n \geq 6$. Also note that a gadget Z for some graph G is at the same time a gadget O

for $\overline{G}$. By the theorems 9, 11 and 14 we then have that the P_n-reconstruction problem is NP-complete for each $n \geq 6$.

Theorem 15. *Let G be a graph on $k \geq 4$ vertices with one edge. Then a gadget Z for G exists.*

Proof. Let H be the hypergraph with the vertex set $V = \{x_0, x_1, \ldots, x_k\}$ and two hyperedges $h_1 = \{x_0, \ldots, x_{k-1}\}$ and $h_2 = \{x_1, \ldots, x_k\}$. Note that H is the G-structure of the graph with two edges $\{x_0, x_1\}$ and $\{x_1, x_k\}$. We prove that if F is a graph whose G-structure is H then the vertices $\{x_1, \ldots, x_{k-1}\}$ are an independent set in F. For contradiction, assume that WLOG $\{x_1, x_2\}$ is an edge of F. Then neither h_1 nor h_2 can induce another edge in F. The only pair of vertices that we have to consider is the pair $\{x_0, x_k\}$. If this pair of vertices is connected by an edge of F, then the set $V \setminus \{x_1\}$ must be a hyperedge of H. On the other hand, if $\{x_0, x_k\}$ is not an edge of F, then $V \setminus \{x_3\}$ must be a hyperedge of H. In both cases we get a contradiction. $\square$

If G is a graph with one edge and at least five vertices then $\kappa(\overline{G}) > 2$, so the G-reconstruction problem is NP-complete.

Conclusion

We have proved that for many graphs G, the G-reconstruction problem is NP-complete. On the other hand, the P_4-reconstruction problem is polynomial and so we might hope that for some other small graphs the problem could be solvable in polynomial time. However, the genericity of Theorem 9 suggests that the problem is usually hard. We are confident enough to state the following

Conjecture 1. *There exists a constant k_0 such that for any graph G with $|V(G)| > k_0$, that is not edgeless or complete, the G-reconstruction problem is NP-complete.*

The results on existence of O gadget (resp. N gadget for self-complementary graphs) might also be interesting by themselves. Very likely such a gadget exists for every graph G.

Acknowledgements

We would like to thank Jan Kratochvíl and Pavel Valtr for introducing the problem to us and for useful discussions about it.

References

[U60] Ulam S. M.: *A Collection of Mathematical Problems*, Wiley, New York, 1960.

[BH77] Bondy J. A. and Hemminger R. L.: Graph Reconstruction – a survey, *J. Graph Theory* **1** (1977), 227–268.

[GJ79] M. R. Garey and D. S. Johnson. *Computers and Intractability: A Guide to the Theory of NP-Completeness.* W.H. Freeman, New York, 1979.

[BES80] L. Babai, P. Erdos, and S. M. Selkow: Random graph isomorphism, *SIAM J. Computing*, **9** (1980), 628–635.

[CH85] V. Chvátal, C.T. Hoàng: On the P_4-structure of perfect graphs I. Even decompositions, *J. Combin. Theory Ser. B* **39** (1985), 209–219.

[H85] C.T. Hoàng: On the P_4-structure of perfect graphs II. Odd decompositions, *J. Combin. Theory Ser. B* **39** (1985), 220–232.

[R87] B. A. Reed: A semi-strong perfect graph theorem, *J. Combin. Theory B* **43** (1987), 223–240.

[C87] V. Chvátal: On the P_4-structure of perfect graphs III. Partner decompositions, *J. Combin. Theory Ser. B* **43** (1987), 349–353.

[H96] R. Hayward: Recognizing P_3-structure: A Switching Approach, *Journal Comb. Theory (Series B)* **66** No. 2 (1996), 247–262.

[HHR02] R. Hayward, S. Hougardy and B. Reed: Polynomial time recognition of P_4-structure, *Proceedings of 13th Annual ACM-SIAM Symp. On Discrete Algorithms* (2002), 382–389.