Sandwich and probe problems for excluding paths

Celina M. H. de Figueiredo[1a], Sophie Spirkl[b]

[a] *COPPE, Universidade Federal do Rio de Janeiro, Brazil*
Email: `celina@cos.ufrj.br`
[b] *Princeton University, Princeton, NJ 08544*
Email: `sspirkl@princeton.edu`

Abstract

Let P_k denote an induced path on k vertices. For $k \geq 5$, we show that the P_k-free sandwich problem, partitioned probe problem, and unpartitioned probe problem are NP-complete. For $k \leq 4$, it is known that the P_k-free sandwich problem, partitioned probe problem, and unpartitioned probe problem are in P.

Keywords: graph sandwich problem, probe problem, partitioned probe problem, unpartitioned probe problem

1. Introduction

All graphs in this paper are finite and simple. Let G be a graph. G^C denotes the *complement* of G, obtained from G by replacing each edge with a non-edge and vice versa. For $X \subseteq V(G)$, $G|X$ denotes the induced subgraph of G with vertex set X. For $X, Y \subseteq V(G)$ with $X \cap Y = \emptyset$, we say that X is *complete* to Y if for all $x \in X, y \in Y$, $xy \in E(G)$; we say that X is *anticomplete* to Y if for all $x \in X, y \in Y$, $xy \notin E(G)$. For $v \in V(G)$, $X \subseteq V(G) \setminus \{v\}$, we say that v is *complete* (*anticomplete*) to X if $\{v\}$ is complete (anticomplete) to X.

Let $G_1 = (V_1, E_1), G_2 = (V_2, E_2)$, then G_2 is a *supergraph* of G_1 if $V_1 = V_2$ and $E_1 \subseteq E_2$. A pair (G_1, G_2) of graphs so that G_2 is a supergraph of G_1 is called a *sandwich instance*. The edges in E_1 are called *forced*, while the edges in $E_2 \setminus E_1$ are *optional*. A graph G is called a *sandwich graph* for the

[1]Supported by Conselho Nacional de Desenvolvimento Científico e Tecnológico CNPq Grant 303622/2011-3.

sandwich instance (G_1, G_2) if G_2 is a supergraph of G and G is a supergraph of G_1. For a graph G and a set E' of edges with both endpoints in $V(G)$, $G \cup E'$ denotes the supergraph $G' = (V(G), E(G) \cup E')$ of G, and $G \setminus E'$ denotes the graph $G'' = (V(G), E(G) \setminus E')$, and G is a supergraph of G''.

Let $\mathcal{P}$ be a graph property. We define the complementary property $\mathcal{P}^C$ by saying that G satisfies $\mathcal{P}^C$ if and only if G^C satisfies $\mathcal{P}$.

The $\mathcal{P}$ RECOGNITION PROBLEM is the problem of deciding whether a given graph G satisfies $\mathcal{P}$. The $\mathcal{P}$ SANDWICH PROBLEM is the following: For a given sandwich instance (G_1, G_2), does there exist a sandwich graph G for (G_1, G_2) so that G satisfies $\mathcal{P}$? This generalization of the recognition problem was introduced by Golumbic and Shamir [4]. The sandwich problem becomes the recognition problem when $G_1 = G_2$, and thus, if the $\mathcal{P}$ recognition problem is NP-complete, so is the $\mathcal{P}$ sandwich problem.

Let G, G' be a pair of graphs such that G' is a supergraph of G. Then G' is a (P, N)*-probe graph* for G if (P, N) is a partition of $V(G)$, N is a stable set in G, and every edge in $E(G') \setminus E(G)$ has both of its endpoints in N.

For a graph property $\mathcal{P}$, a graph $G = (V, E)$ is a $\mathcal{P}$ *probe graph with partition* (P, N) if there exists a (P, N)-probe graph G' for G such that G' satisfies $\mathcal{P}$. A graph G is a $\mathcal{P}$ *probe graph* if there exists a partition (P, N) of its vertex set such that G is a $\mathcal{P}$ probe graph with partition (P, N). The vertices in P are called *probes*, and the vertices in N are called *non-probes*.

For a graph property $\mathcal{P}$, the $\mathcal{P}$ PARTITIONED PROBE PROBLEM is the following: Given a graph $G = (V, E)$, and a stable set $N \subseteq V$, is G a $\mathcal{P}$ probe graph with partition $(V \setminus N, N)$? The partitioned probe problem was first introduced in [6, 7] for interval graphs because of its applications to the physical mapping of DNA.

The $\mathcal{P}$ partitioned probe problem with input graph $G = (V, E)$ and stable set $N \subseteq V$ is a special case of the $\mathcal{P}$ sandwich problem in which $E(G_1) = E$ and the edges in $E(G_2) \setminus E(G_1)$ are precisely the edges between all pairs of distinct vertices in N.

Note that the sandwich problem and the partitioned probe problem are invariant under taking complements in the following sense. The first statement is shown in [5].

Lemma 1. *The $\mathcal{P}^C$ sandwich problem is NP-complete if and only if the $\mathcal{P}$ sandwich problem is. The $\mathcal{P}^C$ partitioned probe problem is NP-complete if and only if the $\mathcal{P}$ partitioned probe problem is.*

Proof. An instance (G_1, G_2) is a YES instance for $\mathcal{P}^C$ sandwich problem if and only if (G_2^C, G_1^C) is a YES instance for the $\mathcal{P}$ sandwich problem. The same is true for the $\mathcal{P}$ partitioned probe problem: A graph G with partition (P, N) is a YES instance for the $\mathcal{P}$ partitioned probe problem if and only if the graph G' arising from G^C by removing all edges with both endpoints in N with the partition (P, N) is a YES instance for the $\mathcal{P}^C$ partitioned probe problem. $\square$

Let $\mathcal{P}$ be a graph property. The $\mathcal{P}$ UNPARTITIONED PROBE PROBLEM is the following: Given a graph G, is G a $\mathcal{P}$ probe graph? We also consider the $\mathcal{P}$ UNPARTITIONED PROBE PROBLEM IN THE COMPLEMENT: Given a graph G, is G^C a $\mathcal{P}^C$ probe graph? In other words, in the unpartitioned probe problem, the goal is to decide whether there is a stable set N in G and a set of edges E' with both endpoints in N such that $G \cup E'$ satisfies $\mathcal{P}$, whereas in the unpartitioned probe problem in the complement, the goal is to decide whether there is a clique N in G and a set of edges E' with both endpoints in N such that $G \setminus E'$ satisfies $\mathcal{P}$. Therefore, these problems are not equivalent in general.

Forbidden induced subgraph sandwich problems have been considered in [3] and then further studied in [2], which considered the complexities of partitioned and unpartitioned probe problems as well. In several papers, probe problems have been considered with respect to subclasses of perfect graphs, and the perfect graph sandwich problem is the only remaining open sandwich problem in the seminal paper by Golumbic et al. [5].

In the present paper, we consider sandwich and probe problems for excluding paths. We let P_k denote an induced path on k vertices. The graph P_5^C is called a *house*. For $k \geq 5$, we show that the P_k-free sandwich problem, partitioned probe problem, and unpartitioned probe problem are NP-complete. For $k \leq 4$, it is known that the P_k-free sandwich problem, partitioned probe problem, and unpartitioned probe problem are in P. The paper is organized as follows. In Section 2, we give some reductions. In Section 3, we consider the sandwich and probe problems for P_3, P_4, P_5 and P_6; and in Section 4, we consider P_k, $k \geq 6$.

2. Reductions

Let G be a graph and let $x, y \in V(G)$ be distinct vertices. Then x and y are *twins* if $N(x) \setminus \{y\} = N(y) \setminus \{x\}$. They are *adjacent twins* if x is adjacent to y, and *non-adjacent twins* otherwise. Note that if x, y are adjacent twins in G, then they are non-adjacent twins in G^C. For $k \geq 4$, P_k does not contain twins.

We will use the following results from [2].

Theorem 2 ([2])**.** *For 3-connected graphs $F \neq K_n$, if the F-free sandwich problem is NP-complete, so is the F-free partitioned probe problem.*

Theorem 3 ([2])**.** *For 2-connected graphs $F \neq K_n$, if the F-free partitioned probe problem is NP-complete, so is the F-free unpartitioned probe problem.*

By taking complements in Theorem 2 and applying Lemma 1, we also obtain the following.

Corollary 4 ([2])**.** *Let F be a graph such that $F^C \neq K_n$ and F^C is 3-connected. If the F-free sandwich problem is NP-complete, so is the F-free partitioned probe problem.*

Theorem 5. *Let H be a graph that does not contain any adjacent twins. If the H-free partitioned probe problem is NP-complete, so is the H-free unpartitioned probe problem.*

Proof. Let G with partition (P, N) be an instance for the H-free partitioned probe problem, and let G' arise from G by adding an adjacent twin for every vertex in P; call this set of new vertices P'. For every vertex $v \in P$, we let v' denote its adjacent twin in P'. We claim that G is a YES instance for the H-free partitioned probe problem if and only if G' is a YES instance for the H-free unpartitioned probe problem.

Let E be a set of edges with both endpoints in N so that $(V(G), E(G) \cup E)$ is H-free, and suppose for a contradiction that $G'' = (V(G'), E(G') \cup E)$ is not H-free. Let $S \subseteq V(G')$ induce a subgraph isomorphic to H in G''. Since H contains no adjacent twins, it follows that S does not contain v and v' for any $v \in P$, since v and v' are adjacent twins in G''. Therefore, we may assume that if S contains a vertex in $\{v, v'\}$ for some $v \in P$, then S contains v and not v', and so $S \subseteq N \cup P$. But then $S \subseteq V(G)$, and so S induces a subgraph isomorphic to H in $(V(G), E(G) \cup E)$. This is a contradiction and it follows that G'' is H-free.

For the converse direction, let $N' \subseteq V(G')$ be a stable set, and let E' be a set of edges with both endpoints in N' so that $G'' = (V(G'), E(G') \cup E')$ is H-free. Since N' is a stable set in G', N' does not contain v and v' for any $v \in P$, and since they are adjacent twins, we may assume (by adjusting E' accordingly) that $N' \cap P = \emptyset$. Let E denote the set of edges in E' with both endpoints in N. It follows that $(V(G), E(G) \cup E) = G''|V(G)$, since E' contains no edge with an endpoint in P. Thus $(V(G), E(G) \cup E)$ is H-free, which proves the claim.

Since the size of G' is at most twice the size of G, this reduction takes polynomial time; the result of the theorem follows. □

By taking complements in Theorem 5, we obtain the following:

Corollary 6. *Let H be a graph that does not contain any non-adjacent twins. If the H-free partitioned probe problem is NP-complete, so is the H-free unpartitioned probe problem in the complement.*

Proof. Let H be a graph that does not contain any non-adjacent twins. Then H^C does not contain any adjacent twins. Suppose that the H-free partitioned probe problem is NP-complete. Then, by Lemma 1, the H^C-free partitioned probe problem is NP-complete. Now by Theorem 5, the H^C-free unpartitioned probe problem is NP-complete, which means that the H-free unpartitioned probe problem in the complement is NP-complete. □

3. Short paths

We first consider the P_3-free case and the P_4-free case. Here, the sandwich problem and both probe problems can be solved in polynomial time, as Theorem 7, Lemma 8 and Theorem 9 show.

Theorem 7 ([3])**.** *The P_3-free sandwich problem and partitioned probe problem can be solved in polynomial time.*

We prove the following simple lemma for completeness.

Lemma 8. *The P_3-free unpartitioned probe problem can be solved in polynomial time.*

Proof. Let G be a graph, and consider the following algorithm: Let G' be a copy of G. Let $N = \emptyset$. While there is an induced P_3 with vertices x, y, z in order in G', add x, z to N and xz to $E(G')$. If N is a stable set in G, then this algorithm yields a $(V(G) \setminus N, N)$-probe graph G' for G which is P_3-free. For the converse direction, note that if there is a P_3-free probe graph G''

for G with partition (P'', N''), then at every step of the algorithm, G'' is a supergraph of G', and hence $N \subseteq N''$. Thus, if such a G'' exists, then N is stable. This concludes the proof. $\square$

Theorem 9 ([5, 1])**.** *The P_4-free sandwich problem, partitioned probe problem and unpartitioned probe problem can be solved in polynomial time.*

From [2], we know the following:

Theorem 10 ([2])**.** *The C_4-free partitioned probe problem is NP-complete.*

We now review a construction given in [3] which we will use to prove that the house-free sandwich problem is NP-complete, and which is used in [3] to prove the following in the case that $k = 4$.

Theorem 11 ([3])**.** *The C_k-free sandwich problem is NP-complete for all fixed $k \geq 4$.*

Let (X, C) be a 3Sat instance with a set $X = \{x_1, \ldots, x_n\}$ of variables and a set $C = \{c_1, \ldots, c_m\}$ of clauses such that each clause contains exactly three variables. We follow the notation of [3], but we use ℓ_q^j for a literal in the 3Sat instance and l_q^j for a vertex of the constructed gadget from the 3Sat instance (instead of using l_q^j for both).

Before giving a detailed definition of the construction, we briefly describe how it works. For every variable x_i of (X, C), we will define a set X_i, which consists of a four-cycle of forced edges, along with two optional edges, which form the diagonals of this four-cycle. At least one of these optional edges is present in every C_4-free sandwich graph for our instance, and which of the diagonal edges is present will correspond to whether x_i is true or false. For every clause c_j of (X, C), and every literal $\ell_q^j \in \{x_i, \overline{x_i}\}$ in c_j, we add two gadgets $\left\{r_{q1}^j, \ldots, r_{q4}^j\right\}$ and $\left\{s_{q1}^j, \ldots, s_{q4}^j\right\}$, which are designed to provide a copy of x_i if x_i is true and if x_i is false, respectively. For c_j, we add a four-cycle with vertex set $\left\{p_1^j, \ldots, p_4^j\right\}$, which has one forced edge, and one optional edge for each literal. It follows that if there is a set of optional edges that we can add to create a C_4-free sandwich graph G, then there is a truth assignment in which for every clause c_j, not all three optional edges among $\left\{p_1^j, \ldots, p_4^j\right\}$ are present in G, and hence c_j contains a true literal.

Later, we will slightly modify the 3Sat instance in order to be able to say more about the structure of the constructed sandwich instance.

We let $(G_1(X, C), G_2(X, C))$ be a sandwich instance with vertex set V, where V contains

- for each variable $x_i \in X$ a set $X_i = \{x_1^i, x_2^i, \overline{x}_1^i, \overline{x}_2^i\}$;
- for each clause $c_j = (\ell_1^j \vee \ell_2^j \vee \ell_3^j)$ in C, a set $C_j = A_j \cup B_j$ where $A_j = \left\{p_1^j, \ldots, p_4^j\right\}$ and $B_j = \bigcup_{q=1,2,3} \left\{l_q^j, t_q^j, s_{q1}^j, \ldots, s_{q4}^j, r_{q1}^j, \ldots, r_{q4}^j\right\}$.

We define the following sets of edges for $1 \leq i \leq n, 1 \leq j \leq m, q \in \{1, 2, 3\}$.

- $X^i = \{x_1^i\overline{x}_1^i, x_2^i\overline{x}_1^i, x_2^i\overline{x}_2^i, x_1^i\overline{x}_2^i\}$;
- $L^j = \left\{p_1^j l_1^j, l_1^j p_2^j, p_2^j l_2^j, l_2^j p_3^j, p_3^j l_3^j, l_3^j p_4^j\right\}$;
- $T^j = \left\{p_1^j t_1^j, t_1^j p_2^j, p_2^j t_2^j, t_2^j p_3^j, p_3^j t_3^j, t_3^j p_4^j\right\}$;
- $P^j = \left\{p_4^j p_1^j\right\}$;
- $A_q^j = \left\{p_q^j s_{q1}^j, p_{q+1}^j s_{q3}^j, l_q^j r_{q1}^j, t_q^j r_{q3}^j\right\}$;
- $S_q^j = \left\{s_{q1}^j s_{q2}^j, s_{q2}^j s_{q3}^j, s_{q3}^j s_{q4}^j, s_{q4}^j s_{q1}^j\right\}$;
- $R_q^j = \left\{r_{q1}^j r_{q2}^j, r_{q2}^j r_{q3}^j, r_{q3}^j r_{q4}^j, r_{q4}^j r_{q1}^j\right\}$;
- if $\ell_q^j = x_i$, $B_q^j = \left\{x_1^i s_{q2}^j, x_2^i s_{q4}^j, \overline{x}_1^i r_{q2}^j, \overline{x}_2^i r_{q4}^j\right\}$;
- if $\ell_q^j = \overline{x}_i$, $B_q^j = \left\{x_1^i r_{q2}^j, x_2^i r_{q4}^j, \overline{x}_1^i s_{q2}^j, \overline{x}_2^i s_{q4}^j\right\}$;
- $X_*^i = \{x_1^i x_2^i, \overline{x}_1^i\overline{x}_2^i\}$;
- $P_*^j = \left\{p_1^j p_2^j, p_2^j p_3^j, p_3^j p_4^j\right\}$;
- $Q_q^j = \left\{l_q^j t_q^j, s_{q1}^j s_{q3}^j, s_{q2}^j s_{q4}^j, r_{q1}^j r_{q3}^j, r_{q2}^j r_{q4}^j\right\}$.

We let

$$E(G_1(X,C)) = \bigcup_{1 \leq i \leq n} X^i \cup \bigcup_{1 \leq j \leq m, q \in \{1,2,3\}} (L^j \cup T^j \cup P^j \cup A_q^j \cup S_q^j \cup R_q^j \cup B_q^j)$$

and

$$E(G_2(X,C)) \setminus E(G_1(X,C)) = \bigcup_{1 \leq i \leq n} X_*^i \cup \bigcup_{1 \leq j \leq m, q \in \{1,2,3\}} (P_*^j \cup Q_q^j).$$

This construction is shown in Figure 1. The edges in set $E(G_1(X,C))$ correspond to the solid edges, and the edges in set $E(G_2(X,C)) \setminus E(G_1(X,C))$ correspond to the dotted edges. A C_4-free sandwich graph corresponds to a selection of the dotted edges that, when added to the solid edges, makes the graph C_4-free.

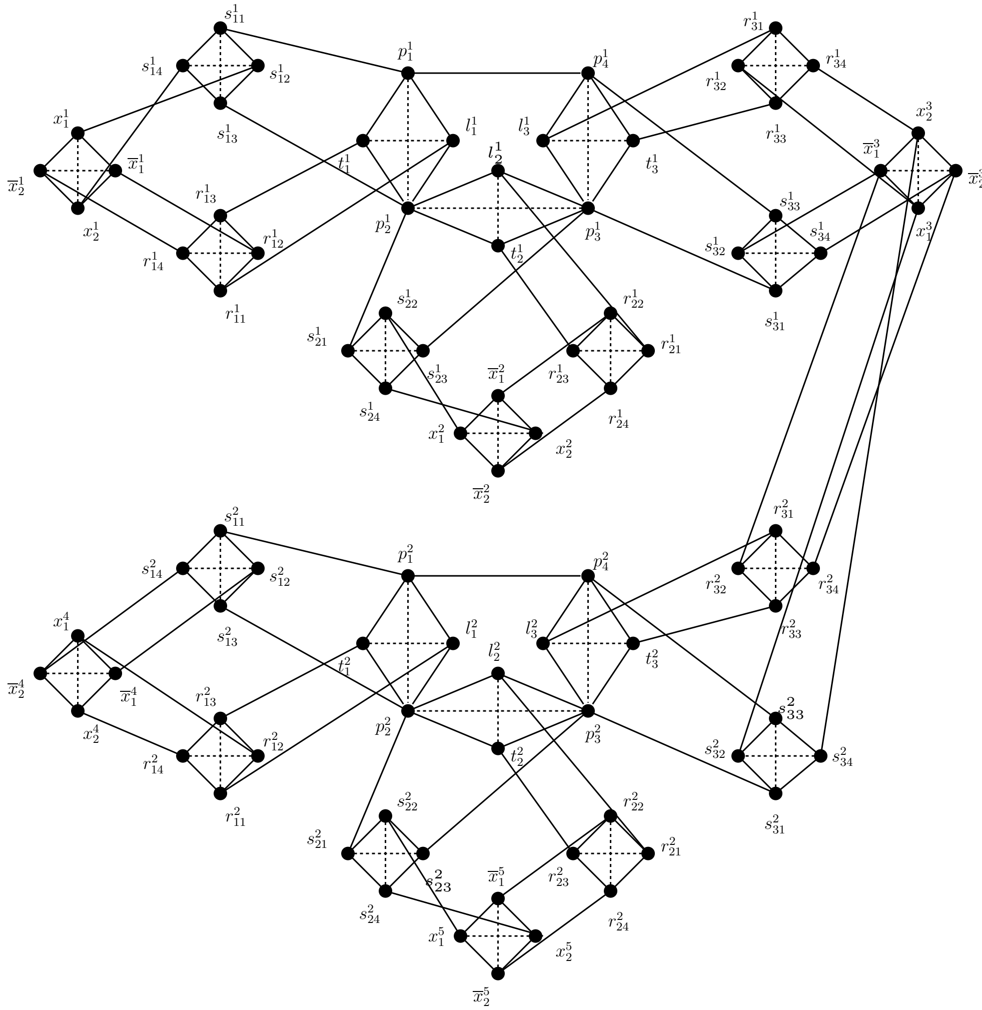

Figure 1: Example from [3] of the constructed instance $(G_1(X,C), G_2(X,C))$ for $X = \{x_1, \ldots, x_5\}, C = \{c_1, c_2\}$ and $c_1 = (x_1 \vee x_2 \vee \overline{x}_3), c_2 = (\overline{x}_4 \vee x_5 \vee x_3)$. Solid edges are forced, and dotted edges are optional.

The following follows immediately from the construction.

Lemma 12. *For every* 3SAT *instance in which each clause contains exactly three variables, the graph* $G = (V(G_1(X, C)), E(G_2(X, C)) \setminus E(G_1(X, C)))$ *of optional edges is a forest, and each connected component consists either of an edge, a single vertex, or a three-edge path.*

Theorem 13 ([3])**.** *Let* (X, C) *be a* 3SAT *instance in which each clause contains exactly three variables. The instance* $(G_1(X, C), G_2(X, C))$ *is a* YES *instance for the* C_4*-free sandwich problem if and only if* (X, C) *is a* YES *instance for* 3SAT*.*

A house can be constructed from a four-cycle by adding a new vertex adjacent to two consecutive vertices of the cycle. By adding such a new vertex to every possible C_4 in the sandwich instance we constructed, we obtain the following result:

Theorem 14. *The house-free sandwich problem is* NP*-complete. It is* NP*-complete to decide if* (G_1, G_2) *is a* YES *instance for the* C_4*-free sandwich problem or a* NO *instance for the house-free sandwich problem.*

Proof. Let (X, C) be an instance of 3SAT in which each clause contains exactly three variables. Let (G'_1, G'_2) arise from $(G_1(X, C), G_2(X, C))$ by adding a new vertex of degree two for each edge $uv \in E(G_1(X, C))$, which defines a new set W. The new set W contains $|E(G_1(X, C))|$ vertices, each $w \in W$ is associated with an edge $e(w) \in E(G_1(X, C))$, and for each $w \in W$, $e(w) = uv$, we add edges uw and wv to G'_1 and G'_2. It follows that since $e(w)$ is in every sandwich graph for every $w \in W$, no $w \in W$ is in a four-cycle in any sandwich graph for (G'_1, G'_2).

Suppose that (G'_1, G'_2) is a YES instance for the house-free sandwich problem, and let G be a house-free sandwich graph for (G'_1, G'_2). Suppose that $G \setminus W$ contains a four-cycle with vertex set $\{a, b, c, d\}$. By Lemma 12, it follows that one of its edges, say ab, is in $E(G_1(X, C))$. Let $w \in W$ with $e(w) = ab$, then $\{a, b, c, d, w\}$ induces a house in G, a contradiction. It follows that $G \setminus W$ is a C_4-free sandwich graph for $(G_1(X, C), G_2(X, C))$, proving that (G'_1, G'_2) is a YES instance for the C_4-free sandwich problem. By Theorem 13, it follows that (X, C) is a YES instance for 3SAT.

Conversely, suppose that (X, C) is a YES instance for 3SAT. Then, by Theorem 13, it follows that there exists a sandwich graph G for $(G_1(X, C), G_2(X, C))$ such that G is C_4-free. Let $E = E(G) \setminus E(G_1(X, C))$, and let $G' = G'_1 \cup E$. Suppose that G' contains a four-cycle. Since no vertex in W is in a four-cycle as observed above, it follows that $G' \setminus W$ contains a four-cycle. But

$G' \setminus W = G$, a contradiction. Thus G' is C_4-free, and hence house-free. This proves that (G'_1, G'_2) is a Yes instance for the C_4-free and the house-free sandwich problem.

We have proved that (X, C) is a Yes instance for 3Sat if and only if (G'_1, G'_2) is a Yes instance for the house-free sandwich problem if and only if (G'_1, G'_2) is a Yes instance for the C_4-free sandwich problem. Since 3Sat is NP-complete, it follows that the house-free sandwich problem is NP-complete. $\square$

To prove the following statement, we replace the vertex set of our previous instance by a stable set N, and add two types of gadgets corresponding to forced edges and non-edges of the previous instance, respectively:

Theorem 15. *The house-free partitioned probe problem is NP-complete.*

Proof. In Theorem 14, we proved that the problem of deciding if (G_1, G_2) is a Yes instance for the C_4-free sandwich problem or a No instance for the house-free sandwich problem is NP-complete.

Let (G_1, G_2) be a sandwich instance. Let $E_1 = E(G_1), E_2 = E(G_2)$. We construct an instance G', N for the house-free partitioned probe problem as follows. We start from a stable set N which is in bijection with the vertices of G_1, and for every pair u, v of distinct vertices in N, we add the following:

- if $uv \in E_1$, we add vertices a, b, c with $N(a) = \{u, v, c\}, N(b) = \{u, v\}$ and $N(c) = \{a, u\}$;
- if $uv \notin E_2$, we add vertices a, b, c with $N(a) = \{u, b, c\}$, $N(b) = \{a, c\}$, $N(c) = \{a, b, v\}$.

Let $P = V(G') \setminus N$.

Suppose that (G_1, G_2) is a Yes instance for the C_4-free sandwich problem, and let H be a C_4-free sandwich graph for (G_1, G_2). Let $H' = G' \cup E(H)$. Then $H'|N$ is C_4-free. Suppose that H' is not house-free. Then H' contains a four-cycle using a vertex x in P. Suppose that x was added by the first bullet, for $uv \in E_1$, then $uv \in E(H')$. If $x = b$ or $x = c$, then x has exactly two neighbors, and they are adjacent; consequently x is not in a four-cycle. Therefore, $x = a$, but neither b nor c are in a four-cycle, but two neighbors of a are in a four-cycle. This is a contradiction, since a has degree three. It follows that x is not in a four-cycle, a contradiction. Thus x was added by the second bullet, and it follows that $x \neq b$, so we may assume that $x = a$.

Since b is not in a four-cycle, it follows that the four-cycle containing x also contains c, u and v. But $uv \notin E_2$ and so $uv \notin E(H')$, a contradiction. This proves that H' is C_4-free, and in particular, house-free. So G', N is a YES instance for the house-free partitioned probe problem.

Now suppose that G', N is a YES instance for the house-free partitioned probe problem, and let H' be a house-free (P, N)-probe graph for G'. Let $uv \in E_1$, then $uv \in E(H')$, because otherwise u, v together with the vertices added for u, v in the first bullet induce a house in H'. Let $uv \notin E_2$, then $uv \notin E(H')$, because otherwise u, v together with the vertices added for u, v in the second bullet induce a house in H'. It follows that $H'|N$ is a house-free sandwich graph for (G_1, G_2), and so (G_1, G_2) is not a NO instance for the house-free sandwich problem.

Thus, we have reduced the problem from Theorem 14 to the house-free partitioned probe problem, and it follows that the house-free partitioned probe problem is NP-complete. □

By taking complements in Theorem 15, we obtain the following.

Corollary 16. *The P_5-free partitioned probe problem is NP-complete.*

Our next goal is to show that the P_6-free sandwich problem is NP-complete.

Lemma 17. *The* 3SAT *problem is NP-complete even when restricted to instances in which no negations appear in clauses of length three (but negations may occur in clauses of length two).*

Proof. Let (X, C) be a 3SAT instance and suppose that there exist i, j such that $\overline{x}_i$ occurs in C_j, where C_j has length three. We replace $\overline{x}_i$ by a new variable y in C_j, and add two clauses $(x_i \vee y)$ and $(\overline{x}_i \vee \overline{y})$. The two new clauses are satisfied if and only if $y = \overline{x}_i$, and so (X, C) is satisfiable if and only if the new instance is. This does not create any new clauses of length three. Thus, by adding at most $3|C|$ new variables and $6|C|$ new clauses, we can transform (X, C) into an equivalent 3SAT instance with the required restrictions. This proves the result. □

We use the following result, a tool in the proof of Theorem 13:

Lemma 18 ([3])**.** *Let (X, C) be a* 3SAT *instance in which every clause contains exactly three variables. Let $(x_1, \ldots, x_n)$ be an assignment of the variables. Then the following are equivalent:*

- *(X,C) is satisfied by the assignment $(x_1,\ldots,x_n)$;*
- *there is a C_4-free sandwich graph G for $(G_1(X,C),G_2(X,C))$ with $x_1^ix_2^i \in E(G)$ if and only if x_i is true in the assignment $(x_1,\ldots,x_n)$.*

The following is a list of every possible four-cycle in a sandwich graph for $(G_1(X,C),G_2(X,C))$:

- $p_1^jp_2^jp_3^jp_4^j$, $p_q^jt_q^jp_{q+1}^jl_q^j$, $p_q^js_{q1}^js_{q3}^jp_{q+1}^j$, $l_q^jr_{q1}^jr_{q3}^jt_q^j$;
- $s_{q1}^js_{q2}^js_{q3}^js_{q4}^j, r_{q1}^jr_{q2}^jr_{q3}^jr_{q4}^j, x_1^i\overline{x}_1^ix_2^i\overline{x}_2^i$;
- if x_i occurs in position q in C_j, $x_1^is_{q2}^js_{q4}^jx_2^i$, $\overline{x}_1^ir_{q2}^jr_{q4}^j\overline{x}_2^i$;
- if $\overline{x}_i$ occurs in position q in C_j, $x_1^ir_{q2}^jr_{q4}^jx_2^i$, $\overline{x}_1^is_{q2}^js_{q4}^j\overline{x}_2^i$.

The following lemma shows how to modify our construction in order to use it to decide if a 3Sat instance (X,C) with $x_i \in X$ has a satisfying assignment in which x_i is false. We accomplish this by removing all vertices designed to represent x_i and its copies, and by turning every edge $p_q^jp_{q+1}^j$ corresponding to the literal $\ell_q^j = x_i$ from an optional edge into a forced edge.

Lemma 19. *Let (X,C) be a 3Sat instance in which every clause contains exactly three variables, and let $x_i \in X$ such that every occurrence of x_i in C is non-negated (i. e. the literal $\overline{x}_i$ does not occur). We let V_i denote the set of vertices that are either in X_i or of the form l_q^j, t_q^j, r_{qk}^j or s_{qk}^j for a clause $c_j = (\ell_1^j \vee \ell_2^j \vee \ell_3^j)$ where $k \in \{1,2,3,4\}$, $\ell_q^j = x_i$, and $j \in \{1,\ldots,m\}$. We let E_i denote the set of edges $p_q^jp_{q+1}^j$ for a clause $c_j = (\ell_1^j \vee \ell_2^j \vee \ell_3^j)$ with $\ell_q^j = x_i$ and $j \in \{1,\ldots,m\}$. We define a sandwich instance $(G_1^i(X,C),G_2^i(X,C))$ with $V(G_d^i(X,C)) = V(G_d(X,C)) \setminus V_i$ and $G_d^i(X,C) = G_d(X,C)|V(G_d^i(X,C)) \cup E_i$ for $d = 1,2$. Then (X,C) is satisfied by an assignment such that x_i is false if and only if there is a C_4-free sandwich graph for $(G_1^i(X,C),G_2^i(X,C))$.*

Proof. By Lemma 18, (X,C) is satisfied by an assignment such that x_i is false if and only if there is a C_4-free sandwich graph G for $(G_1(X,C),G_2(X,C))$ with $x_1^ix_2^i \notin E(G)$. Since G is C_4-free, this implies that $\overline{x}_1^i\overline{x}_2^i \in E(G)$. Let $c_j = (\ell_1^j \vee \ell_2^j \vee \ell_3^j)$ such that $\ell_q^j = x_i$ for some $q \in \{1,2,3\}$ and $j \in \{1,\ldots,m\}$. Then $r_{q2}^jr_{q4}^j \notin E(G)$, for otherwise $\left\{\overline{x}_1^i, r_{q2}^j, r_{q4}^j, \overline{x}_2^i\right\}$ induces a four-cycle in G. Therefore, $r_{q1}^jr_{q3}^j \in E(G)$, and so $l_q^jt_q^j \notin E(G)$. This implies that $p_q^j p_{q+1}^j \in E(G)$. Consequently, $G|V(G_1^i(X,C))$ is a C_4-free sandwich graph for $(G_1^i(X,C), G_2^i(X,C))$.

For the converse direction suppose that G is a C_4-free sandwich graph for $(G_1^i(X,C), G_2^i(X,C))$ and let F denote the set of edges $\overline{x}_1^i\overline{x}_2^i$ and $r_{q1}^j r_{q3}^j$, $s_{q2}^j s_{q4}^j$ for $j \in \{1,\dots,m\}$, $q \in \{1,2,3\}$ such that $c_j = (l_1^j \vee l_2^j \vee l_3^j)$ such that $l_q^j = x_i$. Then G' with $V(G') = V(G_1(X,C))$ and $E(G') = E(G_1(X,C)) \cup E(G) \cup F$ is a sandwich graph for $(G_1(X,C), G_2(X,C))$ with $x_1^i x_2^i \notin E(G)$. Moreover, G' is C_4-free by construction. By Lemma 18, it follows that (X,C) has a satisfying assignment in which x_i is false. □

We now consider the P_6-free sandwich problem. Our strategy is as follows. We switch to the P_6^C-free sandwich problem. Every possible P_6^C in our instance will consist of a possible four-cycle with vertex set $\{a,b,c,d\}$ in the instance (G_1, G_2) as in Lemma 19, along with a vertex e adjacent to b and c, where bc is a forced edge and b-d-e-a-c is a P_5^C, and a vertex f adjacent to either $\{d,e,a,c\}$ or $\{b,d,e,a\}$. Note that if there are two vertices, f and f', such that f is adjacent to $\{d,e,a,c\}$ and f' is adjacent to $\{b,d,e,a\}$, then there is a possible P_6^C that does not use all vertices of the four-cycle. To avoid this problem, our main goal is to identify sets E^*, V^* such that E^* is a set of forced edges and contains an edge of every possible four-cycle of (G_1, G_2), and V^* is a stable set of vertices in G_2 that contains an endpoint of every edge in E^*. Then we add a vertex e to every edge bc in E^*, and add a vertex f with neighbors $\{d,e,a,c\}$ if $b \in V^*$, and with neighbors $\{b,d,e,a\}$ if $c \in V^*$.

Theorem 20. *The P_6-free sandwich problem is NP-complete.*

Proof. Let (X', C') be a 3SAT instance as in Lemma 17. We may assume that (X', C') contains no clauses containing exactly one variable. Let (X, C) arise from (X', C') by adding a new variable x_0 to X, and by replacing every clause $c_j = (\ell_1^j \vee \ell_2^j)$ of length two in C' by $(x_0 \vee \ell_1^j \vee \ell_2^j)$ if ℓ_1^j and ℓ_2^j are either both negated or both non-negated, and by $(\ell_1^j \vee x_0 \vee \ell_2^j)$ otherwise. We remark that the order matters for the remaining of the proof (for the construction of set V^* below). Then (X, C) has a satisfying assignment in which x_0 is false if and only if (X', C') has a satisfying assignment. Furthermore, every clause in C contains exactly three variables.

Let $V_0, E_0, (G_1^0(X,C), G_2^0(X,C))$ as in Lemma 19. Then $(G_1^0(X,C), G_2^0(X,C))$ has a C_4-free sandwich graph if and only if (X', C') is satisfiable. Let $|C| = m$;

and for $j \in \{1, \ldots, m\}$, let $c_j = (\ell_1^j \vee \ell_2^j \vee \ell_3^j)$ denote the jth clause of C. We let E^* denote the following set of edges:

- for $x_i \in X'$, $\overline{x}_1^i x_1^i$;
- for $x_i \in X'$, and q, j such that $x_i = \ell_q^j$, $\overline{x}_1^i r_{q2}^j$, $r_{q1}^j r_{q4}^j$, $r_{q1}^j l_q^j$, $s_{q4}^j x_2^i$, $s_{q4}^j s_{q1}^j$;
- for $x_i \in X'$, and q, j such that $\overline{x}_i = \ell_q^j$, $\overline{x}_1^i s_{q2}^j$, $s_{q1}^j s_{q4}^j$, $s_{q1}^j p_q^j$, $r_{q4}^j x_2^i$, $r_{q4}^j r_{q1}^j$;
- if x_0 does not occur in c_j, $p_1^j p_4^j$, $p_1^j l_1^j$, $p_3^j l_2^j$, $p_3^j l_3^j$, $p_1^j s_{11}^j$, $p_3^j s_{31}^j$, $p_3^j s_{23}^j$;
- if $c_j = (x_0 \vee x_i \vee x_k)$, $p_1^j p_4^j$, $p_3^j l_2^j$, $p_3^j l_3^j$, $p_3^j s_{31}^j$, $p_3^j s_{23}^j$;
- if $c_j = (x_0 \vee \overline{x}_i \vee \overline{x}_k)$, $p_1^j p_4^j$, $l_2^j p_3^j$, $l_3^j p_3^j$, $l_2^j r_{21}^j$, $l_3^j r_{31}^j$;
- if $c_j = (x_i \vee x_0 \vee \overline{x}_k)$, $p_1^j p_4^j$, $p_1^j l_1^j$, $l_3^j p_3^j$, $p_1^j s_{11}^j$, $l_3^j r_{31}^j$; and
- if $c_j = (\overline{x}_i \vee x_0 \vee x_k)$, $p_4^j p_1^j$, $p_4^j l_3^j$, $l_1^j p_2^j$, $l_1^j r_{11}^j$, $p_4^j s_{33}^j$.

We let V^* denote the following set of vertices:

- for $x_i \in X'$, $\overline{x}_1^i$;
- for $x_i \in X'$, and q, j such that $x_i = \ell_q^j$, r_{q1}^j and s_{q4}^j;
- for $x_i \in X'$, and q, j such that $\overline{x}_i = \ell_q^j$, s_{q1}^j and r_{q4}^j;
- if x_0 does not occur in c_j, p_1^j and p_3^j;
- if $c_j = (x_0 \vee x_i \vee x_k)$, p_1^j and p_3^j;
- if $c_j = (x_0 \vee \overline{x}_i \vee \overline{x}_k)$, p_1^j, l_2^j and l_3^j;
- if $c_j = (x_i \vee x_0 \vee \overline{x}_k)$, l_3^j and p_1^j; and
- if $c_j = (\overline{x}_i \vee x_0 \vee x_k)$, l_1^j and p_4^j.

By construction, E^* and V^* have the following properties:

- $E^* \subseteq E(G_1^0(X, C))$;
- for every four-cycle in a sandwich graph for $(G_1^0(X, C), G_2^0(X, C))$, E^* contains an edge of the cycle;

- V^* is a stable set in $G_2^0(X,C)$; and
- every edge in E^* has an endpoint in V^*.

These properties are easily verified from the construction using the additional restrictions on (X,C). Let (G'_1, G'_2) arise from $(G_1^0(X,C), G_2^0(X,C))$ by adding a vertex $w(e)$ for every edge $e \in E^*$ with $N_{G'_1}(w(e)) = N_{G'_2}(w(e)) = \{u,v\}$ where $e = uv$. Let (G_1, G_2) arise from (G'_1, G'_2) by adding a vertex v' for every vertex $v \in V^*$ with $N_{G_1}(v') = N_{G_2}(v') = V(G_1) \setminus \{v, v'\}$, i. e. the set $V' = \{v' : v \in V^*\}$ is a clique.

We claim that (G_1, G_2) is a Yes instance for the P_6^C-free sandwich problem if and only if $(G_1^0(X,C), G_2^0(X,C))$ is a Yes instance for the C_4-free sandwich problem (which, by Lemma 19 is true if and only if (X', C') is a Yes instance for 3Sat).

Suppose that G is a P_6^C-free sandwich graph for (G_1, G_2). Let $G' = G|V(G_1^0(X,C))$. Suppose that G' contains an induced four-cycle with vertex set $\{a_1, a_2, a_3, a_4\}$ and edge set $\{a_1a_2, a_2a_3, a_3a_4, a_4a_1\}$. Then E^* contains an edge of this four-cycle; without loss of generality, say $a_1a_2 \in E^*$. This implies that one of $a_1, a_2 \in V^*$, say $a_1 \in V^*$. But then a'_1-a_1-a_3-$w(a_1a_2)$-a_4-a_2 is an induced P_6 in G^C, a contradiction. So G' is a C_4-free sandwich graph for $(G_1^0(X,C), G_2^0(X,C))$.

Conversely, suppose that G is a C_4-free sandwich graph for $(G_1^0(X,C), G_2^0(X,C))$. Let $G' = G_1 \cup E(G)$. Suppose that the complement of G' contains an induced P_6 with vertices $a_1, \ldots, a_6$ in this order along the path. Then $V' \cap \{a_1, \ldots, a_6\} \subseteq \{a_1, a_6\}$, since every interior vertex has at least two non-neighbors. Suppose first that $V' \cap \{a_1, \ldots, a_6\} = \{a_1, a_6\}$. Then $a_2, a_5 \in V^*$, but a_2 is adjacent to a_5, a contradiction, since V^* is a stable set. It follows that $|V' \cap \{a_1, \ldots, a_6\}| \le 1$, and so $G' \setminus V'$ contains a P_5^C, and thus an induced four-cycle. Since G is C_4-free, it follows that there is a vertex in $V(G') \setminus (V' \cup V(G))$ which is contained in an induced four-cycle in $G' \setminus V'$. Thus there exists an edge $e \in E^*$ such that $w(e)$ is in an induced four-cycle in G. But $w(e)$ has exactly two neighbors in $G' \setminus V'$, and they are adjacent in every sandwich graph $G' \setminus V'$ for (G'_1, G'_2), it follows that $w(e)$ is not in an induced four-cycle in $G' \setminus V'$. This is a contradiction, and it follows that G' is P_6^C-free.

This implies that the P_6^C-free sandwich problem is NP-complete, and by taking complements, it follows that the P_6-free sandwich problem is NP-complete. $\square$

4. From short paths to longer paths

Lemma 21. *For $k \geq 5$, if the P_k-free sandwich is NP-complete, then the P_{k+2}-free sandwich problem is NP-complete.*

Proof. Let (G_1, G_2) be an instance of the P_k-free sandwich problem, and for $i \in \{1, 2\}$ let G'_i arise from G_i by adding, for each vertex v in $V(G_i)$, a new vertex v' with $N(v') = v$.

Suppose that (G_1, G_2) is a YES instance for the P_k-free sandwich problem. Let H be a P_k-free sandwich graph for (G_1, G_2). Let $H' = G'_1 \cup (E(H) \setminus E(G_1)) = G'_1 \cup E(H)$. Then $H'|V(G_1)$ is P_k-free. Suppose that H' is not P_{k+2}-free, and let Q be an induced path on $k+2$ vertices in H'. Then $V(H') \setminus V(G_1)$ contains a vertex in $V(Q)$ of degree at least two. But every vertex in $V(H') \setminus V(G_1)$ has degree one in H', a contradiction. This proves that H' is P_{k+2}-free, and so (G'_1, G'_2) is a YES instance for the P_{k+2}-free sandwich problem.

Conversely, suppose that (G'_1, G'_2) is a YES instance for the P_{k+2}-free sandwich problem. Let H' be a P_{k+2}-free sandwich graph for (G'_1, G'_2). Then $H'|V(G_1)$ is a P_k-free sandwich graph for (G_1, G_2), because if there was a P_k in $H'|V(G_1)$ with endpoints $u, v \in V(G_1)$, then adding u', v' to this path would yield a P_{k+2} in H', a contradiction. So (G_1, G_2) is a YES instance for the P_k-free sandwich problem. □

Theorem 22. *For $k \geq 5$, the P_k-free sandwich problem, partitioned probe problem, and unpartitioned probe problem in the graph and in the complement are NP-complete.*

Proof. For the sandwich problem, this follows from Corollary 16 for $k = 5$, since the partitioned probe problem is a special case of the sandwich problem, and from Theorem 20 and Lemma 21. For the partitioned probe problem, this follows from Corollary 16 for $k = 5$, and from Corollary 4 for $k \geq 6$, since P_k^C is 3-connected for $k \geq 6$.

Since P_k, $k \geq 5$, does not contain twins, the NP-completeness for the unpartitioned problem in the graph and in the complement follows from Theorem 5 and Corollary 6, respectively. □

5. Concluding remarks

We have resolved the complexity of the sandwich problem, partitioned probe problem, unpartitioned probe problem, and unpartitioned probe problem in the complement for the property of being P_k-free for all $k \in \mathbb{N}$. This is a special case of the more general question for which graphs H the abovementioned problems for the property of being H-free is NP-complete, and for which they can be solved in polynomial time.

We further compare the complexities of the sandwich problem, partitioned probe problem, unpartitioned probe problem, and unpartitioned probe problem in the complement, and establish a new sufficient condition for the hardness of the partitioned probe problem to imply that the other problems are also hard.

Acknowledgments

We are thankful to Maria Chudnovsky for many helpful discussions. We thank the anonymous referees for several remarks which helped improve the presentation of this paper. The first author was supported by Conselho Nacional de Desenvolvimento Científico e Tecnológico CNPq Grant 303622/2011-3.

[1] Chandler, David B., Maw-Shang Chang, Ton Kloks, Jiping Liu, and Sheng-Lung Peng. *Recognition of probe cographs and partitioned probe distance hereditary graphs.* In International Conference on Algorithmic Applications in Management, pp. 267–278. Springer Berlin Heidelberg, 2006.

[2] Couto, Fernanda, Luerbio Faria, Sylvain Gravier, and Sulamita Klein. *On the forbidden induced subgraph probe and sandwich problems.* Discrete Applied Mathematics 234, (2018): 56–66.

[3] Dantas, Simone, Celina M. H. de Figueiredo, Murilo V. G. da Silva, and Rafael B. Teixeira. *On the forbidden induced subgraph sandwich problem.* Discrete Applied Mathematics 159, no. 16 (2011): 1717–1725.

[4] Golumbic, Martin Charles, and Ron Shamir. *Complexity and algorithms for reasoning about time: A graph-theoretic approach.* Journal of the ACM (JACM) 40.5 (1993): 1108–1133.

[5] Golumbic, Martin Charles, Haim Kaplan, and Ron Shamir. *Graph sandwich problems.* Journal of Algorithms 19, no. 3 (1995): 449–473.

[6] McMorris, Fred R., Chi Wang, and Peisen Zhang. *On probe interval graphs.* Discrete Applied Mathematics 88, no. 1 (1998): 315–324.

[7] Zhang, Peisen, Eric A. Schon, Stuart G. Fischer, Eftihia Cayanis, Janie Weiss, Susan Kistler, and Philip E. Bourne. *An algorithm based on graph theory for the assembly of contigs in physical mapping of DNA*. Computer applications in the biosciences: CABIOS 10, no. 3 (1994): 309–317.