

## KERNELIZATION OF GRAPH HAMILTONICITY: PROPER $H$ -GRAPHS\*

STEVEN CHAPLICK<sup>†</sup>, FEDOR V. FOMIN<sup>‡</sup>, PETR A. GOLOVACH<sup>‡</sup>, DUŠAN KNOP<sup>§</sup>,  
AND PETER ZEMAN<sup>¶</sup>

**Abstract.** We obtain new polynomial kernels and compression algorithms for PATH COVER and CYCLE COVER, the well-known generalizations of the classical HAMILTONIAN PATH and HAMILTONIAN CYCLE problems. Our choice of parameterization is strongly influenced by the work of Biró, Hujter, and Tuza, who in 1992 introduced  $H$ -graphs, intersection graphs of connected subgraphs of a subdivision of a fixed (multi-)graph  $H$ . In this work, we turn to proper  $H$ -graphs, where the containment relationship between the representations of the vertices is forbidden. As the treewidth of a graph measures how similar the graph is to a tree, the size of graph  $H$  is the parameter measuring the closeness of the graph to a proper interval graph. We prove the following results. PATH COVER admits a kernel of size  $\mathcal{O}(\|H\|^8)$ , where  $\|H\|$  is the size of graph  $H$ . In other words, we design an algorithm that for an  $n$ -vertex graph  $G$  and integer  $k \geq 1$ , in time polynomial in  $n$  and  $\|H\|$ , outputs a graph  $G'$  of size  $\mathcal{O}(\|H\|^8)$  and  $k' \leq |V(G')|$  such that the vertex set of  $G$  is coverable by  $k$  vertex-disjoint paths if and only if the vertex set of  $G'$  is coverable by  $k'$  vertex-disjoint paths. HAMILTONIAN CYCLE admits a kernel of size  $\mathcal{O}(\|H\|^8)$ . CYCLE COVER admits a polynomial kernel. We prove it by providing a compression of size  $\mathcal{O}(\|H\|^{10})$  into another NP-complete problem, namely, PRIZE COLLECTING CYCLE COVER, that is, we design an algorithm that, in time polynomial in  $n$  and  $\|H\|$ , outputs an equivalent instance of PRIZE COLLECTING CYCLE COVER of size  $\mathcal{O}(\|H\|^{10})$ . In all our algorithms we assume that a proper  $H$ -decomposition is given as a part of the input.

**Key words.** cycle cover, path cover, proper  $H$ -graphs, kernelization

**AMS subject classifications.** 68Q25, 68R05, 68R10, 68W01, 05C85

**DOI.** 10.1137/19M1299001

**1. Introduction.** The HAMILTONIAN CYCLE problem, an old mathematical puzzle whose study can be traced back to the 19th century, is still a topic of active research. Our results about the HAMILTONIAN CYCLE problem are at the intersection of two research areas: kernelization and algorithms on special graph classes. In both areas HAMILTONIAN CYCLE has been intensively investigated.

**Parameterized algorithms and kernelization.** The most popular generalization of the HAMILTONIAN CYCLE problem studied in parameterized complexity is known under the name LONGEST CYCLE. This problem is to decide whether a graph contains a cycle of length at least  $k$ , where  $k$  is an integer parameter. LONGEST CYCLE

---

\*Received by the editors November 18, 2019; accepted for publication (in revised form) December 7, 2020; published electronically April 26, 2021. An extended abstract of this paper appeared at WADS 2019 [11] under the same title.

<https://doi.org/10.1137/19M1299001>

**Funding:** The research leading to these results has been supported by the Research Council of Norway via the projects “CLASSIS” (grant 249994) and “MULTIVAL” (grant 263317). The fourth author was supported by OP VVV MEYS project CZ.02.1.01/0.0/0.0/16.019/0000765 (“Research Center for Informatics”) and by grant 17-20065S of the Czech Science Foundation. The fifth author was supported by grant 19-17314J of the Czech Science Foundation.

<sup>†</sup>Department of Data Science and Knowledge Engineering, Maastricht University, Maastricht, The Netherlands (s.chaplick@maastrichtuniversity.nl).

<sup>‡</sup>Department of Informatics, University of Bergen, Bergen, Norway (fedor.fomin@uib.no, petr.golovach@uib.no).

<sup>§</sup>Department of Theoretical Computer Science, Faculty of Information Technology, Czech Technical University, Prague, Czech Republic (dusan.knop@fit.cvut.cz).

<sup>¶</sup>Department of Applied Mathematics, Faculty of Mathematics and Physics, Charles University, Prague, Czech Republic (zeman@kam.mff.cuni.cz).

and its close relative LONGEST PATH are important representatives of the so-called family of “nonlocal” problems, and this is why these problems served as test-beds in the development of various fundamental techniques in areas such as color coding [1], algebraic methods [31, 37, 4], and Cut & Count [16], to name a few. We refer the reader to the book of Cygan et al. [15] for an overview of these techniques. From the perspective of kernelization, the framework developed by Bodlaender et al. [5] excludes the existence of a polynomial kernel (up to some reasonable assumption from complexity theory) for LONGEST CYCLE with the natural parameter  $k$ . This lower bound initiated the development of kernelization algorithms for HAMILTONIAN CYCLE with “structural kernelization.” Fellows et al. [20] proved that HAMILTONIAN CYCLE parameterized by the *max leaf number* of the input graph  $G$ , that is, the maximum number of leaves in a spanning tree of  $G$ , admits a kernel of polynomial size. A systematic approach in the study of structural kernelization of HAMILTONIAN CYCLE (and other related problems) was taken by Bodlaender, Jansen, and Kratsch [6], who considered kernelization of HAMILTONIAN CYCLE parameterized by the size of the modulator to some nice graph property. More precisely, for a graph  $G$  the modulator to a graph property  $\Pi$  is a set of vertices or edges such that after removing this set from graph  $G$ , the resulting graph has property  $\Pi$ . In particular, Bodlaender, Jansen, and Kratsch [6] have shown that HAMILTONIAN CYCLE admits a polynomial kernel when parameterized by the size of a minimum vertex cover (a minimum modulator to an independent set) or by the size of a minimum modulator to the cluster graph, that is, the disjoint union of complete graphs. They also provided a number of lower bounds on the structural kernelization of the problem by showing, for example, that the problem does not admit a polynomial kernel when the parameter is the minimum size of a modulator to an outerplanar graph.

**Graph classes.** There is a large research area in graph algorithms, where the structural properties of graphs, like being interval or chordal, are exploited for developing efficient algorithms for problems intractable on general graphs. We refer the reader to the books [8, 25] for the introduction and survey of the known results. Without a doubt, the oldest and the most studied class of intersection graphs is the class of interval graphs, and there is a long history of research on the HAMILTONIAN CYCLE and HAMILTONIAN PATH problems on interval, circular-arc, and related graph classes. It was shown by Keil [30] in 1985 that HAMILTONIAN CYCLE can be solved in linear time for interval graphs (see also [9, 10, 17, 33]). The problem for circular-arc graphs proved to be much more involved, and the first polynomial-time algorithm for HAMILTONIAN CYCLE on circular-arc graphs was given by Shih, Chern, and Hsu [36] in 1992 (see also [28]). On the other hand, for proper interval graphs, it was already shown by Bertossi [2] that every connected proper interval graph has a Hamiltonian path, and a proper interval graph has a Hamiltonian cycle if and only if it is a 2-connected graph with at least three vertices (see also [14, 29]). This immediately implies a linear time algorithm for the problem. It follows from the results of Brandstädt, Dragan, and Köhler [7] that HAMILTONIAN CYCLE can be solved in linear time for circular-arc graphs. Thus, HAMILTONIAN CYCLE can be solved in linear time for (proper) interval graphs. For chordal graphs, HAMILTONIAN CYCLE is well known to be NP-complete and is even NP-complete for strongly chordal split graphs [34].

**Our results.** In this paper we follow the main question of structural kernelization: if a computational problem can be solved in polynomial time on instances with some structural properties, does it admit a polynomial kernel parameterized by some

“distance” to this nice structural property? In our setting the nice structural property is being a proper interval graph. However, the “distance” we use is quite different from the commonly used size of a modulator.

Our measure of similarity with proper interval graphs is based on the beautiful concept of  $H$ -graphs introduced by Biró, Hujter, and Tuza [3] in the context of the precoloring extension problem. An *intersection representation* of a graph  $G$  assigns a set  $S_v$  to every vertex  $v \in V(G)$  such that  $S_u \cap S_v \neq \emptyset$  if and only if  $uv \in E(G)$ . In the case when the sets  $S_u$  are intervals of the real line, this defines an interval graph. From a different perspective, every interval graph can be viewed as an intersection graph of subpaths of some (sufficiently long) path. Similarly, *circular-arc graphs*, a natural generalization of interval graphs, are the intersection graphs of subpaths of some cycle. It is also a well-known fact that a graph is *chordal* if and only if it is an intersection graph of subtrees of some tree. All of these classes are known to have efficient algorithms for various computational problems. We refer the reader to [8, 25] for the introduction and survey of the known results. A natural generalization of these classes is intersection graphs of subgraphs of some subdivision of an arbitrary *underlying graph*  $H$ . For a fixed graph  $H$ , we say that a graph  $G$  is an  $H$ -graph if it is an intersection graph of connected subgraphs of a subdivision of  $H$ . In this language, interval graphs are  $K_2$ -graphs, circular-arc graphs are  $K_3$ -graphs, and every chordal graph is a  $T$ -graph for some tree  $T$ .

An *intersection representation*  $\{S_v\}_{v \in V(G)}$  of a graph  $G$  is a *proper* representation if  $S_u \subseteq S_v$  implies  $u = v$ . Then a graph  $G$  is a *proper  $H$ -graph* if it admits a proper intersection representation by connected subgraphs of a subdivision of  $H$ . For example, proper  $K_2$ -graphs are proper interval graphs, that is, the graphs admitting a proper representation by intervals of the real line. Various aspects of proper interval and proper circular-arc representations have been well studied, and our goal is again to study how these carry to general proper  $H$ -graphs. Clearly, all positive algorithmic results obtained for  $H$ -graphs in [12, 13, 21] are valid for proper  $H$ -graphs, but since we consider a more restricted graph class, we hope that the tractability area can be expanded.

We consider the following fundamental generalizations of HAMILTONIAN CYCLE and HAMILTONIAN PATH problems, whose tasks are to cover the vertices of a graph by the minimum number of disjoint cycles and paths, respectively (see section 2 for the formal definition).

#### CYCLE COVER

*Input:* A graph  $G$  and a positive integer  $k$ .  
*Task:* Decide whether  $G$  has a cycle cover  $\mathcal{C}$  with at most  $k$  cycles.

and

#### PATH COVER

*Input:* A graph  $G$  and a positive integer  $k$ .  
*Task:* Decide whether  $G$  has a path cover  $\mathcal{P}$  with at most  $k$  paths.

Note that for  $k = 1$ , CYCLE COVER is HAMILTONIAN CYCLE and PATH COVER is HAMILTONIAN PATH.

The main results of this paper are the following theorems about kernelization of CYCLE COVER and PATH COVER. In both theorems we assume that a proper  $H$ -representation of the input graph  $G$  is given.

**THEOREM 1.** *PATH COVER admits a kernel of size  $\mathcal{O}(h^8)$ , where  $h$  is the size of*

the graph  $H$  in a proper  $H$ -representation of the input graph  $G$ .

For CYCLE COVER we succeed only in constructing a polynomial compression of explicitly given size. (Roughly speaking, the difference between kernelization and compression is that a kernelization algorithm outputs an equivalent instance of the same parameterized problem, while a compression algorithm maps an instance of a parameterized problem to an equivalent instance of another nonparameterized problem. We refer the reader to section 2 with preliminaries, where we define kernelization and compressing algorithms.) Note that since we compress into an NP-complete problem, the standard trick involving the Cook–Levin theorem (see, e.g., [22]) implies the existence of a polynomial-in- $h$  kernel for CYCLE COVER, but we are unable to give the exact size of such a kernel.

**THEOREM 2.** *CYCLE COVER admits a compression of size  $\mathcal{O}(h^{10})$ , where  $h$  is the size of the graph  $H$  in a proper  $H$ -representation of the input graph  $G$ .*

However, for the special case of  $k = 1$ , namely, HAMILTONIAN CYCLE, we also are able to obtain a kernel of size  $\mathcal{O}(h^8)$ .

It is not clear whether the requirement that a proper  $H$ -representation is given in the input of the considered problems on proper  $H$ -graphs could be avoided. The hardness result of Chaplick et al. [12] can be extended for proper  $H$ -graphs, and it can be shown that the recognition problem for proper  $H$ -graphs is NP-hard even for small fixed graphs  $H$ . This indicates that the aforementioned requirement may be unavoidable.

**Organization of the paper.** The remaining part of the paper is organized as follows. In section 2, we introduce the notions and notation used throughout the paper. In section 3, we give an informal description of our kernelization and compression algorithms before going into detail. In section 4, we obtain intermediate kernelization results for the aforementioned problems parameterized by the size of a *clique cover*, that is, by the size of a family of pairwise disjoint cliques that cover the vertices of the input graph. In section 5, we introduce additional notation for proper  $H$ -graphs and obtain a number of structural and auxiliary algorithmic results about path and cycle covers in  $H$ -graphs. In section 6, we show that HAMILTONIAN CYCLE and PATH COVER on proper  $H$ -graphs admit polynomial kernels, and we prove that CYCLE COVER admits a polynomial compression. We conclude the paper in section 7 by stating some open problems.

**2. Preliminaries.**

**Graphs.** All graphs considered in this paper are assumed to be finite and simple, that is, finite undirected graphs without loops or multiple edges, unless it is said explicitly that we consider a multigraph. In this respect, the following basic definitions are given for simple graphs. For each of the graph problems considered in this paper, we let  $n = |V(G)|$  and  $m = |E(G)|$  denote the number of vertices and edges, respectively, of the input graph  $G$  if it does not create confusion;  $\|G\| = |E(G)|$  is the *size* of  $G$ . For a graph  $G$  and a subset  $X \subseteq V(G)$  of vertices, we write  $G[X]$  to denote the subgraph of  $G$  induced by  $X$ ; for  $X = \{x_1, \dots, x_k\}$ , we also write  $G[x_1, \dots, x_k]$  instead of  $G[\{x_1, \dots, x_k\}]$ . We write  $G - X$  to denote the subgraph of  $G$  induced by  $V(G) \setminus X$ , and we write  $G - u$  instead of  $G - \{u\}$  for a single element set. Similarly, for an edge  $e$ ,  $G - e$  denotes the graph obtained from  $G$  by the deletion of  $e$ . For a vertex  $v$ , we denote by  $N_G(v)$  the (*open*) *neighborhood* of  $v$ , i.e., the set of vertices that are adjacent to  $v$  in  $G$ . The *closed neighborhood*  $N_G[v]$  is  $N_G(v) \cup \{v\}$ . For a set

of vertices  $X \subseteq V(G)$ ,  $N_G[X]$  is  $\bigcup_{v \in X} N_G[v]$  and  $N_G(X)$  is  $N_G[X] \setminus X$ . The *degree* of a vertex  $v$  is  $d_G(v) = |N_G(v)|$ . A *matching* is a set of edges with pairwise distinct end-vertices. A *vertex cover* of a graph is a set of vertices such that every edge of the graph is incident to a vertex of the set.

A set of vertices  $X$  is said to be *connected* if  $G[X]$  is a connected graph. It is said that a set of vertices  $X \subset V(G)$  is a *cut-set* of a graph  $G$  if  $G - X$  has more connected components than  $G$ ; a vertex  $u$  is a *cut-vertex* if  $\{u\}$  is a cut-set. A connected graph  $G$  is *2-connected* if it has no cut-vertex. A *block*  $B$  of a connected graph  $G$  is a maximal subgraph which does not contain a cut-vertex, and we use  $n_B$  to denote the number of vertices in  $B$ . Clearly, a block of a connected graph with at least one edge is either a single edge (i.e., *trivial*) or is 2-connected with at least three vertices (i.e., *non-trivial*), and we use  $\text{bl}(G)$  and  $\text{tbl}(G)$  to respectively denote the number of nontrivial and trivial blocks of  $G$ . For a connected graph  $G$ , the *block-cutpoint decomposition*  $BC(G)$  of  $G$  is a bipartite graph consisting of one node  $x_B$  for each block  $B$  of  $G$  and one node  $x_v$  for each cut-vertex  $v$  of  $G$  such that  $x_v$  is connected to  $x_B$  if and only if  $v$  is a vertex of  $B$ . It is well known that  $BC(G)$  is a tree and that it can be constructed in linear time [26].

A *path*  $P$  in a graph  $G$  is a connected subgraph whose vertices except at most two of them, called its *end-vertices*, have degree two and the end-vertices have degree one or zero if  $P$  is a single-vertex graph (called *trivial*). We refer to the degree two vertices of a path as *internal*. We write  $P = v_1 \cdots v_s$  to denote the path with the vertices  $v_1, \dots, v_s$  such that  $v_{i-1}v_i \in E(P)$  for  $i \in \{2, \dots, s\}$ . Clearly,  $v_s \cdots v_1$  denotes the same path, but we use  $P^{-1}$  to denote this sequence if we need to be specific. Then  $v_1$  and  $v_s$  are the end-vertices of the path and  $v_2, \dots, v_{s-1}$  are *internal* vertices. A path  $P$  with end-vertices  $u$  and  $v$  is called a  $(u, v)$ -*path*. A *subpath* is a connected subgraph of a path  $P$ ; a subpath is *proper* if it is distinct from  $P$ . A *cycle*  $C$  in a graph  $G$  is a connected subgraph where each vertex has degree two. We write  $C = v_0 \cdots v_s$  to denote that  $C$  is the cycle consisting of distinct vertices  $v_1, \dots, v_s$  such that  $v_0 = v_s$  and  $v_{i-1}v_i \in E(C)$  for  $i \in \{2, \dots, s\}$ . Clearly,  $v_i \cdots v_s v_1 \cdots v_i$  for  $i \in \{1, \dots, s\}$  denotes the same cycle, and we can reverse the ordering. A proper connected subgraph of a cycle  $C$  is called a *segment*. Trivially, a segment is a path. Observe that a cycle has two segments with the same pair of end-vertices. To distinguish such a pair of segments for  $C = v_0 \cdots v_s$ , we use the following convention: For  $i, j \in \{1, \dots, s\}$  such that  $i < j$ , we say that the paths  $v_i \cdots v_j$  and  $v_i \cdots v_s v_1 \cdots v_j$  are the  $(v_i, v_j)$ -*segment* and  $(v_j, v_i)$ -*segment* of  $C$ , respectively. A path  $P = v_1 \cdots v_s$  (a cycle  $C = v_0 \cdots v_s$ ) in  $G$  is *Hamiltonian* if  $\{v_1, \dots, v_s\} = V(G)$ . A family of paths  $\mathcal{P} = \{P_1, \dots, P_k\}$  (a family of cycles  $\mathcal{C} = \{C_1, \dots, C_k\}$ ) is a *path cover* (*cycle cover*, respectively) if the paths (cycles, respectively) are pairwise disjoint and the union of their vertices is  $V(G)$ . The *size* of a path or cycle cover is the number of paths or cycles in it.

A *clique* in a graph  $G$  is a set of pairwise adjacent vertices. A family  $\mathcal{Q} = \{Q_1, \dots, Q_s\}$  of cliques is said to be a (*vertex*) *clique cover* if the cliques are pairwise disjoint and  $\bigcup_{i=1}^s Q_i = V(G)$ . Note that we consider only vertex clique covers in our paper.

Let  $\mathcal{S}$  be a collection of sets. The *intersection graph* of  $\mathcal{S}$  has  $\mathcal{S}$  as its vertex set and two distinct vertices  $X, Y \in \mathcal{S}$  are adjacent if and only if  $X \cap Y \neq \emptyset$ . For an intersection graph  $G$ ,  $\mathcal{S}$  is called an (*intersection*) *model* of  $G$ . The intersection graph of a family of intervals of the real line is called an *interval graph*; it is also said that  $G$  is an interval graph if there is a family of intervals (called *interval model* or *representation*) such that  $G$  is isomorphic to the intersection graph of this family. Throughout the paper it is assumed that the intervals of an interval model are closed.

An interval graph is *proper* if it has an interval model such that no interval is a subinterval of another one.

Let  $H$  be a multigraph. As for simple graphs,  $\|H\| = |E(H)|$  is the *size* of  $H$ . We say that  $H'$  is obtained from  $H$  by the *subdivision* of an edge  $e = xy$  if, to construct  $H'$ , we delete  $e$ , add a new vertex  $z$ , and introduce two new edges  $zx$  and  $zy$ . Similarly,  $H'$  is a *subdivision* of  $H$  if  $H'$  is obtained from  $H$  by consecutively subdividing its edges. The *dissolution* of a vertex  $z$  that is incident to exactly two edges  $zx$  and  $zy$  for  $z \neq x, y$  (note that it could happen that  $x = y$ ) is the opposite operation, that is, we delete  $z$  and add a new edge  $xy$ .

For a multigraph  $H$ , a simple graph  $G$  is an  $H$ -*graph* if  $G$  is an intersection graph of connected subgraphs of some subdivision  $H'$  of  $H$  or, equivalently,  $G$  is an intersection graph of connected subsets of vertices of  $H$ . Throughout the paper we only allow the  $H$ 's in  $H$ -graphs to be multigraphs and all other graphs are assumed to be simple. To distinguish the vertices of  $H$  and  $H'$  from the vertices of  $G$ , we refer to the vertices of  $H$  and  $H'$  as *nodes*. We also say the nodes of  $H$  are *branching* nodes of  $H'$  and the other nodes are *subdivision* nodes. A pair  $(H', \mathcal{M})$ , where  $\mathcal{M} = \{M_v\}_{v \in V(G)}$  is a collection of connected vertex sets of  $H'$  such that  $G$  is the intersection graph of  $\mathcal{M}$ , is called an  $H$ -*representation* of  $G$ . A representation  $(H', \mathcal{M})$  is *proper* if, for every two distinct  $u, v \in V(G)$ , neither  $M_u \subseteq M_v$  nor  $M_v \subseteq M_u$ . In this sense,  $G$  is a *proper*  $H$ -graph if it has a proper  $H$ -representation. It is straightforward to see that interval graphs are precisely the  $K_2$ -graphs and the proper interval graphs are the proper  $K_2$ -graphs.

Note that every graph has the following trivial representation. For a graph  $G$ , let  $I(G)$  denote the *incidence* graph of  $G$ , that is, the graph obtained by subdividing each edge of  $G$  exactly once.

OBSERVATION 1. *Every graph  $G$  is a proper  $G$ -graph. Its trivial proper  $G$ -representation is  $(I(G), \{N_{I(G)}[v]\}_{v \in V(G)})$ .*

**Parameterized complexity and kernelization.** We refer the reader to the books [15, 18, 22] for a detailed introduction to the field. Here we only briefly review the basic notions.

Parameterized complexity is a two-dimensional framework for studying the computational complexity of a problem. One dimension is the input size  $n$  and the other is a *parameter*  $k$  associated with the input. The main goal is to confine the combinatorial explosion in the running time of an algorithm for a (typically, NP-hard) problem to depend only on  $k$ . In this sense, a parameterized problem is said to be *fixed parameter tractable* (or FPT) if it can be solved in time  $f(k) \cdot n^{\mathcal{O}(1)}$  for some function  $f$ .

A *compression* of a parameterized problem  $\Pi_1$  into a (nonparameterized) problem  $\Pi_2$  is a polynomial-time algorithm that maps each instance  $(I, k)$  of  $\Pi_1$  with the input  $I$  and the parameter  $k$  to an instance  $I'$  of  $\Pi_2$  such that

- (i)  $(I, k)$  is a yes-instance of  $\Pi_1$  if and only if  $I'$  is a yes-instance of  $\Pi_2$ , and
- (ii)  $|I'|$  is bounded by  $f(k)$  for a computable function  $f$ .

The output  $I'$  is also called a *compression*. The function  $f$  is said to be the *size* of the compression. A compression is *polynomial* if  $f$  is polynomial. A *kernelization* algorithm for a parameterized problem  $\Pi$  is a polynomial-time algorithm that maps each instance  $(I, k)$  of  $\Pi$  to an instance  $(I', k')$  of  $\Pi$  such that

- (i)  $(I, k)$  is a yes-instance of  $\Pi$  if and only if  $(I', k')$  is a yes-instance of  $\Pi$ , and
- (ii)  $|I'| + k'$  is bounded by  $f(k)$  for a computable function  $f$ .

Respectively,  $(I', k')$  is a *kernel* and  $f$  is its *size*. A kernel is *polynomial* if  $f$  is polynomial. While it can be shown that every decidable parameterized problem is

FPT if and only if it admits a kernel, it is unlikely that every problem in FPT has a polynomial kernel (see, e.g., [15, 22] for details).

For CYCLE COVER, we show that it admits a polynomial compression into a special problem called PRIZE COLLECTING CYCLE COVER which we define here.

Let  $G$  be a graph and let  $\omega: E(G) \rightarrow \mathbb{N}_0$  be a *weight* function; note that we allow zero weights. For a cycle  $C$ ,  $\omega(C)$  is the sum of the weights of its edges. Let  $\alpha: \mathbb{N} \rightarrow \mathbb{N}$  be a nondecreasing *penalty* function. For a cycle cover  $\mathcal{C} = \{C_1, \dots, C_k\}$  of  $G$ , the weight of  $\mathcal{C}$  is  $\omega(\mathcal{C}) = \sum_{i=1}^k \omega(C_i)$  and the *cost* of  $\mathcal{C}$  is  $c_{\alpha, \omega}(\mathcal{C}) = \omega(\mathcal{C}) - \alpha(|\mathcal{C}|)$ . Observe that the cost may be negative.

PRIZE COLLECTING CYCLE COVER

*Input:* A graph  $G$  with a weight function  $\omega: E(G) \rightarrow \mathbb{N}_0$ , a nondecreasing penalty function  $\alpha: \{1, \dots, |V(G)|\} \rightarrow \mathbb{N}$ , and an integer  $r$ .

*Task:* Decide whether  $G$  has a cycle cover  $\mathcal{C}$  of cost  $c_{\alpha, \omega}(\mathcal{C}) \geq r$ .

Notice that if  $G$  is a graph with zero edge-weights and the penalty function  $\alpha(x) = x$  for  $x \in \mathbb{N}$ , then  $G$  has a cycle cover with at most  $k$  cycles if and only if  $G$  has a cycle cover of cost at least  $r = -k$ , that is, PRIZE COLLECTING CYCLE COVER generalizes CYCLE COVER. We prove that CYCLE COVER admits a polynomial compression to PRIZE COLLECTING CYCLE COVER of size  $\mathcal{O}(h^{10})$  when parameterized by the size  $h$  of  $H$  if a proper  $H$ -representation is given in the input.

**3. Description of algorithms.** In this section, we give an informal high-level description of our kernelization and compression algorithms to both outline how they work and provide the general flow of the subsequent technical sections of the paper.

Our first step towards the kernelization of PATH COVER and compression of CYCLE COVER is a kernelization algorithm for CYCLE COVER, PATH COVER, and PRIZE COLLECTING CYCLE COVER being parameterized by the size of a clique cover of the input graph. These results are of independent interest. The parameterization of HAMILTONIAN CYCLE by the clique cover size was considered by Lampis et al. [32], who proved that the problem is FPT for this parameterization. We extend their result by showing the following theorem.

**THEOREM 3.** *CYCLE COVER, PATH COVER, and HAMILTONIAN CYCLE admit kernels of size  $\mathcal{O}(s^8)$ , where  $s$  is the size of a clique cover. PRIZE COLLECTING CYCLE COVER admits a kernel of size  $\mathcal{O}((s + \ell)^{10})$ , where  $s$  is the size of a clique cover and  $\ell$  is the number of edges of the input graph with nonzero weights. In all kernels we assume that a clique cover of size  $s$  is given in the input.*

We sketch the main ideas of the kernelization for CYCLE COVER, which is the easiest among these problems, and then explain how to modify the algorithm for the other problems.

Recall that a clique cover is a collection  $\mathcal{Q} = \{Q_1, \dots, Q_s\}$  of pairwise disjoint cliques such that  $V(G) = \bigcup_{i=1}^s Q_i$ . First, we show that there is always an optimal solution to CYCLE COVER with very specific properties. We call a cycle cover  $\mathcal{C}$  *regular* with respect to  $\mathcal{Q}$  if for every distinct  $i, j \in \{1, \dots, s\}$ ,

- (i) at most one cycle of  $\mathcal{C}$  has an edge between  $Q_i$  and  $Q_j$ ,
- (ii) the number of edges between  $Q_i$  and  $Q_j$  in each of the cycles of  $\mathcal{C}$  is at most 2.

It is possible to prove that every cycle cover can be transformed into a regular one without increasing its size. Informally, if two distinct cycles have edges between two

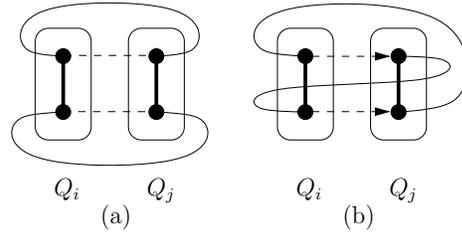


FIG. 1. Rerouting cycles; the deleted edges are shown by dashed lines, and the added edges are shown by thick lines.

cliques  $Q_i$  and  $Q_j$ , we can “glue” them together (as shown in Figure 1(a)), and if a cycle has at least three edges between the cliques, then we can pick two of them that are in the “same direction” according to an arbitrary orientation of the cycle and reroute the cycle (see Figure 1(b)).

Because the cycles of a regular cycle cover have a limited number of edges that are between the cliques of  $\mathcal{Q}$ , it is possible to modify and/or reroute them using the fact that the vertices of the same clique are pairwise adjacent. The regularity of a cycle cover allows us to apply the following reduction rules.

REDUCTION RULE 3.1.

- If there is a clique  $Q_i \in \mathcal{Q}$  and  $v \in Q_i$  such that  $N_G[v] = Q_i$  and  $|Q_i| \geq s + 3$ , then set  $G = G - v$  and  $Q_i = Q_i \setminus \{v\}$ .
- If there are distinct  $i, j \in \{1, \dots, s\}$  such that the bipartite graph  $G_{ij}$ , whose vertex set is  $Q_i \cup Q_j$  and whose edges are the edges of  $G$  between  $Q_i$  and  $Q_j$ , has a matching  $M$  of size at least  $4s - 3$ , then select (arbitrarily) an edge  $e \in M$ , set  $G = G - e$ .
- If there is a clique  $Q_i \in \mathcal{Q}$  and  $v \in V(G) \setminus Q_i$  such that  $|N_G(v) \cap Q_i| \geq 2s + 1$ , then for an arbitrary edge  $e = uv$  with  $u \in Q_i$ , set  $G = G - e$ .

The first item in Reduction Rule 3.1 asserts that if a sufficiently large clique of the clique cover has a simplicial vertex, then this vertex is irrelevant. Similarly, if there is a large matching between two cliques, then one edge of this matching can be deleted safely. Finally, if there is a vertex outside a clique which sees many vertices of the clique, any edge between this vertex and a vertex of the clique can be removed. We apply the rules exhaustively. We prove that any irreducible instance has  $\mathcal{O}(s^4)$  vertices, that is, the size of the obtained instance of CYCLE COVER is  $\mathcal{O}(s^8)$ , and this implies the claim of Theorem 3 for the problem.

As HAMILTONIAN CYCLE is the special case of CYCLE COVER when  $k = 1$  and Reduction Rule 3.1 does not modify  $k$ , the kernelization algorithm for HAMILTONIAN CYCLE is the same. For PATH COVER, we need a tiny adjustment to reroute the paths of a path cover in a slightly different way. However, PRIZE COLLECTING CYCLE COVER requires additional work.

Let  $(G, \omega, \alpha, r)$  be an instance of PRIZE COLLECTING CYCLE COVER and let  $S$  be the set of edges of  $G$  with nonzero weights,  $\ell = |S|$ . First, we modify the clique cover  $\mathcal{Q}$  of  $G$  by making the end-vertices of the edges of  $S$  trivial cliques of size one. Thus, we obtain the clique cover  $\hat{\mathcal{Q}}$  of size  $t \leq s + 2\ell$ . Then we observe that the modifications of the cycles of a cycle cover that were used for CYCLE COVER never affect edges of  $G$  with both end-vertices in trivial cliques. In particular, if the cycles of a cycle cover contain  $e \in S$ , one of the cycles of the cycle cover obtained by the

reroutings still contains  $e$ . Also, we do not increase the number of cycles in cycle covers by such reroutings. This implies that we still can use Reduction Rule 3.1. It is possible to show that an irreducible instance of PRIZE COLLECTING CYCLE COVER obtained by the exhaustive application of the rules has  $\mathcal{O}(t^4)$  vertices.

Note that this is not a polynomial kernel yet because we still have to compress the edge-weights as well as the values of the penalty function  $\alpha$  and  $r$ . For this, we apply the approach proposed by Etscheid et al. [19] for constructing kernels for weighted problems. These techniques are based on the classical algorithm for compressing numbers given by Frank and Tardos in [23]. This allows us to encode the value of the weight function for each  $e \in S$  and the value of the penalty function for each  $i = \mathcal{O}(t^4)$  by a binary string of length  $\mathcal{O}((s+\ell)^6)$ . Summarizing, we obtain an instance of PRIZE COLLECTING CYCLE COVER of size  $\mathcal{O}((s+\ell)^{10})$ . This completes the sketch of the kernelization algorithm.

Note that Theorem 3 requires that a clique cover of the input graph is given. This seems to be unavoidable as it is already NP-complete to decide whether a graph has a clique cover of size 3 [24] (the problem is equivalent to 3-COLORING for the complement of the graph).

**From proper  $H$ -models to small clique covers.** Now we use Theorem 3 to construct kernelization and compression algorithms for PATH COVER and CYCLE COVER on proper  $H$ -graphs, i.e., we build equivalent instances with small clique covers by using a given proper  $H$ -representation.

Suppose that  $G$  is a proper  $H$ -graph given together with its proper  $H$ -representation  $(H', \mathcal{M})$ . Notice that for every node  $x \in V(H')$ , the set  $K_x = \{v \in V(G) \mid x \in M_v\}$  is a clique of  $G$ . Observe also that the graph  $G - \bigcup_{x \in V(H')} K_x$  can be seen as a union of proper interval graphs  $G_e$  corresponding to the edges  $e \in E(H)$ . More formally, let  $e = xy \in E(H)$  and consider the  $(x, y)$ -path  $P$  in  $H'$  obtained from  $e$  by the subdivisions. We denote by  $G_e$  the subgraph of  $G$  induced by  $V_e = \{v \in V(G) \mid M_v \subseteq V(P_e) \setminus \{x, y\}\}$ . Clearly,  $G_e$  is a proper interval graph and the sets  $M_v$  for  $v \in V_e$  form a proper interval representation of it. This representation defines a corresponding total ordering of its vertices (see [35]) by choosing a direction of  $e$  and reading the left endpoints of the individual intervals from left to right. We assume that these orderings are fixed for every  $G_e$ . In particular, whenever we speak about leftmost and rightmost vertices of  $G_e$ , we mean the leftmost and rightmost vertices with respect to this ordering. Notice that for  $e = xy$ ,  $N_G(V_e) \subseteq K_x \cup K_y$ , that is, paths or cycles that cover the vertices in  $G_e$  are either completely in  $G_e$  or enter  $G_e$  via the vertices of  $K_x$  or  $K_y$  that we call the *left* and *right* cliques, respectively.

The graphs  $G_e$  could be huge but, since they are proper interval graphs, they have a relatively simple structure. We exploit this structure in order to replace them by small gadget graphs while maintaining the equivalence of the instances of the considered problems. Since the vertices of  $\bigcup_{x \in V(H')} K_x$  can be covered by at most  $|V(H)|$  cliques and the set of vertices of each gadget replacing  $G_e$  can be covered by a constant number of cliques, we obtain a graph that has a clique cover of size  $\mathcal{O}(|V(H)| + |E(H)|)$ .

To simplify the arguments, we show that we can assume that the considered  $H$ -representation  $(H', \mathcal{M})$  of  $G$  has no redundancies, that is, for every node  $x \in V(H')$ , there is a vertex  $v \in V(G)$  with  $x \in M_v$  and, moreover, for every edge  $xy \in E(H')$ , there is  $v \in V(G)$  with  $x, y \in M_v$ . We call such a representation *nice*. To achieve this niceness, we first observe that if the input graph  $G$  has a component  $F$  that is a proper interval graph, we can find the minimum number of paths or cycles that cover

$F$  depending on the considered problem, and then delete  $F$  and modify the parameter  $k$  of PATH COVER or CYCLE COVER, respectively. Somewhat surprisingly, to the best of our knowledge CYCLE COVER was not studied on proper interval graphs. Therefore, we design a linear time algorithm for the problem. Note that it may happen that we solve the problem by applying the reduction rules. Otherwise, we obtain an induced subgraph  $G'$  of  $G$  such that every component of  $G'$  has a vertex  $v$  with  $M_v$  containing a branching node of  $H'$ . Then we modify  $H'$  by removing irrelevant nodes and edges. This procedure can create new nodes of degree one from some subdivision nodes of  $H'$ , but the number of such vertices is at most  $2|E(H)|$ . From this, we derive that  $G'$  is an  $\hat{H}$ -graph for some  $\hat{H}$  with at most  $3|E(H)|$  nodes and at most  $2|E(H)|$  edges, and we construct the corresponding nice proper  $\hat{H}$ -representation.

From now on we can concentrate only on nice representations. In particular, we assume that every graph  $G_e$  for  $e = xy \in E(H)$  is connected and that the leftmost and the rightmost vertices of  $G_e$  have neighbors in the left and the right cliques, respectively.

Recall that for PATH COVER, we prove the following theorem.

**THEOREM 1.** PATH COVER admits a kernel of size  $\mathcal{O}(h^8)$ , where  $h$  is the size of the graph  $H$  in a proper  $H$ -representation of the input graph  $G$ .

Let  $G$  be a proper  $H$ -graph given together with its nice proper representation  $(H', \mathcal{M})$ . Let  $\mathcal{P}$  be a path cover of  $G$ . For  $e \in E(H)$ , let  $\mathcal{P}_e$  denote the family of inclusion-maximal subpaths of the paths  $P \in \mathcal{P}$  with all their vertices in  $V_e$ . In other words,  $\mathcal{P}_e$  contains the subpaths of every path  $P$  obtained by the deletion of the vertices that are outside  $V_e$ . We say that  $\mathcal{P}_e$  is the *projection* of  $\mathcal{P}$  on  $G_e$ . Since  $\mathcal{P}$  is a path cover of  $G$ ,  $\mathcal{P}_e$  is a path cover of  $G_e$ . It is possible to show that if  $G$  has a path cover of size at most  $k$ , then  $G$  has a path cover of size at most  $k$  such that the paths in each projection  $\mathcal{P}_e$  have a very special structure in the case when the vertices of the graph  $G_e$  cannot be covered by two cliques. We call such a cover *tamed* (this is a slightly simplified definition which we use only for the high-level description of the algorithm). We prove the following properties of  $\mathcal{P}_e$ .

- If  $G_e$  is 2-connected, then either
  - $\mathcal{P}_e$  consists of one Hamiltonian path of  $G_e$  such that its end-vertices are the two leftmost vertices of  $G_e$  (symmetrically, the two rightmost vertices), or
  - $\mathcal{P}_e$  consists of two paths such that each of them has one of its end-vertices among the two leftmost vertices of  $G_e$  and the second end-vertex is among the two rightmost vertices of  $G_e$ , and these paths are proper subpaths of the same path  $P$  of  $\mathcal{P}$  that occur in  $P$  in “opposite directions” for an arbitrary orientation of  $P$ .
- If  $G_e$  has a cut-vertex, then  $\mathcal{P}_e$  consists of two paths such that one of them has its end-vertices in the two leftmost vertices or just in the leftmost vertex if the path is trivial, and the second path behaves symmetrically.

The structure of paths in the projection of a tamed path cover is shown in Figure 2; the vertices of  $G_e$  are denoted by  $v_1^e, \dots, v_{p(e)}^e$  in the figure according to their proper interval ordering. Note that every path of  $\mathcal{P}$  that enters  $G_e$  uses the (one or two) leftmost and rightmost vertices as entry points.

We use this structural result for our kernelization. For each  $G_e$  that cannot be covered by two cliques, we analyze the possible structure of paths in  $\mathcal{P}_e$  for a tamed path cover  $\mathcal{P}_e$ . It appears that the types of paths in  $\mathcal{P}_e$  are defined by cut-vertices of  $G_e$  and the adjacencies of the second leftmost and the second rightmost vertices of

$G_e$  to the corresponding left and right cliques (if, say, the second leftmost vertex is not adjacent to the left clique, then the leftmost vertex “cuts” in a special sense this clique from the remaining part of  $G_e$ ). Then we replace  $G_e$  by a gadget from Figure 4 which has the same structure with respect to how they can be covered by a tamed path cover. Since each of these gadgets can be covered by at most two cliques, in the end we obtain an equivalent instance of PATH COVER such that the input graph can be covered by at most  $|V(H)| + 2|E(H)|$  cliques.

Then we can apply Theorem 3 where  $h \leq |V(H)| + 2|E(H)|$ . Notice that the kernelization from Theorem 3 can destroy the proper  $H$ -representation. Thus we have to be a bit careful here to specify the value of the parameter. We do it by using Observation 1 and output the trivial proper  $\hat{G}$ -representation for the obtained graph  $\hat{G}$ .

CYCLE COVER is more complicated. While the general idea follows the one for PATH COVER, there are several nontrivial differences, which we underline below. We first recall the statement of the main result for CYCLE COVER.

**THEOREM 2.** *CYCLE COVER admits a compression of size  $\mathcal{O}(h^{10})$ , where  $h$  is the size of the graph  $H$  in a proper  $H$ -representation of the input graph  $G$ .*

Let  $G$  be a proper  $H$ -graph given together with its nice proper representation  $(H', \mathcal{M})$ . Let  $\mathcal{C}$  be a cycle cover of  $G$ . Similarly to path covers, for each  $e \in E(H)$ , we define the *projection*  $\mathcal{C}_e$  of  $\mathcal{C}$  on  $G_e$  that is now a family of paths and cycles of  $G_e$ . We show that it suffices to only consider cycle covers of a special structure that again are called *tamed*; the structure of paths and cycles in the projection of a tamed cycle cover can be seen in Figure 3. Note that the crucial difference between projections of tamed path and cycle covers is that the number of elements of the projection of a tamed cycle cover is not bounded by any constant. In particular, if  $G$  has at least three blocks, then either  $\mathcal{C}_e$  contains a Hamiltonian path with its end-vertices in the leftmost and the rightmost vertices of  $G_e$  or each nontrivial middle block should contain a cycle of  $\mathcal{C}_e$ . This implies that we cannot replace  $G_e$  by a gadget which both has the same number of cycles as the original projection and can be covered by cliques whose number is any function of the size of  $H$ .

To deal with this situation we introduce weights that encode the number of cycles that we need to cover  $G_e$  if we do not use a Hamiltonian path between the leftmost and the rightmost vertices. For each  $G_e$ , we construct a gadget with at most three edges of positive weight. The remaining edges of the considered graph receive zero weights. To give a rough idea how this works, we observe that the nonzero weights are assigned to the edges of a gadget in such a way that (i) there is a Hamiltonian path between the leftmost and rightmost vertices that contains all these edges, and (ii) for any cycle cover whose projection has no such path, the cycles of the cover miss some edges of nonzero weights. The simplest way to achieve this property is to use bridges in the replacement gadgets for the assignment of nonzero weights, but this is not always possible and we have to use also more complicated gadgets. We replace the leftmost (rightmost) block by a copy of  $K_5$  if it has size at least 6 and leave it intact otherwise. The replacement gadgets are attached to the graph by the two leftmost (rightmost) vertices of  $G_e$  and the unique cut-vertices of the corresponding blocks. The same replacement is done for the middle part if  $G_e$  has a unique middle block. If  $G_e$  has at least two middle blocks, we replace them by one of the graphs  $F_1$ – $F_{11}$  shown in Figure 8; the edges with nonzero weights are shown by thick lines, and the gadgets are attached via vertices  $s$  and  $t$ .

This way we construct an instance of PRIZE COLLECTING CYCLE COVER where at most  $3|E(H)|$  edges have nonzero weights. Then we apply Theorem 3.

Notice that for HAMILTONIAN CYCLE, we have no such difficulties, because we are looking for a single cycle. This allows us to construct a kernel of size  $\mathcal{O}(h^8)$ .

**4. Parameterization by the size of a clique cover.** In this section we show that CYCLE COVER and PATH COVER admit a polynomial kernel when parameterized by the size  $s$  of a clique cover if such a cover is given. For PRIZE COLLECTING CYCLE COVER, we show that it admits a polynomial kernel when parameterized by  $s$  and the number of edges of the input graph with nonzero weights. We need some auxiliary terminology.

Let  $\mathcal{Q} = \{Q_1, \dots, Q_s\}$  be a clique cover of  $G$ . We say that a clique  $Q_i$  is *trivial* if  $|Q_i| = 1$ . An edge  $e \in E(G)$  is *trivial* if both of its end-vertices are in trivial cliques. For distinct  $i, j \in \{1, \dots, s\}$ , we say that an edge  $uv \in E(G)$  is *between*  $Q_i$  and  $Q_j$  if  $u \in Q_i$  and  $v \in Q_j$ , or vice versa. Let  $\mathcal{P}$  be a path cover ( $\mathcal{C}$  be a cycle cover of  $G$ , respectively). It is said that  $\mathcal{P}$  ( $\mathcal{C}$ , respectively) is *regular* with respect to  $\mathcal{Q}$  if for every distinct  $i, j \in \{1, \dots, s\}$ ,

- (i) at most one path of  $\mathcal{P}$  (cycle of  $\mathcal{C}$ , respectively) has an edge between  $Q_i$  and  $Q_j$ ,
- (ii) the number of edges between  $Q_i$  and  $Q_j$  in each of the paths of  $\mathcal{P}$  (cycles of  $\mathcal{C}$ , respectively) is at most 2.

LEMMA 1. *Let  $\mathcal{Q} = \{Q_1, \dots, Q_s\}$  be a clique cover of  $G$  and let  $S$  be a set of trivial edges of  $G$  with respect to  $\mathcal{Q}$ . If  $G$  has a path cover (cycle cover, respectively) of size at most  $k$  whose paths (cycles, respectively) contain the edges of  $S$ , then  $G$  has a regular path cover (cycle cover, respectively) with respect to  $\mathcal{Q}$  of size at most  $k$  with paths (cycles, respectively) containing the edges of  $S$ .*

*Proof.* Let  $\mathcal{P} = \{P_1, \dots, P_k\}$  ( $\mathcal{C} = \{C_1, \dots, C_k\}$ , respectively) be a path cover (cycle cover, respectively) of  $G$  with the paths (cycles, respectively) containing the edges of  $S$  such that the number of edges in the paths of  $\mathcal{P}$  (cycles of  $\mathcal{C}$ , respectively) that are between the cliques of  $\mathcal{Q}$  is minimum. We show that  $\mathcal{P}$  ( $\mathcal{C}$ , respectively) is regular.

First, we prove the claim for  $\mathcal{P}$ . To obtain a contradiction, assume that  $\mathcal{P}$  is not regular. Assume that condition (i) does not hold, that is, there are two distinct paths  $P_i = u_1 \cdots u_p$  and  $P_{i'} = v_1 \cdots v_q$  in  $\mathcal{P}$  that contain edges between two distinct cliques  $Q_j$  and  $Q_{j'}$  of  $\mathcal{Q}$ . Let  $u_{t-1}u_t$  and  $v_{t'-1}v_{t'}$  be edges of  $P_i$  and  $P_{i'}$ , respectively, that are between  $Q_j$  and  $Q_{j'}$ . By symmetry, assume without loss of generality that  $u_{t-1}, v_{t'-1} \in Q_j$  and  $u_t, v_{t'} \in Q_{j'}$ . Clearly,  $u_{t-1}v_{t'-1}, u_tv_{t'} \in E(G)$ . We replace  $P_i$  and  $P_{i'}$  by two new paths,  $\hat{P}_i = u_1 \cdots u_{t-1}v_{t'-1} \cdots v_1$  and  $\hat{P}_{i'} = u_p \cdots u_tv_{t'} \cdots v_q$ . Note that we replaced the edges  $u_{t-1}u_t, v_{t'-1}v_{t'}$  that are between  $Q_j$  and  $Q_{j'}$  by two edges with their end-vertices in  $Q_j$  and  $Q_{j'}$ , respectively. Since these paths cover the same vertices as  $P_i$  and  $P_{i'}$ ,  $\hat{\mathcal{P}} = (\mathcal{P} \setminus \{P_i, P_{i'}\}) \cup \{\hat{P}_i, \hat{P}_{i'}\}$  is a path cover of  $G$  of the same size  $k$ . It is straightforward to see that the edges of  $S$  are in the paths of  $\hat{\mathcal{P}}$ . Since the number of edges that are between the cliques of  $\mathcal{Q}$  in  $\hat{\mathcal{P}}$  is less than in  $\mathcal{P}$ , we obtain a contradiction with the choice of  $\mathcal{P}$ .

Suppose now that condition (ii) is broken, that is, there is a path  $P = v_1 \cdots v_p$  in  $\mathcal{P}$  that contains at least three edges between some cliques  $Q_i$  and  $Q_j$ . Then  $P$  contains two edges that go between  $Q_i$  and  $Q_j$  in the “same direction,” that is, there are distinct edges  $v_{t-1}v_t$  and  $v_{t'-1}v_{t'}$  of  $P$  such that either  $v_{t-1}, v_{t'-1} \in Q_i$  and  $v_t, v_{t'} \in Q_j$  or  $v_{t-1}, v_{t'-1} \in Q_j$  and  $v_t, v_{t'} \in Q_i$ . Assume without loss of generality

that  $t < t'$  and it holds that  $v_{t-1}, v_{t'-1} \in Q_i$  and  $v_t, v_{t'} \in Q_j$ . We create the new path  $\hat{P} = v_1 \cdots v_{t-1} v_{t'-1} \cdots v_t v_{t'} \cdots v_p$ . Observe that we replaced  $v_{t-1} v_t$  and  $v_{t'-1} v_{t'}$  that are between  $Q_i$  and  $Q_j$  by two edges with the end-vertices in  $Q_i$  and  $Q_j$ , respectively. Since  $\hat{P}$  is the path with the same vertices as  $P$ ,  $\hat{\mathcal{P}} = (\mathcal{P} \setminus \{P\}) \cup \{\hat{P}\}$  is a path cover of  $G$  of size  $k$ . Clearly, the edges of  $S$  are not affected by the rerouting, that is,  $\hat{\mathcal{P}}$  contains the edges of  $S$ . Because the number of edges of  $\hat{\mathcal{P}}$  that are between the cliques of  $\mathcal{Q}$  is reduced by the rerouting of  $P$ , we obtain a contradiction with the choice of  $\mathcal{P}$ .

This completes the proof for  $\mathcal{P}$ . The proof of the claim for  $\mathcal{C}$  is similar. We again consider two cases corresponding to (i) and (ii) and reroute cycles to reduce the number of edges that are between the cliques. The rerouting is essentially the same, the only difference being that in the rerouting that was used to obtain a contradiction if condition (i) is not fulfilled, we obtain one cycle out of two.  $\square$

LEMMA 2. *Let  $\mathcal{Q} = \{Q_1, \dots, Q_s\}$  be a clique cover of a graph  $G$  and let  $S$  be a set of trivial edges of  $G$  with respect to  $\mathcal{Q}$ . The minimum size of a path cover (cycle cover, respectively) whose paths (cycles, respectively) contain the edges of  $S$  is at most  $s(s+1)/2$ .*

*Proof.* We show the lemma for path covers. For cycle covers, the proof is the same.

Let  $\mathcal{P}$  be a path cover whose paths contain the edges of  $S$  of minimum size. By Lemma 1, we can assume that  $\mathcal{P}$  is regular with respect to  $\mathcal{Q}$ . For every clique  $Q_i$  of  $\mathcal{Q}$ ,  $\mathcal{P}$  contains at most one path with all the vertices in  $Q_i$  by minimality. Hence,  $\mathcal{P}$  contains at most  $s$  paths with their vertices included in one clique. By regularity, the number of paths with edges between the cliques is at most  $\binom{s}{2}$ . Therefore, the total number of paths is at most  $\binom{s}{2} + s = s(s+1)/2$ .  $\square$

Since PRIZE COLLECTING CYCLE COVER is a weighted problem, we need special kernelization tools to deal with integers in the input. We follow the approach of Etscheid et al. [19] that is based on the algorithm for compressing numbers given by Frank and Tardos in [23]. We state the result of Frank and Tardos in the form given in [19].

LEMMA 3 (see [23]). *There is an algorithm that, given a vector  $w \in \mathbb{Q}^h$  and an integer  $N$ , in polynomial time finds a vector  $\bar{w} \in \mathbb{Z}^h$  with  $\|\bar{w}\|_\infty \leq 2^{4h^3} N^{h(h+2)}$  such that  $\text{sign}(w \cdot b) = \text{sign}(\bar{w} \cdot b)$  for all vectors  $b \in \mathbb{Z}^h$  with  $\|b\|_1 \leq N - 1$ .*

Now we are ready to prove the main result of the section.

THEOREM 3. CYCLE COVER, PATH COVER, and HAMILTONIAN CYCLE admit kernels of size  $\mathcal{O}(s^8)$ , where  $s$  is the size of a clique cover. PRIZE COLLECTING CYCLE COVER admits a kernel of size  $\mathcal{O}((s+\ell)^{10})$ , where  $s$  is the size of a clique cover and  $\ell$  is the number of edges of the input graph with nonzero weights. In all kernels we assume that a clique cover of size  $s$  is given in the input.

*Proof.* We construct a kernelization algorithm for PRIZE COLLECTING CYCLE COVER as this problem demands the most complicated analysis and then explain how the algorithm can be adapted for CYCLE COVER and PATH COVER.

Let  $(G, \omega, \alpha, r)$  be an instance of PRIZE COLLECTING CYCLE COVER and let  $\mathcal{Q}$  be a clique cover of  $G$  of size  $s$ . Denote by  $S$  the set of edges of  $G$  with nonzero weights and let  $\ell = |S|$ . We modify  $\mathcal{Q}$  by making the end-vertices of  $S$  trivial cliques, that is, we let

$$\hat{\mathcal{Q}} = \{Q \setminus X \mid Q \text{ is a clique of } \mathcal{Q}\} \cup \{x \mid x \in X\},$$

where  $X$  is the set of end-vertices of  $S$ . We assume that  $\hat{Q} = \{Q_1, \dots, Q_t\}$ . Note that  $t \leq s + 2\ell$  and every edge of  $S$  is a trivial edge with respect to  $\hat{Q}$ .

We say that a cycle cover  $\mathcal{C}$  of  $G$  of cost at most  $r$  is a *regular solution* if  $\mathcal{C}$  is regular with respect to  $\hat{Q}$ . We need the following property.

CLAIM 4.1. *If  $(G, \omega, \alpha, r)$  is a yes-instance of PRIZE COLLECTING CYCLE COVER, then it has a regular solution.*

*Proof of Claim 4.1.* To see this, consider a solution for the instance, that is, a cycle cover  $\mathcal{C}$  of cost at most  $r$ . Let  $S'$  be the set of edges of the cycles of  $\mathcal{C}$  with nonzero weights that are between the cliques of  $\hat{Q}$ . Clearly,  $S' \subseteq S$ . Then by Lemma 1, there is a regular cycle cover  $\hat{\mathcal{C}}$  whose size is at most the size of  $\mathcal{C}$  and whose cycles contain the edges of  $S'$ . Because  $\omega$  has nonnegative values and  $\alpha$  is nondecreasing, we have that

$$c_{\alpha, \omega}(\mathcal{C}) = \omega(S') - \alpha(|\mathcal{C}|) \leq c_{\alpha, \omega}(\hat{\mathcal{C}}),$$

that is,  $\hat{\mathcal{C}}$  is a regular solution. □

We say that a regular solution for  $(G, \omega, \alpha, r)$  is *minimal* if it is a regular solution with the minimum number of cycles. Then we have the following property.

CLAIM 4.2. *Let  $\mathcal{C} = \{C_1, \dots, C_h\}$  be a minimal regular solution. Then for every  $i \in \{1, \dots, t\}$ , at most one cycle of  $\mathcal{C}$  has edges with both end-vertices in  $Q_i$ , and the total number of vertices of  $Q_i$  that are in the cycles without edges with both end-vertices in  $Q_i$  is at most  $t - 1$ .*

*Proof of Claim 4.2.* To show the claim, assume that there are two distinct cycles  $C_j = u_0 \dots u_p$  and  $C_{j'} = v_0 \dots v_{p'}$  such that  $u_{q-1}, u_q, v_{q'-1}, v_{q'} \in Q_i$  for some  $i \in \{1, \dots, t\}$  and  $q \in \{1, \dots, p\}$  and  $q' \in \{1, \dots, p'\}$ . Denote by  $R$  the  $(v_{q'}, v_{q'-1})$ -segment of  $C_{j'}$ . Then we can replace  $C_j$  and  $C_{j'}$  in  $\mathcal{C}$  by the single cycle  $C' = u_0 \dots u_{q-1} R u_q \dots u_p$ . In other words, we delete the edges  $u_{q-1}u_q$  and  $v_{q-1}v_q$  from the cycles and add  $u_{q-1}v_{q'}$  and  $u_q v_{q'-1}$  to form a single cycle. It is straightforward to verify that  $\hat{\mathcal{C}} = (\mathcal{C} \setminus \{C_j, C_{j'}\}) \cup \{C'\}$  is a regular solution of cost at most  $r$  since  $\omega(u_{q-1}u_q) = \omega(u_{q'-1}u_{q'}) = 0$  and  $\alpha$  is nondecreasing. But  $|\hat{\mathcal{C}}| = |\mathcal{C}| - 1$ , contradicting the minimality of  $\mathcal{C}$ . Hence, at most one cycle of  $\mathcal{C}$  has edges in each clique. If a cycle  $C_j$  has no edge in a clique  $Q_i$ , then for every vertex  $v$  of  $C_j$  that is in  $Q_i$ , the edges of  $C_j$  incident to  $v$  are edges between  $Q_i$  and other cliques. By regularity, at most two edges between  $Q_i$  and another clique of  $\hat{Q}$  can be in the cycles of  $\mathcal{C}$ . Therefore, the total number of vertices of  $Q_i$  that are in the cycles without edges with both end-vertices in  $Q_i$  is at most  $t - 1$ . □

For a solution  $\mathcal{C}$  for  $(G, \omega, \alpha, r)$ , we say that a vertex  $v \in V(G)$  is *marked* if  $v$  is incident to an edge between some cliques of  $\hat{Q}$  that is in a cycle of  $\mathcal{C}$ ; the other vertices are *unmarked*. The following claim is straightforward to see.

CLAIM 4.3. *For every  $i \in \{1, \dots, t\}$ ,  $Q_i$  has at most  $2(t - 1)$  marked vertices with respect to a regular solution  $\mathcal{C}$ .*

We apply a number of reduction rules to reduce the size of the graph. These rules delete some nontrivial vertices and edges. Whenever we apply these rules and delete edges, we always assume that the weight function  $\omega$  is adjusted by restricting it to the set of remaining edges. Observe that since we never delete trivial vertices and edges, the edges of nonzero weight are not affected, and Claims 4.1–4.3 hold for the obtained instances.

**REDUCTION RULE 4.1.** *If there is a clique  $Q_i \in \hat{\mathcal{Q}}$  and  $v \in Q_i$  such that  $N_G[v] = Q_i$  and  $|Q_i| \geq t + 3$ , then set  $G = G - v$  and  $Q_i = Q_i \setminus \{v\}$ .*

To show that the rule is safe, assume that the rule is applied for  $v \in Q_i$  and let  $G' = G - v$  and  $Q'_i = Q_i \setminus \{v\}$ .

Suppose that  $(G, \omega, \alpha, r)$  is a yes-instance of PRIZE COLLECTING CYCLE COVER and let a cycle cover  $\mathcal{C} = \{C_1, \dots, C_k\}$  be a minimal regular solution of cost at most  $r$  that exists because of Claim 4.1. Then there is  $C_j \in \mathcal{C}$  containing  $v$ . Let  $x$  and  $y$  be the neighbors of  $v$  in  $C_j$ . We claim that  $C_j$  is not a triangle. Otherwise,  $C_j$  contains the edge  $xy$  and, by Claim 4.2, the other cycles have at most  $t - 1$  vertices in  $Q_i$ . Since  $C_i$  is a triangle,  $|Q_i| \leq t + 2$ , contradicting the condition that  $|Q_i| \geq t + 3$ . Hence, if we replace the segment  $xvy$  in  $C_j$  by  $xy$ , we obtain a cycle. Denote it by  $C'_j$ . As  $\omega(vx) = \omega(vy) = 0$ , we have that  $\hat{\mathcal{C}} = (\mathcal{C} \setminus \{C_j\}) \cup \{C'_j\}$  is a solution for  $(G', \omega, \alpha, r)$ , that is, this is a yes-instance.

Assume that  $(G', \omega, \alpha, r)$  is a yes-instance of PRIZE COLLECTING CYCLE COVER. By Claim 4.1, there is a minimal regular solution  $\mathcal{C} = \{C_1, \dots, C_k\}$  for the instance. Since  $|Q'_i| = |Q_i| - 1 \geq t + 2$ , there is a cycle  $C_j = u_0 \dots u_p$  in  $\mathcal{C}$  such that  $u_{q-1}, u_q \in Q'_i$  by Claim 4.2. Let  $C'_j = u_0 \dots u_{q-1} v u_q \dots u_p$  and  $\hat{\mathcal{C}} = (\mathcal{C} \setminus \{C_j\}) \cup \{C'_j\}$ . We have that  $\hat{\mathcal{C}}$  is a solution for  $(G, \omega, \alpha, r)$  and, therefore,  $(G, \omega, \alpha, r)$  is a yes-instance. This completes the safeness proof.

For distinct  $i, j \in \{1, \dots, t\}$ , denote by  $G_{ij}$  the bipartite graph with the set of vertices  $Q_i \cup Q_j$  whose edges are the edges of  $G$  between  $Q_i$  and  $Q_j$ .

**REDUCTION RULE 4.2.** *If there are distinct  $i, j \in \{1, \dots, t\}$  such that  $G_{ij}$  has a matching  $M$  of size at least  $4t - 3$ , then for an arbitrarily chosen edge  $e \in M$ , set  $G = G - e$ .*

Assume that  $G' = G - e$  is obtained by the application of the rule for  $i, j \in \{1, \dots, t\}$ . Clearly, if  $(G', \omega, \alpha, r)$  is a yes-instance of PRIZE COLLECTING CYCLE COVER, then  $(G, \omega, \alpha, r)$  is a yes-instance as well. Hence, to show safeness, we assume that  $(G, \omega, \alpha, r)$  is a yes-instance and prove that  $(G', \omega, \alpha, r)$  is a yes-instance. Let  $\mathcal{C} = \{C_1, \dots, C_k\}$  be a minimal regular solution for  $(G, \omega, \alpha, r)$ . If  $e$  is not an edge of a cycle of  $\mathcal{C}$ ,  $\mathcal{C}$  is a solution for  $(G', \omega, \alpha, r)$ . Suppose that  $e = uu'$  is an edge of some cycle  $C_p$  of  $\mathcal{C}$ . Since  $|M| \geq 4t - 3$ , there is an edge  $vv' \in M$  such that  $v$  and  $v'$  are unmarked with respect to  $\mathcal{C}$  by Claim 4.3. We assume that  $u, v \in Q_i$  and  $u', v' \in Q_j$ . The vertices  $v$  and  $v'$  are included in some cycles  $C_q$  and  $C_{q'}$  of  $\mathcal{C}$  (note that they could be the same and may coincide with  $C_p$ ). Let  $x, y$  be the neighbors of  $v$  in  $C_q$  and let  $x', y'$  be the neighbors of  $v'$  in  $C_{q'}$ . As  $v$  and  $v'$  are unmarked,  $x, y \in Q_i$  and  $x', y' \in Q_j$ . Observe that  $C_q$  and  $C_{q'}$  contain edges with both end-vertices in  $Q_i$  and  $Q_j$ , respectively. Therefore, each of these cycles contains at least  $(4t - 3) - (t - 1) = 3t - 2 \geq 4$  vertices of  $Q_i$  and  $Q_j$ , respectively, by Claim 4.2. We use this and modify  $C_q$  and  $C_{q'}$  by replacing the segments  $xvy$  and  $x'v'y'$  by  $xy$  and  $x'y'$ , respectively. Next, we replace  $uu'$  in  $C_p$  by  $uvv'u'$ . Clearly, this modification leads to the new cycle cover  $\hat{\mathcal{C}}$  of the same size as  $\mathcal{C}$ . Since the edges of  $G$  with both their end-vertices in  $Q_i \cup Q_j$  have zero weight, the cost of  $\hat{\mathcal{C}}$  is the same as the cost of  $\mathcal{C}$ . As  $uu'$  is not included in the cycles of  $\hat{\mathcal{C}}$ , we have that  $\hat{\mathcal{C}}$  is a solution for  $(G', \omega, \alpha, r)$ , that is,  $(G', \omega, \alpha, r)$  is a yes-instance.

**REDUCTION RULE 4.3.** *If there is a clique  $Q_i \in \hat{\mathcal{Q}}$  and  $v \in V(G) \setminus Q_i$  such that  $|N_G(v) \cap Q_i| \geq 2t + 1$ , then for an arbitrary edge  $e = uv$  with  $u \in Q_i$ , set  $G = G - e$ .*

The safeness of the rule is proved along the same lines as the proof for Reduction

Rule 4.2. Assume that  $G' = G - e$  is obtained by the application of the rule for  $Q_i \in \mathcal{Q}$  and  $v \notin Q_i$ . Again, it is sufficient to prove that if  $(G, \omega, \alpha, r)$  is a yes-instance of PRIZE COLLECTING CYCLE COVER, then  $(G', \omega, \alpha, r)$  is a yes-instance. Assume that  $(G, \omega, \alpha, r)$  is a yes-instance and let  $\mathcal{C} = \{C_1, \dots, C_k\}$  be a minimal regular solution. Trivially,  $(G', \omega, \alpha, r)$  is a yes-instance with the same solution  $\mathcal{C}$  if  $e$  is not included in cycles of  $\mathcal{C}$ . Suppose that  $e = uv$  is an edge of  $C_p \in \mathcal{C}$ . By Claim 4.3, there is unmarked  $w \in Q_i \cap N_{G'}(v)$ . We have that  $w$  is a vertex of some cycle  $C_q \in \mathcal{C}$ ; it can happen that  $q = p$ . Let  $x$  and  $y$  be the neighbors of  $w$  in  $C_q$ . Notice that  $C_q$  has edges with both their end-vertices in  $Q_i$ . By Claim 4.2,  $C_q$  contains at least  $(2t + 1) - (t - 1) = t + 2 \geq 4$  vertices, that is,  $C_q$  is not a triangle. Then we can modify  $C_q$  by replacing the segment  $xwy$  by  $xy$  and then replacing  $uv$  in  $C_p$  by  $uvw$ . We obtain the cycle cover  $\hat{\mathcal{C}}$  of the same cost as  $\mathcal{C}$  whose cycles do not contain  $e$ . Hence,  $\hat{\mathcal{C}}$  is a solution for  $(G', \omega, \alpha, r)$  and  $(G', \omega, \alpha, r)$  is a yes-instance.

We apply Reduction Rules 4.1–4.3 exhaustively whenever possible. The following claim shows that we obtain a graph of bounded size.

CLAIM 4.4. *If no Reduction Rules 4.1–4.3 can be applied for  $(G, \omega, \alpha, r)$ , then  $|V(G)| = \mathcal{O}(t^4)$ .*

*Proof of Claim 4.4.* Let  $i \in \{1, \dots, t\}$ . Our goal is to find an upper-bound on  $|Q_i|$ . Consider  $j \in \{1, \dots, t\}$ . Recall that  $G_{ij}$  denotes the bipartite graph with the vertex set  $Q_i \cup Q_j$  whose edge set is the set of edges of  $G$  between  $Q_i$  and  $Q_j$ . There is a vertex cover of size at most  $8t - 8$  in  $G_{ij}$  having at most  $4t - 4$  vertices in  $Q_i$  and at most  $4t - 4$  vertices in  $Q_j$  composed of the end-vertices of a maximum matching in  $G_{ij}$ , as Reduction Rule 4.2 is not applicable. We say that these are *covering* vertices.

Observe that each covering vertex in  $Q_j$  has at most  $2t$  neighbors in  $Q_i$ , since Reduction Rule 4.3 is not applicable to those vertices. Notice also that the covering vertices in  $Q_i$  are neighbors of the covering vertices in  $Q_j$  by the construction of the vertex cover. Summing over all  $j \neq i$ , we obtain that  $Q_i$  contains at most  $2t(4t - 4)(t - 1)$  neighbors of covering vertices. Therefore, we conclude that there are at most  $2t(4t - 4)(t - 1)$  vertices in  $Q_i$  that are incident to the edges of  $G$  between  $Q_i$  and other cliques. Since Reduction Rule 4.1 is not applicable,  $Q_i$  has at most  $2t(4t - 4)(t - 1)$  vertices. Summing up over all cliques in the cover yields the claimed bound  $|V(G)| = \mathcal{O}(t^4)$ .  $\square$

Denote by  $\hat{G}$  the graph obtained from the input graph  $G$  by Reduction Rules 4.1–4.3. By Claim 4.4,  $\hat{G}$  has  $\mathcal{O}(t^4)$  vertices and, therefore,  $\mathcal{O}(t^8)$  edges. As the rules are safe, we have that the instance  $(\hat{G}, \omega, \alpha, r)$  of PRIZE COLLECTING CYCLE COVER is equivalent to  $(G, \omega, \alpha, r)$ . Note also that in parallel we get the clique cover of  $\hat{G}$  obtained from the original cover  $\hat{\mathcal{Q}}$  by the vertex deletions, that is, the obtained graph has a clique cover of size at most  $t$ .

Our next aim is to compress the weights and penalties. For this, we use Lemma 3. Recall that  $S$  is the set of edges of  $G$  with nonzero weights and  $\ell = |S|$ . Note that  $S \subseteq E(\hat{G})$ . Let  $S = \{e_1, \dots, e_\ell\}$  and  $p = \min\{|V(\hat{G})|, t(t + 1)/2\}$ . We construct the vector  $w \in \mathbb{Z}^h$  for  $h = \ell + p + 1$  by setting  $w = (\omega(e_1), \dots, \omega(e_\ell), \alpha(1), \dots, \alpha(p), r)$  and define  $N = \ell + 3 \geq 3$ . We apply the algorithm of Frank and Tardos from Lemma 3 and this algorithm constructs the vector  $\bar{w} = (\bar{\omega}_1, \dots, \bar{\omega}_\ell, \bar{\alpha}_1, \dots, \bar{\alpha}_p, \hat{r})$  that has the property that

$$(4.1) \quad \text{sign}(w \cdot b) = \text{sign}(\bar{w} \cdot b) \text{ for all vectors } b \in \mathbb{Z}^h \text{ with } \|b\|_1 \leq N - 1.$$

We define the weight function  $\hat{\omega}: E(\hat{G}) \rightarrow \mathbb{Z}$  by setting  $\hat{\omega}(e_i) = \bar{\omega}_i$  for  $i \in \{1, \dots, \ell\}$

and  $\hat{\omega}(e) = 0$  for  $e \in E(\hat{G}) \setminus S$ , and we define  $\hat{\alpha}: \{1, \dots, |V(\hat{G})|\} \rightarrow \mathbb{Z}$  by setting  $\hat{\alpha}(i) = \bar{\alpha}_i$  for  $i \in \{1, \dots, p\}$  and  $\hat{\alpha}(i) = \max_{j \in \{1, \dots, p\}} \bar{\alpha}_j$  for  $j \in \{p+1, \dots, |V(\hat{G})|\}$ . Note that the number of edges of nonzero weight is at most  $\ell$ .

CLAIM 4.5. *The instance  $(\hat{G}, \hat{\omega}, \hat{\alpha}, \hat{r})$  is a feasible instance of PRIZE COLLECTING CYCLE COVER and  $(\hat{G}, \hat{\omega}, \hat{\alpha}, \hat{r})$  is equivalent to  $(\hat{G}, \omega, \alpha, r)$ .*

*Proof of Claim 4.5.* Observe that for every  $i \in \{1, \dots, \ell\}$ ,  $\text{sign}(\hat{\omega}(e_i)) = \text{sign}(\omega(e_i))$ , because (4.1) holds for vectors  $b$  with one nonzero element that is equal to one. Hence,  $\hat{\omega}$  has nonnegative values. By the same arguments,  $\text{sign}(\hat{\alpha}(i)) = \text{sign}(\alpha(i))$  for  $i \in \{1, \dots, p\}$ . This implies that  $\hat{\alpha}$  has positive values, since  $\hat{\alpha}(i)$  for  $i > p$  is positive by the definition if  $\bar{\alpha}_j > 0$  for  $j \leq p$ . Since (4.1) holds for vectors  $b$  with two nonzero elements, one of which is 1 and the other is  $-1$ , we obtain that for  $i \in \{2, \dots, p\}$ ,  $\text{sign}(\hat{\alpha}(i-1) - \hat{\alpha}(i)) = \text{sign}(\alpha(i-1) - \alpha(i))$ . Also we have that  $\hat{\alpha}(j)$  for  $j > p$  has the same values and  $\hat{\alpha}(j) \geq \hat{\alpha}(i)$  for  $i \leq p$  by the definition. Therefore,  $\hat{\alpha}$  is a nondecreasing function. We conclude that  $(\hat{G}, \hat{\omega}, \hat{\alpha}, \hat{r})$  is a feasible instance of PRIZE COLLECTING CYCLE COVER.

To show that  $(\hat{G}, \hat{\omega}, \hat{\alpha}, \hat{r})$  and  $(\hat{G}, \omega, \alpha, r)$  are equivalent, assume first that  $(\hat{G}, \omega, \alpha, r)$  is a yes-instance of PRIZE COLLECTING CYCLE COVER. By Claim 4.1, there is a minimal regular solution  $\mathcal{C}$  for the instance. Let  $S'$  be the set of edges with nonzero weights included in the cycles of  $\mathcal{C}$ . Clearly,  $S' \subseteq S$ . By Lemma 2, we have that  $q = |C| \leq t(t+1)/2 \leq p$ . Let  $\ell' = |S'|$ . Because  $N = \ell + 3$ , (4.1) holds for vectors  $b$  with  $\ell' + 2$  nonzero elements such that  $\ell'$  elements are 1's and 2 elements are  $-1$ 's. Then we have that

$$\text{sign} \left( \sum_{e \in S'} \omega(e) - \alpha(q) - r \right) = \text{sign} \left( \sum_{e \in S'} \hat{\omega}(e) - \hat{\alpha}(q) - \hat{r} \right).$$

Then  $c_{\alpha, \omega}(\mathcal{C}) = \sum_{e \in S'} \omega(e) - \alpha(q) \leq r$  if and only if  $c_{\hat{\alpha}, \hat{\omega}}(\mathcal{C}) = \sum_{e \in S'} \hat{\omega}(e) - \hat{\alpha}(q) \leq \hat{r}$ . Hence,  $\mathcal{C}$  is a solution for  $(\hat{G}, \hat{\omega}, \hat{\alpha}, \hat{r})$  and, therefore,  $(\hat{G}, \hat{\omega}, \hat{\alpha}, \hat{r})$  is a yes-instance. For the opposite direction, the arguments are the same. If  $(\hat{G}, \hat{\omega}, \hat{\alpha}, \hat{r})$  is a yes-instance of PRIZE COLLECTING CYCLE COVER, it has a minimal regular solution  $\mathcal{C}$ . Then we repeat the same arguments as above and conclude that  $\mathcal{C}$  is a solution for  $(\hat{G}, \omega, \alpha, r)$ . Then  $(\hat{G}, \omega, \alpha, r)$  is a yes-instance.  $\square$

By Lemma 3, we have that for every  $e \in E(\hat{G})$ ,  $\hat{\omega}(e) \leq 2^{4h^3} N^{h(h+2)}$ . Hence, the weight of  $e$  can be encoded in binary by a string of length  $\mathcal{O}(h^3 + h^2 \log N)$ . Since  $h \leq \ell + t(t+1)/2 + 1$  and  $N = \ell + 3$ , the weight can be encoded by a string of length  $\mathcal{O}((\ell + t^2)^3)$ . Recall that  $t = s + \ell$  and  $\ell$  edges of  $\hat{G}$  have nonzero weights. It follows that the weights can be encoded by  $\mathcal{O}(\ell(s + \ell)^6)$  symbols. Similarly, we have that for each  $i \in \{1, \dots, |V(\hat{G})|\}$ ,  $\hat{\alpha}(i) \leq 2^{4h^3} N^{h(h+2)}$  and  $\alpha(i)$  can be encoded by a string of length  $\mathcal{O}((s + \ell)^6)$ . As  $\hat{G}$  has  $\mathcal{O}((s + \ell)^4)$  vertices, the penalty function  $\hat{\alpha}$  can be encoded by  $\mathcal{O}((s + \ell)^{10})$  symbols. Taking into account that  $\hat{G}$  has  $\mathcal{O}((s + \ell)^4)$  vertices and  $\mathcal{O}((s + \ell)^8)$  edges, we obtain that the size of the instance  $(\hat{G}, \hat{\omega}, \hat{\alpha}, \hat{r})$  is  $\mathcal{O}((s + \ell)^{10})$ .

It remains to evaluate the running time of our kernelization algorithm. Observe that Reduction Rules 4.1–4.3 can be applied in polynomial time. In particular, to apply Reduction Rule 4.2, we have to find a maximum matching in a bipartite graph, and this can be done, e.g., by the classical algorithm of Hopcroft and Karp [27]. Since the algorithm of Frank and Tardos is polynomial (see Lemma 3), we conclude that the kernelization algorithm is polynomial.

This completes the proof of the theorem for PRIZE COLLECTING CYCLE COVER. Now we explain how to adapt the proof for CYCLE COVER (HAMILTONIAN CYCLE)

and PATH COVER.

For CYCLE COVER, we just apply our kernelization algorithm for PRIZE COLLECTING CYCLE COVER for  $S = \emptyset$  and skip the compression of weights and penalties. Recall that an instance  $(G, k)$  of CYCLE COVER is equivalent to the instance  $(G, \omega, \alpha, r)$  of PRIZE COLLECTING CYCLE COVER with zero edge weights,  $\alpha(x) = x$  and  $r = -k$ . Respectively, given an instance  $(G, k)$  of CYCLE COVER together with a clique cover  $\mathcal{Q}$  of size  $s$ , we apply Reduction Rules 4.1–4.3. This immediately implies that we obtain a kernel of size  $\mathcal{O}(s^8)$ . For HAMILTONIAN CYCLE, it is sufficient to observe that a graph  $G$  has a Hamiltonian cycle if and only if  $(G, k)$  is a yes-instance of CYCLE COVER for  $k = 1$ . Since Reduction Rules 4.1–4.3 do not modify  $k$ , we immediately conclude that the kernelization algorithm for  $(G, 1)$ , in fact, produces an instance of HAMILTONIAN CYCLE.

For PATH COVER, we use the same approach, but we have to adjust the reduction rules and the safeness proofs.

Let  $(G, k)$  be an instance of PATH COVER and let  $\mathcal{Q} = \{Q_1, \dots, Q_s\}$  be a clique cover of  $G$ .

We say that a path cover  $\mathcal{P}$  is a *regular solution* if  $|\mathcal{P}| \leq k$  and  $\mathcal{P}$  is regular with respect to  $\mathcal{Q}$ . Lemma 1 immediately implies the following claim.

CLAIM 4.6. *If  $(G, k)$  is a yes-instance of PATH COVER, then it has a regular solution.*

In exactly the same way as in Claim 4.2, we show the following claim.

CLAIM 4.7. *Let  $\mathcal{P}$  be a regular solution. Then for every  $i \in \{1, \dots, s\}$ , the total number of vertices of  $Q_i$  that are internal vertices of the paths of  $\mathcal{P}$  without edges with both end-vertices in  $Q_i$  is at most  $s - 1$ .*

For a solution  $\mathcal{P}$ , we say that a vertex  $v \in V(G)$  is *marked* if  $v$  is incident to an edge between some cliques of  $\mathcal{Q}$  that is in a path of  $\mathcal{P}$ ; the other vertices are *unmarked*. Then we have the next claim.

CLAIM 4.8. *For every  $i \in \{1, \dots, s\}$ ,  $Q_i$  has at most  $2(s - 1)$  marked vertices with respect to a regular solution  $\mathcal{P}$ .*

We use Claims 4.7 and 4.8 to construct the analogues of Reduction Rules 4.1–4.3 for PATH COVER. Notice that  $\hat{\mathcal{Q}} = \mathcal{Q}$  in this case.

REDUCTION RULE 4.4. *If there is a clique  $Q_i \in \mathcal{Q}$  and  $v \in Q_i$  such that  $N_G[v] = Q_i$  and  $|Q_i| \geq s + 1$ , then set  $G = G - v$  and  $Q_i = Q_i \setminus \{v\}$ .*

To show that the rule is safe, assume that the rule is applied for  $v \in Q_i$  and let  $G' = G - v$  and  $Q'_i = Q_i \setminus \{v\}$ .

Suppose that  $(G, k)$  is a yes-instance of PATH COVER. Then there is a regular solution  $\mathcal{P} = \{P_1, \dots, P_r\}$  for the instance by Claim 4.6. Since  $\mathcal{P}$  is a path cover, there is  $P_j = u_1 \cdots u_t$  in  $\mathcal{P}$  that contains  $v$ . If  $t = 1$ , that is,  $v = u_1 = u_t$ , we define  $\hat{\mathcal{P}} = \mathcal{P} \setminus \{P_j\}$ . Otherwise, if  $t \geq 2$ , we define  $\hat{\mathcal{P}}$  as follows. If  $v = u_1$  or  $v = u_t$ , we set  $P'_i = u_2 \cdots u_t$  or  $u_1 \cdots u_{t-1}$ , respectively, and if  $v = u_i$  for  $i \in \{2, \dots, t - 2\}$ , we set  $P'_i = u_1 \cdots u_{i-1} u_{i+1} \cdots u_t$ . Then we define  $\hat{\mathcal{P}} = (\mathcal{P} \setminus \{P_j\}) \cup \{P'_i\}$ . It is straightforward to verify that  $\hat{\mathcal{P}}$  is a solution for  $(G', k)$ , that is,  $(G', k)$  is a yes-instance.

Assume that  $(G', k)$  is a yes-instance of PATH COVER. By Claim 4.6, there is a regular solution  $\mathcal{P} = \{P_1, \dots, P_r\}$  for the instance. Because  $|Q'_i| \geq s$ , there is a path  $P_j = u_1 \cdots u_t$  in  $\mathcal{P}$  that either has an end-vertex in  $Q'_i$  or an edge with both end-vertices in  $Q'_i$  by Claim 4.7. If  $u_1 \in Q'_i$  or  $u_t \in Q'_i$ , let  $P'_i = v u_1 \cdots u_t$  or  $P'_i = u_1 \cdots u_t v$ , respectively. If  $u_1, u_t \notin Q'_i$  and  $u_{h-1}, u_h \in Q'_i$  for some  $h \in \{2, \dots, t\}$ , then

$P'_i = u_1 \cdots u_{h-1} v u_h \cdots u_s$ . Let  $\hat{\mathcal{P}} = (\mathcal{P} \setminus \{P_i\}) \cup \{P'_i\}$ . We have that  $\hat{\mathcal{P}}$  is a solution for  $(G, k)$  and, therefore, this is a yes-instance. This completes the safeness proof.

REDUCTION RULE 4.5. *If there are distinct  $i, j \in \{1, \dots, s\}$  such that  $G_{ij}$  has a matching  $M$  of size at least  $4s - 3$ , then for an arbitrarily chosen edge  $e \in M$ , set  $G = G - e$ .*

Assume that  $G' = G - e$  is obtained by the application of the rule for  $i, j \in \{1, \dots, t\}$ . It is sufficient to show that if  $(G, k)$  is a yes-instance of PATH COVER, then  $(G', k)$  is a yes-instance. Let  $\mathcal{P} = \{P_1, \dots, P_r\}$  be a regular solution for  $(G, k)$ . If  $e$  is not an edge of a path of  $\mathcal{P}$ ,  $\mathcal{P}$  is a solution for  $(G', k)$ . Suppose that  $e = uu'$  is an edge of some path  $P_h$  of  $\mathcal{P}$ . Since  $|M| \geq 4s - 3$ , there is an edge  $vv' \in M$  such that  $v$  and  $v'$  are unmarked with respect to  $\mathcal{P}$  by Claim 4.8. We assume that  $u, v \in Q_i$  and  $u', v' \in Q_j$ . The vertices  $v$  and  $v'$  are included in some paths  $P_t$  and  $P_{t'}$  of  $\mathcal{P}$  that could be the same and may coincide with  $P_h$ . We modify  $P_t$  and  $P_{t'}$  and exclude  $v$  and  $v'$  from the paths. Let  $P_t = x_1 \cdots x_p$ . If  $p = 1$ , we simply delete  $P_t$ . If  $v = x_1$  or  $v = x_p$ , we delete  $w$  from the path. If  $v = x_\ell$  for  $\ell \in \{2, \dots, p - 2\}$ , we replace  $P_t$  by  $x_1 \cdots x_{\ell-1} x_{\ell+1} \cdots x_p$ . This replacement can be done, because  $v$  is unmarked and, therefore,  $x_{\ell-1}, x_{\ell+1} \in Q_i$ . The path  $P_{t'}$  is modified in the same way (if  $P_t$  and  $P_{t'}$  are the same, we modify the path obtained from  $P_t$  by the exclusion of  $v$ ). Next, we modify  $P_h$  (or the path obtained from  $P_h$  by the modification of  $P_t$  and/or  $P_{t'}$ ) by replacing the subpath  $uu'$  by the subpath  $uvv'u'$ . Clearly, this modification creates the new path cover  $\hat{\mathcal{P}}$  of the same size as  $\mathcal{P}$ , but now  $e$  is not included in a path of  $\hat{\mathcal{P}}$ , that is,  $\hat{\mathcal{P}}$  is a solution for  $(G', k)$ . Therefore,  $(G', k)$  is a yes-instance.

REDUCTION RULE 4.6. *If there is a clique  $Q_i \in \mathcal{Q}$  and  $v \in V(G) \setminus Q_i$  such that  $|N_G(v) \cap Q_i| \geq 2s - 1$ , then for an arbitrary edge  $e = uv$  with  $u \in Q_i$ , set  $G = G - e$ .*

The safeness of the rule is proved similarly to the safeness proof of Reduction Rule 4.5. Assume that  $G' = G - e$  is obtained by the application of the rule for  $Q_i \in \mathcal{Q}$  and  $v \notin Q_i$ . Again, it is sufficient to prove that if  $(G, k)$  is a yes-instance of PATH COVER, then  $(G', k)$  is a yes-instance. Assume that  $\mathcal{P} = \{P_1, \dots, P_r\}$  is a regular solution. If  $e$  is not included in any path of  $\mathcal{P}$ , then  $\mathcal{P}$  is a solution for  $(G', k)$ . Suppose that  $e = uv$  is an edge of  $P_h \in \mathcal{P}$ . By Claim 4.8, there is unmarked  $w \in Q_i \cap N_G(v)$ . We have that  $w$  is a vertex of some path  $P_t \in \mathcal{P}$ ; it can happen that  $t = h$ . We modify  $P_t = x_1 \cdots x_p$  to exclude  $w$  from it. If  $p = 1$ , we simply delete  $P_t$ . If  $w = x_1$  or  $w = x_p$ , we delete  $w$  from the path. If  $w = x_\ell$  for  $\ell \in \{2, \dots, p - 2\}$ , we replace  $P_t$  by  $x_1 \cdots x_{\ell-1} x_{\ell+1} \cdots x_p$ . This replacement can be done, because  $w$  is unmarked and, therefore,  $x_{\ell-1}, x_{\ell+1} \in Q_i$ . Next, we modify  $P_h$  (or the path obtained from  $P_h$  by the modification of  $P_t$  if  $h = t$ ) by replacing the subpath  $vu$  by the subpath  $vwu$ . Clearly, this modification creates the new path cover  $\hat{\mathcal{P}}$  of the same size as  $\mathcal{P}$ , but now  $e$  is not included in a path of  $\hat{\mathcal{P}}$ , that is,  $\hat{\mathcal{P}}$  is a solution for  $(G', k)$ . We obtain that  $(G', k)$  is a yes-instance.

We apply Reduction Rules 4.4–4.6 exhaustively whenever possible. Then we have the following claim that is proved by the same arguments as Claim 4.4

CLAIM 4.9. *If no Reduction Rules 4.4–4.6 can be applied for  $(G, k)$ , then  $|V(G)| = \mathcal{O}(s^4)$ .*

Let  $\hat{G}$  be the graph obtained by the exhaustive application of Reduction Rules 4.4–4.6. We have that  $|V(\hat{G})| = \mathcal{O}(s^4)$  and, therefore,  $|E(\hat{G})| = \mathcal{O}(s^8)$ . We also have that the instances  $(G, k)$  and  $(\hat{G}, k)$  are equivalent. The final step is to replace  $k$  by  $\hat{k} = \min\{k, s(s + 1)/2\}$ . By Lemma 2, the instances  $(\hat{G}, k)$  and  $(\hat{G}, \hat{k})$  are equivalent.

Clearly, the size of  $(\hat{G}, \hat{k})$  is  $\mathcal{O}(s^8)$ . Since Reduction Rules 4.4–4.6 can be applied in polynomial time in the same way as Reduction Rules 4.1–4.3, we conclude that PATH COVER admits a polynomial kernel of size  $\mathcal{O}(s^8)$ .  $\square$

**5. Structure of path and cycle covers in proper  $H$ -graphs.** In this section we introduce additional notation and obtain a number of structural and auxiliary algorithmic results about path and cycle covers in proper  $H$ -graphs.

**5.1. Path and cycle covers in proper interval graphs.** Our kernelization algorithms for proper  $H$ -graphs use the algorithms for PATH COVER and CYCLE COVER on proper interval graphs. PATH COVER on proper interval graphs is easy. Recall that every connected proper interval graph has a Hamiltonian path, and a proper interval graph has a Hamiltonian cycle if and only if it has at least three vertices and is 2-connected [2, 14, 29]. This immediately implies the following observation.

OBSERVATION 2. *For a proper interval graph  $G$ , the minimum size of a path cover is exactly the number of components of  $G$ .*

It is also useful to make some further observations about the structure of Hamiltonian paths in proper interval graphs. Let  $G$  be a proper interval graph and let  $\mathcal{I} = \{[\ell_v, r_v] \mid v \in V(G)\}$  be its proper interval representation. Notice that by properness,  $\ell_u \neq \ell_v$  ( $r_u \neq r_v$ ) for distinct  $u, v \in V(G)$  and  $\ell_u \leq \ell_v$  implies that  $r_u \leq r_v$ . In particular, this means that  $\mathcal{I}$  provides a total order  $\pi_{\mathcal{I}} = v_1, \dots, v_n$  on  $V(G)$ . Throughout this subsection we consider proper interval graph  $G$  with a given representation denoted by  $\mathcal{I} = \{[\ell_v, r_v] \mid v \in V(G)\}$  and fix the ordering  $\pi_{\mathcal{I}} = v_1, \dots, v_n$  of  $V(G)$ . Observe that  $v_1$  and  $v_n$  are *simplicial* vertices, that is,  $N_G(v_1)$  and  $N_G(v_n)$  are cliques. A vertex  $v$  is the *leftmost* (*rightmost*, respectively) vertex if  $\ell_v$  is minimum (respectively, maximum). We extend this definition in a natural way for two leftmost (rightmost, respectively) vertices. It is said that a path  $x_1 \cdots x_k$  in  $G$  is *monotone* if  $\ell_{x_1} < \cdots < \ell_{x_k}$ . It is well known [2] that every connected proper interval graph has a monotone Hamiltonian path. Clearly, the leftmost and the rightmost vertices are the end-vertices of such a path. Also every 2-connected proper interval graph with at least three vertices has a Hamiltonian cycle that is the concatenation of two monotone paths [2]. This leads to the following lemma.

LEMMA 4. *Let  $G$  be a connected proper interval graph with at least two vertices. Then  $G$  has a Hamiltonian  $(v_1, v_n)$ -path. Moreover, if  $G$  is 2-connected, then  $G$  has a Hamiltonian  $(v_1, v_2)$ -path (a Hamiltonian  $(v_{n-1}, v_n)$ -path) and, provided additionally that  $|V(G)| \geq 3$ ,  $G$  has a path cover of size two formed by  $(a_1, b_1)$ - and  $(a_2, b_2)$ -paths with  $\{a_1, a_2\} = \{v_1, v_2\}$  and  $\{b_1, b_2\} = \{v_{n-1}, v_n\}$ .*

*Proof.* The first claim trivially follows from the results of Bertossi [2]. To show the second claim, assume that  $|V(G)| \geq 3$  and  $G$  is 2-connected. Then  $G$  has a Hamiltonian cycle formed by two monotone paths  $x_1 \cdots x_s$  and  $y_1 \cdots y_t$  with  $x_1 = y_1 = v_1$  and  $x_s = y_t = v_n$  [2]. Since the paths are monotone, either  $v_2 = x_2$  or  $v_2 = y_2$ . In the first case, we have that  $y_1 \cdots y_t x_{s-1} \cdots x_2$  is a Hamiltonian  $(v_1, v_2)$ -path, and  $x_1 \cdots x_s y_{t-1} \cdots y_2$  is a Hamiltonian  $(v_1, v_2)$ -path in the other case. Similarly, either  $v_{n-1} = x_{s-1}$  or  $v_{n-1} = y_{t-1}$ . If  $v_2 = x_2$  and  $v_{n-1} = x_{s-1}$ , we have the path cover  $\{x_2 \cdots x_{s_1}, y_1 \cdots y_t\}$ , and if  $v_2 = x_2$  and  $v_{n-1} = y_{t-1}$ , we obtain the path cover  $\{x_2 \cdots x_s, y_1 \cdots y_{t-1}\}$ . All other cases are symmetric with these two. Clearly, if  $|V(G)| = 2$ , then  $v_1 v_2 = v_{n-1} v_n$  is a Hamiltonian path.  $\square$

Even though many algorithms are known for dozens of problems when the input is restricted to (proper) interval graphs, the complexity of CYCLE COVER seems to

be unknown. We provide a linear time algorithm using a greedy approach based on a standard decomposition of a graph into 2-connected components in order to compute the minimum cycle cover of a proper interval graph.

First, we need some properties of the blocks of a proper interval graph.

**OBSERVATION 3.** *Let  $B$  be a block of  $G$  and  $v_i \in V(B)$  be a cut-vertex of  $G$ . Then either  $j \leq i$  for every  $v_j \in V(B)$  or  $i \leq j$  for every  $v_j \in V(B)$ .*

*Proof.* Suppose otherwise; then there are vertices  $v_j$  and  $v_{j'}$  in  $B$  with  $j < i < j'$ . Since  $B$  is a block, there is a path  $P$  in  $B$  connecting  $v_j$  and  $v_{j'}$  which does not use  $v_i$ . In particular, in the interval representation, the intervals of the vertices on  $P$  cover the interval  $I_{v_i}$ , contradicting the fact that  $v_i$  is a cut-vertex.  $\square$

Observation 3 directly implies the following lemma.

**LEMMA 5.** *Let  $G$  be a connected proper interval graph. Then the following hold:*

- (i) *For every block  $B$  of  $G$ , the vertices of  $B$  occur consecutively in  $\pi_{\mathcal{I}}$ , say as  $v_1^B, \dots, v_{n_B}^B$ .*
- (ii) *The block-cutpoint decomposition graph  $BC(G)$  is a path.*
- (iii) *For each block  $B$  of  $G$ , if  $B$  is not a leaf of  $BC(G)$ , then precisely  $v_1^B$  and  $v_{n_B}^B$  are the cut-vertices contained in  $B$ .*
- (iv) *If  $G$  contains at least one cut-vertex and  $B$  is the block containing  $v_1$  ( $v_n$ , respectively), then  $v_{n_B}^B$  ( $v_1^B$ , respectively) is the only cut-vertex in  $B$ .*

**LEMMA 6.** *If  $G$  is a 2-connected proper interval graph, then  $G[v_1, \dots, v_{n-1}]$  is also 2-connected (symmetrically,  $G[v_2, \dots, v_n]$  is also 2-connected).*

*Proof.* We first observe that  $N(v_n)$  is a clique. Namely, let  $v_i$  be a neighbor of  $v_n$ . Clearly, in order for  $I_{v_i}$  to intersect  $I_{v_n}$ , it must also intersect  $I_{v_j}$  for every  $j$  between  $i$  and  $n$ , i.e.,  $N(v_n)$  is a clique.

Now consider any path  $P$  in  $G$ . Suppose  $P = P_1 v_n P_2$  for some paths  $P_1$  and  $P_2$ , i.e.,  $P$  uses  $v_n$  as an internal node. Since  $N(v_n)$  is a clique, it is clear that  $P' = P_1 P_2$  is also a path in  $G$  and, in particular,  $P'$  is a path in  $G' = G[v_1, \dots, v_{n-1}]$ . However, if there is a cut-vertex  $v$  in  $G'$ , it would need to belong to every  $(u, u')$ -path in  $G'$  for some pair of vertices  $u, u' \in V(G')$  and, as such, it would also belong to every  $(u, u')$ -path in  $G$ . This would contradict the 2-connectedness of  $G$ .  $\square$

Notice that if  $G$  contains a trivial block, then the edge in this block cannot be used in any cycle cover of  $G$ . Namely, to compute a minimum cycle cover, it suffices to compute a minimum cycle cover of each of the two induced subgraphs of  $G$  obtained by deleting the edge of a trivial block. This together with Lemma 5 (ii) provides the following observation.

**OBSERVATION 4.** *For any proper interval graph  $G$ , if two trivial blocks share a cut-vertex, then  $G$  has no cycle cover.*

We use the above structural observations to construct a linear time algorithm to solve the cycle cover problem on proper interval graphs. Note that in this proof we will at times implicitly use the properties (i)–(iv) given in Lemma 5. Recall that  $\text{bl}(G)$  denotes the number of blocks of  $G$  and  $\text{tbl}(G)$  is the number of trivial blocks.

**THEOREM 4.** *Let  $G$  be a proper interval graph. In linear time, we can decide whether  $G$  has a cycle cover, and when it does a minimum cycle cover can be constructed. Additionally, this cover will use precisely  $\text{bl}(G) - \text{tbl}(G)$  cycles.*

*Proof.* Recall that if  $G$  contains a trivial block, then the edge in this block cannot be used in any cycle cover of  $G$ . Namely, to compute a minimum cycle cover, it

suffices to compute a minimum cycle cover of each of the two induced subgraphs of  $G$  obtained by deleting the edge of a trivial block. Therefore, we delete the edges from the trivial blocks and consider each connected subgraph which remains. At this point, we also apply Observation 4 and conclude that  $G$  has no cycle cover if we have created any isolated vertices.

Recall that when a proper interval graph is 2-connected (and has at least three vertices), it has a Hamiltonian cycle. In particular, every nontrivial block of  $G$  has a Hamiltonian cycle. From now on we assume that  $G$  is a connected graph without trivial blocks.

We proceed inductively on the number of blocks in  $G$ . Namely, we will see that when  $G$  has a cycle cover, it has a cycle cover by  $\text{bl}(G)$  cycles. Such a cycle cover is optimal, since we only have nontrivial blocks, i.e., each block has a private vertex which cannot be covered simultaneously with the private vertex of another block.

If  $G$  is a block, then, since  $G$  has no trivial blocks, it has at least three vertices and in particular has a Hamiltonian cycle.

So, suppose  $G$  contains at least two blocks and let  $B$  be the block containing  $v_1$ . This gives us two cases to consider.

*Case 1:*  $B$  contains exactly three vertices. Then  $V(B) = \{v_1, v_2, v_3\}$ .

Note that  $B$  is a triangle, and this triangle is the only cycle which contains  $v_1$ , i.e., if  $G$  has a cycle cover, then  $B$  must be a cycle in every cycle cover of  $G$ . Moreover,  $G$  has a cycle cover if and only if  $G' = G - V(B) = G[v_4, \dots, v_n]$  has a cycle cover.

Now, since  $G$  only contains nontrivial blocks,  $v_4$  is not a cut-vertex of  $G$ , i.e., there is a block  $B'$  in  $G$  which contains  $v_4$ . If this block has only three vertices, then the block containing  $v_4$  in  $G'$  has only two vertices. In particular,  $G$  has no cycle cover because we cannot simultaneously cover both  $v_1$  and  $v_4$  by cycles. On the other hand, if  $B'$  contains at least four vertices, then in  $G'$  the block containing  $v_4$ , which is precisely  $B' - v_3$  by Lemma 6, has at least three vertices, i.e.,  $G'$  is a proper interval graph with no trivial blocks, and  $\text{bl}(G') = \text{bl}(G) - 1$ , and applying induction on  $G'$  completes this case.

*Case 2:*  $B$  contains at least four vertices.

Now, by Lemma 6,  $G[v_1, \dots, v_{n_B-1}]$  is 2-connected and as such has a Hamiltonian cycle  $C_B$  which can be computed in linear time [29]. Moreover,  $G' = G[v_{n_B}, \dots, v_n]$  consists of only nontrivial blocks, and as such we can apply induction to see that if  $G'$  has a cycle cover, then it has a cycle cover using  $\text{bl}(G') = \text{bl}(G) - 1$  blocks. Namely, if  $G'$  has a cycle cover  $C'$ , then  $C' \cup \{C_B\}$  is a cycle cover of  $G$  using  $\text{bl}(G)$  cycles.

So, it only remains to argue that when  $G$  has a cycle cover,  $G'$  also has a cycle cover. Let  $\mathcal{C}$  be a minimum cycle cover of  $G$ , and let  $C$  be the cycle which covers  $v_{n_B}$  (i.e., the cycle covering the cut-vertex contained in  $B$ ).

Suppose  $C$  is not contained in  $B$ . Now, the only vertex of  $B$  covered by  $C$  is  $v_{n_B}$ . Thus, removing all cycles contained in  $B$  from  $\mathcal{C}$  provides a cycle cover of  $G'$ .

Finally, suppose that  $C$  is contained in  $B$ . Since  $G$  does not contain trivial blocks, there is a well-defined block  $B'$  of  $G$  which contains  $v_{n_B+1}$ . Let  $\mathcal{C}_{B'}$  be the cycles in  $\mathcal{C}$  contained in  $B'$ . Note that  $\mathcal{C}_{B'}$  must contain at least one cycle, namely, the cycle which covers  $v_{n_B+1}$ . Moreover,  $\mathcal{C}_{B'}$  covers either  $B' - v_{n_B}$  or  $B' - \{v_{n_B}, v_{n_B+n_{B'}-1}\}$ . In either case, by Lemma 6, the subgraph of  $G$  induced by covered vertices together with  $v_{n_B}$  is 2-connected and as such has a Hamiltonian cycle  $C_{B'}$ . Namely, replacing  $\mathcal{C}_{B'}$  with  $C_{B'}$  in  $\mathcal{C}$  and then removing  $C$  results in a cycle cover of  $G'$ .

The algorithm implicitly contained in the above proof is extremely simple and outlined as follows. We first compute the block-cutpoint decomposition of the given

proper interval graph. We then delete any edges corresponding to trivial blocks. For each of the connected components that remain, we proceed as follows. We iteratively consider a leaf block and either add this cycle to the cover (when  $B$  has only three vertices) or construct a Hamiltonian cycle of this block without its cut-vertex (say, using [29]). In the former case we also check if the neighboring block of this leaf block becomes trivial after deleting the cut-vertex. If it does,  $G$  has no cycle cover. Otherwise, we remove the vertices covered by our constructed cycle and repeat. It is straightforward to verify that this algorithm runs in linear time.  $\square$

Theorem 4 implies the following useful lemma.

**LEMMA 7.** *Let  $G$  be a connected proper interval graph. Then if  $G$  and  $G - v_1$  ( $G$  and  $G - v_n$ , respectively) both have cycle covers, then the minimum sizes of cycle covers of these graphs are the same. Moreover, if  $G$  and  $G - \{v_1, v_n\}$  both have cycle covers, then the minimum sizes of cycle covers of these graphs are the same.*

*Proof.* Suppose that  $G$  and  $G' = G - v_1$  have cycle covers. Then the minimum sizes of cycle covers are  $\text{bl}(G) - \text{tbl}(G)$  and  $\text{bl}(G') - \text{tbl}(G')$ . If  $\text{bl}(G) - \text{tbl}(G) \neq \text{bl}(G') - \text{tbl}(G')$ , the nontrivial block of  $G$  that contains  $v_1$  is  $G[v_1, v_2, v_3]$  and  $G[v_2, v_3]$  is a block of  $G'$ . This means that  $v_2$  has a unique neighbor in  $G'$  and, clearly, that  $G'$  has no cycle cover, contradicting the assumption of the lemma. The proof of the second claim uses the same arguments.  $\square$

For a connected proper interval graph  $G$ , we define the *signature* of  $G$  as the sequence  $\sigma(G) = \langle A_1, A_2, A_3, A_4 \rangle$  of yes- or no-answers on the question of whether  $G$ ,  $G - v_1$ ,  $G - v_n$ , or  $G - \{v_1, v_n\}$  has a cycle cover, respectively.

We need the following lemma.

**LEMMA 8.** *There is no connected proper interval graph with one of the following signatures:*

- (i)  $\langle \text{yes}, \text{yes}, \text{no}, \text{yes} \rangle$ .
- (ii)  $\langle \text{yes}, \text{no}, \text{yes}, \text{yes} \rangle$ .
- (iii)  $\langle \text{yes}, \text{no}, \text{no}, \text{yes} \rangle$ .
- (iv)  $\langle \text{no}, \text{yes}, \text{yes}, \text{yes} \rangle$ .
- (v)  $\langle \text{no}, \text{yes}, \text{yes}, \text{no} \rangle$ .

*Proof.* Let  $G$  be a connected proper interval graph  $G$  with  $n \geq 3$ . Let  $\sigma(G) = \langle A_1, A_2, A_3, A_4 \rangle$ . If  $G$  is 2-connected, then  $A_1 = \text{yes}$  by the results of [2], and if  $A_2 = \text{no}$  or  $A_3 = \text{no}$ , then  $n = 3$ . Clearly, we have that  $A_4 = \text{no}$  in these cases. We conclude that  $G$  cannot have signatures (i)–(v).

Assume that  $G$  has a cut-vertex. Suppose that  $A_1 = A_4 = \text{yes}$ , that is,  $G$  and  $G - \{v_1, v_n\}$  have cycle covers. Observe that  $v_1$  is included in a nontrivial block of size at least 4. Otherwise, by Lemma 5, either  $G[v_1, v_2]$  or  $G[v_1, v_2, v_3]$  is a block and either  $v_2$  or  $v_3$ , respectively, is a unique cut-vertex of  $G$  in the block. In the first case,  $v_1$  is a vertex of degree one in  $G$  and  $A_1 = \text{no}$ . In the second case,  $v_2$  is a vertex of degree one in  $G - \{v_1, v_n\}$  and, therefore,  $A_4 = \text{no}$ . In both cases, we get a contradiction. By symmetry,  $v_n$  is included in a nontrivial block of size at least 4. Then a cycle cover  $\mathcal{C}$  of  $G - \{v_1, v_n\}$  of minimum size has a cycle  $C_1$  with the vertices  $v_2, \dots, v_i$  for some  $i \geq 3$  and a cycle  $C_2$  with the vertices  $v_j, \dots, v_{n-1}$  for some  $i < j \leq n - 3$ . By Lemma 4, we obtain that  $G[v_2, \dots, v_i]$  has a Hamiltonian  $(v_2, v_3)$ -path. Since  $v_1$  is adjacent to  $v_2$  and  $v_3$ , we have that there is a cycle  $C'_1$  with the vertices  $v_1, \dots, v_i$ . By symmetry, there is a cycle  $C'_2$  with the vertices  $v_j, \dots, v_n$ . By replacing  $C_1$  by  $C'_1$  and  $C_2$  by  $C'_2$ , we can construct cycle covers of  $G - v_1$  and  $G - v_n$ , that is,  $A_2 = A_3 = \text{yes}$ . This implies, that the signatures (i)–(iii) do not occur.

Suppose that  $A_2 = A_3 = \text{yes}$ . Similarly to the above case, we have that  $v_1$  is included in a nontrivial block of size at least 4: if  $G[v_1, v_2]$  is a block, then  $v_1$  is a vertex of degree one and  $G - v_n$  has no cycle cover contradicting  $A_3 = \text{yes}$ , and if  $G[v_1, v_2, v_3]$  is a block, then  $v_2$  is a vertex of degree one in  $G - v_1$ , contradicting  $A_2 = \text{yes}$ . Since  $A_2 = \text{yes}$ ,  $G - v_1$  has a cycle cover. Then a cycle cover  $\mathcal{C}$  of  $G - v_1$  of minimum size has a cycle  $C$  with the vertices  $v_2, \dots, v_i$  for some  $i \geq 3$  and, therefore,  $G[v_2, \dots, v_i]$  has a Hamiltonian  $(v_2, v_3)$ -path. Also we obtain that  $G$  has a cycle  $C'$  with the vertices  $v_1, \dots, v_i$ , because  $v_1$  is adjacent to  $v_2$  and  $v_3$ . By replacing  $C$  by  $C'$  in  $\mathcal{C}$ , we obtain a cycle cover of  $G$ , that is,  $A_1 = \text{yes}$ . Hence, (iv) and (v) do not occur.  $\square$

For a connected proper interval graph  $G$  given together with its proper interval representation  $\mathcal{I}$  and the corresponding ordering of the vertices  $\pi_{\mathcal{I}} = v_1, \dots, v_n$ , we define  $\text{cover}(G)$  to be the minimum size of a cycle cover of the graphs  $G$ ,  $G - v_1$ ,  $G - v_n$ , and  $G - \{v_1, v_n\}$ , and we assume that  $\text{cover}(G) = +\infty$  if none of these graphs has a cycle cover. Note that the minimum sizes of vertex covers in the graphs that have cycle covers are the same by Lemma 7.

**5.2. Nice  $H$ -representations.** To simplify the construction of our kernelization and compression algorithms on  $H$ -graphs, it is convenient to restrict the considered  $H$ -representations. We say that an  $H$ -representation  $(H', \mathcal{M})$  with  $\mathcal{M} = \{M_v\}_{v \in V(G)}$  is *nice* if  $H'$  is a simple graph and for every  $xy \in E(H')$  there is  $v \in V(G)$  such that  $x, y \in M_v$ . We show the following lemma.

LEMMA 9. *There is a polynomial-time algorithm that, given an instance  $(G, k)$  of PATH COVER (CYCLE COVER, respectively) for a proper  $H$ -graph  $G$  with its proper  $H$ -representation  $(H', \mathcal{M})$ , either solves the problem or constructs an equivalent instance  $(\hat{G}, \hat{k})$  of PATH COVER (CYCLE COVER, respectively) and a nice proper  $\hat{H}$ -representation  $(\hat{H}', \hat{\mathcal{M}})$  of  $\hat{G}$  such that the following hold:*

- (i)  $\hat{G}$  is an induced subgraph of  $G$  and  $\hat{k} \leq k$ .
- (ii)  $|V(\hat{H}')| \leq 3|E(H)|$  and  $|E(\hat{H}')| \leq 2|E(H)|$ .

*Proof.* To be able to switch to nice representations, we apply the following reduction rules for the considered instance  $(G, k)$  of PATH COVER and CYCLE COVER, respectively.

REDUCTION RULE 5.1. *If  $G$  has a connected component  $F$  that is a proper interval graph, then set  $G = G - V(F)$  and set  $k = k - 1$ .*

For CYCLE COVER, the rule is slightly more complicated.

REDUCTION RULE 5.2. *If  $G$  has a connected component  $F$  that is a proper interval graph, then find the minimum size  $h$  of a cycle cover of  $F$  if it exists. If  $F$  has no cycle cover, then stop and return a no-answer. Otherwise, set  $G = G - V(F)$  and set  $k = k - h$ .*

It is straightforward to see that the rules are safe. Moreover, they could be applied in polynomial time. In particular, for Reduction Rule 5.2, we use Theorem 4.

Then we apply the following trivial stopping rule.

REDUCTION RULE 5.3. *If  $G$  is empty and  $k \geq 0$ , then stop and return a yes-answer, and if  $k < 0$ , then stop and return a no-answer.*

We apply the rules exhaustively and either solve the problem or obtain an equivalent instance  $(\hat{G}, \hat{k})$  of PATH COVER or CYCLE COVER, respectively, such that no component of  $\hat{G}$  is a proper interval graph. It is trivial to see that (i) holds for  $(\hat{G}, \hat{k})$ .

It is also clear that if  $G$  is a proper  $H$ -graph and  $(H', \mathcal{M})$  is its representation, then  $\hat{G}$  is a proper  $H$ -graph as well and its representation  $(H', \hat{\mathcal{M}})$  is obtained from the representation of  $G$  by deleting from  $\mathcal{M}$  the sets corresponding to the vertices of the deleted components. Our next aim is to reduce  $H'$ . We do it by deleting irrelevant nodes and edges.

REDUCTION RULE 5.4.

- If  $H'$  has a node  $x$  such that  $x \notin M_v$  for every  $v \in V(\hat{G})$ , then set  $H' = H' - x$ .
- If  $H'$  has an edge  $xy$  such that either  $x \notin M_v$  or  $y \notin M_v$  for every  $v \in V(\hat{G})$ , then set  $H' = H' - xy$ .
- If  $H'$  has a loop  $e$ , then set  $H' = H' - e$ .
- If  $H'$  has at least two parallel edges  $e, e' = xy$ , then set  $H' = H' - e$ .

Let  $\hat{H}'$  be the graph obtained from  $H'$  by exhaustively applying the subrules of Reduction Rule 5.4. It is straightforward to verify that every  $M_v$  is a connected subset of  $V(\hat{H}')$ . It is also easy to see that  $\hat{H}'$  is a simple graph, and for every  $xy \in E(\hat{H}')$  there is  $v \in V(G)$  such that  $x, y \in M_v$ . Now we construct  $\hat{H}$  by dissolving some nodes by the exhaustive application of the next rule.

REDUCTION RULE 5.5. *If  $\hat{H}'$  has a node  $z$  incident exactly to two edges  $zx$  and  $zy$  such that  $z \neq x, y$ , then dissolve  $z$ .*

Denote by  $\hat{H}$  the obtained graph. We have that  $\hat{H}'$  is a subdivision of  $\hat{H}$  and, therefore,  $(\hat{H}', \hat{\mathcal{M}})$  is a nice proper  $\hat{H}$ -representation of  $G$ . Notice that some subdivision nodes of  $H'$  could be nodes of  $\hat{H}$  because of the deletion of some nodes and edges of  $H'$ . Since  $\hat{G}$  has no connected component that is a proper interval graph, we have that for every component  $F$  of  $\hat{G}$ , there is a vertex  $v \in V(F)$  such that  $M_v$  contains at least one branching node of  $H'$ . This immediately implies that at most two subdivision nodes that are on the same edge  $e$  of  $H$  could become nodes of  $\hat{H}$ . Hence, at most  $2|E(H)|$  subdivision nodes of  $H'$  could be nodes of  $\hat{H}$ . Observe that an isolated vertex of  $H$  is deleted by Reduction Rule 5.4. Therefore,  $|V(\hat{H})| \leq 3|E(H)|$ . By these arguments, we also have that  $|E(\hat{H})| \leq 2|E(H)|$ . We conclude that (ii) is fulfilled.

To complete the proof, it is sufficient to observe that the construction of  $\hat{G}$ ,  $\hat{H}$ , and  $(\hat{H}', \hat{\mathcal{M}})$  can be done in polynomial time.  $\square$

**5.3. Tamed path and cycle covers.** In this subsection we show that it suffices to consider path and cycle covers of proper  $H$ -graphs that have special structure. To do so, we need additional notation that will be used also in the next section of the paper.

Let  $G$  be a proper  $H$ -graph and let  $(H', \mathcal{M})$  with  $\mathcal{M} = \{M_v\}_{v \in V(G)}$  be a nice proper  $H$ -representation of  $G$ . For an edge  $e = xy$  of  $H$ , denote by  $S_e = x_0^e \cdots x_{h(e)+1}^e$  the  $(x, y)$ -path of  $H'$  corresponding to  $e$ . We have that  $x_0^e = x$ ,  $x_{h(e)+1}^e = y$  and  $x_1^e, \dots, x_{h(e)}^e$  are the subdivision nodes of  $H'$ ; for each  $e$ , we fix this ordering of the subdivision nodes on  $e$ . Denote  $X_e = \{x_1^e, \dots, x_{h(e)}^e\}$ . We denote by  $G_e$  the subgraph of  $G$  induced by the vertices  $V_e = \{v \in V(G) \mid M_v \subseteq X_e\}$ . Note that  $G_e$  is a proper interval graph and for each  $v \in V(G_e)$ ,  $M_v = \{x_{\ell_v}, \dots, x_{r_v}\}$  for some  $1 \leq \ell_v \leq r_v \leq h(e)$ . Observe also that  $\ell_v$  and  $r_v$  are distinct, because there is  $u \in V(G)$  such that  $x_{\ell_v}, x_{\ell_v+1} \subseteq M_u$ . Also, for distinct  $u, v \in V_e$ , either  $\ell_u < \ell_v$  and  $r_u < r_v$  or  $\ell_u > \ell_v$  and  $r_u > r_v$  by properness. This imposes the total ordering  $v_1^e, \dots, v_{p(e)}^e$  of the vertices of  $V_e$  according to the ordering of  $\ell_v$  for  $v \in V_e$ . In accordance with proper interval graphs, we refer to the first vertex (vertices) in

the ordering as *leftmost* vertex (vertices) and similarly define the *rightmost* vertex (vertices) as the last in the ordering. We say that a vertex  $v \in V_e$  is *left-attached* (*right-attached*, respectively) if there is  $u \in V(G)$  such that  $\{x_0^e, \dots, x_{\ell_v}^e\} \subseteq M_u$  ( $\{x_{r_v}^e, \dots, x_{h(e)+1}^e\} \subseteq M_u$ , respectively). Denote by  $L_e$  and  $R_e$ , respectively, the sets of left- and right-attached vertices of  $V_e$ . The useful properties of  $G_e$  are summarized in the following lemma.

LEMMA 10. For every  $e \in E(H)$ , the following hold if  $V_e \neq \emptyset$ :

- (i)  $G_e$  is a connected graph.
- (ii)  $L_e$  and  $R_e$  are nonempty cliques.
- (iii) If  $L_e \cap R_e \neq \emptyset$ , then  $V_e = L_e \cup R_e$ .
- (iv) If  $L_e \cap R_e = \emptyset$ , then  $N_G(L_e) \setminus V_e$  and  $N_G(R_e) \setminus V_e$  are cliques.
- (v) If  $u, v \in L_e$  ( $u, v \in R_e$ , respectively) and  $\ell_u \leq \ell_v$ , then  $N_G(u) \setminus V_e \subseteq N_G(v) \setminus V_e$  ( $N_G(u) \setminus V_e \subseteq N_G(v) \setminus V_e$ , respectively).

*Proof.* Suppose that  $G_e$  is disconnected. Then there is  $i \in \{2, \dots, h(e)\}$  such that for every  $v \in V_e$ , either  $x_{i-1}^e \notin M_v$  or  $x_i^e \notin M_v$  and there are two vertices  $u_1, u_2 \in V_e$  such that  $M_{u_1} \subseteq \{x_1^e, \dots, x_{i-1}^e\}$  and  $M_{u_2} \subseteq \{x_i^e, \dots, x_{h(e)}^e\}$ . Since  $(H', \mathcal{M})$  is a nice proper  $H$ -representation of  $G$ , there is  $w \in V(G)$  such that  $x_{i-1}^e, x_i^e \in M_w$ . Clearly,  $w \notin V_e$ . Then either  $M_{u_1} \subseteq \{x_1^e, \dots, x_{i-1}^e\} \subseteq M_w$  or  $M_{u_2} \subseteq \{x_i^e, \dots, x_{h(e)}^e\} \subseteq M_w$ , contradicting the properness of the representation. This proves (i).

We prove (ii) for  $L_e$ ; the claim for  $R_e$  is symmetric. Let  $v$  be the leftmost vertex of  $G_e$ . Since  $(H', \mathcal{M})$  is a nice proper  $H$ -representation of  $G$ , there is  $w \in V(G)$  such that  $x_{\ell_v-1}^e, x_{\ell_v}^e \in M_w$ . Clearly,  $w \notin V_e$ . As  $M_w$  is connected, we have that either  $\{x_0^e, \dots, x_{\ell_v}^e\} \subseteq M_w$  or  $\{x_{\ell_v-1}^e, \dots, x_{h(e)+1}^e\} \subseteq M_w$ . But in the second case, we have that  $M_v \subseteq \{x_{\ell_v-1}^e, \dots, x_{h(e)+1}^e\} \subseteq M_w$  by the choice of  $v$  contradicting properness. Therefore,  $\{x_0^e, \dots, x_{\ell_v}^e\} \subseteq M_w$  and  $v$  is left-attached. Hence,  $L_e \neq \emptyset$ . To show that  $L_e$  is a clique, assume that  $u$  and  $v$  are distinct vertices of  $L_e$  and let  $\ell_u < \ell_v$ . Since  $v$  is left-attached, there is  $w \in V(G)$  such that  $\{x_0^e, \dots, x_{\ell_v}^e\} \subseteq M_w$ . If  $u$  is not adjacent to  $v$ , then  $M_u \subseteq \{x_1^e, \dots, x_{\ell_u-1}^e\} \subseteq M_w$ , but this contradicts the properness of the representation. Hence,  $uv \in E(G)$ .

To show (iii), assume that there is  $v \in V_e$  such that  $v \in L_e \cap R_e$ . Then there are  $u_1, u_2 \in V(G)$  such that  $\{x_0^e, \dots, x_{v_\ell}^e\} \subseteq M_{u_1}$  and  $\{x_{v_r}, \dots, x_{h(e)+1}^e\} \subseteq M_{u_2}$ . Suppose that there is  $w \in V_e$  such that  $w \notin L_e \cup R_e$ . Then  $M_w \subseteq \{x_{\ell_v+1}^e, \dots, x_{r_v-1}^e\} \subseteq M_v$ , which is a contradiction. Hence,  $V_e = L_e \cup R_e$ .

To see that (iv) hold, notice that if  $L_e \cap R_e = \emptyset$ , then for every  $u \in N_G(L_e) \setminus V_e$ , we have that  $x_0^e \in M_u$ , and this immediately implies that  $N_G(L_e) \setminus V_e$  is a clique. For  $N_G(R_e) \setminus V_e$ , the claim holds by symmetry.

Finally, to prove (v), observe that if  $v \in L_e$  and  $w \in V(G) \setminus V_e$  is adjacent to  $v$ , then  $\{x_0^e, \dots, x_{v_\ell}^e\} \subseteq M_w$ . Therefore, if  $\ell_u \leq \ell_v$ , then  $M_v \cap M_w \neq \emptyset$  and, therefore,  $uw \in E(G)$ . Hence,  $N_G(v) \setminus V_e \subseteq N_G(u) \setminus V_e$ . The claim for vertices of  $R_e$  is again symmetric.  $\square$

We say that a vertex  $u \in V_e$  is an *e-cut-vertex* if one of the following holds:

- (i)  $u = v_1^e$  and  $v_2^e$  is not left-attached or, symmetrically,  $u = v_{p(e)}^e$  and  $v_{p(e)-1}^e$  is not right-attached.
- (ii)  $u$  is a cut-vertex of  $G_e$ .

Observe that if  $u = v_i^e$  is an *e-cut-vertex*, then  $v_{i+1}^e$  is not left-attached and  $v_{i-1}^e$  is not right-attached. Otherwise, we would get that either  $M_{v_i^e} \subseteq M_w$  or  $M_{p(e)}^e$  for some  $w \in V(G) \setminus V_e$ . This implies the following observation.

OBSERVATION 5. *The  $e$ -cut-vertices are exactly the cut-vertices of the intersection graph  $G_e^*$  of the family of all inclusion maximal nonempty connected subsets of  $M_v \cap \{x_0^e, \dots, x_{h(e)+1}^e\}$  for all  $v \in V(G)$ .*

We say that a  $(u, v)$ -path  $P$  in  $G_e$  is

- *straight* if  $u = v_1^e$  and  $v = v_{p(e)}^e$ ,
- a *left  $U$ -path* if either  $P = v_1^e$  or  $\{u, v\} = \{v_1^e, v_2^e\} \subseteq L_e$ ,
- a *right  $U$ -path* if either  $P = v_{p(e)}^e$  or  $\{u, v\} = \{v_{p(e)-1}^e, v_{p(e)}^e\} \subseteq R_e$ .

We also say that a pair of vertex disjoint paths—an  $(a_1, b_1)$ -path and an  $(a_2, b_2)$ -path—is a *straight pair* if  $\{a_1, a_2\} = \{v_1^e, v_2^e\} \subseteq L_e$  and  $\{b_1, b_2\} = \{v_{p(e)-1}^e, v_{p(e)}^e\} \subseteq R_e$ .

Let  $\mathcal{P}$  be a path cover of  $G$ . Denote by  $\mathcal{P}_e$  the family of paths of inclusion maximal subpaths of the paths  $P \in \mathcal{P}$  with all their vertices in  $V_e$ . We say that  $\mathcal{P}_e$  is the *projection* of  $\mathcal{P}$  on  $G_e$ . Since  $\mathcal{P}$  is a path cover of  $G$ ,  $\mathcal{P}_e$  is a path cover of  $G_e$ .

We say that an end-vertex  $v$  of a path  $P \in \mathcal{P}_e$  is *saturated* if  $P$  is a proper subpath of a path  $Q \in \mathcal{P}$  and  $v$  has a neighbor in  $Q$  that is outside of  $P$ . For the special case of a trivial single-vertex path, we say that its unique end-vertex is saturated if it has two distinct neighbors in  $Q$ . Let  $P$  be a proper subpath of a path  $Q$ . We say that the path  $Q'$  is obtained from  $Q$  by the *truncation with respect to  $P$*  if  $Q'$  is constructed by the deletion of the subsequence of the vertices of  $P$  from  $Q$ . Notice that this operation can be performed only in the following two cases: either if one of the end-vertices of  $P$  is an end-vertex of  $Q$ , that is,  $Q - V(P)$  is a path, or if all the vertices of  $P$  are internal vertices of  $Q$  but both neighbors of the end-vertices of  $P$  in  $Q$  are adjacent in  $G$ , that is,  $Q - V(P)$  is a disjoint union of two paths that are reconnected via the neighbors of the end-vertices of  $P$ .

Let  $P$  be an  $(u, v)$ -path and let  $P'$  be  $(u', v')$ -path such that  $P, P' \in \mathcal{P}_e$  for some  $e \in E(H)$ ,  $u, u' \in L_e$ ,  $v, v' \in R_e$ , and  $P$  and  $P'$  are subpaths of the same path  $Q \in \mathcal{P}$ . We say that  $P$  and  $P'$  are in the *same direction* in  $Q$  if either  $v$  and  $u'$  are vertices of the  $(u, v')$ -subpath of  $Q$  or, symmetrically,  $v'$  and  $u$  are vertices of the  $(u', v)$ -subpath of  $Q$ . If  $P$  and  $P'$  are not in the same direction, we say that  $P$  and  $P'$  are in *opposite directions* in  $Q$ .

We say that a path cover  $\mathcal{P}$  of  $G$  is *tamed* if the following hold for every  $e \in E(H)$  such that  $|V_e| \geq 3$  and  $L_e \cap R_e = \emptyset$ :

- If  $G_e$  is 2-connected, then one of the following is fulfilled:
  - $\mathcal{P}_e$  consists of a single straight path (see Figure 2(a)).
  - $\mathcal{P}_e$  consists of a single left  $U$ -path or, symmetrically, a single right  $U$ -path and the end-vertices of the path are saturated (see Figure 2(c)).
  - $\mathcal{P}_e$  is a straight pair of paths, whose end-vertices are saturated, and these paths are subpaths of the same path  $Q \in \mathcal{P}$  and are in opposite directions in  $Q$  (see Figure 2(b)).
- If  $G_e$  has a cut-vertex, then one of the following is fulfilled:
  - $\mathcal{P}_e$  contains a single straight path (see Figure 2(a)).
  - $\mathcal{P}_e$  contains two paths, whose end-vertices are saturated, such that one is a left  $U$ -path and the other is a right  $U$ -path whose end-vertices are saturated (see Figure 2(d)).

The definition of a tamed path cover is rather technical as it is tailored for use in the proof of our main result about PATH COVER on proper  $H$ -graphs, but essentially it says that the paths that cover  $G_e$  behave as shown in Figure 2. In particular, the vertices of  $G_e$  are covered by at most two paths.

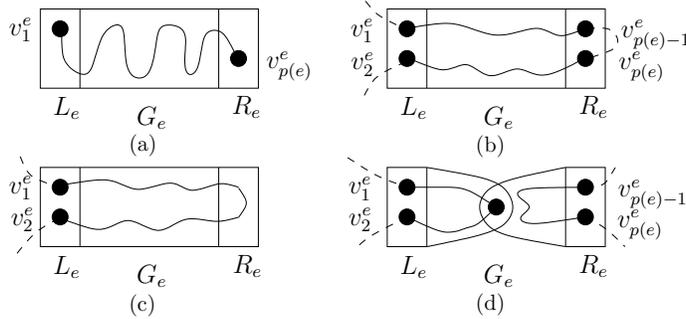


FIG. 2. The types of covering (up to symmetry) of  $G_e$  by a tamed path cover.

LEMMA 11. Let  $G$  be a proper  $H$ -graph such that  $G$  has a nice proper  $H$ -representation. Now if  $G$  has a path cover of size at most  $k$ , then  $G$  has a tamed path cover of size at most  $k$ .

*Proof.* Let  $G$  be a proper  $H$ -graph and let  $(H', \mathcal{M})$  with  $\mathcal{M} = \{M_v\}_{v \in V(G)}$  be a nice proper  $H$ -representation of  $G$ . Let  $\mathcal{P}$  be a path cover of  $G$  of size at most  $k$  such that

- (i) the total number of paths in the projections  $\mathcal{P}_e$  for  $e \in E(H)$  is minimum, and
- (ii) subject to (i), the total number of paths in the projections such that their end-vertices are either the leftmost and the rightmost vertices, or two leftmost vertices, or two rightmost vertices of  $G_e$  for each  $e$ , is maximum.

We claim that  $\mathcal{P}$  is tamed.

Let  $e \in E(H)$  such that  $|V_e| \geq 3$  and  $L_e \cap R_e = \emptyset$ . We prove the following series of claims about  $\mathcal{P}_e$  and show that the structure of paths in  $\mathcal{P}_e$  satisfies the definition of tamed path cover.

CLAIM 5.1. If there is a  $(u, v)$ -path  $P \in \mathcal{P}_e$  such that  $u \in L_e$  and  $v \in R_e$ , then either  $\mathcal{P}_e$  consists of a single straight path or  $\mathcal{P}_e$  is a straight pair of paths, whose end-vertices are saturated, and these paths are subpaths of the same path  $Q \in \mathcal{P}$  and are in opposite directions in  $Q$ .

*Proof of Claim 5.1.* Suppose that  $\mathcal{P}_e = \{P\}$ . Then to prove that claim, we have to show that  $P$  is straight.

To obtain a contradiction, assume that this is not the case, that is,  $\{u, v\} \neq \{v_1^e, v_{p(e)}^e\}$ . Then we can replace  $P$  in  $\mathcal{P}_e$  by a Hamiltonian  $(v_1^e, v_{p(e)}^e)$ -path  $\hat{P}$  in  $G_e$  that exists by Lemmas 4 and 10(i). The replacement is possible by Lemma 10(v). Observe that the replacement is local in the sense that the other projections  $\mathcal{P}_{e'}$  for  $e' \neq e$  remain the same. However, by this replacement we increase the number of paths with leftmost and rightmost end-vertices in the projections, and this leads to a contradiction of condition (ii) of the choice of  $\mathcal{P}$ . Therefore,  $P$  is straight.

Thus, if  $\mathcal{P}_e = \{P\}$ , then the claim holds. From here, assume that  $|\mathcal{P}_e| \geq 2$ . We now prove that  $\mathcal{P}_e$  is a straight pair of paths, whose end-vertices are saturated, and these paths are subpaths of the same path  $Q \in \mathcal{P}$  such that the paths of  $\mathcal{P}_e$  are in opposite directions.

The first step is to show that every path  $P' \in \mathcal{P}_e$  is a  $(u', v')$ -path for some  $u' \in L_e$  and  $v' \in R_e$  such that  $u'$  and  $v'$  are saturated.

To obtain a contradiction, assume first that  $\mathcal{P}_e$  contains a  $(w, w')$ -path  $P' \neq P$  such that  $w \in L_e$  and either  $w' \notin R_e$  or  $w' \in R_e$  but is not saturated. Then we modify  $\mathcal{P}$  as follows. By Lemma 10(iv), if  $P'$  is a proper subpath of a path  $Q' \in \mathcal{P}$ ,  $Q'$  can be truncated with respect to  $P'$ : if  $w' \in L_e$  and both  $w$  and  $w'$  are saturated, then the neighbors of  $w$  and  $w'$  in  $Q'$  that are outside of  $P'$  are adjacent. Then we replace  $Q'$  by the truncated path. If  $P'$  is a path of  $\mathcal{P}$ , we simply delete  $P'$  from  $\mathcal{P}$ . Let  $U = V(P) \cup V(P')$ . By Lemma 10(ii),  $U \subseteq V_e$  is a connected set. Let  $a$  and  $b$  be, respectively, the leftmost and rightmost vertices of  $U$ . By Lemma 4,  $G[U]$  has a Hamiltonian  $(a, b)$ -path  $\hat{P}$ . By the choice of  $a$  and  $b$  and Lemma 10(v), we can replace  $P$  by  $\hat{P}$  in  $Q \in \mathcal{P}$  if  $P$  is a proper subpath of  $Q$ . If  $P \in \mathcal{P}$ , we replace  $P$  by  $\hat{P}$  in  $\mathcal{P}$ . It is straightforward to see that by these modifications we obtain the path cover  $\hat{\mathcal{P}}$  of  $G$  such that the number of paths in the projection  $\hat{\mathcal{P}}_e$  has decreased and the paths in the other projections are staying intact. This contradicts condition (i) of the choice of  $\mathcal{P}$ . We conclude that  $\mathcal{P}_e$  has no  $(w, w')$ -path such that  $w \in L_e$  and either  $w' \notin R_e$  or  $w' \in R_e$  but is not saturated. By symmetry, we also have that  $\mathcal{P}_e$  has no  $(w, w')$ -path such that  $w \in R_e$  and either  $w' \notin L_e$  or  $w' \in L_e$  but is not saturated.

Suppose now that  $\mathcal{P}_e$  contains a  $(w, w')$ -path  $P' \neq P$  such that  $w, w' \notin L_e \cup R_e$ . We use similar arguments to modify  $\mathcal{P}$ . As  $w, w' \notin L_e \cup R_e$ ,  $P' \in \mathcal{P}$  and we simply delete this path. Let  $U = V(P) \cup V(P')$ . It is easy to see that  $U \subseteq V_e$  is a connected set. Let  $a$  and  $b$  be, respectively, the leftmost and rightmost vertices of  $U$ . Then we have a Hamiltonian  $(a, b)$ -path in  $G[U]$  that is used to modify the path of  $\mathcal{P}$  that either contains  $P$  as a proper subpath or coincides with  $P$ . This leads to a contradiction with the choice of  $\mathcal{P}$ . We conclude that  $\mathcal{P}$  has no such paths.

We obtain that if  $\mathcal{P}_e$  contains  $P' \neq P$ , then  $P'$  is a  $(u', v')$ -path for some  $u' \in L_e$  and  $v' \in R_e$  and the end-vertices of  $P'$  are saturated. Since  $P'$  is a  $(u', v')$ -path for some  $u' \in L_e$  and  $v' \in R_e$ , we have, by symmetry between  $P$  and  $P'$ , that the end-vertices of  $P$  are saturated as well.

Now we have that every path in  $\mathcal{P}_e$  joins a vertex in  $L_e$  with a vertex of  $R_e$  and the end-vertices of every path are saturated. Further, we will show that the paths of  $\mathcal{P}_e$  are subpaths of the same path of  $\mathcal{P}$ .

Suppose, for the sake of contradiction, that there are distinct  $P', P'' \in \mathcal{P}_e$  such that  $P'$  and  $P''$  are proper subpaths of distinct paths  $Q$  and  $Q'$  of  $\mathcal{P}$ , respectively. Assume that  $P'$  is a  $(u', v')$ -path and  $P''$  is a  $(u'', v'')$ -path, respectively, for  $u', u'' \in L_e$  and  $v', v'' \in R_e$ . Then we can write that  $Q = a_1 \cdots a_{i-1} P' a_i \cdots a_r$  and  $Q' = b_1 \cdots b_{j-1} P'' b_j \cdots b_t$  for some vertices  $a_1, \dots, a_r$  and  $b_1, \dots, b_t$  such that  $a_{i-1}, b_{j-1} \in N_G(L_e) \setminus V_e$ . Note that  $a_{i-1}, b_{j-1}$  are adjacent by Lemma 10(iv) and  $u', u''$  are adjacent by Lemma 10(ii). Then we can replace  $Q$  and  $Q'$  by the path  $a_1 \cdots a_{i-1} b_{j-1} \cdots b_1$  and  $a_r \cdots a_i (P')^{-1} P'' b_j \cdots b_t$ , where  $(P')^{-1}$  is the path  $P'$  traversed in the opposite direction. Then the projection of the obtained path cover  $\hat{\mathcal{P}}$  of  $G$  contains the path  $(P')^{-1} P''$  instead of  $P'$  and  $P''$ . Again, we obtain the path cover such that the number of paths in the projection  $\hat{\mathcal{P}}_e$  has decreased and the paths in the other projections are staying intact—a contradiction of the choice of  $\mathcal{P}$ . Hence,  $P'$  and  $P''$  are subpaths of the same path of  $\mathcal{P}$ . This proves that the paths of  $\mathcal{P}_e$  are subpaths of the same  $Q$  path of  $\mathcal{P}$ .

The next step is to prove that the paths of  $\mathcal{P}_e$  are in opposite directions in  $Q$ .

Again, the proof is by contradiction. Suppose that there are distinct  $P', P'' \in \mathcal{P}_e$  that are in the same direction. By symmetry, we can assume that  $Q = a_1 \cdots a_{i-1} P' a_i \cdots a_{j-1} P'' a_{j+1} \cdots a_t$  for some vertices  $a_1, \dots, a_t \in V(G)$ . Assume that  $P'$  is a  $(u', v')$ -path and  $P''$  is a  $(u'', v'')$ -path, respectively, for  $u', u'' \in L_e$  and  $v', v'' \in R_e$ . By Lem-

mas 10(ii) and 10(iv), we have that  $u', u''$  are adjacent and  $a_{i-1}, a_{j-1}$  are adjacent. We reroute  $Q$  and replace it by  $Q' = a_1 \cdots a_{i-1} a_{j-1} \cdots a_{i+1} (P')^{-1} P'' a_{j+1} \cdots a_t$ . As before, the replacement decreases the number of paths in the projections, leading to a contradiction.

Observing that since every two paths are in opposite directions in  $Q$ , then  $|\mathcal{P}_e| = 2$ . Indeed, if  $|\mathcal{P}_e| \geq 3$ , then there are two paths that are in the same direction, contradicting the proved property.

Summarizing, we conclude that if  $|\mathcal{P}_e| \geq 2$ , then  $|\mathcal{P}_e| = 2$  and  $\mathcal{P}_e$  is a straight pair of paths, whose end-vertices are saturated, and these paths are subpaths of the same path  $Q \in \mathcal{P}$  and are in opposite directions in  $Q$ . Let  $\mathcal{P}_e = \{P, P'\}$  and assume that  $P'$  is a  $(u', v')$ -path for  $u \in L_e$  and  $v' \in R_e$ . Recall that  $u, u'$  and  $v, v'$  are adjacent by Lemma 10(ii). This implies that  $G_e$  is 2-connected. Therefore, to complete the proof of Claim 5.1, it remains to show that  $\{u, u'\} = \{v_1^e, v_2^e\}$  or  $\{v, v'\} = \{v_{p(e)-1}^e, v_{p(e)}^e\}$ .

To get a contradiction, assume that  $\{u, u'\} \neq \{v_1^e, v_2^e\}$  or  $\{v, v'\} \neq \{v_{p(e)-1}^e, v_{p(e)}^e\}$ . Recall that  $G_e$  is 2-connected. We also have that  $|V_e| \geq 4$  as  $u, u', v, v'$  are distinct. Then by Lemma 4  $G_e$  has a path cover of size two formed by  $(a_1, b_1)$ - and  $(a_2, b_2)$ -paths  $\hat{P}$  and  $\hat{P}'$  such that  $\{a_1, a_2\} = \{v_1^e, v_2^e\}$  and  $\{b_1, b_2\} = \{v_{p(e)-1}^e, v_{p(e)}^e\}$ . The paths  $P$  and  $P'$  are subpaths of the same path  $Q \in \mathcal{P}$ . Moreover,  $P$  and  $P'$  are in opposite directions in  $Q$ . This means that either  $(u, u')$ - or  $(v, v')$ -subpath  $Q'$  of  $Q$  has no inner vertices in  $V_e$ . By symmetry, assume without loss of generality that  $Q'$  is such a path with the end-vertices  $v$  and  $v'$  and denote by  $Q''$  the path obtained from  $Q'$  by the deletion of  $v$  and  $v'$ . Assume also that  $u$  is before  $u'$  in the ordering of the vertices of  $G_e$ . We replace  $P$  and  $P'$  by  $\hat{P}$  and  $\hat{P}'$  in  $Q$  as follows. We replace  $P$  by the path with its leftmost end-vertex  $v_1$  and  $P'$  by the path with its leftmost end-vertex  $v_2'$ . By Lemma 10(v), we have that one end-vertex of  $Q''$  is adjacent to the rightmost vertex of the first path and the second end-vertex of  $Q''$  is adjacent to the rightmost vertex of the second path. This implies that replacing  $P$  and  $P'$  by  $\hat{P}$  and  $\hat{P}'$  creates a path with the same vertices as  $Q$ . It remains to observe that this replacement increases the number of pairs of leftmost and rightmost end-vertices in the projections. This contradicts condition (ii) of the choice of  $\mathcal{P}$ . We conclude that  $\{u, u'\} = \{v_1^e, v_2^e\}$  or  $\{v, v'\} = \{v_{p(e)-1}^e, v_{p(e)}^e\}$ .  $\square$

By Claim 5.1, we have that if there is a  $(u, v)$ -path  $P \in \mathcal{P}_e$  such that  $u \in L_e$  and  $v \in R_e$ , then the paths of  $\mathcal{P}_e$  satisfy the conditions of the definition of a tamed path cover. Assume from now on that there is no  $(u, v)$ -path in  $\mathcal{P}_e$  with  $u \in L_e$  and  $v \in R_e$ .

CLAIM 5.2. *For every  $(u, v)$ -path  $P \in \mathcal{P}_e$ ,  $u$  and  $v$  are saturated.*

*Proof of Claim 5.2.* The proof is by contradiction. Suppose that there is a  $(u, v)$ -path  $P \in \mathcal{P}_e$  such that  $u$  or  $v$  is not saturated. We assume that  $\ell_u \leq \ell_v$ .

Suppose that  $\mathcal{P}_e$  contains some other  $(w, w')$ -path  $P'$ . Observe that if  $P'$  is a proper subpath of a path  $Q' \in \mathcal{P}$ , then  $Q'$  can be truncated with respect to  $P'$ . If  $w$  or  $w'$  is not saturated, then this claim is trivial. If both  $w$  and  $w'$  are saturated, then we have that either  $w, w' \in L_e$  or  $w, w' \in R_e$  and the truncation is possible by Lemma 10(iv). We use this observation and modify  $\mathcal{P}$  as follows. We truncate all the paths of  $\mathcal{P}$  that have proper subpaths in  $\mathcal{P}_e$  except the path containing  $P$  (if it exists) and delete the paths of  $\mathcal{P}_e \setminus \{P\}$  that are in  $\mathcal{P}$ . Then we replace  $P$  by a Hamiltonian  $(v_1^e, v_{p(e)}^e)$ -path  $\hat{P}$  of  $G_e$  that exists by Lemmas 4 and 10(i). The replacement is possible by Lemma 10(v). This replacement leads to a decrease in the

number of paths in the projection and contradicts the choice of  $\mathcal{P}$ . Therefore,  $P$  is the unique path of  $\mathcal{P}_e$ .

Notice that  $P$  cannot be a  $(v_1^e, v_{p(e)}^e)$ -path by the assumption that there is no  $(u, v)$ -path in  $\mathcal{P}_e$  with  $u \in L_e$  and  $v \in R_e$ . However, if  $P$  is not a  $(v_1^e, v_{p(e)}^e)$ -path, we can replace  $P$  by a Hamiltonian  $(v_1^e, v_{p(e)}^e)$ -path  $\hat{P}$  of  $G_e$  and increase the number of leftmost and rightmost end-vertices. This contradicts condition (ii) of the choice of  $\mathcal{P}$  and completes the proof of the claim.  $\square$

By Claim 5.2, the end-vertices of every path of  $\mathcal{P}_e$  are saturated. Recall also that we assume that the end-vertices are either in  $L_e$  or in  $R_e$ .

CLAIM 5.3. *If  $|\mathcal{P}_e| = 1$ , then  $G_e$  is 2-connected and  $\mathcal{P}_e$  consists of a single left  $U$ -path or, symmetrically, a single right  $U$ -path with saturated end-vertices.*

*Proof of Claim 5.3.* Let  $P \in \mathcal{P}_e$ , and let  $u$  and  $v$  be the end-vertices of  $P$ . By Claim 5.2,  $u$  and  $v$  are saturated. By symmetry, assume without loss of generality that  $u, v \in L_e$ . Notice that because  $R_e \neq \emptyset$  by Lemma 10(ii),  $u \neq v$ . By Lemma 10(ii),  $u$  and  $v$  are adjacent. Recall that  $|V_e| \geq 3$ . Then  $G_e$  is a 2-connected graph with at least 3 vertices.

Assume that  $\{u, v\} \neq \{v_1^e, v_2^e\}$ . By Lemma 10(ii),  $u$  and  $v$  are adjacent. Since  $|V_e| \geq 3$  as  $L_e \neq \emptyset$ ,  $G_e$  is a 2-connected graph with at least 3 vertices. By Lemma 4,  $G_e$  has a Hamiltonian  $(v_1^e, v_2^e)$ -path  $\hat{P}$ . Then by Lemma 10(v) we can replace  $P$  by  $\hat{P}$  and increase the number of pairs of leftmost end-vertices in the paths of the projections, contradicting the choice of  $\mathcal{P}$ . Therefore, we have that  $P$  is a  $(v_1^e, v_2^e)$ -path, and the claim holds.  $\square$

CLAIM 5.4. *If  $|\mathcal{P}_e| \geq 2$ , then  $G_e$  has a cut-vertex and  $\mathcal{P}_e$  contains two paths with saturated end-vertices such that one is a left  $U$ -path and the other is a right  $U$ -path.*

*Proof of Claim 5.4.* By symmetry, assume without loss of generality that  $P \in \mathcal{P}_e$  is a  $(u, v)$ -paths with  $u, v \in L_e$ . Let  $P' \in \mathcal{P}_e$  be a  $(u', v')$ -path distinct from  $P$ . By Claim 5.2, the end-vertices of both paths are saturated.

We show that  $u', v' \in R_e$ .

To obtain a contradiction, suppose that  $u', v' \in L_e$ . Assume without loss of generality that  $\ell_{u'} < \ell_u$ , as otherwise we can exchange  $P$  and  $P'$ . Then we modify  $\mathcal{P}$  as follows. We truncate the path of  $\mathcal{P}$  containing  $P'$  as a proper subpath using Lemma 10(iv). Let  $U = V(P) \cup V(P')$ . By Lemma 10(ii),  $U \subseteq V_e$  is a connected set. Moreover,  $G[U]$  is 2-connected. Let  $a$  and  $b$  be the leftmost vertices of  $U$ . Notice that  $G[U]$  has a Hamiltonian  $(a, b)$ -path  $\hat{P}$  by Lemma 4. By the choice of  $a, b$  and the assumption that  $\ell_{u'} < \ell_u$ , we can use Lemma 10(v) and replace  $P$  by  $\hat{P}$  in  $Q \in \mathcal{P}$  containing  $P$  as a proper subpath. These modifications decrease the number of paths in the projections and therefore contradict the choice of  $\mathcal{P}$ . We conclude that  $u', v' \in R_e$ .

By symmetry, this means that  $\mathcal{P}_e = \{P, P'\}$ , because every path in  $\mathcal{P}_e$  distinct from  $P'$  should have its end-vertices in  $L_e$ .

Because  $|V_e| \geq 3$ , either  $P$  or  $P'$  has at least two vertices. By symmetry, we assume without loss of generality that  $P$  has this property. In particular, this means that  $u \neq v$ .

Assume that  $G_e$  is 2-connected. Then we can modify  $\mathcal{P}$  as follows. First, we truncate the path of  $\mathcal{P}$  containing  $P'$ . By Lemma 4, we have that  $G_e$  has a Hamiltonian  $(v_1^e, v_2^e)$ -path  $\hat{P}$ . Because  $u, v$  are distinct vertices of  $P$ , we can replace the subpath  $P$  is the path  $Q \in \mathcal{P}$  containing  $P$  by Lemma 10(v). As before, these modifications

decrease the number of paths in the projections and lead to a contradiction. Hence,  $G_e$  has a cut-vertex.

By Lemma 5(ii), we have that  $v_1^e$  and  $v_2^e$  are vertices of  $P$  and  $v_{p(e)}^e$  is a vertex of  $P'$ . Suppose that  $\{u, v\} \neq \{v_1^e, v_2^e\}$ . Then we can modify  $P$ . By Lemma 10(ii),  $u$  and  $v$  are adjacent. Hence,  $G[V(P)]$  is a 2-connected graph. By Lemma 4, there is a  $(v_1^e, v_2^e)$ -path  $\hat{P}$  with the same vertices as  $P$ . Using Lemma 10(v) we can replace  $P$  in the path  $Q \in \mathcal{P}$  that contains  $P$  by  $\hat{P}$ . This increases the number of pairs of leftmost vertices in the projections and contradicts the choice of  $\mathcal{P}$ . Using symmetry, we conclude that  $P$  is a  $(v_1^e, v_2^e)$ -path and for  $P'$  it holds that it is either the trivial  $(v_{p(e)}^e, v_{p(e)}^e)$ -path or a  $(v_{p(e)-1}^e, v_{p(e)}^e)$ -path. This concludes the proof.  $\square$

Combining Claim 5.1 with Claims 5.3 and 5.4, we conclude that all the conditions of the definition of a tamed path cover are fulfilled for the paths of  $\mathcal{P}_e$ . We obtain that  $\mathcal{P}$  is a tamed path cover.  $\square$

Our next aim is to introduce *tamed* cycle covers.

Let  $\mathcal{C}$  be a cycle cover. We denote by  $\mathcal{C}_e$  the family of subgraphs of  $G_e$  that contains the cycles of  $\mathcal{C}$  that are cycles of  $G_e$  and the inclusion-maximal segments with their vertices in  $V_e$  of the cycles of  $\mathcal{C}$  that have vertices in both  $V_e$  and  $V(G) \setminus V_e$ . In the same way as with path covers, we say that  $\mathcal{C}_e$  is the *projection* of  $\mathcal{C}$  on  $G_e$ . Notice that  $\mathcal{C}_e$  may contain both cycles and paths. Clearly, the paths and cycles of  $\mathcal{C}_e$  cover the vertices of  $G_e$ .

Let  $C$  be a cycle and let  $P$  be a  $(u, v)$ -segment of  $C$  such that  $|V(C) \setminus V(P)| \geq 3$ . We say that the cycle  $C'$  is obtained from  $C$  by the *truncation with respect to  $P$*  if  $C'$  is constructed by the deletion of the vertices of  $P$  and making the neighbors of  $u$  and  $v$  in  $C$  that are outside  $P$  adjacent. Note that the truncation is only possible if the neighbors of  $u$  and  $v$  are adjacent in  $G$ .

Let  $P$  be a  $(u, v)$ -path and let  $P'$  be a  $(u', v')$ -path such that  $P, P' \in \mathcal{C}_e$  for some  $e \in E(H)$ ,  $u, u' \in L_e$ ,  $v, v' \in R_e$ , and  $P$  and  $P'$  are subpaths of the same cycle  $C \in \mathcal{C}$ . We say that  $P$  and  $P'$  are in the *same direction* in  $C$  if either  $v$  and  $u'$  are vertices of the  $(u, v')$ -segment of  $C$  containing both  $P$  and  $P'$  or, symmetrically,  $v'$  and  $u$  are vertices of the  $(u', v)$ -segment of  $C$  containing both  $P$  and  $P'$ . If  $P$  and  $P'$  are not in the same direction, we say that  $P$  and  $P'$  are in *opposite directions* in  $C$ .

Recall that by Lemma 5 the block-cutpoint decomposition graph of each  $G_e$  is a path and the blocks of  $G_e$  could be ordered according to the ordering of its vertices. We say that a block  $B$  is the *leftmost* block if it contains  $v_1^e$  and  $B$  is the *rightmost* block if it contains  $v_{p(e)}^e$ .

We say that a cycle cover  $\mathcal{C}$  of  $G$  is *tamed* if the following hold for every  $e \in E(H)$  such that  $|V_e| \geq 3$  and  $L_e \cap R_e = \emptyset$ :

- If  $G_e$  is 2-connected, then one of the following is fulfilled:
  - $\mathcal{C}_e$  consists of a single straight path (see Figure 3(a)).
  - $\mathcal{C}_e$  is a straight pair of paths such that these paths are segments of the same cycle  $C$  of  $\mathcal{C}$  and are in opposite directions (see Figure 3(b)).
  - $\mathcal{C}_e$  consists of a single left  $U$ -path or, symmetrically, a single right  $U$ -path (see Figure 3(c)).
  - $\mathcal{C}_e$  consists of a left  $U$ -path and the trivial right  $U$ -path or, symmetrically, a right  $U$ -path and the trivial left  $U$ -path (see Figure 3(d)).
  - $\mathcal{C}_e$  consists of a single cycle and, possibly, the trivial left and/or trivial right  $U$ -path (see Figures 3(e)–(g)).
- If  $G_e$  has a cut-vertex, then either  $\mathcal{C}_e$  consists of a single straight path or the following hold (see Figures 3(h) and 3(i)):

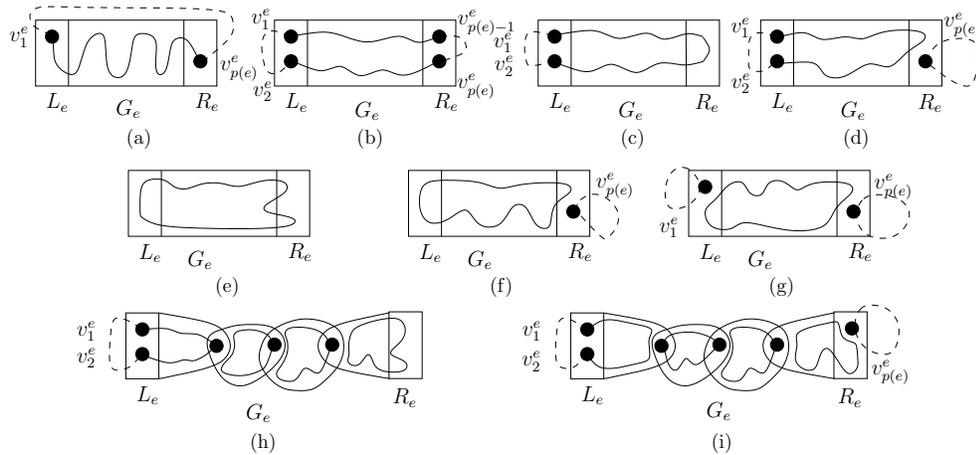


FIG. 3. The types of covering (up to symmetry) of  $G_e$  by a tamed cycle cover.

- The paths in  $\mathcal{C}_e$  are left and right  $U$ -paths,  $\mathcal{C}_e$  contains at most one left  $U$ -path and at most one right  $U$ -path, and these paths are paths in the leftmost and the rightmost blocks, respectively.
- Each block contains at most one cycle of  $\mathcal{C}_e$ .
- If the leftmost block contains a nontrivial left  $U$ -path, then it does not contain a cycle of  $\mathcal{C}_e$ , and, symmetrically, if the rightmost block contains a nontrivial right  $U$ -path, then it does not contain a cycle of  $\mathcal{C}_e$ .

The definition is overloaded by technical details to make it easier to use tamed cycle covers in the proof of our compression result, but the structure of paths and cycles in  $\mathcal{C}_e$  for a tamed cycle cover  $\mathcal{C}$  can be seen in Figure 3.

LEMMA 12. *Let  $G$  be a proper  $H$ -graph such that  $G$  has a nice proper  $H$ -representation. Then if  $G$  has a cycle cover of size at most  $k$ , then  $G$  has a tamed cycle cover of size at most  $k$ . Moreover, given a cycle cover  $\mathcal{C}$ , one can find a tamed cycle cover  $\hat{\mathcal{C}}$  of size at most  $|\mathcal{C}|$  such that for every  $e \in E(H)$  with  $|V_e| \geq 3$  and  $L_e \cap R_e = \emptyset$  if  $G_e$  has a cut-vertex and  $\mathcal{C}_e$  contains a  $(u, v)$ -path with  $u \in L_e$  and  $v \in R_e$ , then  $\hat{\mathcal{C}}_e$  consists of a single straight path.*

*Proof.* The proof uses the same approach as the proof of Lemma 11. Let  $G$  be a proper  $H$ -graph and let  $(H', \mathcal{M})$  with  $\mathcal{M} = \{M_v\}_{v \in V(G)}$  be a nice proper  $H$ -representation of  $G$ . Let  $\mathcal{C}$  be a cycle cover of  $G$  of size at most  $k$ . Denote by  $S$  the set of edges of  $H$  such that the following conditions  $(*)$  hold for every  $e \in S$ :

- $|V_e| \geq 3$ .
- $L_e \cap R_e = \emptyset$ .
- $G_e$  has a cut-vertex.
- $\mathcal{C}_e$  contains a  $(u, v)$ -path with  $u \in L_e$  and  $v \in R_e$ .

- Assume that  $\mathcal{C}$  is a cycle cover of  $G$  such that  $(*)$  are fulfilled for the edges of  $S$  and
- (i) the total number of paths and cycles in the projections  $\mathcal{C}_e$  for  $e \in E(H)$  is minimum,
  - (ii) subject to (i), the number of paths with one end-vertex in  $L_e$  and the second in  $R_e$  is maximum, and
  - (iii) subject to (i) and (ii), the total number of paths in the projections such that their end-vertices are either the leftmost and the rightmost vertices, or two

leftmost vertices or two rightmost vertices of  $G_e$  for each  $e$ , is maximum. We claim that  $\mathcal{C}$  is a tamed cycle cover such that for every  $e \in S$ ,  $G_e$  is covered by a straight path.

Let  $e \in E(H)$  such that  $|V_e| \geq 3$  and  $L_e \cap R_e = \emptyset$ . Notice that every path in  $\mathcal{C}_e$  has its end-vertices in  $L_e \cup R_e$  as they are segments of cycles. Similarly to the proof of Lemma 11, we prove a series of claims showing that the required conditions hold for  $\mathcal{C}_e$ .

CLAIM 5.5. *If there is a  $(u, v)$ -path  $P \in \mathcal{C}_e$  such that  $u \in L_e$  and  $v \in R_e$ , then either  $\mathcal{C}_e$  consists of a single straight path or  $\mathcal{C}_e$  is a straight pair of paths such that these paths are segments of the same cycle  $C$  of  $\mathcal{C}$  and are in opposite directions in  $C$ . Moreover, in the second case,  $G_e$  is 2-connected and  $e \notin S$ .*

*Proof of Claim 5.5.* Suppose that  $\mathcal{C}_e = \{P\}$ . We claim that  $P$  is straight in this case.

Suppose that this is not the case, that is,  $\{u, v\} \neq \{v_1^e, v_{p(e)}^e\}$ . Then exactly as in the proof of Claim 5.1, we can replace  $P$  by a Hamiltonian  $(v_1^e, v_{p(e)}^e)$ -path  $\hat{P}$  in  $G_e$  that exists by Lemmas 4 and 10(i). The replacement is possible by Lemma 10(v). Note that by this replacement we increase the number of paths with leftmost and rightmost end-vertices in the projections, and this leads to a contradiction of condition (ii) of the choice of  $\mathcal{C}$ .

This proves the claim if  $\mathcal{C}_e = \{P\}$ . Notice, in particular, that if  $G_e$  has a cut-vertex, that is,  $G_e$  satisfies (\*), we have that  $G_e$  is covered by a straight path.

Assume from now on that  $|\mathcal{C}_e| \geq 2$ . We will show that  $\mathcal{C}_e$  consists of paths with one end-vertex in  $L_e$  and the other in  $R_e$ .

Suppose that  $\mathcal{P}_e$  contains a  $(w, w')$ -path  $P'$  distinct from  $P$ . Since  $P'$  is a segment of a cycle of  $\mathcal{C}$ ,  $w, w' \in L_e \cup R_e$ . We show that one end-vertex of  $P'$  is in  $L_e$  and the other in  $R_e$ .

To obtain a contradiction, assume by symmetry that  $w, w' \in L_e$ . If  $P'$  is a segment of a cycle  $C' \in \mathcal{C}$  that has at least 3 vertices outside  $P'$ , then we modify  $\mathcal{C}$  as follows. By Lemma 10(iv),  $C'$  can be truncated with respect to  $P'$  as the neighbors of the end-vertices of  $P'$  in  $C'$  that are outside  $P'$  are adjacent. We replace  $C'$  by the truncated cycle. Let  $U = V(P) \cup V(P')$ . By Lemma 10(ii),  $U \subseteq V_e$  is a connected set. Let  $a$  and  $b$  be, respectively, the leftmost and rightmost vertices of  $U$ . By Lemma 4,  $G[U]$  has a Hamiltonian  $(a, b)$ -path  $\hat{P}$ . By the choice of  $a$  and  $b$  and Lemma 10(v), we can replace  $P$  by  $\hat{P}$  in the cycle containing  $P$ . Clearly, we obtain the cycle cover  $\hat{\mathcal{C}}$  of  $G$  satisfying (\*) such that the number of paths in the projection  $\hat{\mathcal{C}}_e$  has decreased and the paths and cycles in the other projections are staying intact. This contradicts condition (i) of the choice of  $\mathcal{C}$ . Assume that  $P'$  is a segment of a cycle  $C' \in \mathcal{C}$  that has at most 2 vertices outside  $P'$  and denote the neighbors of  $w$  and  $w'$  in  $C'$  that are outside  $P'$  by  $t$  and  $t'$ , respectively. Notice that it may happen that  $t = t'$ . Observe also that  $C' = tP't$  if  $t = t'$  and  $C' = t'tP't'$  otherwise. Also, let  $s$  be the neighbor of  $u$  in the cycle  $C \in \mathcal{C}$  containing  $P$  that is outside  $P$ . Clearly,  $C \neq C'$ . We also have that  $s, t, t' \in N_G(L_e) \setminus V_e$ , and therefore these vertices are pairwise adjacent by Lemma 10(iv). Let  $U = V(P) \cup V(P')$ . In the same way as above,  $G[U]$  has a Hamiltonian  $(a, b)$ -path  $\hat{P}$ . By the choice of  $a$  and  $b$  and Lemma 10(v), we can now delete  $C'$  from  $\mathcal{C}$  and replace  $P$  by  $t't\hat{P}$  or by  $t\hat{P}$ , depending on whether  $t = t'$ . Again, we obtain the cycle cover  $\hat{\mathcal{C}}$  of  $G$  satisfying (\*) such that the number of paths in the projection  $\hat{\mathcal{C}}_e$  has decreased and the paths and cycles in the other projections are staying intact, contradicting the choice of  $\mathcal{C}$ . We conclude that  $\mathcal{C}_e$  has no  $(w, w')$ -path

with  $w, w' \in L_e$  and, by symmetry,  $\mathcal{C}_e$  has no  $(w, w')$ -path with  $w, w' \in R_e$ .

Now we prove that  $\mathcal{C}_e$  does not contain cycles. For the sake of contradiction, suppose that  $\mathcal{C}_e$  contains a cycle  $C'$ . Then we can modify  $\mathcal{C}$  as follows. First, we delete  $C'$  from  $\mathcal{C}$ . Let  $U = V(P) \cup V(C')$ . Clearly,  $U \subseteq V_e$  is a connected set. Let  $a$  and  $b$  be, respectively, the leftmost and rightmost vertices of  $U$ . By Lemma 4,  $G[U]$  has a Hamiltonian  $(a, b)$ -path  $\hat{P}$ . By the choice of  $a$  and  $b$  and Lemma 10(v), we can replace  $P$  by  $\hat{P}$  in the cycle containing  $P$ . Using the same arguments as before, we obtain a contradiction with the choice of  $\mathcal{C}$ .

We conclude that  $\mathcal{C}_e$  consists of paths with one end-vertex in  $L_e$  and the other in  $R_e$ . Since  $|\mathcal{C}_2| \geq 2$ , we have that  $G_e$  is 2-connected and, in particular,  $e \notin S$ .

Further, we show that all the paths in  $\mathcal{C}_e$  are segments of the same cycle of  $\mathcal{C}$ .

As before, the proof is by contradiction. Assume that  $\mathcal{C}_e$  contains a  $(u', v')$ -path  $P'$  that is a segment of a cycle  $C' \in \mathcal{C}$  that is distinct from the cycle  $C$  containing  $P$ . Then we can “glue”  $C$  and  $C'$  together. We have that  $u$  and  $u'$  are in the clique  $L_e$  by Lemma 10(ii) and the neighbors  $w$  and  $w'$  of  $u$  and  $u'$  in  $C$  and  $C'$  that are outside  $P$  and  $P'$ , respectively, are adjacent by Lemma 10(iv). We delete the edges  $wu$  and  $w'u'$  in  $C$  and  $C'$  and replace them by  $ww'$  and  $uu'$ . This way we reduce the number of paths in the projections in the obtained cycle cover, contradicting the choice of  $\mathcal{C}$ .

Hence, all the paths in  $\mathcal{C}_e$  are segments of the same cycle  $C \in \mathcal{C}$ .

Next, we prove that every two paths of  $\mathcal{C}_e$  are in opposite directions in  $C$ .

To get a contradiction, suppose that there are distinct  $P', P'' \in \mathcal{C}_e$  that are in the same direction in  $C$ , that is, by symmetry,  $C$  can be written as  $a_0 a_1 \cdots a_{i-1} P' a_i \cdots a_{j-1} P'' a_j \cdots a_s$ , where  $a_0, \dots, a_s$  are vertices of  $G$  and  $a_0 = a_s$ . Assume that  $P'$  and  $P''$  are  $(u', v')$ - and  $(u'', v'')$ -paths, respectively, for  $u', u'' \in L_e$  and  $v', v'' \in R_e$ . Note that  $a_{i-1} a_{j-1}, u' u'' \in E(G)$  by Lemmas 10(ii) and 10(iv). Then  $C$  can be replaced by  $C' = a_0 \cdots a_{i-1} a_{j-1} \cdots a_i (P')^{-1} P'' a_j \cdots a_s$  in the cycle cover. Clearly,  $C'$  is a cycle and  $V(C') = V(C)$ . Then the obtained cycle cover has fewer paths in the projections, contradicting the choice of  $\mathcal{C}$ . Hence,  $P'$  and  $P''$  are in opposite directions in  $C$ , that is,  $C$  has  $(u', u'')$ - and  $(v', v'')$ -segments with the inner vertices outside  $V(P) \cup V(P')$ .

Observe that, in particular, this implies that  $|\mathcal{C}_e| = 2$ . Otherwise, if  $\mathcal{C}_e$  contains three paths, then we have that at least two of them are in the same direction in a cycle of  $\mathcal{C}$ , and this is impossible, as was shown above.

We obtain that  $\mathcal{C}_e$  consists of  $P$  and some  $(u', v')$ -path  $P'$  such that  $P, P'$  are segments of the same cycle  $C \in \mathcal{C}$ . Moreover,  $P$  and  $P'$  are in opposite directions in  $C$ . To show that  $\{P, P'\}$  is a straight pair, we prove that  $\{u, u'\} = \{v_1^e, v_2^e\}$  and  $\{v, v'\} = \{v_{p(e)-1}^e, v_{p(e)}^e\}$ .

Suppose that  $\{u, u'\} \neq \{v_1^e, v_2^e\}$  or  $\{v, v'\} \neq \{v_{p(e)-1}^e, v_{p(e)}^e\}$ , that is,  $P$  and  $P'$  do not compose a straight pair. Recall that  $u, u'$  and  $v, v'$  are adjacent by Lemma 10(ii). This implies that  $G_e$  is 2-connected. We also have that  $|V_e| \geq 4$  as  $u, u', v, v'$  are distinct. Then by Lemma 4  $G_e$  has a path cover of size two formed by  $(a_1, b_1)$ - and  $(a_2, b_2)$ -paths  $\hat{P}$  and  $\hat{P}'$  such that  $\{a_1, a_2\} = \{v_1^e, v_2^e\}$  and  $\{b_1, b_2\} = \{v_{p(e)-1}^e, v_{p(e)}^e\}$ . Recall that  $P$  and  $P'$  are subpaths of the same cycle  $C \in \mathcal{C}$ . Moreover,  $C$  has  $(u, u')$ - and  $(v, v')$ -segments with the inner vertices outside  $V_e = V(P) \cup V(P')$ . Denote these segments by  $Q$  and  $Q'$ . Let  $\hat{Q}$  and  $\hat{Q}'$  be the paths obtained from  $Q$  and  $Q'$  by the deletion of their end-vertices. By Lemma 10(v), we have that one end-vertex of  $\hat{Q}$  is adjacent to  $a_1$  and the other to  $a_2$ . Similarly, we can order the end-vertices of  $\hat{Q}'$  in such a way that the first end-vertex is adjacent to  $b_1$  and the second to  $b_2$ . It follows

that we can combine  $P, P', \hat{Q}$ , and  $\hat{Q}'$  into a cycle with the same vertices as  $C$ . It remains to observe that this replacement increases the number of pairs of leftmost and rightmost end-vertices in the projections—a contradiction.

We conclude that  $\mathcal{C}_e$  is a straight pair of paths such that these paths are segments of the same cycle  $C$  of  $\mathcal{C}$  and are in opposite directions in  $C$ . Moreover,  $G_e$  is 2-connected and  $e \notin S$ .  $\square$

Claim 5.5 proves the desired properties for  $\mathcal{C}_e$  if the projection has a  $(u, v)$ -path with  $u \in L_e$  and  $v \in R_e$ . Assume from now on that there is no  $(u, v)$ -path in  $\mathcal{C}_e$  with  $u \in L_e$  and  $v \in R_e$ .

We need the following auxiliary observation.

CLAIM 5.6. *The projection  $\mathcal{C}_e$  contains at most one path with the end-vertices in  $L_e$  and at most one path with the end-vertices in  $R_e$ .*

*Proof of Claim 5.6.* To obtain a contradiction, assume that there are  $P, P' \in \mathcal{C}_e$  such that  $P$  is a  $(u, v)$ -path and  $P'$  is a  $(u', v')$ -path for  $u, v, u', v' \in L_e$ ; it may happen that  $u = v$  or  $u' = v'$  as the paths may be trivial. Denote by  $C$  and  $C'$  the cycles of  $\mathcal{C}$  that contain  $P$  and  $P'$ , respectively; it may happen that  $C = C'$ . Denote by  $s$  and  $t$  the neighbors of  $u$  and  $v$  in  $C$  that are outside  $P$ . Similarly, let  $s'$  and  $t'$  be the neighbors of  $u'$  and  $v'$  in  $C'$  that are outside  $P'$ . Some of these vertices  $s, t, s', t'$  could be the same. Note that  $s, t, s', t' \in N_G(L_e) \setminus V_e$  and, by Lemma 10(iv), every two vertices are either the same or adjacent. Observe also that by Lemma 10(ii),  $u, v, u', v'$  are pairwise adjacent.

Suppose that  $C$  and  $C'$  are distinct. Then we can “glue” them together into one cycle: delete the edges  $su$  and  $s'u'$  and replace them by  $ss'$  and  $uu'$ . This way we obtain a cycle cover with fewer paths in the projections, contradicting the choice of  $\mathcal{C}$ . Hence,  $C = C'$ . In particular, it means that  $C$  has at least 3 vertices outside  $P'$ . We reroute  $C$  as follows. We truncate  $C$  with respect to  $P'$ . Let  $U = V(P) \cup V(P')$ . Notice that  $G[U]$  is 2-connected. Let  $a$  and  $b$  be the pair of leftmost vertices of  $U$ . By Lemma 4,  $G[U]$  has a Hamiltonian  $(a, b)$ -path  $\hat{P}$ . By the choice of  $a$  and  $b$  and Lemma 10(v), we can replace  $P$  by  $\hat{P}$  in  $C$ . We decrease the number of paths in the projections by modifying  $\mathcal{C}$ , and this contradicts the choice of  $\mathcal{C}$ . This means that  $\mathcal{C}_e$  contains at most one path with its end-vertices in  $L_e$ . The claim for  $R_e$  is symmetric.  $\square$

Our next aim is to show the lemma for the case when  $G_e$  is 2-connected but  $\mathcal{C}_e$  has no path with one end-vertex in  $L_e$  and the other in  $R_e$ .

CLAIM 5.7. *If  $G_e$  is 2-connected, then one of the following holds:*

- $\mathcal{C}_e$  consists of a single left  $U$ -path or, symmetrically, a single right  $U$ -path.
- $\mathcal{C}_e$  consists of a left  $U$ -path and the trivial right  $U$ -path or, symmetrically, a right  $U$ -path and the trivial left  $U$ -path.
- $\mathcal{C}_e$  consists of a single cycle and, possibly, the trivial left and/or right  $U$ -path.

*Proof of Claim 5.7.* Suppose that  $\mathcal{C}_e$  contains two nontrivial paths: a  $(u, v)$ -path  $P$  and a  $(u', v')$ -path  $P'$ . By Claim 5.6, we can assume that  $u, v \in L_e$  and  $u', v' \in R_e$ . Let  $C, C' \in \mathcal{C}$  be the cycles containing  $P$  and  $P'$ , respectively; it can happen that  $C = C'$ .

Assume that  $C \neq C'$ . Then we modify  $\mathcal{C}$  as follows. First, we delete every cycle in  $\mathcal{C}_e$ . Recall that  $G_e$  is 2-connected and has at least 3 vertices. By Lemma 4, it has a path cover formed by an  $(a_1, b_1)$ -path  $Q$  and an  $(a_2, b_2)$ -path  $Q'$  such that  $\{a_1, a_2\} = \{v_1^e, v_2^e\}$  and  $\{b_1, b_2\} = \{v_{p(e)-1}^e, v_{p(e)}^e\}$ . We delete  $P$  and  $P'$  from  $C$  and

$C'$  and replace them by  $Q$  and  $Q'$ . Lemma 10(v) guarantees that this way we can construct a new cycle  $\tilde{C}$  by traversing the segments of  $C$  and  $C'$  that are outside  $P$  and  $P'$ , respectively, in an appropriate direction. This way we either decrease the number of cycles in the projections or increase the number of paths in the projection having one end-vertex on  $L_e$  and the second in  $R_e$ , contradicting the choice of  $\mathcal{C}$ .

Suppose that  $C = C'$ . In particular, this means that  $C$  has at least 3 vertices outside  $P'$ . By Lemma 10(iv),  $C$  can be truncated with respect to  $P'$ . We modify  $\mathcal{C}$  as follows. First, we delete every cycle in  $\mathcal{C}_e$ . Recall that  $G_e$  is 2-connected and has at least 3 vertices. By Lemma 4, it has a Hamiltonian  $(v_1^e, v_2^e)$ -path  $Q$ . We truncate  $C$  with respect to  $P'$  and replace  $P$  by  $Q$  using Lemma 10(v). Again, we either decrease the number of cycles in the projections or decrease the number of paths in the projections, and this gives a contradiction of the choice of  $\mathcal{C}$ .

We conclude that if  $\mathcal{C}_e$  contains two paths, then one of them is trivial.

Suppose that  $\mathcal{C}_e$  includes a nontrivial  $(u, v)$ -path  $P$  and a trivial path  $P' = w$ . By Claim 5.6, we let  $u, v \in L_e$  and  $w \in R_e$  using symmetry. If  $\mathcal{C}_e$  contains some other elements, or  $\{u, v\} \neq \{v_1^e, v_2^e\}$ , or  $w \neq v_{p(e)}^e$ , we modify  $\mathcal{C}$ . First, we delete all cycles of  $\mathcal{C}$  that are in  $\mathcal{C}_e$ . If  $w \neq v_{p(e)}^e$ , we replace  $w$  by  $v_{p(e)}^e$  in the corresponding cycle of  $\mathcal{C}$  using Lemma 10(v). Note that by Lemmas 6 and 4,  $G_e - v_{p(e)}^e$  has a Hamiltonian  $(v_1^e, v_2^e)$ -path  $Q$  and we use it to replace  $P$  by making use of Lemma 10(v). It remains to observe that these modifications contradict the choice of  $\mathcal{C}$ . We use the same arguments for the case when  $\mathcal{C}_e$  contains a nontrivial path and the other elements are cycles. We conclude that if  $\mathcal{C}_e$  contains a nontrivial path, then either (a)  $\mathcal{C}_e$  consists of a single left  $U$ -path or, symmetrically, a single right  $U$ -path, or (b)  $\mathcal{C}_e$  consists of a left  $U$ -path and the trivial right  $U$ -path or, symmetrically, a right  $U$ -path and the trivial left  $U$ -path, as required by the claim.

From now on we assume that the paths in  $\mathcal{C}_e$  are trivial (if they exist). Since  $|V_e| \geq 3$ ,  $\mathcal{C}_e$  contains a cycle  $C$ .

Suppose that  $\mathcal{C}_e$  contains two trivial paths  $P = u$  and  $P' = v$ . Assume by Claim 5.6 that  $u \in L_e$  and  $v \in R_e$ . If  $\mathcal{C}_e$  contains some other cycles except  $C$ , or  $u \neq v_1^e$ , or  $v \neq v_{p(e)}^e$ , we modify  $\mathcal{C}$ . First, we delete all cycles of  $\mathcal{C}$  that are in  $\mathcal{C}_e$ . If  $u \neq v_1^e$ , we replace  $u$  by  $v_1^e$  in the corresponding cycle of  $\mathcal{C}$  using Lemma 10(v) and do the same for  $v \neq v_{p(e)}^e$ . Since  $|V_e \setminus \{u, v\}| \geq |V(C)| \geq 3$ , we have that  $G_e - \{v_1^e, v_{p(e)}^e\}$  has a Hamiltonian cycle. We use it to replace  $C$ . Again, the modifications lead to a contradiction of the choice of  $\mathcal{C}$ . We use basically the same (in fact, simplified) arguments for the case when  $\mathcal{C}$  contains at most one trivial path and the other elements are cycles.

We conclude that if  $G_e$  is 2-connected and  $\mathcal{C}_e$  contains a cycle, then the cycle is unique and, besides, the cycle  $\mathcal{C}_e$  can contain only the trivial left  $U$ -path or the trivial right  $U$ -path or both. This concludes the analysis of the structure of  $\mathcal{C}_e$  if  $G_e$  is 2-connected.  $\square$

Finally, we consider the case when  $G_e$  contains a cut-vertex. Recall that we assume that  $\mathcal{C}_e$  has no path with one end-vertex in  $L_e$  and the other in  $R_e$ .

CLAIM 5.8. *If  $G_e$  has a cut-vertex, then the following hold:*

- *The paths in  $\mathcal{C}_e$  are left and right  $U$ -paths,  $\mathcal{C}_e$  contains at most one left  $U$ -path and at most one right  $U$ -path, and these paths are paths in the leftmost and the rightmost blocks, respectively.*
- *Each block contains at most one cycle of  $\mathcal{C}_e$ .*
- *If the leftmost block contains a nontrivial left  $U$ -path, then it does not contain*

a cycle of  $\mathcal{C}_e$ , and, symmetrically, if the rightmost block contains a nontrivial right  $U$ -path, then it does not contain a cycle of  $\mathcal{C}_e$ .

*Proof of Claim 5.8.* By Claim 5.6,  $\mathcal{C}_e$  contains at most one path with its end-vertices in  $L_e$  and at most one path with its end-vertices in  $R_e$ . Clearly, such paths are paths in the leftmost block  $B_\ell$  and the rightmost block  $B_r$  of  $G_e$ , respectively, by Lemma 5 and Observation 5. Observe also that every cycle in  $\mathcal{C}_e$  has all its vertices in some block of  $G_e$ .

Suppose that  $\mathcal{C}_e$  contains a nontrivial  $(u, v)$ -path  $P$  such that  $u, v \in L_e$ . Clearly,  $P$  is a path in  $B_\ell$ . Assume that  $\{u, v\} \neq \{v_1^e, v_2^e\}$  or there is a cycle of  $\mathcal{C}$  that has all its vertices in  $B_\ell$ . In this case we modify  $\mathcal{C}$  as follows. First, we delete every cycle of  $\mathcal{C}$  that has all its vertices in  $B_\ell$ . Let  $v_i^e$  be the unique cut-vertex of  $G_e$  in  $B_\ell$  (see Lemma 5). If  $v_i^e$  is included in an element of  $\mathcal{C}_e$  that has all its vertices in the block of  $G_e$  that is next after  $B_\ell$ , we have that  $B_\ell - v_i^e$  has at least 2 vertices. Since  $B_\ell - v_i^e$  is 2-connected by Lemma 6,  $B_\ell - v_i^e$  has a Hamiltonian  $(v_1^e, v_2^e)$ -path  $\hat{P}$  by Lemma 4. We replace  $P$  by  $\hat{P}$  using Lemma 10(v). Similarly, if  $v_i^e$  is a vertex of an element of  $\mathcal{C}_e$  that has all its vertices in  $B_\ell$ , we replace  $P$  by a Hamiltonian  $(v_1^e, v_2^e)$ -path in  $B_\ell$ . By these modifications we either reduce the number of paths and cycles in the projections or increase the number of paths with the leftmost pairs of end-vertices. In both cases, we obtain a contradiction of the choice of  $\mathcal{C}$ . Hence, if  $\mathcal{C}_e$  contains a nontrivial path that is either a left  $U$ -path or a right  $U$ -path that is, respectively, in the leftmost block or the rightmost block, then this block does not contain cycles of  $\mathcal{C}$ .

Suppose that  $\mathcal{C}_e$  contains a trivial path  $P = u$  with  $u \in L_e$ . Clearly,  $u \in V(B_\ell)$ . Assume that  $u \neq v_1^e$ . Since  $v_1^e$  should be covered, there is a cycle  $C \in \mathcal{C}$  that has all its vertices in  $B_\ell$ . Let  $v_i^e$  be the unique cut-vertex of  $G_e$  in  $B_\ell$  (see Lemma 5). If  $v_i^e$  is included in an element of  $\mathcal{C}_e$  that has all its vertices in the block of  $G_e$  that is next after  $B_\ell$ , we have that  $B_\ell - v_i^e$  has at least 2 vertices. Since  $B_\ell - \{v_1^e, v_i^e\}$  is 2-connected by Lemma 6 and has at least 3 vertices as  $B_\ell - \{u, v_i^e\}$  contains a cycle,  $B_\ell - \{v_1^e, v_i^e\}$  has a Hamiltonian cycle  $\hat{C}$  (see [2]). We modify  $\mathcal{C}$  as follows. First, we replace  $u$  by  $v_1^e$  in the cycle of  $\mathcal{C}$  containing  $u$  using Lemma 10(v), then we replace the cycles that have all their vertices in  $B_\ell$  by  $\hat{C}$ . If  $v_i^e$  is a vertex of an element of  $\mathcal{C}_e$  that has all its vertices in  $B_\ell$ , we use a Hamiltonian cycle in  $B_\ell - v_1^e$  for the replacement. In both cases we obtain a contradiction of the choice of  $\mathcal{C}$ . Using symmetry, we conclude that if  $\mathcal{C}$  contains a trivial path, then this is either the trivial left or right  $U$ -path.

Suppose that  $\mathcal{C}_e$  contains a cycle  $C$  that has all its vertices in a block  $B = G[v_i^e, \dots, v_j^e]$  of  $G_e$ . Assume that there is a cycle  $C' \in \mathcal{C}$  distinct from  $C$  that has all its vertices in  $B$ . Let  $F$  be the induced subgraph of  $G_e$  whose vertices are covered by the cycles of  $\mathcal{C}$  with all their vertices in  $B$ . Notice that either  $F = B$ , or  $F = B - v_i^e$ , or  $F = B - v_j^e$ , or  $F = B - \{v_i^e, v_j^e\}$ . In all cases  $F$  is 2-connected by Lemma 6 and has at least 3 vertices as  $F$  contains a cycle. Hence,  $F$  has a Hamiltonian cycle  $\hat{C}$ . We modify  $\mathcal{C}$  by removing the cycles with their vertices in  $B$  and replacing them by  $\hat{C}$ . This contradicts the choice of  $\mathcal{C}$  as we decrease the number of cycles in the projections. We conclude that each block contains at most one cycle of  $\mathcal{C}$ .  $\square$

Combining Claims 5.5, 5.7, and 5.8, we conclude that the conditions required by the definition of a tamed cycle cover are fulfilled for  $\mathcal{C}_e$ . This immediately implies the claim of the lemma.  $\square$

**6. Kernelization for PATH and CYCLE COVER.** In this section we construct a kernel for PATH COVER and a compression for CYCLE COVER on proper  $H$ -graphs.

**6.1. Kernel for PATH COVER.** In this section we prove the first main result of the paper, namely, Theorem 1, about kernelization of PATH COVER. For the reader's convenience, we repeat the statement of the theorem here.

**THEOREM 1.** *PATH COVER admits a kernel of size  $\mathcal{O}(h^8)$ , where  $h$  is the size of the graph  $H$  in a proper  $H$ -representation of the input graph  $G$ .*

Our kernelization algorithm for PATH COVER is based on the following crucial lemma.

**LEMMA 13.** *There is a polynomial-time algorithm that, given an instance  $(G, k)$  of PATH COVER for a proper  $H$ -graph  $G$  with its nice proper  $H$ -representation  $(H', \mathcal{M})$ , constructs an equivalent instance  $(\hat{G}, k)$  of PATH COVER and a clique cover  $\mathcal{Q}$  of  $\hat{G}$  with  $|\mathcal{Q}| \leq |V(H)| + 2|E(H)|$ .*

*Proof.* Let  $G$  be a proper  $H$ -graph and let  $(H', \mathcal{M})$  with  $\mathcal{M} = \{M_v\}_{v \in V(G)}$  be a nice proper  $H$ -representation of  $G$ . We apply the notation introduced in subsection 5.3 for  $G$  and its representation.

For every  $e \in E(H)$  with  $|V_e| \geq 4$  and  $L_e \cap R_e = \emptyset$ , we apply the following reduction rule using the first condition that could be applied. The general idea is to replace the proper interval graph  $G_e$  by a small gadget (consequently, the gadget admits a small clique cover).

**REDUCTION RULE 6.1.**

- (i) *If  $G_e$  has no  $e$ -cut-vertex, then delete the vertices  $v_3^e, \dots, v_{p(e)-2}^e$  and make  $v_1^e, v_2^e, v_{p(e)-1}^e, v_{p(e)}^e$  pairwise adjacent.*
- (ii) *If there are  $i, j \in \{1, \dots, p(e)\}$  such that  $i < j - 1$  and  $v_i$  and  $v_j$  are  $e$ -cut-vertices, that is,  $G_e$  contains two nonconsecutive  $e$ -cut-vertices, then delete  $v_2^e, \dots, v_{p(e)-1}^e$ , construct a new vertex  $u^e$ , and make it adjacent to  $v_1^e$  and  $v_{p(e)}^e$ .*
- (iii) *If  $v_2^e$  and  $v_{p-1}^e$  are  $e$ -cut-vertices or there is  $i \in \{3, \dots, p(e) - 2\}$  such that  $v_i^e$  is an  $e$ -cut-vertex, then delete  $v_3^e, \dots, v_{p(e)-2}^e$ , construct a new vertex  $u^e$ , and make it adjacent to  $v_1^e, v_2^e, v_{p(e)-1}^e, v_{p(e)}^e$ .*
- (iv) *If  $v_2^e$  is an  $e$ -cut-vertex, then delete  $v_3^e, \dots, v_{p(e)-2}^e$  and make  $v_2^e$  adjacent to  $v_{p(e)-1}^e$  and  $v_{p(e)}^e$ .*
- (v) *If  $v_{p(e)-1}^e$  is an  $e$ -cut-vertex, then delete  $v_3^e, \dots, v_{p(e)-2}^e$  and make  $v_1^e$  and  $v_2^e$  adjacent to  $v_{p(e)-1}^e$ .*
- (vi) *If  $v_1^e$  is an  $e$ -cut-vertex, then delete  $v_2^e, \dots, v_{p(e)-2}^e$  and make  $v_1^e$  adjacent to  $v_{p(e)-1}^e$  and  $v_{p(e)}^e$ .*
- (vii) *If  $v_{p(e)}^e$  is an  $e$ -cut-vertex, then delete  $v_3^e, \dots, v_{p(e)-1}^e$  and make  $v_{p(e)}^e$  adjacent to  $v_1^e$  and  $v_2^e$ .*

The application of the rule is shown in Figure 4. Notice that the list of conditions in (i)–(vii) is exhaustive and one of them is always applied. Observe also that the modifications result in a proper  $H$ -graph that has a nice  $H$ -representation. The construction of such a representation is shown in Figure 4. Now we prove that Reduction Rule 6.1 is safe. By symmetry, it is sufficient to prove safeness for (i)–(iv) and (vi). Denote by  $G'$  the graph obtained from  $G$  by the application of Reduction Rule 6.1 for an edge  $e \in E(H)$  with  $|V_e| \geq 4$  and  $L_e \cap R_e = \emptyset$ . Let also  $G'_e$  be the graph obtained

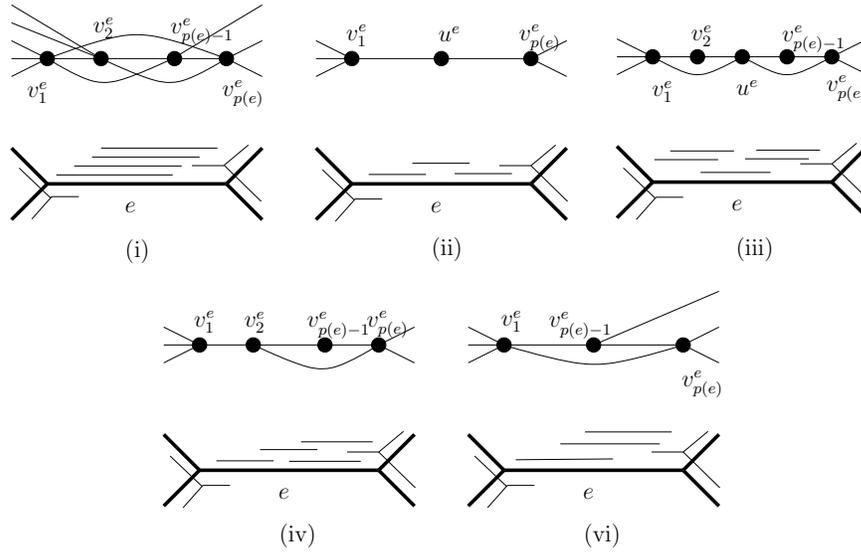


FIG. 4. The application of Reduction Rule 6.1 and the construction of the  $H$ -representation; cases (v) and (vii) are omitted as they are symmetric to cases (iv) and (vi), respectively.

from  $G_e$  by the rule. We show that  $(G, k)$  is a yes-instance of PATH COVER if and only if  $(G', k)$  is a yes-instance.

Suppose that  $(G, k)$  is a yes-instance. By Lemma 11,  $G$  has a tamed path cover  $\mathcal{P}$  of size at most  $k$ . We consider 5 cases corresponding to the cases of Reduction Rule 6.1. To show that  $(G', k)$  is a yes-instance, we consider the projection  $\mathcal{P}_e$  of the tamed path cover  $\mathcal{P}$  and replace each path in  $\mathcal{P}_e$  by a path with the same end-vertices in such a way that the obtained paths form a path cover of  $G'_e$ . Observe that if  $\mathcal{P}_e$  consist of a straight path, that is, a  $(v_1^e, v_{p(e)}^e)$ -path, then we always can replace it by a Hamiltonian  $(v_1^e, v_{p(e)}^e)$ -path in  $G'_e$  (see Figure 4). Hence, we exclude this case from the analysis and assume that  $\mathcal{P}_e$  does not contain a straight path. Observe that if  $\mathcal{P}_e$  contains a  $(v_1^e, v_2^e)$ -path or a  $(v_{p(e)-1}^e, v_{p(e)}^e)$ -path, then the vertices of such a path are in the same block of  $G_e$ , because  $v_1^e v_2^e, v_{p(e)-1}^e v_{p(e)}^e \in E(G)$ . We use this observation in our case analysis.

*Case (i).* If  $G_e$  has no  $e$ -cut-vertex, then  $G_e$  is 2-connected. Clearly, we have the following path covers of  $G'_e$ :  $\{v_1^e v_{p(e)-1}^e v_{p(e)}^e v_2^e\}$ ,  $\{v_2^e v_{p(e)-1}^e v_{p(e)}^e v_1^e\}$ ,  $\{v_1^e v_{p(e)-1}^e, v_2^e v_{p(e)}^e\}$ , and  $\{v_2^e v_{p(e)-1}^e, v_1^e v_{p(e)}^e\}$ , whose paths could trivially be used for the replacement of the paths in  $\mathcal{P}_e$  with the same end-vertices.

*Case (ii).* Suppose first that  $G_e$  is 2-connected. Then  $v_1^e$  and  $v_{p(e)}^e$  are the  $e$ -cut-vertices of  $G_e$ . Then  $v_2^e$  is not left-attached and  $v_{p(e)-1}^e$  is not right-attached. This implies that these vertices cannot be saturated, but this contradicts the saturation condition for the paths in  $\mathcal{P}_e$ . Hence,  $G_e$  has a cut-vertex  $v_i^e$  for  $i \in \{2, \dots, p(e) - 1\}$ . Suppose that  $v_1^e$  is an  $e$ -cut-vertex and  $i \geq 3$ . Then  $\mathcal{P}_e$  does not contain a left  $U$ -path as  $v_2^e$  is not left-attached and, therefore, cannot be saturated. Hence,  $\mathcal{P}_e$  contains the trivial path  $v_1^e$ . It follows that  $v_2^e$  is in the right  $U$ -path of  $\mathcal{P}_e$ , but the vertices of this path should be in the block of  $G_e$  that contains  $v_{p(e)-1}^e$  and  $v_{p(e)}^e$  but  $v_2^e$  does not belong to this block. This leads to a contradiction, and therefore  $v_1^e$  is not an  $e$ -cut-vertex. By symmetry, we have that  $v_{p(e)}^e$  is not an  $e$ -cut-vertex. Then there

are  $i, j \in \{2, \dots, p(e) - 1\}$  such that  $i < j - 1$  and  $v_i^e, v_j^e$  are cut-vertices of  $G_e$ . But then  $v_{j-1}^e$  belongs neither to the block of  $G_e$  containing  $v_1^e$  nor the block containing  $v_{p(e)}^e$ . Hence,  $v_{j-1}^e$  cannot be in any left or right  $U$ -path. This means that (ii) does not occur.

*Case (iii).* If  $v_2^e$  and  $v_{p(e)-1}^e$  are  $e$ -cut-vertices or there is  $i \in \{3, \dots, p(e) - 2\}$  such that  $v_i^e$  is an  $e$ -cut-vertex, then these vertices are cut-vertices of  $G_e$  and we have that  $v_2^e$  and  $v_{p(e)-1}^e$  are in distinct blocks of  $G_e$ . If  $\mathcal{P}_e$  contains the trivial path  $v_1^e$ , then  $v_2^e$  cannot be in a  $(v_{p(e)-1}^e, v_{p(e)}^e)$ -path. Therefore this is impossible. By symmetry,  $\mathcal{P}_e$  cannot contain  $v_{p(e)}^e$ . Hence,  $\mathcal{P}_e$  consists of a nontrivial left  $U$ -path and a nontrivial right  $U$ -path. We replace them by  $v_1^e v_2^e$  and  $v_{p(e)-1}^e u^e v_{p(e)}^e$ , respectively.

*Case (iv).* Since  $|V_e| \geq 4$ ,  $v_3^e \neq v_{p(e)}^e$ . Suppose that  $\mathcal{P}_e$  contains the trivial right  $U$ -path  $v_{p(e)}^e$ . Then  $v_3^e$  cannot be in any left  $U$ -path, because  $v_2^e$  is a cut-vertex of  $G_e$ . This means that  $\mathcal{P}_e$  consists of either the trivial left  $U$ -path and a nontrivial right  $U$ -path or a nontrivial left  $U$ -path and a nontrivial right  $U$ -path. We replace them by  $v_1^e$  and  $v_{p(e)-1}^e v_2^e v_{p(e)}^e$  or by  $v_1^e v_2^e$  and  $v_{p(e)-1}^e v_{p(e)}^e$ , respectively.

*Case (vi).* Note that since the conditions of (i)–(v) could not be applied,  $G_e$  is 2-connected and  $v_1^e$  is the unique  $e$ -cut-vertex of  $G_e$ . Since  $v_2^e \notin L_e$ ,  $\mathcal{P}_e$  cannot contain a nontrivial left  $U$ -path. Hence  $\mathcal{P}_e$  contains a unique nontrivial right  $U$ -path that can be replaced by  $v_{p(e)-1}^e v_2^e v_{p(e)}^e$ .

This completes the case analysis, and we conclude that  $G'$  has a path cover of size at most  $k$ , that is,  $(G', k)$  is a yes-instance of PATH COVER.

Suppose now that  $(G', k)$  is a yes-instance of PATH COVER. We show that  $(G, k)$  is a yes-instance as well. Recall that  $G'$  is a proper  $H$ -graph that has a nice  $H$ -representation. By Lemma 11,  $G'$  has a tamed path cover  $\mathcal{P}'$  of size at most  $k$ . We show that we can construct a path cover of  $G$  by modifying the paths of  $\mathcal{P}'$ . The idea is the same as above: we replace the paths in the projection  $\mathcal{P}'_e$  by the paths in  $G_e$  in such a way that these paths cover  $G_e$ . Note that since  $G_e$  is connected,  $G_e$  has a Hamiltonian  $(v_1^e, v_{p(e)}^e)$ -path  $P$  by Lemma 4. Hence, if  $\mathcal{P}'_e$  contains a unique straight path, then this path can be replaced by  $P$ . Assume that this is not the case. Again, we consider 5 cases corresponding to (i)–(iv) and (vi).

*Case (i).* Recall that  $G_e$  is 2-connected. Since  $|V_e| \geq 4$ , by Lemma 4 we have that  $G_e$  has the path covers formed by a single  $(v_1^e, v_2^e)$ -path or a single  $(v_{p(e)-1}^e, v_{p(e)}^e)$ -path or two  $(a_1, b_1)$ - and  $(a_2, b_2)$ -paths for  $\{a_1, a_2\} = \{v_1^e, v_2^e\}$  and  $\{b_1, b_2\} = \{v_{p(e)-1}^e, v_{p(e)}^e\}$ . Clearly, we can use these paths to replace the paths of  $\mathcal{P}'_e$  of the same structure. In particular, if  $\mathcal{P}'_e$  contains two paths, say,  $v_1^e v_{p(e)-1}^e$  and  $v_2^e v_{p(e)}^e$ , we use the property that they are subpaths of the same path  $Q$  of  $\mathcal{P}'$  and are in opposite directions in  $Q$ . This implies that the corresponding replacement creates a path.

*Case (ii).* Note that  $\mathcal{P}'_e = \{v_1^e u^e v_{p(e)}^e\}$  is the only possibility in this case, and therefore (ii) does not occur.

*Case (iii).* Clearly, we have that either  $\mathcal{P}'_e = \{v_1^e v_2^e, v_{p(e)-1}^e u^e v_{p(e)}^e\}$  or, symmetrically,  $\mathcal{P}'_e = \{v_1^e u^e v_2^e, v_{p(e)-1}^e v_{p(e)}^e\}$ , that is,  $\mathcal{P}'_e$  consists of a single left and single right  $U$ -path. Recall that (ii) could not be applied. Hence,  $G_e$  has at most two  $e$ -cut-vertices, and if it has exactly two such vertices, they are consecutive. Assume that  $v_{i-1}^e, v_i^e$  are the  $e$ -cut-vertices. By the conditions of (iii),  $i \in \{3, \dots, p(e) - 1\}$ . It follows that  $G_e[v_1^e, \dots, v_{i-1}^e]$  and  $G_e[v_i^e, \dots, v_{p(e)}^e]$  are blocks of  $G_e$  containing the vertices  $v_1^e, v_2^e$  and  $v_{p(e)-1}^e, v_{p(e)}^e$ , respectively. By Lemma 4, these blocks can be covered

by a  $(v_1^e, v_2^e)$ -path and a  $(v_{p(e)-1}^e, v_{p(e)}^e)$ -path, respectively. Then we use these paths to replace the paths of  $\mathcal{P}'_e$ . Assume now that  $G_e$  has a unique  $e$ -cut-vertex  $v_i$ . Then by the conditions of (iii),  $i \in \{3, \dots, p(e) - 2\}$ , and  $G_e[v_1^e, \dots, v_i^e]$  and  $G_e[v_i^e, \dots, v_{p(e)}^e]$  are the blocks of  $G_e$  containing the vertices  $v_1^e, v_2^e, v_i^e$  and  $v_i^e, v_{p(e)-1}^e, v_{p(e)}^e$ , respectively. By Lemma 6,  $G_e[v_1^e, \dots, v_{i-1}^e] = G_e[v_1^e, \dots, v_i^e] - v_i^e$  is 2-connected. Hence, by Lemma 4,  $G_e[v_1^e, \dots, v_{i-1}^e]$  has a Hamiltonian  $(v_1^e, v_2^e)$ -path. Similarly,  $G_e[v_i^e, \dots, v_{p(e)}^e]$  has a Hamiltonian  $(v_{p(e)-1}^e, v_{p(e)}^e)$ -path. We use these paths to replace the paths of  $\mathcal{P}'_e$ .

*Case (iv).* Observe that either  $\mathcal{P}'_e = \{v_1^e, v_{p(e)-1}^e v_2^e v_{p(e)}^e\}$  or  $\mathcal{P}'_e = \{v_1^e v_2^e, v_{p(e)-1}^e v_{p(e)}^e\}$ . Suppose that  $\mathcal{P}'_e = \{v_1^e, v_{p(e)-1}^e v_2^e v_{p(e)}^e\}$ . Since (ii) and (iii) could not be applied, we have that  $G_e[v_2^e, \dots, v_{p(e)}^e]$  is a block of  $G_e$ . By Lemma 4,  $G_e[v_2^e, \dots, v_{p(e)}^e]$  has a Hamiltonian  $(v_{p(e)-1}^e, v_{p(e)}^e)$ -path that is used to replace  $v_{p(e)-1}^e v_2^e v_{p(e)}^e$ . Assume that  $\mathcal{P}'_e = \{v_1^e v_2^e, v_{p(e)-1}^e v_{p(e)}^e\}$ . Since  $G_e[v_2^e, \dots, v_{p(e)}^e]$  is 2-connected and  $|V_e| \geq 4$ , we have that  $G_e[v_3^e, \dots, v_{p(e)}^e]$  is a 2-connected graph by Lemma 4 that has a Hamiltonian  $(v_{p(e)-1}^e, v_{p(e)}^e)$ -path. We use this path to replace  $v_{p(e)-1}^e v_{p(e)}^e$ .

*Case (vi).* Notice that  $G'_e$  is 2-connected. Observe also that  $v_{p(e)-1}^e$  is not left-attached in the proper  $H$ -representation of  $G'$ . Hence,  $\mathcal{P}'_e$  cannot contain a  $(v_1^e, v_{p(e)-1}^e)$ -path. We obtain that  $\mathcal{P}'_e = \{v_{p(e)-1}^e v_1^e v_{p(e)}^e\}$ . Since  $G_e$  is 2-connected and  $|V_e| \geq 4$ ,  $G_e$  has a Hamiltonian  $(v_{p(e)-1}^e, v_{p(e)}^e)$ -path that is used to replace  $v_{p(e)-1}^e v_1^e v_{p(e)}^e$ .

This completes the proof that  $G$  has a path cover of size at most  $k$ , and therefore  $(G, k)$  is a yes-instance of PATH COVER. We conclude that  $(G, k)$  is a yes-instance of PATH COVER if and only if  $(G', k)$  is a yes-instance. Hence, Reduction Rule 6.1 is safe.

Denote by  $\hat{G}$  the graph obtained from  $G$  by the application of Reduction Rule 6.1 for all  $e \in E(H)$  with  $|V_e| \geq 4$  and  $L_e \cap R_e = \emptyset$ .

We construct a clique cover of  $\hat{G}$  as follows. For each  $e \in E(H)$ , we construct a set  $\mathcal{K}^e$  of at most two cliques such that every vertex of  $\hat{G}_e$  is included in at least one of the cliques. If  $|V_e| \leq 3$ , then it can be done since  $\hat{G}_e = G_e$  is connected. If  $L_e \cap R_e \neq \emptyset$ , then by Lemmas 10(ii) and 10(iii), we have that  $L_e$  and  $R_e$  are nonempty cliques and  $V_e = L_e \cup R_e$ . Otherwise, if  $|V_e| \geq 3$  and  $L_e \cap R_e = \emptyset$ , Reduction Rule 6.1 was applied for  $e$ . It is straightforward to observe (see Figure 4) that every gadget used to replace  $G_e$  can be covered by at most two cliques. For every node  $x \in V(H)$ , let  $K_x = \{v \in V(G) \mid x \in M_v\}$ . Clearly, each  $K_x$  is a clique and we have that  $\mathcal{K} = \{K_x \mid x \in V(H)\} \cup \bigcup_{e \in E(H)} \mathcal{K}^e$  is a family of cliques of  $\hat{G}$  such that every vertex of  $\hat{G}$  is included in at least one clique of  $\mathcal{K}$ . Observe that  $|\mathcal{K}| \leq |V(H)| + 2|E(H)|$ . It may happen that  $\mathcal{K}$  is not a clique cover as the cliques can have common vertices. We construct the clique cover  $\mathcal{Q}$  from  $\mathcal{K}$  by the following greedy procedure: we select an arbitrary nonempty clique  $Q$  in  $\mathcal{K}$ , include it in  $\mathcal{Q}$ , and update the cliques of  $\mathcal{K}$  by deleting the vertices of  $Q$  from them. It is straightforward to verify that  $\mathcal{Q}$  is a clique cover of  $\hat{G}$  and  $|\mathcal{Q}| \leq |V(G)| + 2|E(H)|$ .

Finally, we observe that Reduction Rule 6.1 can be applied in polynomial time and the construction of  $\mathcal{Q}$  is also polynomial.  $\square$

Finally, putting it all together, we are ready to prove the main result of this section, Theorem 1.

*Proof of Theorem 1.* Let  $(G, k)$  be an instance of PATH COVER where  $G$  is a proper  $H$ -graph given together with its proper  $H$ -representation  $(H', \mathcal{M})$ . We use the algorithm from Lemma 9 that either solves the problem or constructs an equivalent instance  $(G', k')$  of PATH COVER and a nice proper  $\hat{H}$ -representation  $(\hat{H}', \hat{\mathcal{M}})$  of  $G'$  such that  $|V(\hat{H}')| \leq 3|E(H)|$  and  $|E(\hat{H}')| \leq 2|E(H)|$ . If the algorithm solves the problem, our kernelization algorithm returns a trivial yes- or no-instance, respectively. Otherwise, we apply the algorithm from Lemma 13 for  $(G', k')$  and  $(\hat{H}', \hat{\mathcal{M}})$ . This way we obtain an equivalent instance  $(G'', k')$  of PATH COVER together with a clique cover  $\mathcal{Q}$  of  $G''$  such that  $|\mathcal{Q}| \leq |V(\hat{H}')| + 2|E(\hat{H}')| \leq 7|E(H)|$ . Finally, we use the algorithm from Theorem 3 for  $(G'', k')$  and  $\mathcal{Q}$  that produces an equivalent instance  $(\hat{G}, \hat{k})$  of size  $\mathcal{O}(|\mathcal{Q}|^8) = \mathcal{O}((|E(H)|)^8)$ . Observe that the used algorithms are polynomial. This gives us a polynomial compression of PATH COVER parameterized by the size of  $H$  into the nonparameterized PATH COVER problem. Formally, to claim that we get a kernel, we have to specify the value of the parameter. We do this by using Observation 1 and obtain that  $\hat{G}$  is a proper  $\hat{G}$ -graph, that is, the value of the parameter can be set equal to the size of  $\hat{G}$ .  $\square$

**6.2. Kernel for CYCLE COVER.** In this subsection we prove Theorem 2 by constructing a compression of CYCLE COVER to PRIZE COLLECTING CYCLE COVER and a polynomial kernel for HAMILTONIAN CYCLE.

First, we state the variant of Theorem 2 that includes the claim for HAMILTONIAN CYCLE.

**THEOREM 2.** *CYCLE COVER on proper  $H$ -graphs admits a compression into PRIZE COLLECTING CYCLE COVER of size  $\mathcal{O}(h^{10})$  and HAMILTONIAN CYCLE has a kernel of size  $\mathcal{O}(h^8)$  when parameterized by the size  $h$  of  $H$  if a proper  $H$ -representation is given.*

We start the proof from the following lemma.

**LEMMA 14.** *There is a polynomial-time algorithm that, given an instance  $(G, k)$  of CYCLE COVER for a proper  $H$ -graph  $G$  with its nice proper  $H$ -representation  $(H', \mathcal{M})$ , constructs an equivalent instance  $(\hat{G}, \omega, \alpha, r)$  of PRIZE COLLECTING CYCLE COVER such that at most  $3|E(H)|$  edges of  $\hat{G}$  have nonzero weights and a clique cover  $\mathcal{Q}$  of  $\hat{G}$  with  $|\mathcal{Q}| \leq |V(H)| + 6|E(H)|$ . Moreover,  $\alpha(x) = x$  for  $x \in \mathbb{N}$ ,  $r \geq -k$ , and  $\omega(e) \leq k - 1$  for each  $e \in E(\hat{G})$ .*

*Proof.* Let  $G$  be a proper  $H$ -graph and let  $(H', \mathcal{M})$  with  $\mathcal{M} = \{M_v\}_{v \in V(G)}$  be a nice proper  $H$ -representation of  $G$ . We apply the notation introduced in subsection 5.3 for  $G$  and its representation. Denote additionally by  $B_\ell^e$  and  $B_r^e$ , respectively, the leftmost and the rightmost blocks of  $G_e$  for  $e \in E(H)$ . We denote the cut-vertices of  $G_e$  in these blocks by  $v_{\ell(e)}^e$  and  $v_{r(e)}^e$ , respectively, that are unique by Lemma 5. Also let  $F^e = G[v_{\ell(e)}^e, \dots, v_{r(e)}^e]$ . In the same way as in the kernelization for PATH COVER, we replace the sufficiently big graphs  $G_e$  by gadgets.

For every  $e \in E(H)$  with  $L_e \cap R_e = \emptyset$ , we apply a series of reduction rules.

**REDUCTION RULE 6.2.** *If  $|V_e| \geq 6$  and  $G_e$  is 2-connected, then delete  $v_3^e, \dots, v_{p(e)-2}^e$ , make  $v_1^e, v_2^e, v_{p(e)-1}^e, v_{p(e)}^e$  pairwise adjacent, create a new vertex  $u^e$ , and make it adjacent to  $v_1^e, v_2^e, v_{p(e)-1}^e, v_{p(e)}^e$ .*

The application of the rule is shown in Figure 5. Observe also that the modifications result in a proper  $H$ -graph that has a nice  $H$ -representation shown in Figure 5.

Denote by  $G'$  the graph obtained from  $G$  by the application of Reduction Rule 6.2 for an edge  $e \in E(H)$  with  $|V_e| \geq 6$  and  $L_e \cap R_e = \emptyset$ . Also let  $G'_e$  be the graph obtained

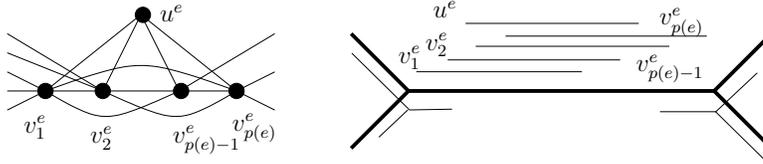


FIG. 5. The application of Reduction Rule 6.2 and the construction of the  $H$ -representation.

from  $G_e$  by the rule. We show that  $(G, k)$  is a yes-instance of CYCLE COVER if and only if  $(G', k)$  is a yes-instance.

Suppose that  $(G, k)$  is a yes-instance. By Lemma 12,  $G$  has a tamed cycle cover  $\mathcal{C}$  of size at most  $k$ . We consider the projection  $\mathcal{C}_e$  and replace its elements by the elements of the same structure. If the projection  $\mathcal{C}_e$  consists of a straight path  $P$ , we replace it by  $v_1^e v_2^e u^e v_{p(e)-1}^e v_{p(e)}^e$ . If  $\mathcal{C}_e$  contains a straight pair, we replace it by  $\{v_1^e v_{p(e)-1}^e, v_2^e u^e v_{p(e)}^e\}$  or  $\{v_1^e v_{p(e)}^e, v_2^e u^e v_{p(e)-1}^e\}$  to obtain a pair of paths with the same end-vertices as in  $\mathcal{C}_e$ . If  $\mathcal{C}_e$  consists of a left  $U$ -path and a right  $U$ -path, then because  $|V_e| \geq 6$ , one of them is nontrivial. We replace a nontrivial left  $U$ -path either by  $v_1^e v_{p(e)-1}^e v_{p(e)}^e u^e v_2^e$  or by  $v_1^e v_{p(e)-1}^e u^e v_2^e$  depending on whether  $\mathcal{C}_e$  contains the trivial right  $U$ -path or not. The replacement for the case when  $\mathcal{C}_e$  contains a nontrivial right  $U$ -path is symmetric. The trivial paths are not replaced. Finally, if  $\mathcal{C}_e$  contains a cycle, it is replaced by one of the cycles  $u^e v_1^e v_2^e v_{p(e)-1}^e v_{p(e)}^e u^e$ ,  $u^e v_1^e v_2^e v_{p(e)-1}^e u^e$ ,  $u^e v_2^e v_{p(e)-1}^e v_{p(e)}^e u^e$ , or  $u^e v_2^e v_{p(e)-1}^e u^e$  depending on whether  $\mathcal{C}_e$  has trivial left or right  $U$ -paths that are not replaced. It is straightforward to verify that the replacement of the cycles and the replacement of the segments of the cycles of  $\mathcal{C}$  that form paths in  $\mathcal{C}_e$  gives a cycle cover of the same size. Therefore,  $(G', k)$  is a yes-instance.

Suppose now that  $(G', k)$  is a yes-instance. We essentially repeat the same arguments: by Lemma 12,  $G'$  has a tamed cycle cover  $\mathcal{C}'$  of size at most  $k$  and we replace the elements of the projection  $\mathcal{C}'_e$  by the elements of the same structure. Note that since  $|V_e| \geq 6$  and  $G_e$  is 2-connected,  $G_e - v_1^e$ ,  $G_e - v_{p(e)}^e$ ,  $G_e - \{v_1^e, v_2^e\}$  are 2-connected by Lemma 6 and have at least 3 vertices. Therefore, by [2] and Lemma 4,  $G_e$  has a Hamiltonian  $(v_1^e, v_{p(e)}^e)$ -path,  $G_e - v_1^e$ ,  $G_e - v_{p(e)}^e$ ,  $G_e - \{v_1^e, v_2^e\}$  have Hamiltonian cycles,  $G_e$  and  $G_e - v_{p(e)}^e$  have Hamiltonian  $(v_1^e, v_2^e)$ -paths,  $G_e$  and  $G_e - v_1^e$  have Hamiltonian  $(v_{p(e)-1}^e, v_{p(e)}^e)$ -paths. We use these paths and cycles to replace the paths and cycles of  $\mathcal{C}'_e$ . We also have that by Lemma 4,  $G_e$  has a path cover composed of a pair of paths with one of the end-vertices in  $\{v_1^e, v_2^e\}$  and the other in  $\{v_{p(e)-1}^e, v_{p(e)}^e\}$ . We use these paths to replace a straight pair of paths in  $\mathcal{C}'_e$ , and we use additionally that the paths of the pair are segments of the same cycle of  $\mathcal{C}'$  and are in opposite directions in the cycle. Applying Lemma 10(v), we can do a replacement that results in a cycle of  $G$ . This completes the safeness proof for the rule.

For simplicity, we use the same notation  $(G, k)$  for the instance obtained from the original instance of CYCLE COVER by the exhaustive application of Reduction Rule 6.2. Our next reduction rule deals with the leftmost and rightmost blocks of  $G_e$  with cut-vertices.

REDUCTION RULE 6.3. *If  $G_e$  has a cut-vertex, then*

- (i) *if  $|V(B_\ell^e)| \geq 6$ , then delete  $v_3^e, \dots, v_{\ell(e)-1}^e$ , make  $v_1^e, v_2^e, v_{\ell(e)}^e$  pairwise adjacent, create two new adjacent vertices  $u_1^e$  and  $u_2^e$ , and make them adjacent to  $v_1^e, v_2^e, v_{\ell(e)}^e$ ;*

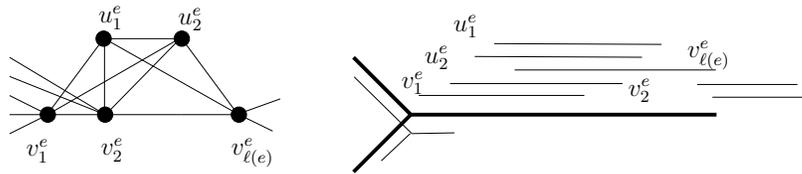


FIG. 6. The application of Reduction Rule 6.3 for  $B_\ell^e$  and the construction of the  $H$ -representation.

- (ii) if  $|V(B_r^e)| \geq 6$ , then delete  $v_{r(e)+1}^e, \dots, v_{p(e)-2}^e$ , make  $v_{r(e)}^e, v_{p(e)-1}^e, v_{p(e)}^e$  pairwise adjacent, create two new adjacent vertices  $u_1^e$  and  $u_2^e$ , and make them adjacent to  $v_{r(e)}^e, v_{p(e)-1}^e, v_{p(e)}^e$ .

The application of the rule is shown in Figure 6. Observe also that the modifications result in a proper  $H$ -graph that has a nice  $H$ -representation, as shown in Figure 6.

Denote by  $G'$  the graph obtained from  $G$  by applying Reduction Rule 6.3(i) for an edge  $e \in E(H)$  such that  $|V(B_\ell^e)| \geq 6$ . Let also  $G'_e$  be the graph obtained from  $G_e$  by the rule. We show that  $(G, k)$  is a yes-instance of CYCLE COVER if and only if  $(G', k)$  is a yes-instance. The proof uses the same arguments as the proof for Reduction Rule 6.2.

Suppose that  $(G, k)$  is a yes-instance. By Lemma 12,  $G$  has a tamed cycle cover  $\mathcal{C}$  of size at most  $k$ . We consider the projection  $\mathcal{C}_e$  and replace its elements by the elements of the same structure. If the projection  $\mathcal{C}_e$  consists of a straight path, we replace its  $(v_1^e, v_{\ell(e)}^e)$ -subpath by  $v_1^e v_2^e u_1^e u_2^e v_{\ell(e)}^e$ . If  $\mathcal{C}_e$  contains a nontrivial left  $U$ -path, this path covers the vertices of  $B_\ell^e$  or  $B_\ell^e - v_{\ell(e)}^e$  and we replace it by  $v_1^e u_1^e u_2^e v_{\ell(e)}^e v_2^e$  or  $v_1^e u_1^e u_2^e v_2^e$ . If  $\mathcal{C}_e$  does not contain a nontrivial left  $U$ -path, then  $\mathcal{C}_e$  contains a cycle that covers the vertices of  $B_\ell^e$  or  $B_\ell^e - v_{\ell(e)}^e$  or  $B_\ell^e - v_1^e$  or  $B_\ell^e - \{v_1^e, v_{\ell(e)}^e\}$  and we replace it by similar cycles in the constructed gadget in a straightforward way.

Suppose now that  $(G', k)$  is a yes-instance. By Lemma 12,  $G'$  has a tamed cycle cover  $\mathcal{C}'$  of size at most  $k$  and we replace the elements of the projection  $\mathcal{C}'_e$  by the elements of the same structure. It is sufficient to observe that since  $|V(B_\ell^e)| \geq 6$ , the graphs  $B_\ell^e - v_1^e$ ,  $B_\ell^e - v_{\ell(e)}^e$ , and  $B_\ell^e - \{v_1^e, v_{\ell(e)}^e\}$  are 2-connected by Lemma 6 and have at least 3 vertices. Therefore, by [2] and Lemma 4, these graphs and  $B_\ell^e$  have Hamiltonian cycles,  $B_\ell^e$  has a Hamiltonian  $(v_1^e, v_{\ell(e)}^e)$ -path, and  $B_\ell^e$  and  $B_\ell^e - v_{\ell(e)}^e$  have Hamiltonian  $(v_1^e, v_2^e)$ -paths. Clearly, these cycles and paths can be used to replace the corresponding elements of  $\mathcal{C}'_e$ . This completes the safeness proof for the rule.

As before, assume for simplicity that  $(G, k)$  is the instance of CYCLE COVER obtained by the exhaustive application of Reduction Rule 6.3. Our next aim is to deal with  $F^e$ . It is convenient to consider first the case when  $F^e$  is 2-connected, that is, it is a block of  $G_e$ .

**REDUCTION RULE 6.4.** *If  $F^e$  is 2-connected and  $|V(F^e)| \geq 6$ , then delete  $v_{\ell(e)+1}^e, \dots, v_{r(e)-1}^e$ , construct 3 new vertices  $u_1^e, u_2^e, u_3^e$ , and make  $v_{\ell(e)}^e, v_{r(e)}^e, u_1^e, u_2^e, u_3^e$  pairwise adjacent.*

The application of the rule is shown in Figure 7 and, again, the modifications give a proper  $H$ -graph that has a nice  $H$ -representation, as shown in the figure.

To prove safeness, we use the same arguments as before and only briefly sketch them here. Denote by  $G'$  the graph obtained from  $G$  by the application of Reduction

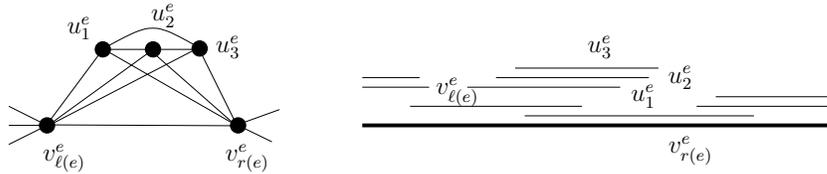


FIG. 7. The application of Reduction Rule 6.3 and the construction of the  $H$ -representation.

Rule 6.4 for an edge  $e \in E(H)$  such that  $|V(F^e)| \geq 6$ . Suppose that  $(G, k)$  is a yes-instance. By Lemma 12,  $G$  has a tamed cycle cover  $\mathcal{C}$  of size at most  $k$ . We consider the projection  $\mathcal{C}_e$  and replace its elements by the elements of the same structure. If the projection  $\mathcal{C}_e$  consists of a straight path  $P$ , we replace its  $(v_{\ell(e)}^e, v_{r(e)}^e)$ -subpath by  $v_{\ell(e)}^e u_1^e u_2^e u_3^e v_{r(e)}^e$ . Otherwise,  $\mathcal{C}_e$  contains a cycle that covers the vertices of  $F^e$  or  $F^e - v_{\ell(e)}^e$  or  $F^e - v_{r(e)}^e$  or  $F^e - \{v_{\ell(e)}^e, v_{r(e)}^e\}$  and we replace it by a similar cycle in the constructed gadget. For the opposite direction, the graphs  $F^e$ ,  $F^e - v_{\ell(e)}^e$ ,  $F^e - v_{r(e)}^e$ , and  $F^e - \{v_{\ell(e)}^e, v_{r(e)}^e\}$  are 2-connected by Lemma 6. Hence, they have Hamiltonian cycles by the results of [2]. Furthermore,  $F^e$  has a Hamiltonian  $(v_{\ell(e)}^e, v_{r(e)}^e)$ -path. We use this observation and replace the elements of the projection  $\mathcal{C}'_e$  for the tamed cycle cover of  $G'$ .

Let  $(G, k)$  be the instance of CYCLE COVER obtained by Reduction Rules 6.2–6.4. Notice that up to now we are reducing an instance of CYCLE COVER to an instance of the same problem. To deal with the case when  $F^e$  contains at least two blocks, we reduce  $(G, k)$  to the instance of PRIZE COLLECTING CYCLE COVER.

We consider the graphs  $F_1$ – $F_{11}$  shown in Figure 8. Observe that these graphs have proper interval representations that define the ordering of their vertices. Notice that each graph has two vertices denoted by  $s$  and  $t$ , respectively, that are the first and the last vertices of the vertex ordering defined by the given interval representations. The graphs  $F_1$ – $F_{10}$  have edges that are shown by thick lines. The graphs  $F_1$  and  $F_3$ – $F_{10}$  have one such edge each and  $F_2$  has three such edges. We call them *marked*. It is straightforward to verify that the graphs  $F_1$ – $F_{11}$  have the signatures given in Table 1. Observe that the signatures are pairwise distinct and the list of signatures contains all feasible signatures by Lemma 8. It is also easy to verify that Table 1 gives the correct values of  $\text{cover}(F_i)$ . The general idea of our final reduction rule is to replace  $F^e$  by one of the gadgets  $F_i$  with the same signature. Then we define the weights of the marked edges to encode  $\text{cover}(F)$ .

We construct the initial instance  $(G, \omega, \alpha, r)$  of PRIZE COLLECTING CYCLE COVER from  $(G, k)$  by setting  $\omega(e) = 0$  for every  $e \in E(G)$ , defining  $\alpha(x) = x$  for  $x \in \mathbb{N}$ , and setting  $r = -k$ . Recall that  $(G, k)$  is a yes-instance of CYCLE COVER if and only if  $(G, \omega, \alpha, r)$  is a yes-instance of PRIZE COLLECTING CYCLE COVER. Note that our reduction rule also uses the parameter  $k$  from the constructed by Reduction Rules 6.2–6.4 instance  $(G, k)$  of CYCLE COVER.

**REDUCTION RULE 6.5.** *If  $|V_e| \geq 6$  and  $F^e$  has a cut-vertex, then do the following:*

- (i) *If  $\text{cover}(F) \geq k$ , then*
  - *delete  $v_{\ell(e)+1}^e, \dots, v_{r(e)-1}^e$ ,*
  - *construct a copy of  $F_{11}$  identifying  $s$  and  $t$  with  $v_{\ell(e)}^e$  and  $v_{r(e)}^e$ , respectively,*
  - *set the weights of the edges of the copy of  $F_{11}$  to be 0.*

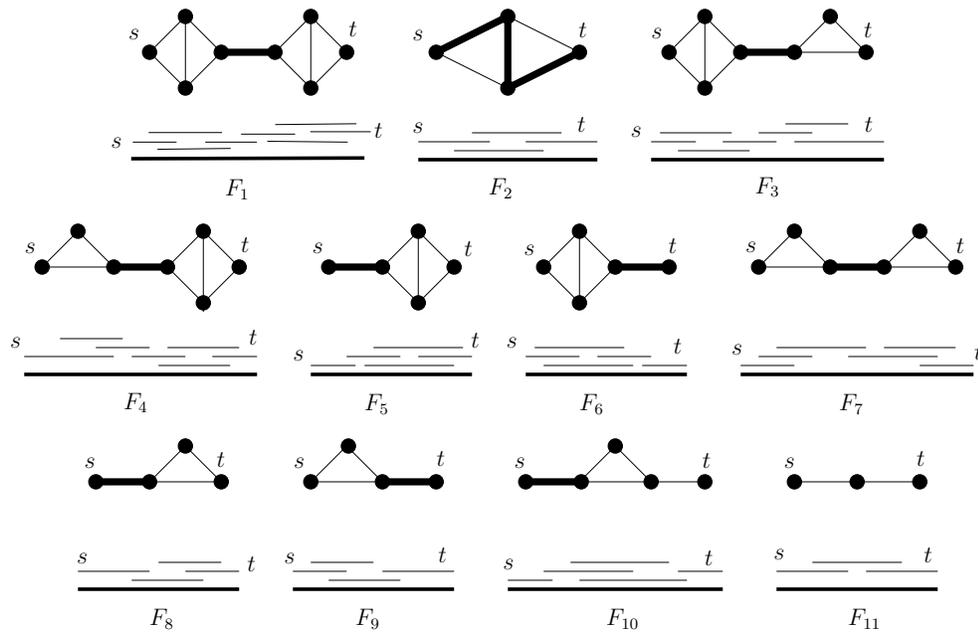


FIG. 8. The construction of  $F_1$ – $F_{11}$  and their proper interval representations; the marked edges are shown by thick lines.

TABLE 1  
The signatures of the graphs  $F_1$ – $F_{11}$  and the values of  $\text{cover } F_i$ .

$F_i$	$\sigma(F_i)$	$\text{cover}(F_i)$	$F_i$	$\sigma(F_i)$	$\text{cover}(F_i)$
$F_1$	$\langle \text{yes, yes, yes, yes} \rangle$	2	$F_7$	$\langle \text{yes, no, no, no} \rangle$	2
$F_2$	$\langle \text{yes, yes, yes, no} \rangle$	1	$F_8$	$\langle \text{no, yes, no, no} \rangle$	1
$F_3$	$\langle \text{yes, yes, no, no} \rangle$	2	$F_9$	$\langle \text{no, no, yes, no} \rangle$	1
$F_4$	$\langle \text{yes, no, yes, no} \rangle$	2	$F_{10}$	$\langle \text{no, no, no, yes} \rangle$	1
$F_5$	$\langle \text{no, yes, no, yes} \rangle$	1	$F_{11}$	$\langle \text{no, no, no, no} \rangle$	$+\infty$
$F_6$	$\langle \text{no, no, yes, yes} \rangle$	1			

(ii) Else, find  $F_i$  with  $i \in \{1, \dots, 10\}$  such that  $\sigma(F_i) = \sigma(F^e)$  and do the following:

- delete  $v_{\ell(e)+1}^e, \dots, v_{r(e)-1}^e$ ,
- construct a copy of  $F_i$  identifying  $s$  and  $t$  with  $v_{\ell(e)}^e$  and  $v_{r(e)}^e$ , respectively,
- set the weights of the marked edges of the copy of  $F_i$  to be  $\text{cover}(F^e) - \text{cover}(F_i)$  and set the weight of the other edges of the copy of  $F_i$  to be 0,
- set  $r = r + \text{cover}(F^e) - \text{cover}(F_i)$  if  $i \neq 2$  and set  $r = r + 3(\text{cover}(F^e) - \text{cover}(F_i))$  if  $i = 2$ .

Observe that Reduction Rule 6.5 produces a feasible instance of PRIZE COLLECTING CYCLE COVER as it assigns nonnegative weights. To see it, observe that  $\text{cover}(F^e) - \text{cover}(F_i) \geq 0$  if Reduction Rule 6.5(ii) is applied, because  $\text{cover}(F^e) \geq 1$  and, moreover,  $\text{cover}(F^e) \geq 2$  if  $F^e$  has a cycle cover since  $F^e$  has at least two blocks. Notice also that the rule does not decrease the value of  $r$ .

Before we show that the rule is safe, we need the following claim.

CLAIM 6.1. Suppose that the instance  $(G', \omega', \alpha, r')$  of PRIZE COLLECTING CY-

CLE COVER is obtained from the initial instance by applying Reduction Rule 6.5. Then for any cycle cover  $\mathcal{C}$  of  $G'$ , there is a tamed cycle cover  $\mathcal{C}'$  such that  $|\mathcal{C}| \geq |\mathcal{C}'|$  and  $c_{\alpha,\omega'}(\mathcal{C}) \leq c_{\alpha,\omega'}(\mathcal{C}')$ .

*Proof of Claim 6.1.* Assume that  $|\mathcal{C}| = k$  and let  $S$  be the set of edges of  $H$  such that for every  $e \in S$ , it holds that  $|V_e| \geq 3$ ,  $L_e \cap R_e = \emptyset$ ,  $G_e$  has a cut-vertex, and  $\mathcal{C}_e$  contains a  $(u, v)$ -path with  $u \in L_e$  and  $v \in R_e$ . By Lemma 12,  $G'$  has a tamed cycle cover  $\mathcal{C}'$  of size at most  $k$  such that for every  $e \in S$ ,  $\mathcal{C}'_e$  consists of a straight path. Let  $\mathcal{C}'$  be such a cycle cover that has maximum weight. We show that  $\omega'(\mathcal{C}') \geq \omega'(\mathcal{C})$ .

Let  $A$  be the set of edges of  $H$  such that for each  $e \in A$ ,  $G_e$  was modified by Reduction Rule 6.5 to one of the gadgets  $F_1$ – $F_{10}$ . Clearly, only edges of  $G'_e$  for  $e \in A$  can have nonzero weights. Let  $e \in A$  and assume that  $F_i$  was used to construct  $G'_e$ . Denote by  $\omega'(\mathcal{C}_e)$  and  $\omega'(\mathcal{C}'_e)$  the total weight of the edges of the elements of the corresponding projections. If  $e \in S$ , then the projection  $\mathcal{C}'_e$  contains a straight path  $P'$ . Note that since  $v_{\ell(e)}^e$  and  $v_{r(e)}^e$  are cut-vertices of  $G'_e$ ,  $P'$  has a  $(v_{\ell(e)}^e, v_{r(e)}^e)$ -subpath of  $P'$  that is an  $(s, t)$ -path in the corresponding gadget  $F_i$ . As  $\mathcal{C}'$  has maximum weight, this subpath should contain the marked edges of the copy of  $F_i$  (see Figure 8). Then  $\omega'(E(G'_e)) = \omega'(E(P'))$  and  $\omega'(\mathcal{C}_e) \leq \omega'(\mathcal{C}'_e)$ . Assume that  $e \in A \setminus S$ . Because  $v_{\ell(e)}^e$  and  $v_{r(e)}^e$  are cut-vertices of  $G'_e$ , the marked edges of  $F_i$  cannot be included in any path of  $\mathcal{C}_e$ . It follows that the edges of the copy of  $F_i$  could only be in cycles of  $\mathcal{C}_e$ . If  $i \neq 2$ , then the marked edges are bridges of  $F_i$ , and therefore no cycle of  $G_e$  can contain them, and we have that  $\omega'(\mathcal{C}_e) = 0$ . Then  $\omega'(\mathcal{C}_e) \leq \omega'(\mathcal{C}'_e)$ . Suppose that  $i = 2$ . Then cycles of  $\mathcal{C}_e$  can contain only two marked edges of the copy of  $F_2$  (see Figure 8). If  $\mathcal{C}'_e$  has a straight path, then this path contains three marked edges. If  $\mathcal{C}'_e$  does not contain a straight path, then the marked edges could only be in cycles of  $\mathcal{C}_e$  and these cycles always contain at least two marked edges. In both cases, we have that  $\omega'(\mathcal{C}_e) \leq \omega'(\mathcal{C}'_e)$ . Since this inequality holds for every  $e \in A$ , we have that  $\omega'(\mathcal{C}) \leq \omega'(\mathcal{C}')$ .

It remains to observe that because  $|\mathcal{C}'| \leq k = |\mathcal{C}|$ ,  $c_{\alpha,\omega'}(\mathcal{C}') = \omega'(\mathcal{C}') - \alpha(|\mathcal{C}'|) \geq \omega'(\mathcal{C}) - \alpha(|\mathcal{C}|) = c_{\alpha,\omega'}(\mathcal{C})$ .  $\square$

Now we prove that Reduction Rule 6.5 is safe. Let  $(G, \omega, \alpha, r)$  be the current instance of PRIZE COLLECTING CYCLE COVER and denote by  $(G', \omega', \alpha, r')$  the instance obtained by application of Reduction Rule 6.5 for an edge  $e \in E(H)$  such that  $F^e$  has a cut-vertex. Assume that the gadget  $F_i$  for  $i \in \{1, \dots, 11\}$  was used for the construction. We show that if  $G$  has a cycle cover  $\mathcal{C}$  of size at most  $k$  with  $c_{\alpha,\omega}(\mathcal{C}) \geq r$ , then  $G'$  has a cycle cover  $\mathcal{C}'$  of size at most  $k$  with  $c_{\alpha,\omega'}(\mathcal{C}') \geq r'$ . For the opposite direction, we show a slightly different claim: if  $G'$  has a cycle cover  $\mathcal{C}'$  with  $c_{\alpha,\omega'}(\mathcal{C}') \geq r'$ , then  $G$  has a cycle cover  $\mathcal{C}$  with  $c_{\alpha,\omega}(\mathcal{C}) \geq r$ . By asymmetry, applying the rule we never increase the size of a cycle cover that provides a solution if the instance  $(G, k)$  of CYCLE COVER obtained by Reduction Rules 6.2–6.4 is a yes-instance. Then it is safe to apply the special case (i) of Reduction Rule 6.5.

Let  $\mathcal{C}$  be a cycle cover of  $G$  of size at most  $k$  such that  $c_{\alpha,\omega}(\mathcal{C}) \geq r$ . We assume that  $\mathcal{C}$  is tamed using Claim 6.1. The general idea of the proof is the same as for the previous rules: we consider  $\mathcal{C}_e$  and modify the elements of the projection to obtain a cycle cover  $\mathcal{C}'$ . We also use the observation that, since  $\mathcal{C}$  is tamed and  $F^e$  contains at least two blocks, either the vertices of  $F^e$  are covered by a straight path or  $\mathcal{C}_e$  contains  $\text{cover}(F^e)$  cycles that cover the vertices of one of the graphs  $F^e$ ,  $F^e - v_{\ell(e)}^e$ ,  $F^e - v_{r(e)}^e$ ,  $F^e - \{v_{\ell(e)}^e, v_{r(e)}^e\}$ .

Assume first that Reduction Rule 6.5(i) was used, that is,  $\sigma(F^e) = \langle no, no, no, no \rangle$  or  $\text{cover}(F^e) \geq k$ . If  $\mathcal{C}_e$  contains a straight path  $P$ , we modify  $P$  using the observation

that  $P$  contains a  $(v_{\ell(e)}^e, v_{r(e)}^e)$ -subpath that contains the vertices of  $F^e$ . We replace this subpath by the unique  $(s, t)$ -path in the copy of  $F_{11}$ . Clearly, this way we obtain a cycle cover  $\mathcal{C}'$  of the same size as  $\mathcal{C}$  such that  $c_{\alpha, \omega'}(\mathcal{C}') = c_{\alpha, \omega}(\mathcal{C}) \geq r = r'$ . Suppose that  $\mathcal{C}_e$  has no straight path. Then  $\mathcal{C}_e$  contains  $\text{cover}(F^e)$  cycles that cover the vertices of one of the graphs  $F^e, F^e - v_{\ell(e)}^e, F^e - v_{r(e)}^e, F^e - \{v_{\ell(e)}^e, v_{r(e)}^e\}$ . Clearly, it is possible only if  $\sigma(F^e) \neq \langle no, no, no, no \rangle$ . But then the minimum number of cycles is at least  $k$ . Since we need at least one more cycle to cover the vertices of  $G$  outside  $F^e$ , we conclude that we use at least  $k + 1$  cycles, contradicting the condition that  $|\mathcal{C}| \leq k$ . Hence,  $\mathcal{C}_e$  consists of a straight path.

Assume now that Reduction Rule 6.5(ii) was used to produce  $(G', \omega', \alpha, r')$ . If  $\mathcal{C}_e$  contains a straight path  $P$ , we replace the  $(v_{\ell(e)}^e, v_{r(e)}^e)$ -subpath of  $P$  by a Hamiltonian  $(s, t)$ -path in the copy of  $F_i$  that contains the marked edges. It is straightforward to verify (see Figure 8) that such a path always exists. We obtain a cycle cover  $\mathcal{C}'$  of the same size as  $\mathcal{C}$  such that  $\omega'(\mathcal{C}') = \omega(\mathcal{C}) + \text{cover}(F^e) - \text{cover}(F_i)$  if  $i \neq 2$ , and  $\omega'(\mathcal{C}') = \omega(\mathcal{C}) + 3(\text{cover}(F^e) - \text{cover}(F_i))$  otherwise. This implies that  $c_{\alpha, \omega'}(\mathcal{C}') - r' = c_{\alpha, \omega}(\mathcal{C}) - r \geq 0$ . Suppose that  $\mathcal{C}_e$  has no straight path. Then  $\mathcal{C}_e$  contains a set  $\mathcal{S}$  of  $\text{cover}(F^e)$  cycles that cover the vertices of one of the graphs  $F^e, F^e - v_{\ell(e)}^e, F^e - v_{r(e)}^e, F^e - \{v_{\ell(e)}^e, v_{r(e)}^e\}$ . Since  $\sigma(F^e) = \sigma(F_i)$ , we can replace the cycles of  $\mathcal{S}$  by a set  $\mathcal{S}'$  of  $\text{cover}(F_i)$  cycles in the copy of  $F_i$  in such a way that  $\mathcal{S}'$  leaves uncovered exactly the same vertices as  $\mathcal{S}$ . Thus, we obtain the cycle cover  $\mathcal{C}'$  of size  $|\mathcal{C}| - (\text{cover}(F^e) - \text{cover}(F_i)) \leq k - (\text{cover}(F^e) - \text{cover}(F_i)) \leq k$ . We also have that  $\omega'(\mathcal{C}') = \omega(\mathcal{C})$  if  $i \neq 2$ , because the cycles of  $\mathcal{S}'$  cannot contain marked edges that are bridges of  $F_i$ . Then  $c_{\alpha, \omega'}(\mathcal{C}') - r' = c_{\alpha, \omega}(\mathcal{C}) - r \geq 0$ . If  $i = 2$ , then  $\mathcal{S}'$  contains a unique cycle and, as can be easily seen from Figure 8, this cycle contains exactly two marked edges. Then  $\omega'(\mathcal{C}') = \omega(\mathcal{C}) + 2(\text{cover}(F^e) - \text{cover}(F_i))$ . Therefore,  $c_{\alpha, \omega'}(\mathcal{C}') - r' = c_{\alpha, \omega}(\mathcal{C}) - r \geq 0$ .

For the opposite direction, let  $\mathcal{C}'$  be a cycle cover of  $G'$  such that  $c_{\alpha, \omega'}(\mathcal{C}') \geq r'$ . Again, we assume that  $\mathcal{C}'$  is tamed using Claim 6.1. We use the same idea as before, and the difference is that now we replace the elements of  $\mathcal{C}'_e$  to produce a cycle cover  $\mathcal{C}$  of  $G$ .

Again, assume first that Reduction Rule 6.5(i) was used. As  $F_{11}$  has no cycle, we immediately obtain that  $\mathcal{C}'_e$  consists of the straight path. We replace the unique  $(s, t)$ -subpath in the copy of  $F_{11}$  by a Hamiltonian  $(v_{\ell(e)}^e, v_{r(e)}^e)$ -path of  $F^e$  which exists by Lemma 4. We obtain a cycle cover  $\mathcal{C}$  of  $G$  of the same size as  $\mathcal{C}'$  such that  $c_{\alpha, \omega}(\mathcal{C}) = c_{\alpha, \omega'}(\mathcal{C}') \geq r' = r$ .

Suppose that we used Reduction Rule 6.5(ii). If  $\mathcal{C}'_e$  consists of a straight path, we replace its  $(v_{\ell(e)}^e, v_{r(e)}^e)$ -subpath  $P$  in the same way as above by a Hamiltonian  $(v_{\ell(e)}^e, v_{r(e)}^e)$ -path of  $F^e$ . We get a cycle cover  $\mathcal{C}$  of  $G$  of the same size as  $\mathcal{C}'$ . Note that  $\omega'(\mathcal{C}') \geq \omega(\mathcal{C}) - (\text{cover}(F^e) - \text{cover}(F_i))$  if  $i \neq 2$  and  $\omega'(\mathcal{C}') \leq \omega(\mathcal{C}) - 3(\text{cover}(F^e) - \text{cover}(F_i))$  if  $i = 2$ , because  $P$  contains at most one and at most three marked edges, respectively. Then  $c_{\alpha, \omega}(\mathcal{C}) - r \geq c_{\alpha, \omega'}(\mathcal{C}') - r' \geq 0$ . Suppose that  $\mathcal{C}'_e$  has no straight path. Then  $\mathcal{C}'_e$  contains a set  $\mathcal{S}'$  of  $\text{cover}(F_i)$  cycles that cover the vertices of one of the graphs  $F_i, F_i - v_{\ell(e)}^e, F_i - v_{r(e)}^e, F_i - \{v_{\ell(e)}^e, v_{r(e)}^e\}$  (recall that  $s$  and  $t$  are identified with  $v_{\ell(e)}^e$  and  $v_{r(e)}^e$ , respectively). Since  $\sigma(F^e) = \sigma(F_i)$ , we can replace the cycles of  $\mathcal{S}'$  by a set  $\mathcal{S}$  of  $\text{cover}(F^e)$  cycles in  $F^e$  in such a way that  $\mathcal{S}$  leaves uncovered exactly the same vertices as  $\mathcal{S}'$ . Thus, we obtain the cycle cover  $\mathcal{C}$  of size  $|\mathcal{C}'| - (\text{cover}(F^e) - \text{cover}(F_i)) \leq k - (\text{cover}(F^e) - \text{cover}(F_i))$ . We also have that  $\omega'(\mathcal{C}') = \omega(\mathcal{C})$  if  $i \neq 2$ , because the cycles of  $\mathcal{S}'$  cannot contain marked edges that are bridges of  $F_i$ . Then  $c_{\alpha, \omega}(\mathcal{C}) - r = c_{\alpha, \omega'}(\mathcal{C}') - r' \geq 0$ . If  $i = 2$ , then  $\mathcal{S}'$  contains a unique cycle and,

as can be easily seen from Figure 8, this cycle contains exactly two marked edges. Then  $\omega'(\mathcal{C}') = \omega(\mathcal{C}) + 2(\text{cover}(F^e) - \text{cover}(F_i))$ . Therefore,  $c_{\alpha, \omega}(\mathcal{C}) - r = c_{\alpha, \omega'}(\mathcal{C}') - r' \geq 0$ .

This completes the safeness proof.

Denote by  $(\hat{G}, \hat{\omega}, \alpha, \hat{r})$  the instance of PRIZE COLLECTING CYCLE COVER obtained by the application of all the rules. Note that we have that  $\hat{\omega}(e) \leq k - 1$  and  $\hat{\omega}(e) \leq -r$ . Observe also that  $\hat{r} \geq -k$  as Reduction Rule 6.5 can only increase the parameter  $r$ . Reduction Rule 6.5 creates at most 3 marked edges for every  $e \in E(H)$ . Therefore, the number of edges with nonzero weights is at most  $3|E(H)|$ .

We construct a clique cover of  $\hat{G}$  as follows.

For each  $e \in E(H)$ , we construct a set  $\mathcal{K}^e$  of at most 6 cliques such that every vertex of  $\hat{G}_e$  is included in at least one of the cliques. Recall that Reduction Rules 6.2–6.5 are applied only for  $e \in E(H)$  only if  $L_e \cap R_e = \emptyset$ . If  $L_e \cap R_e \neq \emptyset$ , then by Lemmas 10(ii) and 10(iii), we have that  $L_e$  and  $R_e$  are nonempty cliques and  $V_e = L_e \cup R_e$ . If  $|V_e| \leq 5$ , then  $\mathcal{K}^e$  can be trivially constructed. If we apply Reduction Rule 6.2 for  $e \in E(H)$ , then  $G_e$  is replaced by a complete graph whose vertices can be covered by a single clique. Suppose that we apply Reduction Rule 6.3. If we do not modify  $B_\ell^e$ , then it has at most 5 vertices that could be covered by at most 2 cliques. The same holds for  $B_r^e$ . If  $B_\ell^e$  ( $B_r^e$ ) is modified, it is replaced by a complete graph that can be covered by a single clique. If  $F^e$  is 2-connected and is not modified by Reduction Rule 6.4, it has at most 5 vertices and can be covered by 2 cliques. Otherwise, again, we can cover the obtained complete graph by a single clique. If  $V^e$  has a cut-vertex but is not modified by Reduction Rule 6.5, we cover its vertices by at most 2 cliques. Assume that  $F^e$  is modified. Consider the graphs  $F_1$ – $F_{11}$  that are shown in Figure 8. Note that for each of them,  $F_i - \{s, t\}$  can be covered by 2 cliques. We conclude that if  $G_e$  has a cut-vertex, then we can cover the graph constructed from  $G_e$  by at most 6 cliques.

For every node  $x \in V(H)$ , let  $K_x = \{v \in V(G) \mid x \in M_v\}$ . Clearly, each  $K_x$  is a clique and we have that  $\mathcal{K} = \{K_x \mid x \in V(H)\} \cup \bigcup_{e \in E(H)} \mathcal{K}^e$  is a family of cliques of  $\hat{G}$  such that every vertex of  $\hat{G}$  is included in at least one clique of  $\mathcal{K}$ . Observe that  $|\mathcal{K}| \leq |V(H)| + 6|E(H)|$ . It may happen that  $\mathcal{K}$  is not a clique cover as the cliques can have common vertices. We construct the clique cover  $\mathcal{Q}$  from  $\mathcal{K}$  by the following greedy procedure: we select an arbitrary nonempty clique  $Q$  in  $\mathcal{K}$ , include it in  $\mathcal{Q}$ , and update the cliques of  $\mathcal{K}$  by deleting the vertices of  $Q$  from them. It is straightforward to verify that  $\mathcal{Q}$  is a clique cover of  $\hat{G}$  and  $|\mathcal{Q}| \leq |V(G)| + 6|E(H)|$ .

Finally, we observe that Reduction Rules 6.2–6.5 can be applied in polynomial time. In particular, we compute  $\text{cover}(F^e)$  in Reduction Rule 6.5 using Theorem 4. Since the construction of  $\mathcal{Q}$  is also polynomial, we conclude that the algorithm runs in polynomial time.  $\square$

Finally, we are ready to prove Theorem 2.

*Proof of Theorem 2.* The proof essentially repeats the proof of Theorem 1. Let  $(G, k)$  be an instance of CYCLE COVER where  $G$  is a proper  $H$ -graph given together with its proper  $H$ -representation  $(H', \mathcal{M})$ . We use the algorithm from Lemma 9, which either solves the problem or constructs an equivalent instance  $(G', k')$  of CYCLE COVER together with a nice proper  $\hat{H}$ -representation  $(\hat{H}', \hat{\mathcal{M}})$  of  $\hat{G}$  such that  $|V(\hat{H}')| \leq 3|E(H)|$  and  $|E(\hat{H}')| \leq 2|E(H)|$ . If the algorithm solved the problem, our algorithm returns a trivial yes- or no-instance, respectively. Otherwise, we apply the algorithm from Lemma 14 for  $(G', k')$  and  $(\hat{H}', \hat{\mathcal{M}})$ . This way we obtain an equivalent instance  $(G'', \omega, \alpha, r)$  of PRIZE COLLECTING CYCLE COVER such that the number of edges with nonzero weight  $\ell \leq 3|E(\hat{H}')| \leq 6|E(H)|$ . The algorithm also constructs a clique cover  $\mathcal{Q}$  of  $G''$  such that  $|\mathcal{Q}| \leq |V(\hat{H}')| + 6|E(\hat{H}')| \leq 15|E(H)|$ . Finally, we use the

algorithm from Theorem 3 for  $(G'', \omega, \alpha, r)$  and  $\mathcal{Q}$  that produces an equivalent instance  $(\hat{G}, \hat{\omega}, \hat{\alpha}, \hat{r})$  of size  $\mathcal{O}((|\mathcal{Q}| + \ell))^{10} = \mathcal{O}(|E(H)|)^{10}$ .

For HAMILTONIAN CYCLE, we write an instance  $G$  of the problem as the instance  $(G, k)$  of CYCLE COVER for  $k = 1$ . Then we combine Lemmas 9 and 14 and obtain an equivalent instance  $(G'', \omega, \alpha, r)$  of PRIZE COLLECTING CYCLE COVER such that  $\alpha(x) = x$  for  $x \in \mathbb{N}$ ,  $r \geq -k = -1$  and  $\omega(e) \leq k - 1 = 0$  for  $e \in E(\hat{G})$ . Then we have that this instance of PRIZE COLLECTING CYCLE COVER is equivalent to the instance  $(G'', -r)$  of CYCLE COVER. If  $r \geq 0$ , we obtain a no-instance and return a trivial no-instance of HAMILTONIAN CYCLE. Otherwise,  $r = -1$  and  $(G'', -r)$  is an equivalent instance of HAMILTONIAN CYCLE. Then we use Theorem 3 to obtain a kernel of size  $\mathcal{O}(|V(H)| + |E(H)|)^8$ .

Finally, in both cases, we use Observation 1 to define the value of the parameter (size of  $G''$ ) as in the proof of Theorem 1.  $\square$

**7. Conclusion.** We obtained compression and kernelization results for HAMILTONIAN PATH and HAMILTONIAN CYCLE and their generalizations (PATH COVER and CYCLE COVER, respectively) for classes of intersection graphs parameterized by their distance from proper interval graphs in a nonstandard way. We proved that HAMILTONIAN CYCLE and PATH COVER on proper  $H$ -graphs admit a polynomial kernel of size  $\mathcal{O}(h^8)$  when parameterized by the size  $h$  of  $H$  if a proper  $H$ -representation is given in the input. For CYCLE COVER, it was shown that it admits a polynomial compression into PRIZE COLLECTING CYCLE COVER. As a by-product, we also established that HAMILTONIAN CYCLE, CYCLE COVER, PATH COVER, and PRIZE COLLECTING CYCLE COVER admit polynomial kernels when parameterized by the clique cover size if a clique cover is given in the input. Here it would be interesting to investigate whether a “robust” approach can be used instead where we would only be given a graph as the input (with the promise of having either a small clique cover or a proper  $H$ -representation for a fixed  $H$ ).

It is natural to ask whether our results for proper  $H$ -graphs can be generalized to (not necessarily proper)  $H$ -graphs. Since HAMILTONIAN CYCLE is NP-complete on strongly chordal split graphs [34], this question is interesting even for special families of graphs  $H$ , like trees or stars. It might also be interesting to consider other covering problems on (proper)  $H$ -graphs. For example, what can be said about CLIQUE COVER? Recall that several other classes of optimization problems have already been considered on  $H$ -graphs [12, 13, 21].

#### REFERENCES

- [1] N. ALON, R. YUSTER, AND U. ZWICK, *Color-coding*, J. ACM, 42 (1995), pp. 844–856.
- [2] A. A. BERTOSSI, *Finding Hamiltonian circuits in proper interval graphs*, Inform. Process. Lett., 17 (1983), pp. 97–101.
- [3] M. BIRÓ, M. HJUTER, AND Z. TUZA, *Precoloring extension. I. Interval graphs*, Discrete Math., 100 (1992), pp. 267–279.
- [4] A. BJÖRKLUND, T. HUSFELDT, P. KASKI, AND M. KOIVISTO, *Narrow sieves for parameterized paths and packings*, J. Comput. Syst. Sci., 87 (2017), pp. 119–139.
- [5] H. L. BODLAENDER, R. G. DOWNEY, M. R. FELLOWS, AND D. HERMELIN, *On problems without polynomial kernels*, J. Comput. System Sci., 75 (2009), pp. 423–434.
- [6] H. L. BODLAENDER, B. M. P. JANSEN, AND S. KRATSCH, *Kernel bounds for path and cycle problems*, Theoret. Comput. Sci., 511 (2013), pp. 117–136.
- [7] A. BRANDSTÄDT, F. F. DRAGAN, AND E. KÖHLER, *Linear time algorithms for Hamiltonian problems on (claw, net)-free graphs*, SIAM J. Comput., 30 (2000), pp. 1662–1677, <https://doi.org/10.1137/S0097539799357775>.

- [8] A. BRANDSTÄDT, V. B. LE, AND J. P. SPINRAD, *Graph Classes: A Survey*, Discrete Math. Appl. 3, SIAM, Philadelphia, 1999, <https://doi.org/10.1137/1.9780898719796>.
- [9] H. BROERSMA, J. FIALA, P. A. GOLOVACH, T. KAISER, D. PAULUSMA, AND A. PROSKUROWSKI, *Linear-time algorithms for scattering number and Hamilton-connectivity of interval graphs*, J. Graph Theory, 79 (2015), pp. 282–299.
- [10] M. CHANG, S. PENG, AND J. LIAW, *Deferred-query: An efficient approach for some problems on interval graphs*, Networks, 34 (1999), pp. 1–10.
- [11] S. CHAPLICK, F. V. FOMIN, P. A. GOLOVACH, D. KNOP, AND P. ZEMAN, *Kernelization of graph hamiltonicity: Proper H-graphs*, in Algorithms and Data Structures, WADS 2019, Lecture Notes in Comput. Sci. 11646, Z. Friggstad, J. R. Sack, and M. Salavatipour, eds., Springer, Cham, 2019, pp. 296–310, [https://doi.org/10.1007/978-3-030-24766-9\\_22](https://doi.org/10.1007/978-3-030-24766-9_22).
- [12] S. CHAPLICK, M. TOEPFER, J. VOBORNÍK, AND P. ZEMAN, *On H-topological intersection graphs*, in the 43rd International Workshop on Graph-Theoretic Concepts in Computer Science (Eindhoven, The Netherlands, 2012), Lecture Notes in Comput. Sci. 10520, Springer, Cham, 2017, pp. 167–179, <https://arxiv.org/abs/1608.02389>.
- [13] S. CHAPLICK AND P. ZEMAN, *Combinatorial problems on H-graphs*, Electron. Notes Discrete Math., 61 (2017), pp. 223–229.
- [14] C. CHEN, C. CHANG, AND G. J. CHANG, *Proper interval graphs and the guard problem*, Discrete Math., 170 (1997), pp. 223–230.
- [15] M. CYGAN, F. V. FOMIN, L. KOWALIK, D. LOKSHTANOV, D. MARX, M. PILIPCZUK, M. PILIPCZUK, AND S. SAURABH, *Parameterized Algorithms*, Springer, Cham, 2015.
- [16] M. CYGAN, J. NEDERLOF, M. PILIPCZUK, M. PILIPCZUK, J. M. M. VAN ROOIJ, AND J. O. WOJTASZCZYK, *Solving connectivity problems parameterized by treewidth in single exponential time*, in Proceedings of the 52nd Annual Symposium on Foundations of Computer Science (FOCS), IEEE, 2011, pp. 150–159.
- [17] P. DAMASCHKE, *Paths in interval graphs and circular arc graphs*, Discrete Math., 112 (1993), pp. 49–64.
- [18] R. G. DOWNEY AND M. R. FELLOWS, *Fundamentals of Parameterized Complexity*, Texts Comput. Sci., Springer, London, 2013.
- [19] M. ETSCHIED, S. KRATSCHE, M. MNICH, AND H. RÖGLIN, *Polynomial kernels for weighted problems*, J. Comput. Syst. Sci., 84 (2017), pp. 1–10.
- [20] M. R. FELLOWS, D. LOKSHTANOV, N. MISRA, M. MNICH, F. A. ROSAMOND, AND S. SAURABH, *The complexity ecology of parameters: An illustration using bounded max leaf number*, Theory Comput. Syst., 45 (2009), pp. 822–848.
- [21] F. V. FOMIN, P. A. GOLOVACH, AND J. RAYMOND, *On the tractability of optimization problems on H-graphs*, Algorithmica, 82 (2020), pp. 2432–2473.
- [22] F. V. FOMIN, D. LOKSHTANOV, S. SAURABH, AND M. ZEHAVID, *Kernelization. Theory of Parameterized Preprocessing*, Cambridge University Press, Cambridge, UK, 2019.
- [23] A. FRANK AND E. TARDOS, *An application of simultaneous diophantine approximation in combinatorial optimization*, Combinatorica, 7 (1987), pp. 49–65.
- [24] M. R. GAREY AND D. S. JOHNSON, *Computers and Intractability: A Guide to the Theory of NP-completeness*, W. H. Freeman, San Francisco, CA, 1979.
- [25] M. C. GOLUBIC, *Algorithmic Graph Theory and Perfect Graphs*, 2nd ed., Ann. Discrete Math. 57, Elsevier Science B.V., Amsterdam, 2004.
- [26] J. HOPCROFT AND R. TARJAN, *Algorithm 447: Efficient algorithms for graph manipulation*, Commun. ACM, 16 (1973), pp. 372–378.
- [27] J. E. HOPCROFT AND R. M. KARP, *An  $n^{5/2}$  algorithm for maximum matchings in bipartite graphs*, SIAM J. Comput., 2 (1973), pp. 225–231, <https://doi.org/10.1137/0202019>.
- [28] R. HUNG AND M. CHANG, *Linear-time certifying algorithms for the path cover and Hamiltonian cycle problems on interval graphs*, Appl. Math. Lett., 24 (2011), pp. 648–652.
- [29] L. IBARRA, *A simple algorithm to find Hamiltonian cycles in proper interval graphs*, Inform. Process. Lett., 109 (2009), pp. 1105–1108.
- [30] J. M. KEIL, *Finding Hamiltonian circuits in interval graphs*, Inform. Process. Lett., 20 (1985), pp. 201–206.
- [31] I. KOUTIS, *Faster algebraic algorithms for path and packing problems*, in Proceedings of the 35th International Colloquium of Automata, Languages and Programming (ICALP), Lecture Notes in Comput. Sci. 5125, Springer, Berlin, 2008, pp. 575–586.
- [32] M. LAMPIS, K. MAKINO, V. MITSOU, AND Y. UNO, *Parameterized edge Hamiltonicity*, Discrete Appl. Math., 248 (2018), pp. 68–78.
- [33] G. K. MANACHER, T. A. MANKUS, AND C. J. SMITH, *An optimum  $\Theta(n \log n)$  algorithm for finding a canonical Hamiltonian path and a canonical Hamiltonian circuit in a set of intervals*, Inform. Process. Lett., 35 (1990), pp. 205–211.

- [34] H. MÜLLER, *Hamiltonian circuits in chordal bipartite graphs*, Discrete Math., 156 (1996), pp. 291–298.
- [35] F. S. ROBERTS, *Indifference graphs*, in Proof Techniques in Graph Theory (Proc. Second Ann Arbor Graph Theory Conf., Ann Arbor, MI, 1968), Academic Press, New York, 1969, pp. 139–146.
- [36] W.-K. SHIH, T. C. CHERN, AND W.-L. HSU, *An  $O(n^2 \log n)$  algorithm for the Hamiltonian cycle problem on circular-arc graphs*, SIAM J. Comput., 21 (1992), pp. 1026–1046, <https://doi.org/10.1137/0221061>.
- [37] R. WILLIAMS, *Finding paths of length  $k$  in  $O(2^k)$  time*, Inform. Process. Lett., 109 (2009), pp. 315–318.