Journal of Combinatorial Optimization
https://doi.org/10.1007/s10878-019-00488-w

Nontrivial path covers of graphs: existence, minimization and maximization

Renzo Gómez[1] · **Yoshiko Wakabayashi**[1]

Abstract

Let G be a graph and $\mathcal{P}$ be a set of pairwise vertex-disjoint paths in G. We say that $\mathcal{P}$ is a *path cover* if every vertex of G belongs to a path in $\mathcal{P}$. In the *minimum path cover problem*, one wishes to find a path cover of minimum cardinality. In this problem, known to be NP-hard, the set $\mathcal{P}$ may contain trivial (single-vertex) paths. We study the problem of finding a path cover composed only of nontrivial paths. First, we show that the corresponding existence problem can be reduced to a matching problem. This reduction gives, in polynomial time, a certificate for both the YES-answer and the NO-answer. When trivial paths are forbidden, for the feasible instances, one may consider either minimizing or maximizing the number of paths in the cover. We show that, the minimization problem on feasible instances is computationally equivalent to the minimum path cover problem: their optimum values coincide and they have the same approximation threshold. We show that the maximization problem can be solved in polynomial time. We also consider a weighted version of the path cover problem, in which we seek a path cover with minimum or maximum total weight (the number of paths do not matter), and we show that while the first is polynomial, the second is NP-hard, but admits a constant-factor approximation algorithm. We also describe a linear-time algorithm on (weighted) trees, and mention results for graphs with bounded treewidth.

Keywords Covering · Min path cover · Max path cover · [1, 2]-factor · Bounded treewidth

Renzo Gómez was financed by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior)—Brasil (CAPES)—Finance Code 001, and MaCLinC Proj. NUMEC/USP.
Yoshiko Wakabayashi was partially supported by CNPq (Proc. 456792/2014-7, 306464/2016-0) and FAPESP (Proc. 2015/11937-9).

✉ Renzo Gómez
rgomez@ime.usp.br

Yoshiko Wakabayashi
yw@ime.usp.br

[1] Institute of Mathematics and Statistics, University of São Paulo, Rua do Matão 1010, São Paulo, SP 05508-090, Brazil

Published online: 30 November 2019

1 Introduction

All graphs considered here are simple and undirected. The *length* of a path in a graph is its number of edges. If a path has length k, we say that it is a k-*path*; and when its length is zero, we say that it is *trivial*. Here, a *path cover* of a graph G means a set of pairwise vertex-disjoint paths that collectively spans $V(G)$.

The Minimum Path Cover (MinPC) problem asks for a path cover of minimum cardinality. Clearly, MinPC is NP-hard on the classes of graphs for which the Hamiltonian path problem is NP-complete. This is known to hold for cubic planar 3-connected graphs (Garey et al. 1976), circle graphs, split graphs, chordal bipartite graphs (Müller 1996), etc. Polynomial-time algorithms have been designed for MinPC on several classes of perfect graphs, such as interval graphs (Arikati and Pandu Rangan 1990), cocomparability graphs (Corneil et al. 2013), trees (Franzblau and Raychaudhuri 2002), etc. No approximation algorithm has been designed for this problem. The cardinality of a minimum path cover of a graph has been studied as a graph parameter (Reed 1996; Magnant and Martin 2009; Yu 2017), and also used to study other parameters such as L(2,1)-labelling (Georges et al. 1994), domination number (Henning and Wash 2017), etc. Some practical applications of MinPC include establishing ring protocols in a network, code optimization and mapping parallel processes to parallel architectures (Moran and Wolfstahl 1991).

We study the problem of finding a path cover without trivial paths. First, we consider the existence problem, and then the corresponding optimization problems: the Minimum Nontrivial Path Cover (MinNtPC) and the Maximum Nontrivial Path Cover (MaxNtPC), both for the cardinality version (number of paths). Furthermore, we study an optimization version of these problems in which we associate weights to the edges of the graph. We consider the Maximum Weight Nontrivial Path Cover (MaxWNtPC) and the Minimum Weight Nontrivial Path Cover (MinWNtPC). In both cases our objective is to optimize the sum of the weights of the edges in the path cover.

In Sect. 2 we show that the existence problem and MaxNtPC are closely related to the maximum matchings of a graph. We show a new characterization of graphs that have a nontrivial path cover, which allows us to obtain an algorithm that solves the existence problem in an interesting way: it returns in polynomial time either (a) a yes-answer which is an optimal solution to the MaxNtPC or (b) a no-answer together with a certificate. We also show that MinNtPC on feasible instances is computationally equivalent to MinPC: we show a polynomial-time algorithm that, from a given minimum path cover, either finds a minimum nontrivial path cover or finds a non-existence certificate. In Sect. 3 we describe a simple linear-time algorithm to solve MinNtPC on trees. In Sect. 4, we show that MaxWNtPC is NP-hard, and that MinWNtPC is solvable in polynomial time. In Sect. 5, we show that for the class of graphs that have nontrivial path covers both problems MinPC and MinNtPC admit the same approximation ratio.

A preliminary version of this work (Gómez and Wakabayashi 2018) was presented at the 44th International Workshop on Graph-Theoretic Concepts in Computer Science (WG 2018).

2 Nontrivial path covers on arbitrary graphs

This section is devoted to the nontrivial path cover problem: both the existence and the optimization version. While every graph admits a path cover, not every graph admits a nontrivial path cover. For example, a star with at least four vertices does not admit a nontrivial path cover. One may naturally ask whether minimum degree two would suffice for a graph to admit a nontrivial path cover. This is not the case, even when we ask for a higher constant degree.

The nontrivial path cover existence problem is, in fact, a special case of classical and intensively studied problems in graph theory. We say that an $[a, b]$-*factor* of a graph G is a spanning subgraph H of G such that each vertex in H has degree at least a and at most b, where a and b are constants. Thus, asking for a $[1, 2]$-factor of a graph is equivalent to asking for a nontrivial path cover. Kano and Saito (1983) proved that, for $r \geq 1$, if a graph G has $\delta(G) \geq r$ and $\Delta(G) \leq r + s$, where $s \in \{1, \ldots, r\}$, then it contains a $[1, 2]$-factor. This means that all regular graphs contain $[1, 2]$-factors. These authors also showed classes of complete bipartite graphs that do not contain $[1, 2]$-factors ($K_{r,r+s}$, with $s > r \geq 1$). A result of Li and Mao-cheng (1998), when specialized to $[1, 2]$-factors guarantees that: if G is an n-vertex-graph of order at least 4 such that $\max\{\deg(x), \deg(y)\} \geq n/3$, for all non-adjacent vertices x and y, then G has a $[1, 2]$-factor. A natural generalization of the concept of $[a, b]$-factor is the concept of (g, f)-factor, in which we replace the constants a and b by integer functions g and f. As in the case of $[a, b]$-factors, there is a large number of structural results on the existence of (g, f)-factors in graphs. In 1952, Tutte characterized graphs that have an (f, f)-factor. Lovász (1970) generalized this result characterizing graphs that have a (g, f)-factor. Later, some results on the algorithmic aspect of this problem were obtained. Anstee (1985) showed a polynomial-time algorithm that finds a (g, f)-factor, if it exists, or a (g, f)-barrier. Heinrich et al. (1990) showed a faster algorithm that finds a (g, f)-factor when $g(x) \leq 1$ and $g(x) < f(x)$ for every vertex x in G. Thus, the nontrivial path cover existence problem has been shown to be solvable in polynomial time.

In what follows, we show that the existence of a nontrivial path cover in a graph is closely related to the structure of its maximum matchings. Clearly, if a graph has a perfect matching, it has a nontrivial path cover consisting solely of 1-paths. And, if a graph has a nontrivial path cover but does not have a perfect matching, we need at least one k-path with $k \geq 2$ to cover it. In fact, we do not need $k > 2$, as such k-paths can be broken into paths of length one or two. This observation indicates that we may focus only on the problem of deciding whether a graph G admits a nontrivial path cover consisting solely of 1-paths or 2-paths.

We denote by $\mathscr{P}_{1,2}(G)$, or simply $\mathscr{P}_{1,2}$, the class of these special types of nontrivial path covers in G, and by $\mathscr{P}^1_{1,2}$ the subclass of $\mathscr{P}_{1,2}$ consisting of path covers with the largest possible number of 1-paths.

Proposition 1 *If a graph G admits a nontrivial path cover, then the cardinality of any path cover in $\mathscr{P}^1_{1,2}(G)$ coincides with the cardinality of a maximum matching in G.*

Proof Consider a path cover $\mathcal{P}' \in \mathscr{P}^1_{1,2}(G)$, and let $\mathcal{P}'_2$ be the set of 2-paths in $\mathcal{P}'$. Let M' be a matching of G obtained by choosing one edge from every path in $\mathcal{P}'$. We claim that M' is a maximum matching of G. Suppose, on the contrary, that this claim does not hold. Let P be an M'-augmenting path that intersects a minimum number of paths in $\mathcal{P}'_2$, and let u and v be its endvertices. Let P_u and P_v be the paths in $\mathcal{P}'_2$ whose endvertices are u and v, respectively. Next, we show that P_u and P_v are the unique paths in $\mathcal{P}'_2$ that intersect P.

Suppose, by contradiction, that there exists a path $Q \in \mathcal{P}'_2$, distinct from P_u and P_v, that intersects P. Consider that $Q := \langle x, y, z \rangle$, and that z is the endvertex of Q not covered by M'. (The proof for the case x is the endvertex of Q not covered by M' is analogous.) Since every internal vertex of P is incident to an edge in M', it follows that $z \notin V(P)$. Therefore, x and y are adjacent vertices in P. Suppose first that $P = \langle u, \ldots, x, y, w, \ldots, v \rangle$. In this case, since $yw \notin M'$, and $w \neq z$, the path $\langle u, \ldots, x, y, z \rangle$ is an M'-augmenting path that contradicts the choice of P. If $P = \langle u, \ldots, y, x, w, \ldots, v \rangle$, then the path $\langle z, y, x, \ldots, v \rangle$ is an M'-augmenting path that contradicts the choice of P.

Since P is an M'-augmenting path from u to v, then P_u and P_v are subpaths (the initial and final part) of P. Now, to conclude the proof, let $E(\mathcal{P}')$ be the set of edges of the path cover $\mathcal{P}'$. Then, the edges in $(E(\mathcal{P}') \setminus (M' \cap E(P))) \cup (E(P) \setminus M')$ define a path cover of G by 1-paths and 2-paths, containing more 1-paths than $\mathcal{P}'$, a contradiction. This concludes the proof that M' is a maximum matching of G, as we claimed. $\square$

Proposition 1 tells us that if a graph G admits a nontrivial path cover, then any path cover in $\mathscr{P}^1_{1,2}(G)$ contains a maximum matching. However, not every maximum matching can be extended to a path cover in $\mathscr{P}^1_{1,2}$ (see Fig. 1).

As we shall see, using the well-known Edmonds–Gallai decomposition (Lovász and Plummer 1986) of G, we can determine whether G admits a nontrivial path cover. This decomposition is defined by the following partition $[D(G), A(G), C(G)]$ of $V(G)$:

- $D(G)$ is the set of vertices in G which are not covered by at least one maximum matching;
- $A(G)$ is the set of vertices in $V(G) \setminus D(G)$ adjacent to at least one vertex in $DG)$; and
- $C(G)$ is the set $V(G) \setminus (A(G) \cup D(G))$.

A *near-perfect* matching of a graph G is one covering all but one vertex of G. A graph G is called *hypomatchable* if $G - v$ has a perfect matching for every vertex v in G. The next theorem describes the structure of every maximum matching in G.

Theorem 1 (Edmonds–Gallai 1965) *If G is a graph and $[D(G), A(G), C(G)]$ is the decomposition of G previously defined, then*

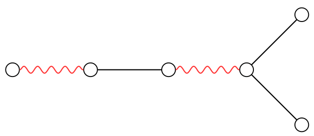

Fig. 1 A maximum matching that cannot be extended to a path cover in $\mathscr{P}^1_{1,2}$

(a) *The components of the subgraph induced by $D(G)$ are hypomatchable;*
(b) *The subgraph induced by $C(G)$ has a perfect matching;*
(c) *Every maximum matching of G contains a near-perfect matching of each component of $D(G)$, a perfect matching of each component of $C(G)$ and matches all the vertices of $A(G)$ with vertices in different components of $D(G)$.*

Next, from a path cover $\mathcal{P}$ in $\mathscr{P}^1_{1,2}(G)$, we define a partition of $V(G)$, as follows:

$$L(\mathcal{P}) := \{v \in V(G) : v \text{ is an endvertex of a 2-path in } \mathcal{P}\};$$
$$R(\mathcal{P}) := \{v \in V(G) : v \text{ is an internal vertex of a 2-path in } \mathcal{P}\};$$
$$S(\mathcal{P}) := V(G) \setminus (L(\mathcal{P}) \cup R(\mathcal{P})).$$

Considering this partition, we prove how it is related to the Edmonds–Gallai decomposition of G.

Proposition 2 *Let G be a graph that admits a nontrivial path cover, and let $\mathcal{P}$ be a path cover in $\mathscr{P}^1_{1,2}(G)$. Moreover, let $[D(G), A(G), C(G)]$ be the Edmonds–Gallai decomposition of G. Then, the following hold:*

(a) $L(\mathcal{P}) \subseteq D(G)$;
(b) *If $u \in S(\mathcal{P}) \cap A(G)$, then the neighbor of u in $\mathcal{P}$ belongs to $D(G)$;*
(c) *Let $u, v \in R(\mathcal{P}) \cup S(\mathcal{P})$ (possibly, $u = v$). Let N be the set of neighbors of u and v in $\mathcal{P}$. If $u, v \in A(G)$, then, each vertex in N belongs to a different component in $D(G)$.*

Proof Denote by P_u the path in $\mathcal{P}$ that contains the vertex $u \in V(G)$. First, we show that (a) holds. Let u be a vertex in $L(\mathcal{P})$. Note that, for every edge in P_u we can find a maximum matching that does not contain that edge, therefore $u \in D(G)$. Now, we prove that (b) and (c) hold. Let M be a maximum matching of G obtained by choosing one edge from every path in $\mathcal{P}$.

Let $u \in S(\mathcal{P}) \cap A(G)$. Since $u \in S(\mathcal{P}) \cap A(G)$, the unique edge of P_u belongs to M. Furthermore, by Theorem 1, the other endvertex of P_u must belong to $D(G)$. Finally, let u and v be vertices in $S(\mathcal{P}) \cup R(\mathcal{P})$. To show (c), we consider three cases depending on whether u or v belong to $S(\mathcal{P})$ or $R(\mathcal{P})$.

Case 1 $u, v \in S(\mathcal{P})$.

In this case, P_u and P_v are 1-paths of $\mathcal{P}$. Let u' and v' be the neighbors of u and v, respectively, in $\mathcal{P}$. Since $uu', vv' \in M$ and $u, v \in A(G)$, then u' and v' must belong to $D(G)$. If $u = v$, then $u' = v'$ and there is nothing to prove. So, suppose that $u \neq v$. Since M is a maximum matching, by Theorem 1, u and v are matched to vertices in different components of $D(G)$. Therefore, u' and v' belong to different components of $D(G)$.

Case 2 $u \in S(\mathcal{P})$ and $v \in R(\mathcal{P})$.

Let u' be the neighbor of u in P_u. Since $v \in R(\mathcal{P})$, it is an internal vertex of a 2-path in $\mathcal{P}$. Let v' and w' be the neighbors of v in P_v, such that $vv' \in M$. We can show that u' and v' belong to different components of $D(G)$ in an analogous way to Case 1. Now, we shall prove that v' and w' belong to different components of $D(G)$. Let K be the component of $D(G)$ that contains v'. Since $vv' \in M$, by Theorem 1, M

contains a near-perfect matching of $K - v'$. This implies that w' must belong to another component of $D(G)$, as w' is not matched in M. Note that an analogous argument shows that v' and w' belong to different components of $D(G)$.

Case 3 $u, v \in R(\mathcal{P})$.

Let u' and t' be the neighbors of u in P_u, such that $uu' \in M$. Also, let v' and w' be the neighbors of v in P_v, such that $vv' \in M$. Using a similar argument as in Case 1, we can show that u' and v' belong to different components of $D(G)$. Also, if we consider a pair of vertices such that one is matched and the other unmatched in M, then we can show that those vertices belong to different components of $D(G)$ using similar arguments as in Case 2. In particular, this shows the case in which $u = v$. Suppose that $u \neq v$. Now, let us prove that t' and w' belong to different components of $D(G)$. Note that, by Theorem 1, there is at most one M-unmatched vertex in each component of $D(G)$. Since w' and t' are unmatched, it follows that they must belong to different components of $D(G)$. □

Proposition 2 shows how the vertices in $A(G) \cup D(G)$ are covered by a path cover in $\mathscr{P}^1_{1,2}(G)$. Next, we show how to reduce the problem of deciding the existence of a nontrivial path cover to a maximum matching problem in a bipartite graph. Let $[D(G), A(G), C(G)]$ be the Edmonds–Gallai decomposition of G. Also, let $T \subseteq D(G)$ be the vertices corresponding to trivial components in $D(G)$. Furthermore, consider a set of vertices, say S, that represent the nontrivial components in $D(G) \setminus T$. Consider the bipartite graph $G' = (U \cup W, F)$ defined in the following way.

$$U = A(G) \cup A' \cup S',$$
$$W = S \cup T,$$

where A' and S' are copies of the sets $A(G)$ and S, respectively. For every vertex $a \in A(G)$ (resp. $s \in S$), we denote by a' (resp. s') its copy in G'. The set F consists precisely of the following edges: (a) an edge between a vertex $a \in A(G) \cup A'$ and a vertex $w \in W$ if there is an edge, in G, incident to a and to a vertex in the component represented by the vertex w; (b) an edge linking s to s', for every $s \in S$. In Fig. 2, we show an example of a graph G' obtained from $A(G) \cup D(G)$.

Now, we state a result from Dulmage and Mendelsohn (1958) that will be useful to prove our reduction.

Theorem 2 (Dulmage and Mendelsohn 1958) *Let G be an (X, Y)-bipartite graph. If there exist two matchings, one that covers $A \subseteq X$ and another that covers $B \subseteq Y$, then there exists a matching that covers $A \cup B$.*

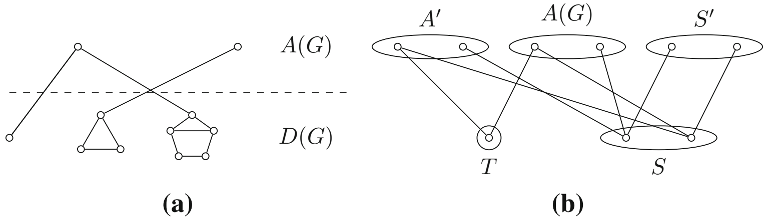

Fig. 2 **a** The subgraph $A(G) \cup D(G)$; **b** the graph G'

Finally, we show that the existence of a nontrivial path cover in a graph G is characterized by the existence of a specific matching in the graph G'.

Theorem 3 *Let G be a graph and let $G' = (U \cup W, F)$ be the bipartite graph defined from G as above. Then, G has a nontrivial path cover if, and only if, there is a matching that covers W in G'.*

Proof First, suppose that G has a nontrivial path cover. Take a path cover $\mathcal{P}$ in $\mathscr{P}^1_{1,2}(G)$. Now, using the structure of $\mathcal{P}$, revealed by Proposition 2, we exhibit a matching M in G' that covers W. We note first that, by this proposition, the 2-paths in $\mathcal{P}$ have both endvertices in $D(G)$. In G', denote by w_K the vertex in W that represents the component K of $D(G)$. Take w_K in W.
Case (1): In $\mathcal{P}$ there is a (unique) 1-path with an endvertex in K and the other endvertex, say a, in $A(G)$. In this case, set $w_K a$ to M. Case (2): In $\mathcal{P}$ there is a 2-path Q that ends in a vertex in K. If Q is entirely contained in K, then set $w_k w'_k$ to M. If the internal vertex of Q, say a, belongs to $A(G)$, then the other endvertex of Q belongs to a component R of $D(G)$ that is distinct from K. In this case, set $w_K a$ and $w_R a'$ to M. Note that, no other 2-path in $\mathcal{P}$ ends in $K \cup R$. Thus, M is a matching in G' that covers W.

Conversely, we show that, if there is a matching in G' that covers W, then there is a nontrivial path cover of G. By Theorem 1, every maximum matching covers the vertices of $A(G)$ with edges whose other endvertices belong to different components of $D(G)$. Note that, these edges induce a matching in G' that covers $A(G)$. Then, by Theorem 2, there exists a matching that covers $A(G) \cup W$ in G'. Let M be such a matching.

Next, we construct a set of edges $F' \subseteq E(G)$ that induces a nontrivial path cover of G. First, let F' be any perfect matching of $C(G)$. Now, let w_K be a vertex of W. We consider two cases.

Case 1 $w_K \in S$ is matched to its copy in S'.

In this case, K is a nontrivial component of $D(G)$. Take a near-perfect matching N of K, and an edge $e \in E(K)$ that is incident to the vertex not covered by N. Add $N \cup \{e\}$ to F'.

Case 2 w_K is matched to a vertex in $A(G) \cup A'$.

Take $u \in A(G) \cup A'$ such that $w_K u \in M$. Let v be a vertex in K that is adjacent to u in G. In this case, consider a perfect matching N of $K - v$ (possibly empty), and add $N \cup \{uv\}$ to F'.

Since M is a matching in G' that covers $A(G) \cup W$, it follows that F' induces a nontrivial path cover of G. □

In Fig. 3, we show a path cover in $\mathscr{P}^1_{1,2}(G)$ obtained from a matching that covers $A(G) \cup W$ in G'. Observe that $\mathscr{P}^1_{1,2}(G)$ consists of path covers of G with the maximum number of paths. Thus, MaxNtPC on G reduces to the problem of finding a path cover in $\mathscr{P}^1_{1,2}(G)$. Theorem 3 shows that the latter can be reduced to a maximum matching problem in a bipartite graph that is obtained using the Edmonds–Gallai decomposition of G. Since both, the Edmonds–Gallai decomposition and a maximum matching, can be computed in polynomial time, we obtain the following result.

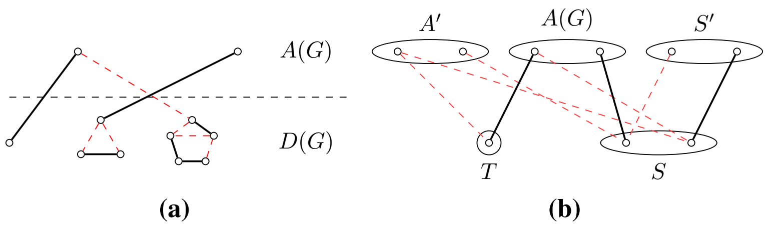

Fig. 3 **a** a nontrivial path cover of $A(G) \cup D(G)$ obtained from **b** a matching in G'

Theorem 4 *The* MAXNTPC *problem can be solved in polynomial time.*

Classical results characterizing graphs containing (g, f)-factors, when specialized to $[1, 2]$-factors, give the following result (see Las Vergnas (1978)).

Theorem 5 (Lovász 1970) *A graph G has a* $[1, 2]$*-factor if, and only if,* $i(G-S) \leq 2|S|$ *for every* $S \subseteq V(G)$*, where* $i(G-S)$ *is the number of isolated vertices in* $G - S$.

Most of the characterization results, except for Anstee (1985), were not concerned with an efficient way to find a NO-certificate (a set S that does not satisfy the condition stated in Theorem 5). Interestingly, our approach of searching for a path cover in $\mathscr{P}^1_{1,2}$ gives an efficient way to find a $[1, 2]$-factor, when it exists, or to find a NO-certificate. The next theorem tells how the latter can be achieved.

Theorem 6 *Let G be a graph,* $D(G)$ *be the set given by the Edmonds–Gallai decomposition of G, and* $T \subseteq D(G)$ *be the set of vertices corresponding to the trivial hypomatchable components in* $D(G)$*. Then the following hold:*

(i) *G has a* $[1, 2]$*-factor if and only if* $|X| \leq 2|N_G(X)|$*, for every* $X \subseteq T$.
(ii) *If G does not have a* $[1, 2]$*-factor, and X is a set that violates the condition stated in* (i)*, then* $S = N_G(X)$ *is a set that violates the condition stated in Theorem 5. Moreover, S can be found in polynomial time.*

Proof First, we prove (i). Let G' be the bipartite graph defined from G as in Theorem 3. By that theorem, it suffices to show that G' has a matching that covers W if and only if $|X| \leq 2|N_G(X)|$, for every subset X of T. First, suppose that G' has a matching that covers W in G'. By Hall's Theorem, for every subset $X \subseteq W$, we have that $|X| \leq |N_{G'}(X)|$. This implies that $|X| \leq |N_{G'}(X)|$, for every $X \subseteq T$, since $T \subseteq W$. By the definition of G', we have that $|N_{G'}(X)| = 2|N_G(X)|$, for every $X \subseteq T$. Now, suppose that $|X| \leq 2|N_G(X)| = |N_{G'}(X)|$, for every $X \subseteq T$. By Hall's Theorem, G' has a matching that covers T. Since $N_{G'}(T) \subseteq A(G) \cup A'$, in G' we can match every vertex $s \in S$ to its copy in S' to obtain a matching that covers W.

Now, to show (ii), note that every subset $X \subseteq T$ is an independent set in G. If we consider $S = N_G(X)$, then the graph $G - S$ has at least $|X|$ isolated vertices. Therefore, if X violates (i), we have that

$$i(G - S) \geq |X| > 2|N_G(X)| = 2|S|.$$

To find a set X that violates (i), we take a maximum matching M' in G', which clearly does not cover some vertex $t \in T$, and construct in G' an M'-alternating tree rooted at t. □

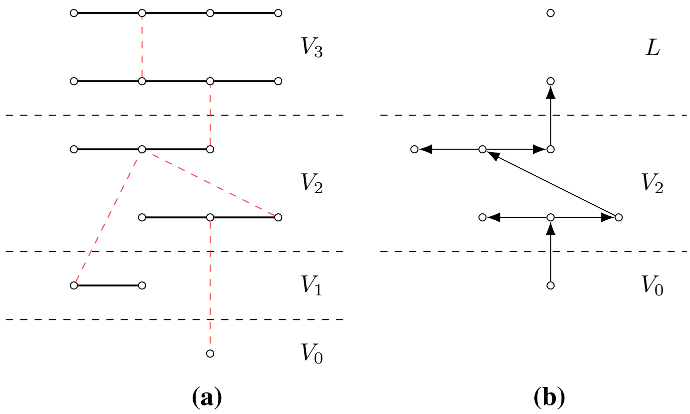

Fig. 4 **a** a graph G and $\mathcal{P}$, and **b** the $\mathcal{P}$-digraph D

Now, we prove an interesting relation between MinPC and MinNtPC. More specifically, given a graph G and a path cover $\mathcal{P}$ of G, we show a polynomial-time algorithm that either obtains a nontrivial path cover of G, if it exists, or finds a set of vertices of G that violates the condition stated in Theorem 5. Moreover, in the former case, the nontrivial path cover obtained by the algorithm has cardinality at most $|\mathcal{P}|$, implying that if $\mathcal{P}$ is a minimum path cover then the nontrivial path cover obtained is also minimum.

If $\mathcal{P}$ denotes a path cover of a graph G, then for every integer $k \geq 0$, we denote by $\mathcal{P}_k$ the set of k-paths in $\mathcal{P}$; and by V_k, the set of vertices spanned by the paths in $\mathcal{P}_k$. In our algorithm, the paths in $\mathcal{P}_2$ play a special role, in particular their internal vertices. The set of such vertices will be denoted by B (as we think of such paths as being of the form $\langle a, b, c \rangle$).

We say that a path cover $\mathcal{P}$ is *good* if no path in $\mathcal{P}$ has an endvertex that is adjacent to an endvertex of another path in $\mathcal{P}$. Clearly, given a path cover that is not good, by joining endvertices of different paths, we can obtain in polynomial time another path cover (with fewer paths) that is good.

Given a graph G and a good path cover $\mathcal{P}$ of G, we will define a directed graph D, called $\mathcal{P}$*-digraph* associated with G, which will help us obtain a nontrivial path cover of G, if it exists.

First, consider the graph $G' = (V', E')$ that results from G by removing the vertices in V_1 and contracting each path in $\bigcup_{k \geq 3} \mathcal{P}_k$ into a single vertex. Let L be the set of vertices that result from the contraction of those (long) paths. Now, let G^* be the graph obtained from G' by removing all edges with both ends in L and also all edges with an end in B and the other in $B \cup L$.

Now we are ready to define the $\mathcal{P}$-digraph D. It is an orientation of G^* defined in the following way. First, for every vertex $v \in V_0$, orient the edges incident to v as going out from that vertex. Second, for every edge $uv \in E(G^*)$ such that $u \in B$ and $v \in V_2 \setminus B$, orient that edge from v to u if u and v belong to different paths in $\mathcal{P}_2$, otherwise orient it from u to v. Third, orient every edge with one end in $V_2 \setminus B$ and the other in L as going from $V_2 \setminus B$ to L. Finally, orient arbitrarily every edge with both ends in $V_2 \setminus B$. For each vertex $v \in V$, we denote by P_v the path in $\mathcal{P}$ that contains v.

We show in Fig. 4 the $\mathcal{P}$-digraph D obtained from a path cover $\mathcal{P}$ represented by a trivial path and the paths consisting of solid edges.

The orientation of the arcs in the $\mathcal{P}$-digraph D is defined in such a way that, the directed paths in D starting at V_0 and ending in a vertex in $B \cup L$, may help finding another path cover of G with fewer trivial paths or confirm the non-existence of such a path cover.

Lemma 1 *Let G be a graph, $\mathcal{P}$ be a good path cover of G, and D be the $\mathcal{P}$-digraph associated with G. Let $\langle a, b, c\rangle$ be a path in $\mathcal{P}_2$ such that $ac \in E(G)$. If D has a directed path from V_0 to b, then G has a path cover $\mathcal{P}'$ such that $|\mathcal{P}'| < |\mathcal{P}|$ and $|\mathcal{P}'_0| < |\mathcal{P}_0|$.*

Proof Let $P = \langle u = r_1, r_2, \dots, r_k = b\rangle$ be a shortest path in D from V_0 to b. We claim that the following hold.

(a) $k = 2m$ for some $m \geq 1$; and
(b) $r_{2i} \in B, r_{2i+1} \in V_2 \setminus B$, for $i = 1, 2, \dots, m-1$;
(c) r_{2i} and r_{2i+1} belong to the same path in $\mathcal{P}_2$, for $i = 1, \dots, m-1$.

In fact, considering the construction of D, we have that

(i) $N^+(v) = \emptyset$, for $v \in L$;
(ii) $N^+(v) \subseteq B \cup L$, for $v \in V_0$;
(iii) $N^+(v) \subseteq V_2 \cup L$, for $v \in V_2 \setminus B$;
(iv) $N^+(v) = V(P_v) \setminus \{v\}$, for $v \in B$.

Since P begins at V_0 and ends at B, by (i), we have that $V(P) \cap L = \emptyset$. Therefore, by (ii), $r_2 \in B$. Moreover, by (iii) and (iv), if an internal vertex of P belongs to $V_2 \setminus B$, then the next vertex in P belongs to B and vice versa, so the claim follows.

Now, we will construct a path cover of G that has fewer paths than $\mathcal{P}$. Let $Q_i \in \mathcal{P}_2$ be the path that contains r_{2i}, and let s_i be the endvertex of Q_i which is different from r_{2i+1}, for $i = 1, \dots, m-1$. Thus, $Q_m = \langle a, b, c\rangle$. Without loss of generality, suppose that the edge $ac \in E(G)$ is oriented from a to c in D. Consider the paths Q_i^* defined in the following way,

$$Q_i^* := \begin{cases} \langle r_{2i-1}, r_{2i}, s_i\rangle, & \text{for } 1 \leq i \leq m-1, \\ \langle r_{2m-1}, b, a, c\rangle, & \text{for } i = m. \end{cases}$$

Let $\mathcal{P}'$ be obtained from $\mathcal{P}$ by replacing the paths $Q_1, Q_2, \dots, Q_m, P_u$ with the paths $Q_1^*, Q_2^*, \dots, Q_m^*$. Clearly, $\mathcal{P}'$ is a path cover of G with $|\mathcal{P}'| < |\mathcal{P}|$ and $|\mathcal{P}'_0| < |\mathcal{P}_0|$. □

The proof of the next theorem contains the description of the steps of the algorithm that either returns a nontrivial path cover of a graph or a non-existence certificate.

Theorem 7 *Let G be a graph and let $\mathcal{P}$ be a path cover of G containing trivial paths. Then, we can obtain in polynomial time either a nontrivial path cover $\mathcal{P}'$ such that $|\mathcal{P}'| \leq |\mathcal{P}|$, if it exists, or a certificate that G does not have a nontrivial path cover.*

Proof We assume that $\mathcal{P}$ is a good path cover (as it suffices to prove for this case). Now, consider the $\mathcal{P}$-digraph D associated with G (that can be constructed in polynomial time). Check whether there is path $\langle a, b, c\rangle$ in $\mathcal{P}_2$ such that $ac \in E(G)$ and D has a directed path from V_0 to b. If this happens, by Lemma 1, we can obtain (in polynomial time) a path cover of G with fewer paths and fewer trivial paths. If the new path cover is nontrivial, the proof is complete; otherwise, we repeat this procedure until we obtain a good path cover, say $\widehat{\mathcal{P}}$, that has no 2-path whose endvertices are adjacent, and whose internal vertex b is the endvertex of a directed path starting at V_0 in the $\widehat{\mathcal{P}}$-digraph D associated with G. We say that such a path cover is *very good*. Clearly, it can be obtained in polynomial time. For ease of notation, let us call $\mathcal{P}$ the very good path cover obtained this way, and consider the $\mathcal{P}$-digraph D associated with G. If $\mathcal{P}$ contains trivial paths, then one of the following two cases must occur.

Case 1 there exists a directed path from V_0 to L in D.

Let P be such a path with minimum length. By arguments similar to those used in the proof of Lemma 1, we can consider that

$$P := \langle u, x_1, y_1, \ldots, x_m, y_m, v\rangle,$$

where $u \in V_0$, $v \in L$, $x_i \in B$, $y_i \in V_2 \setminus B$ and $m \geq 0$. Without loss of generality, consider that $m \geq 1$ (the case in which $m = 0$ can be dealt in a similar way). We recall that P_v denotes the path in $\mathcal{P}$ that was contracted to v. Let w be a vertex of P_v that is a neighbor of y_m in G. Since y_m is an endvertex of a path in $\mathcal{P}$, and $\mathcal{P}$ is good, we conclude that w is an internal vertex of P_v. Let P'_v be a shortest subpath of P_v containing w and an endvertex of P_v. Also, consider that $Q'_v = P_v \setminus V(P'_v)$. Since $|P_v| \geq 3$, we have that $\left|Q'_v\right| \geq 1$.

Let $Q_i \in \mathcal{P}_2$ be the path that contains x_i, and let z_i be the endvertex of Q_i which is different from y_i, for $i = 1, 2, \ldots, m$. Now, we define the paths Q^*_i in the following way,

$$Q^*_i = \begin{cases} \langle u, x_1, z_1\rangle, & \text{if } i = 1,\\ \langle y_{i-1}, x_i, z_i\rangle, & \text{if } 1 < i \leq m,\\ \langle y_m\rangle \cdot P'_v, & \text{if } i = m+1. \end{cases}$$

Let $\mathcal{Q}$ be the path cover of G obtained from $\mathcal{P}$ when we replace the paths in $\{P_u, Q_1, \ldots, Q_m, P_v\}$ by the paths in $\{Q^*_1, \ldots, Q^*_{m+1}, Q'_v\}$. Note that $|\mathcal{Q}| = |\mathcal{P}|$ and $|\mathcal{Q}_0| = |\mathcal{P}_0| - 1$. In Fig. 5 we show an example considering the path cover $\mathcal{P}$ and the $\mathcal{P}$-digraph D shown in Fig. 4.

Case 2 there is no directed path from V_0 to L in D.

Let u be a vertex in V_0. If $N_G(u) = \emptyset$, then $S = \emptyset$ violates the condition stated in Theorem 5. So, suppose that $N_G(u) \neq \emptyset$. Since there is no directed path from V_0 to L, and $\mathcal{P}$ is good, we have that $N^+_D(u) \subseteq B$. Let B_u be the set of vertices of B that are reachable by a directed path from u.

Let $S := B_u$. In what follows, we show that S violates the condition stated in Theorem 5. First, note that u is an isolated vertex in $G - S$. Now, let T_u be the set of the endvertices of the paths in $\mathcal{P}_2$ whose internal vertex belongs to B_u. We will show

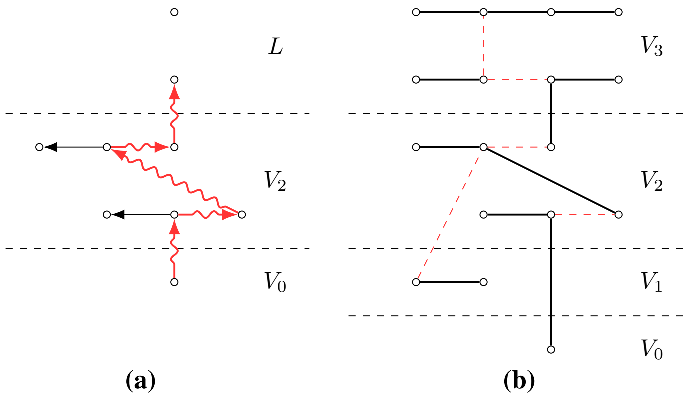

Fig. 5 **a** a path P; and **b** the resulting path cover $\mathcal{Q}$

that T_u is an independent set in $G - S$. Let t be any vertex in T_u. Since there is no directed path from u to L in D, we have that t is not adjacent, in G, to any vertex in V_k, $k \geq 3$. Since $\mathcal{P}$ is very good, we have that t is not adjacent, in G, to any vertex in $V_2 \setminus S$. (Note that, if t is adjacent to an internal vertex b of another path in $\mathcal{P}_2$, then b is also reachable by a directed path from u, and thus, b belongs to S.)

Moreover, since $\mathcal{P}$ is good, the vertex t is not adjacent to any vertex in $V_0 \cup V_1$. Therefore, t is an isolated vertex in $G - S$. Since $G - S$ has at least $2\,|S| + 1$ isolated vertices, by Theorem 5, the graph G does not have a nontrivial path cover.

If Case 1 occurred, and the new path cover still contains trivial paths, we repeat the procedure. As, each time this procedure is repeated we obtain a path cover with fewer trivial paths, we either find a nontrivial path cover of G, or Case 2 occurs and we find a non-existence certificate. All these steps can be done in polynomial time. □

Let us denote by $\mu(G)$ the cardinality of a minimum path cover of G. Also, if G admits a nontrivial path cover, then let $\mu_{nt}(G)$ be the cardinality of a minimum nontrivial path cover of G. By Theorem 7, the following results follow.

Corollary 1 *If a graph G admits a nontrivial path cover, then $\mu_{nt}(G) = \mu(G)$.*

Corollary 2 *Let $\mathcal{G}$ be the class of graphs that admit a nontrivial path cover. The problems* MINPC *and* MINNTPC *have the same computational complexity in $\mathcal{G}$.*

Corollary 1 implies a result on the approximability of MINPC and MINNTPC, which we state in Sect. 5.

3 Minimum nontrivial path covers on trees

As we mentioned, MINPC on trees can be solved in linear time. Thus, given a graph G, we can take a minimum path cover of G and use Theorem 7 to either transform it into a minimum nontrivial path cover of G, if it exists; or exhibit a NO-certificate. Instead of using such a procedure, we show a simpler linear-time algorithm for MINNTPC on trees, which can be easily extended for the edge-weighted version of the problem (presented in the next section).

To understand the idea behind the algorithm, given a tree T, suppose that T has a minimum nontrivial path cover $\mathcal{P}$. Consider that T' is the arborescence obtained when we root T at a vertex r in T. Let T'_u be the subtree of T' rooted at u. Note that, when we consider the subgraph of T' spanned by the edges of $\mathcal{P}$, the vertex u has degree 0, 1 or 2 in T'_u.

Let $f(u, d)$ be the cardinality of a minimum nontrivial path cover of T_u, in which the vertex u has degree d. Algorithm 1, described below, computes the values $f(u, d)$ in post-order using a DFS traversal of T. We choose r to be any leaf of T, and call DFS(r, nil). At the end, if T has a nontrivial path cover, then $f(r, 1) < +\infty$, and it indicates the cardinality of a minimum nontrivial path cover. We may retrieve the edges of such a cover by modifying slightly this algorithm.

Algorithm 1 DFS$(v, parent)$

Input: A tree T
Output: Values $f(v, d)$ for every vertex v in T and integer d in $\{0, 1, 2\}$
1: $deg \leftarrow 0$
2: $m_1 \leftarrow +\infty$
3: $m_2 \leftarrow +\infty$
4: $sum \leftarrow 0$
5: **for** $u \in N(v)$:
6:     **if** $u \neq parent$:
7:         DFS(u, v)
8:         $Y \leftarrow \min(f(u, 1), f(u, 2))$
9:         $X \leftarrow \min(f(u, 1), f(u, 0))$
10:         $sum \leftarrow sum + Y$
11:         $deg \leftarrow deg + 1$
12:         **if** $m_1 > X - Y$:
13:             $m_2 \leftarrow m_1$
14:             $m_1 \leftarrow X - Y$
15:         **else if** $m_2 > X - Y$:
16:             $m_2 \leftarrow X - Y$
17: **for** $d = 0$ **to** 2 :
18:     $f(v, d) \leftarrow +\infty$
19: **if** $deg \geq 0$:
20:     $f(v, 0) \leftarrow sum + 1$
21: **if** $deg \geq 1$:
22:     $f(v, 1) \leftarrow sum + m_1$
23: **if** $deg \geq 2$:
24:     $f(v, 2) \leftarrow sum + m_1 + m_2 - 1$

Theorem 8 *Algorithm 1 correctly computes the values $f(v, d)$ for every vertex v in T and $d \in \{0, 1, 2\}$.*

Proof Let v be any vertex in T. We prove the statement by induction on $n := |V(T_v)|$. If $n = 1$, then v is a leaf of the tree. In this case, at the end of the loop starting at line 5, we have that $sum = deg = 0$. Therefore, at the end of the algorithm $f(v, 0) = 1$ and $f(v, 1) = f(v, 2) = +\infty$.

Now, suppose that $n \geq 2$. Let $v_1, v_2, \ldots, v_k$ be the neighbors of u in T'_u. First, if $d > k$, then there is no solution to $f(u, d)$. Therefore, in line 17, we set $f(u, d) = +\infty$, and this value is not changed afterwards. So, let us suppose that $d \leq k$. Since we are restricting the path covers to those in which u has degree d in T'_u, we have to choose d edges incident to u and v_i, $1 \leq i \leq k$, to belong to the path cover. Observe that for the vertices v_i such that uv_i belongs to the path cover, the degree in its corresponding subtree T'_{v_i} can be zero or one. In case this edge does not belong to the cover, its degree in T'_{v_i} must be one or two. For $i = 1, 2, \ldots, k$, let $X_{v_i} := \min\{f(v_i, 1), f(v_i, 0)\}$ and let $Y_{v_i} := \min\{f(v_i, 1), f(v_i, 2)\}$. Now, we show how to express $f(u, d)$ in terms of X_{v_i} and Y_{v_i}.

First, let us suppose that $d = 2$, and let v_a and v_b be neighbors of u in a minimum path cover where u has degree two in T'_u. By the previous arguments, we have that

$$f(u, 2) = \sum_{\substack{i=1 \\ i \neq a,b}}^{k} Y_{v_i} + X_{v_a} + X_{v_b} - 1.$$

Adding and subtracting Y_{v_a} and Y_{v_b}, we get the following expression

$$f(u, 2) = \sum_{i=1}^{k} Y_{v_i} + (X_{v_a} - Y_{v_a}) + (X_{v_b} - Y_{v_b}) - 1.$$

Since $\sum_{i=1}^{k} Y_{v_i}$ is a constant, to compute the value of $f(u, 2)$, we need to find two neighbors of u that minimize $X_{v_i} - Y_{v_i}$. Observe that at the end of the loop starting at line 5, the variable *sum* is equal to $\sum_{i=1}^{k} Y_{v_i}$, and m_1 and m_2 hold the desired minima. By induction hypothesis our algorithm correctly computes the values of $f(v_i, d)$ for $1 \leq i \leq k$, $1 \leq d \leq 2$. Therefore, the algorithm correctly computes the value $f(u, 2)$. The cases in which $d \in \{0, 1\}$ can be shown by analogous arguments. □

Since we process each vertex u of the tree just once, and we iterate through its neighbors to compute the values of $f(u, d)$, the complexity of our algorithm is $\mathcal{O}(n)$, where n is the number of vertices of the tree. Thus, we obtain the following result.

Corollary 3 MinNtPC *on trees can be solved in linear time.*

4 Weighted path covers

We now focus on the weighted versions of the path cover problems. In these problems, a graph with nonnegative weights on the edges is given, and we seek for a path cover with maximum or minimum total weight.

We note that MinPC is equivalent to the problem of finding a path cover with maximum total weight on a graph with uniform (or unit) weights. To see this, consider a graph $G = (V, E)$ with weight $w(e) = 1$, for all $e \in E$. If $\mathcal{P}$ is a path cover of G, then

$$\sum_{P \in \mathcal{P}} w(E(P)) = \sum_{P \in \mathcal{P}} (|V(P)| - 1) = |V| - |\mathcal{P}|.$$

Thus, the weighted maximization version, called MAXPC, is also NP-hard on the classes of graphs for which MINPC is NP-hard. Moreover, this fact also implies that the MAXWNTPC problem is NP-hard on these classes. Now, we consider the minimization version.

Theorem 9 MINWNTPC *can be solved in polynomial time.*

Proof Let $G = (V, E)$ be a graph that admits a nontrivial path cover. Also, let $w \in \mathbb{R}_+^E$ be the nonnegative weights associated with the edges of G. Let $G_1 = (V_1, E_1)$ and $G_2 = (V_2, E_2)$ be copies of G such that $v_i \in V_i$ is the copy of vertex $v \in V$. From G_1 and G_2, we define the graph $G' = (V', E')$ in the following way:

$$V' = V_1 \cup V_2,$$
$$E' = E_1 \cup E_2 \cup \{v_1 v_2 : v \in V\}.$$

Also, consider the following weights on the edges of G'

$$w'_e = \begin{cases} 0, & \text{if } e = v_1 v_2 \text{ for some } v \in V, \\ w_e, & \text{otherwise.} \end{cases}$$

Now, we prove that finding a minimum weight nontrivial path cover in G is equivalent to finding a 2-factor of minimum weight in G'. First, we show that G' has a 2-factor. Let $\mathcal{P}$ be a nontrivial path cover of G. Consider $\mathcal{P}_i$ as the copy of $\mathcal{P}$ in G_i, for $i = 1, 2$. To obtain a 2-factor of G', we take each path in $\mathcal{P}_1$ and its copy in $\mathcal{P}_2$ and use the two edges linking its corresponding endvertices. In what follows, $\mathcal{P}^*$ (resp. $\mathcal{C}^*$) denotes a minimum weight nontrivial path cover (resp. 2-factor) of G (resp. G'). The previous construction shows that

$$w'(\mathcal{C}^*) \leq 2w(\mathcal{P}^*). \tag{1}$$

Now, we show that $w'(\mathcal{C}^*) \geq 2w(\mathcal{P}^*)$. Let $\mathcal{C}_1^*$ and $\mathcal{C}_2^*$ be the subgraphs induced by the edges of $\mathcal{C}^*$ in G_1 and G_2, respectively. Observe that

$$w'(\mathcal{C}^*) = w'(\mathcal{C}_1^*) + w'(\mathcal{C}_2^*).$$

It is immediate that $w'(\mathcal{C}_1^*) = w'(\mathcal{C}_2^*)$. Otherwise, using the lightest of $\mathcal{C}_1^*$ and $\mathcal{C}_2^*$, we could construct a 2-factor $\widehat{\mathcal{C}}$ with $w'(\widehat{\mathcal{C}}) < w'(\mathcal{C}^*)$.

Since $\mathcal{C}^*$ is a 2-factor, and each vertex in G_1 is adjacent to only one vertex in G_2, we have that $\mathcal{C}_1^*$ is a collection of cycles and nontrivial paths. Therefore, if we remove one edge from every cycle contained in $\mathcal{C}_1^*$, then we obtain a nontrivial path cover of G_1, say $\mathcal{P}'$. Since the weights are nonnegative, this implies that

$$2w(\mathcal{P}^*) \leq 2w'(\mathcal{P}') \leq 2w'(\mathcal{C}_1^*) = w(\mathcal{C}^*). \tag{2}$$

From (1) and (2), we have that $w(\mathcal{C}^*) = 2w(\mathcal{P}^*)$. Furthermore, when we remove one edge from each cycle of $\mathcal{C}_1^*$, we obtain a minimum weight nontrivial path cover of G. As we can find a 2-factor of minimum weight in polynomial time (Schrijver 2003), the result follows. □

Observe that, in the proof of Theorem 9, the subgraph $\mathcal{C}_1^*$ is a minimum-weight $[1, 2]$-factor of G. Moreover, an analogous claim is valid for $\mathcal{C}_1^*$ in the maximization case. Therefore, the following result follows.

Corollary 4 *Let G be a graph with nonnegative weights associated with its edges. Then, an optimum weight* $[1, 2]$*-factor of G can be found in polynomial time.*

5 Approximability

We present now some approximability results for MinPC, MinNtPC, MaxPC and MaxWNtPC (defined in Sect. 4). First, since the cardinality of any path cover is an integer number, then for any $\epsilon > 0$, there cannot exist a $(2 - \epsilon)$-approximation for MinPC or MinNtPC, otherwise, such algorithm would solve the Hamiltonian path problem in polynomial time. Therefore, we have that

Corollary 5 *For every* $\epsilon > 0$*, there is no* $(2 - \epsilon)$*-approximation for* MinPC *and* MinNtPC*, unless* P $\neq$ NP.

Now, consider a special variant of the Minimum Traveling Salesman problem, in which the edge-weights are 1 or 2, denoted as MinTSP-(1,2), known to be Max SNP-hard (Papadimitriou and Yannakakis 1993). Vishwanathan (1992) showed a relation between the approximation ratios of MaxPC and MinTSP-(1, 2).

Proposition 3 (Vishwanathan 1992) *If there exists an* α*-approximation algorithm for* MaxPC*, then there is a* $(2 - \alpha)$*-approximation algorithm for* MinTSP-(1, 2).

In a similar way, we show an analogous relation between MinPC and MinTSP-(1, 2).

Proposition 4 *If there is an* α*-approximation algorithm for* MinPC*, then there is a* $(1 + \alpha)/2$*-approximation algorithm for* MinTSP-(1, 2).

Proof Let (K, c) be a complete graph on n vertices whose edge-weights are 1 or 2. Let G be the spanning subgraph of K that contains only the edges of weight one. Let $\mathcal{A}$ be an α-approximation for MinPC, $\alpha \geq 1$, and let $\mathcal{P}$ be a path cover of G returned by Algorithm $\mathcal{A}$. Thus,

$$|\mathcal{P}| \leq \alpha \, |\mathcal{O}| \, , \tag{3}$$

where $\mathcal{O}$ is a minimum path cover of G. Linking the paths of $\mathcal{P}$ (resp. $\mathcal{O}$) we can obtain a cycle in K of weight at most $n + |\mathcal{P}|$ (resp. $n + |\mathcal{O}|$). From (3), we have that $n + |\mathcal{P}| \leq n + \alpha\,|\mathcal{O}|$. Finally, the approximation ratio of such algorithm is given by the following expression.

$$\inf\left\{r \geq 1 : \frac{n + \alpha\,|\mathcal{O}|}{n + |\mathcal{O}|} \leq r,\ 1 \leq |\mathcal{O}| \leq n\right\},$$

whose optimal value is attained when $r = (1 + \alpha)/2$. □

Now, we show an interesting relation between the approximability of MINPC and MINNTPC.

Corollary 6 *Let $\mathcal{G}$ be the class of graphs that admit a nontrivial path cover. There is an α-approximation for* MINPC *in $\mathcal{G}$, if and only if, there is an α-approximation for* MINNTPC *in $\mathcal{G}$.*

Proof First, observe that if there is an α-approximation for MINNTPC, then by Corollary 1, this is also an α-approximation for MINPC. Now, let $\mathcal{A}$ be an α-approximation for MINPC, and let $\mathcal{P}$ be a path cover returned by $\mathcal{A}$ when applied to a graph $G \in \mathcal{G}$. By Theorem 7, we can obtain in polynomial time a nontrivial path cover $\mathcal{P}'$ of G, such that $|\mathcal{P}'| \leq |\mathcal{P}|$. Let $\mathcal{O}$ be a minimum nontrivial path cover of G. By Corollary 1, we have that $|\mathcal{P}| \leq \alpha|\mathcal{O}|$. Therefore, $|\mathcal{P}'| \leq \alpha|\mathcal{O}|$, and hence $\mathcal{A}$ is also an α-approximation for MINNTPC. □

Finally, let $G = (V, E)$ be a graph with nonnegative weight w_e, for every edge $e \in E$. By Corollary 4, we can find a $[1, 2]$-factor of G of maximum weight, say $\mathcal{C}$, in polynomial time. Note that, if we remove an edge of minimum weight from every cycle in $\mathcal{C}$, we loss at most $1/3$ of $w(\mathcal{C})$. This implies the following result.

Corollary 7 *There is a $(2/3)$-approximation algorithm for* MAXWNTPC.

6 Path cover problems on graphs with bounded treewidth

Many hard problems when restricted to graphs with bounded treewidth can be solved in polynomial time. Thus, it is natural to ask whether this is the case for the hard path cover problems that we have considered here. The answer is yes, and in fact, these problems can be solved in linear time: a result more of theoretical interest, as large constant factors (exponential growth in terms of treewidth) are hidden in the O-notation.

Let $G = (V, E)$ be a graph. We show that the property "G has a nontrivial path cover" can be formulated in monadic second order logic (MSOL). Roughly speaking, for graphs, in *second order logic* we can use quantifications on the vertices and the edges (or subsets of them), and the relation is the incidence relation of the graph. The *monadic second order logic* is a restriction of the second order logic in which quantifications are allowed only on first order variables and on unary relations. We will not elaborate more on this, but we refer the reader to Courcelle and Engelfriet (2012), for a comprehensive work on this subject.

The following theorem, shown by Courcelle (1990), states an important algorithmic consequence for such properties.

Theorem 10 Courcelle (1990) *Every graph property definable in monadic second order logic can be decided in linear time on the class of graphs with bounded treewidth.*

We observe that the property "G has a nontrivial path cover" is equivalent to "G has a subset $F \subseteq E$ that induces an acyclic $[1, 2]$-factor". We give below a formula φ that describes this property in MSOL. It is composed of other formulas that are expressed subsequently. All of them use the incidence relation inc, where $\text{inc}(a, b)$ is *true* if and only if b is an endpoint of a or a is an endpoint of b.

$$\begin{aligned}\varphi \triangleq \exists F \subseteq E,\\ &\forall v \in V[\text{deg1}(v, F) \vee \text{deg2}(v, F)] \wedge\\ &\forall Y \subseteq V[\text{conn}(Y, F) \Rightarrow \exists y \in Y[\text{deg1}(y, F)]],\end{aligned}$$

where $\text{conn}(Y, F)$, $\text{deg1}(v, F)$ and $\text{deg2}(v, F)$ are defined as follows:

$$\begin{aligned}\text{conn}(Y, F) &\triangleq \forall Z \subseteq V\ [(\exists u \in Y[u \in Z] \wedge \exists v \in Y[v \notin Z]) \Rightarrow\\ &\quad (\exists e \in F\ \exists u \in Y\ \exists v \in Y[\text{inc}(u, e) \wedge \text{inc}(v, e) \wedge u \in Z \wedge v \notin Z])].\\ \text{deg1}(v, F) &\triangleq \exists e_1 \in F[\text{inc}(v, e_1) \wedge (\forall e_2 \in F[\text{inc}(v, e_2) \Rightarrow (e_2 = e_1)])].\\ \text{deg2}(v, F) &\triangleq \exists e_1, e_2 \in F[(e_1 \neq e_2) \wedge \text{inc}(v, e_1) \wedge \text{inc}(v, e_2) \wedge\\ &\quad (\forall e_3 \in F[\text{inc}(v, e_3) \Rightarrow (e_3 = e_1) \vee (e_3 = e_2)])].\end{aligned}$$

Since MinNtPC (resp. MaxNtPC) is equivalent to MaxWNtPC (resp. MinWNtPC) on unit-weight graphs, by a variant of Courcelle's Theorem (see Rao 2007, Theorem 6), we obtain the following result.

Corollary 8 MinNtPC*,* MaxNtPC*,* MinWNtPC *and* MaxWNtPC *can be solved in linear time on the class of graphs with bounded treewidth.*

We note that the property "G has a path cover" can also be formulated in MSOL. One can use an expression similar to the one we have exhibited above, and include the possibility of having vertices of degree zero. Since MinPC is equivalent to finding a maximum weight path cover in a unit-weight graph, this means that these problems can also be solved in linear time on the class of graphs with bounded treewidth.

7 Concluding remarks

As far as we know, the MinNtPC and MaxNtPC problems have not been treated in the literature. To deal with these optimization problems, we considered first the corresponding existence problem, which turns out to be a special case of the well-studied (g, f)-factor problem. Our result showing the close relation of the existence problem with the maximum matchings of the graph, contributes with a characterization that gives a polynomial-time algorithm that either finds a nontrivial path cover in a graph or finds a no-certificate. The proof of this characterization (Theorem 6) can also be seen as an alternative proof of Theorem 5.

We showed that MinNtPC on feasible instances is computationally equivalent to MinPC: their optimum values coincide and they have the same approximation threshold. We also studied a weighted version of the problems MinNtPC and MaxNtPC (in which the number of paths do not matter): for the minimization version we showed a polynomial-time algorithm on arbitrary graphs, a constant-factor approximation for the maximization version, and a linear-time algorithm on trees.

We are currently testing some integer programming formulations we have proposed for MinNtPC. The computational results are preliminary, but seem very promising. The design of approximation algorithms for MinNtPC (possibly for special class of graphs) is a challenging problem, as no such results are known.

Acknowledgements We thank the referees for the valuable suggestions.

References

Anstee R (1985) An algorithmic proof of Tutte's f-factor theorem. J Algorithms 6(1):112–131

Arikati SR, Pandu Rangan C (1990) Linear algorithm for optimal path cover problem on interval graphs. Inf Process Lett 35(3):149–153

Corneil DG, Dalton B, Habib M (2013) LDFS-based certifying algorithm for the minimum path cover problem on cocomparability graphs. SIAM J Comput 42(3):792–807

Courcelle B (1990) The monadic second-order logic of graphs. I. Recognizable sets of finite graphs. Inf Comput 85(1):12–75. https://doi.org/10.1016/0890-5401(90)90043-H

Courcelle B, Engelfriet J (2012) Graph structure and monadic second-order logic-a language-theoretic approach, encyclopedia of mathematics and its applications, vol 138. Cambridge University Press, Cambridge

Dulmage A, Mendelsohn N (1958) Coverings of bipartite graphs. Can J Math 10:517–534

Franzblau DS, Raychaudhuri A (2002) Optimal Hamiltonian completions and path covers for trees, and a reduction to maximum flow. ANZIAM J 44(2):193–204

Garey M, Johnson D, Tarjan R (1976) The planar Hamiltonian circuit problem is NP-complete. SIAM J Comput 5(4):704–714

Georges J, Mauro D, Whittlesey M (1994) Relating path coverings to vertex labellings with a condition at distance two. Discret Math 135(1):103–111

Gómez R, Wakabayashi Y (2018) Covering a graph with nontrivial vertex-disjoint paths: existence and optimization. In: Graph-theoretic concepts in computer science—44th international workshop, WG 2018, Cottbus, Germany, June 27–29, 2018. Lecture Notes in Computer Science, vol 11159, pp 228–238. https://doi.org/10.1007/978-3-030-00256-5_19

Heinrich K, Hell P, Kirkpatrick D, Liu G (1990) A simple existence criterion for $(g < f)$-factors. Discret Math 85(3):313–317

Henning M, Wash K (2017) Matchings, path covers and domination. Discret Math 340(1):3207–3216

Kano M, Saito A (1983) $[a, b]$-factors of graphs. Discret Math 47(1):113–116

Las Vergnas M (1978) An extension of Tutte's 1-factor theorem. Discret Math 23(3):241–255

Li Y, Mao-cheng C (1998) A degree condition for a graph to have $[a, b]$-factors. J Graph Theory 27(1):1–6

Lovász L (1970) Subgraphs with prescribed valencies. J Comb Theory 8:391–416

Lovász L, Plummer M (1986) Matching theory, North-Holland mathematics studies, vol 121. North-Holland, Amsterdam

Magnant C, Martin D (2009) A note on the path cover number of regular graphs. Australas J Comb 43:211–217

Moran S, Wolfstahl Y (1991) Optimal covering of cacti by vertex-disjoint paths. Theor Comput Sci 84(2):179–197

Müller H (1996) Hamiltonian circuits in chordal bipartite graphs. Discret Math 156(1–3):291–298

Papadimitriou CH, Yannakakis M (1993) The traveling salesman problem with distances one and two. Math Oper Res 18(1):1–11

Rao M (2007) MSOL partitioning problems on graphs of bounded treewidth and clique-width. Theor Comput Sci 377(1–3):260–267. https://doi.org/10.1016/j.tcs.2007.03.043
Reed B (1996) Paths, stars and the number three. Comb Probab Comput 5(3):277–295
Schrijver A (2003) Combinatorial optimization. Polyhedra and efficiency. Vol A, algorithms and combinatorics, vol 24. Springer, Berlin (Paths, flows, matchings, Chapters 1–38)
Vishwanathan S (1992) An approximation algorithm for the asymmetric travelling salesman problem with distances one and two. Inf Process Lett 44(6):297–302
Yu G (2017) Covering 2-connected 3-regular graphs with disjoint paths. J Graph Theory 88(3):385–401