HAL
open science

On the Complexity of Finding Internally Vertex-Disjoint Long Directed Paths

Júlio Araújo, Victor A Campos, Ana Karolinna Maia de Oliveira, Ignasi Sau, Ana Silva

▶ **To cite this version:**

Júlio Araújo, Victor A Campos, Ana Karolinna Maia de Oliveira, Ignasi Sau, Ana Silva. On the Complexity of Finding Internally Vertex-Disjoint Long Directed Paths. Algorithmica, 2020, 82 (6), pp.1616-1639. 10.1007/s00453-019-00659-5 . lirmm-02989813

HAL Id: lirmm-02989813

https://hal-lirmm.ccsd.cnrs.fr/lirmm-02989813v1

Submitted on 5 Nov 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Distributed under a Creative Commons Attribution 4.0 International License

On the complexity of finding internally vertex-disjoint long directed paths*

Júlio Araújo[1], Victor A. Campos[2], Ana Karolinna Maia[2], Ignasi Sau[1,3], and Ana Silva[1]

1 **Departamento de Matemática, ParGO research group, Universidade Federal do Ceará, Fortaleza, Brazil**
2 **Departamento de Computação, ParGO research group, Universidade Federal do Ceará, Fortaleza, Brazil**
3 **CNRS, AlGCo project team, LIRMM, Université de Montpellier, France**

Abstract

For two positive integers k and ℓ, a $(k \times \ell)$-*spindle* is the union of k pairwise internally vertex-disjoint directed paths with ℓ arcs each between two vertices u and v. We are interested in the (parameterized) complexity of several problems consisting in deciding whether a given digraph contains a subdivision of a spindle, which generalize both the Maximum Flow and Longest Path problems. We obtain the following complexity dichotomy: for a fixed $\ell \geq 1$, finding the largest k such that an input digraph G contains a subdivision of a $(k \times \ell)$-spindle is polynomial-time solvable if $\ell \leq 3$, and NP-hard otherwise. We place special emphasis on finding spindles with exactly two paths and present FPT algorithms that are asymptotically optimal under the ETH. These algorithms are based on the technique of representative families in matroids, and use also color-coding as a subroutine. Finally, we study the case where the input graph is acyclic, and present several algorithmic and hardness results.

1998 ACM Subject Classification F.2.2 Nonnumerical Algorithms and Problems, G.2.2 Graph Theory.

Keywords and phrases digraph subdivision; spindle; parameterized complexity; FPT algorithm; representative family; complexity dichotomy.

1 Introduction

A *subdivision* of a digraph F is a digraph obtained from F by replacing each arc (u, v) of F by a directed (u, v)-path. We are interested in the (parameterized) complexity of several problems consisting in deciding whether a given digraph contains as a subdigraph a subdivision of a *spindle*, defined by Bang-Jensen et al. [3] as follows. For k positive integers $\ell_1, \ldots, \ell_k$, an $(\ell_1, \ldots, \ell_k)$-*spindle* is the digraph containing k paths $P_1, \ldots, P_k$ from a vertex u to a vertex v, such that $|E(P_i)| = \ell_i$ for $1 \leq i \leq k$ and $V(P_i) \cap V(P_j) = \{u, v\}$ for $1 \leq i \neq j \leq k$. If $\ell_i = \ell$ for $1 \leq i \leq k$, an $(\ell_1, \ldots, \ell_k)$-spindle is also called a $(k \times \ell)$-*spindle*. See Figure 1 for an example.

* Emails of authors: julio@mat.ufc.br, campos@lia.ufc.br, karolmaia@lia.ufc.br, ignasi.sau@lirmm.fr (corresponding author), anasilva@mat.ufc.br.
Work supported by projects DEMOGRAPH (ANR-16-CE40-0028) and ESIGMA (ANR-17-CE23-0010), CNPq grants 304576/2017-4 and 304478/2018-0, CNPq Universal Project 401519/2016-3, FUNCAP-PRONEM PNE-0112-00061.01.00/16 and CAPES-PRINT.
A conference version of this article appeared in the *Proc. of the 13th Latin American Theoretical INformatics Symposium (**LATIN**), volume 10807 of LNCS, pages 66-79, Buenos Aires, Argentina, April **2018***. This article is permanently available at https://arxiv.org/abs/1706.09066.

© Júlio Araújo, Victor A. Campos, Ana Karolinna Maia, Ignasi Sau, and Ana Silva;
licensed under Creative Commons License CC-BY

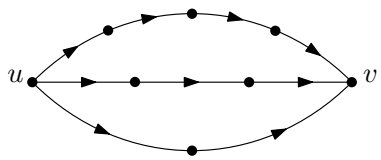

Figure 1 A $(4, 3, 2)$-spindle. This digraph contains a subdivision of a (3×2)-spindle, but not of a (3×3)-spindle.

Note that a digraph G contains a subdivision of a $(k \times 1)$-spindle if and only if there exist two vertices u, v and k internally vertex-disjoint paths from u to v. On the other hand, G contains a subdivision of a $(1 \times \ell)$-spindle if and only if G contains a path of length at least ℓ. Hence, finding a subdivision of a spindle generalizes both the Maximum Flow with unit capacities, and the Longest Path problems.

Subdivisions of spindles were considered by Bang-Jensen et al. [3], who introduced the general problem of finding a subdivision of a *fixed* digraph F in an n-vertex input digraph, and presented NP-hardness results and polynomial-time algorithms for several choices of F. In particular, they proved that when F is a spindle, the problem can be solved in time $n^{O(|V(F)|)}$ by a simple combination of brute force and a flow algorithm. Using terminology from parameterized complexity, this means that the problem is in XP parameterized by the size of F, and they left open whether it is FPT. Note that on undirected graphs the notion of subdivision coincides with that of topological minor, and therefore by the results of Grohe et al. [16] the problem is FPT parameterized by the size of F, for a general graph F. We refer to the introduction of [3] for a more detailed discussion about problems related to containment relations on graphs and digraphs. It is worth mentioning that detecting the existence of a spindle (*not* a subdivision of it) is easier: since the treewidth of the underlying graph of a spindle is two, the classical color-coding technique Alon et al. [1] can detect a spindle on s vertices in an n-vertex digraph in time $2^{O(s)} \cdot n^2$.

We first consider the following two optimization problems about finding subdivisions of spindles:

(1) for a fixed positive integer k, given an input digraph G, find the largest integer ℓ such that G contains a subdivision of a $(k \times \ell)$-spindle, and
(2) for a fixed positive integer ℓ, given an input digraph G, find the largest integer k such that G contains a subdivision of a $(k \times \ell)$-spindle.

We call these problems Max $(k \times \bullet)$-Spindle Subdivision and Max $(\bullet \times \ell)$-Spindle Subdivision, respectively. We prove that the first problem is NP-hard for any integer $k \geq 1$, by a simple reduction from Longest Path. The second problem turns out to be much more interesting, and we achieve the following dichotomy.

▶ **Theorem 1.** *Let $\ell \geq 1$ be a fixed integer.* Max $(\bullet \times \ell)$-Spindle Subdivision *is polynomial-time solvable if $\ell \leq 3$, and* NP*-hard otherwise, even restricted to acyclic digraphs. In addition, the same dichotomy applies to finding a spindle between a given pair of vertices.*

The reduction for the NP-hard cases is inspired by a result of Brewster et al. [7] to prove the NP-hardness of packing vertex-disjoint paths on digraphs. Concerning the polynomial algorithms, to solve the case $\ell = 3$, which is the only nontrivial one, we use a *vertex splitting procedure* that builds on ideas of Schrijver [27] on undirected graphs and by Kriesell [22] on directed graphs (see also [2, Section 5.9]).

As stated in Theorem 1, the same dichotomy holds for finding a spindle between a given pair of vertices. Itai et al. [19] considered the problems of, given a digraph G and two distinct

vertices s and t, finding the maximum number of internally vertex-disjoint (s,t)-paths whose lengths are *at most* or *exactly equal* to a fixed constant ℓ, and achieved dichotomies for both cases. Note that the problem we consider corresponds to a constraint of type '*at least*' on the lengths of the desired paths. Hence, Theorem 1 together with the results of Itai et al. [19] provide a full picture of the complexity of finding a maximum number of length-constrained internally vertex-disjoint directed (s,t)-paths. For future work, one could consider mixed constraints, i.e., some of the paths of type '*at least*', some '*at most*', and some '*exactly equal*'.

Due to the apparent hardness of finding an FPT algorithm for a general spindle, we decided to place special emphasis on finding subdivisions of spindles with exactly two paths, which we call 2-*spindles*. We mention that the existence of subdivisions of 2-spindles has attracted some interest in other contexts. Indeed, Benhocine and Wojda [4] showed that a tournament on $n \geq 7$ vertices always contains a subdivision of an (ℓ_1, ℓ_2)-spindle such that $\ell_1 + \ell_2 = n$. And more recently, Cohen et al. [9] showed that a strongly connected digraph with chromatic number $\Omega((\ell_1 + \ell_2)^4)$ contains a subdivision of an (ℓ_1, ℓ_2)-spindle, and this bound was subsequently improved to $\Omega((\ell_1 + \ell_2)^2)$ by Kim et al. [20], who also provided improved bounds for Hamiltonian digraphs.

We consider two problems concerning the existence of subdivisions of 2-spindles. The first one is, given an input digraph G, find the largest integer ℓ such that G contains a subdivision of an (ℓ_1, ℓ_2)-spindle with $\min\{\ell_1, \ell_2\} \geq 1$ and $\ell_1 + \ell_2 = \ell$. We call this problem Max $(\bullet, \bullet)$-Spindle Subdivision, and we show the following results.

▶ **Theorem 2.** *Given a digraph G and a positive integer ℓ, the problem of deciding whether there exist two strictly positive integers ℓ_1, ℓ_2 with $\ell_1 + \ell_2 = \ell$ such that G contains a subdivision of an (ℓ_1, ℓ_2)-spindle is* NP*-hard and* FPT *parameterized by ℓ. The running time of the* FPT *algorithm is $2^{O(\ell)} \cdot n^{O(1)}$, which is asymptotically optimal unless the* ETH *fails. Moreover, the problem does not admit polynomial kernels unless* NP $\subseteq$ coNP/poly.

The second problem is, for a fixed strictly positive integer ℓ_1, given an input digraph G, find the largest integer ℓ_2 such that G contains a subdivision of an (ℓ_1, ℓ_2)-spindle. We call this problem Max $(\ell_1, \bullet)$-Spindle Subdivision, and we show the following results.

▶ **Theorem 3.** *Given a digraph G and two integers ℓ_1, ℓ_2 with $\ell_2 \geq \ell_1 \geq 1$, the problem of deciding whether G contains a subdivision of an (ℓ_1, ℓ_2)-spindle can be solved in time $2^{O(\ell_2)} \cdot n^{O(\ell_1)}$. When ℓ_1 is a constant, the problem remains* NP*-hard and the running time of the* FPT *algorithm parameterized by ℓ_2 is asymptotically optimal unless the* ETH *fails. Moreover, the problem does not admit polynomial kernels unless* NP $\subseteq$ coNP/poly.

The hardness results of Theorems 2 and 3 are based on a simple reduction from Directed Hamiltonian Cycle. Both FPT algorithms, which are our main technical contribution, are based on the technique of *representative families* in matroids introduced by Monien [24], and in particular its improved version recently presented by Fomin et al. [14]. The FPT algorithm of Theorem 3 also uses the *color-coding* technique of Alon et al. [1] as a subroutine.

Finally, we consider the case where the input digraph G is *acyclic*. We prove the following result by using a standard dynamic programming algorithm.

▶ **Theorem 4.** *Given an acyclic digraph G and two positive integers k, ℓ, the problem of deciding whether G contains a subdivision of a $(k \times \ell)$-spindle can be solved in time $O(\ell^k \cdot n^{2k+1})$.*

The above theorem implies, in particular, that when k is a constant the problem is polynomial-time solvable on acyclic digraphs, which generalizes the fact that Longest Path, which corresponds to the case $k = 1$, is polynomial-time solvable on acyclic digraphs (cf. [28]).

As observed by Bang-Jensen et al. [3], from the fact that the k-Linkage problem is in XP on acyclic digraphs [23], it easily follows that finding a subdivision of a general digraph F is in XP on DAGs parameterized by $|V(F)|$. Motivated by this, we prove two further hardness results about finding subdivisions of spindles on DAGs. Namely, we prove that if F is the disjoint union of (2×1)-spindles, then finding a subdivision of F is NP-complete on planar DAGs, and that if F is the disjoint union of a $(k_1 \times 1)$-spindle and a $(k_2 \times 1)$-spindle, then finding a subdivision of F is W[1]-hard on DAGs parameterized by $k_1 + k_2$. These two results should be compared to the fact that finding a subdivision of a single $(k \times 1)$-spindle can be solved in polynomial time on general digraphs by a flow algorithm.

Organization of the paper. In Section 2 we provide some definitions about (di)graphs, parameterized complexity, and matroids. In Section 3 we prove Theorem 1, and in Section 4 we prove Theorem 2 and Theorem 3. In Section 5 we focus on acyclic digraphs and we prove, in particular, Theorem 4. In Section 6 we present some open problems for further research.

2 Preliminaries

Graphs and digraphs. We use standard graph-theoretic notation, and we refer the reader to the books [11] and [2] for any undefined notation about graphs and directed graphs, respectively.

A *(multi-)directed graph* G, or just *(multi-)digraph*, consists of a non-empty set $V(G)$ of elements called *vertices* and a finite (multi)set $A(G)$ of ordered pairs of distinct vertices called *arcs*. All our positive results hold even for digraphs where multiple arcs between the same pair of vertices are allowed. We denote by (u, v) an arc from a vertex u to a vertex v. Vertex u is called the *tail* and vertex v is called the *head* of an arc (u, v), and we say that (u, v) is an arc *outgoing* from u and *incoming* at v.

For a vertex v in a digraph G, we let $N_G^+(v) = \{u \in V(G) \setminus \{v\} : (v, u) \in A(G)\}$, $N_G^-(v) = \{w \in V(G) \setminus \{v\} : (w, v) \in A(G)\}$, and $N_G(v) = N_G^+(v) \cup N_G^-(v)$, and we call these sets the *out-neighborhood*, *in-neighborhood*, and *neighborhood* of v, respectively. The *out-degree* (resp. *in-degree*) of a vertex v is the number of arcs outgoing from (resp. incoming at) v, and its *degree* is the sum of its out-degree and its in-degree. In all these notations, we may omit the subscripts if the digraph G is clear from the context.

A *subdigraph* of a digraph $G = (V, A)$ is a digraph $H = (V', A')$ such that $V' \subseteq V$ and $A' \subseteq A$. Given vertices $u, v \in V$, a *(u,v)-path* G is a sequence $(u = w_1, \ldots, w_q = v)$ such that $(w_i, w_{i+1}) \in A$ for every $i \in \{1, \ldots, q-1\}$. The *length* of a path is its number of arcs, and by an ℓ-path we denote a path of length ℓ. A *directed acyclic graph*, or DAG for short, is a digraph with no directed cycles. It is easy to prove that a digraph G is a DAG if and only if there exists a total ordering of $V(G)$, called a *topological ordering*, so that all arcs of G go from smaller to greater vertices in this ordering.

For two positive integers k and ℓ, a $(k \times \ell)$-*spindle* is the union of k pairwise internally vertex-disjoint directed (u, v)-paths of length ℓ between two vertices u and v, which are called the *endpoints* of the spindle. More precisely, u is called the *tail* and v the *head* of a spindle. A 2-*spindle* is any (ℓ_1, ℓ_2)-spindle with $\ell_1, \ell_2 \geq 1$.

For an undirected graph G, we denote by $\{u, v\}$ an edge between two vertices u and v. A *matching* in a graph is a set of pairwise disjoint edges. A vertex v is *saturated* by a matching M if v is an endpoint of one of the edges in M. In that case, we say that v is M-*saturated*. Given two matchings M and N in a graph, we let $M \triangle N$ denote their symmetric difference, that is, $M \triangle N = (M \setminus N) \cup (N \setminus M)$.

Parameterized complexity. We refer the reader to [10, 12, 13, 25] for basic background on

parameterized complexity, and we recall here only some basic definitions. A *parameterized problem* is a decision problem whose instances are pairs $(x, k) \in \Sigma^* \times \mathbb{N}$, where k is called the *parameter*. A parameterized problem is *fixed-parameter tractable* (FPT) if there exists an algorithm $\mathcal{A}$, a computable function f, and a constant c such that given an instance $I = (x, k)$, $\mathcal{A}$ (called an FPT *algorithm*) correctly decides whether $I \in L$ in time bounded by $f(k) \cdot |I|^c$. A parameterized problem is *slice-wise polynomial* (XP) if there exists an algorithm $\mathcal{A}$ and two computable functions f, g such that given an instance $I = (x, k)$, $\mathcal{A}$ (called an XP *algorithm*) correctly decides whether $I \in L$ in time bounded by $f(k) \cdot |I|^{g(k)}$.

Within parameterized problems, the class W[1] may be seen as the parameterized equivalent to the class NP of classical optimization problems. Without entering into details (see [10, 12, 13, 25] for the formal definitions), a parameterized problem being W[1]-*hard* can be seen as a strong evidence that this problem is *not* FPT. The canonical example of W[1]-hard problem is Independent Set parameterized by the size of the solution. To transfer W[1]-hardness from one problem to another, one uses a *parameterized reduction*, which given an input $I = (x, k)$ of the source problem, computes in time $f(k) \cdot |I|^c$, for some computable function f and a constant c, an equivalent instance $I' = (x', k')$ of the target problem, such that k' is bounded by a function depending only on k. An equivalent definition of W[1]-hard problem is any problem that admits a parameterized reduction from Independent Set parameterized by the size of the solution.

Matroids. A pair $\mathcal{M} = (E, \mathcal{I})$, where E is a ground set and $\mathcal{I}$ is a family of subsets of E, is a *matroid* if it satisfies the following three axioms:

1. $\emptyset \in \mathcal{I}$.
2. If $A' \subseteq A$ and $A \in \mathcal{I}$, then $A' \in \mathcal{I}$.
3. If $A, B \in \mathcal{I}$ and $|A| < |B|$, then there is $e \in B \setminus A$ such that $A \cup \{e\} \in \mathcal{I}$.

The sets in $\mathcal{I}$ are called the *independent sets* of the matroid. An inclusion-wise maximal set of $\mathcal{I}$ is called a *basis* of the matroid. Using the third axiom, it is easy to show that all the bases of a matroid $\mathcal{M}$ have the same size, which is called the *rank* of $\mathcal{M}$. A pair $\mathcal{M} = (E, \mathcal{I})$ over an n-element ground set E is called a *uniform matroid* if $\mathcal{I} = \{A \subseteq E : |A| \leq k\}$ for some for constant k. For a broader overview on matroids, we refer to [26].

For a positive integer k, we denote by $[k]$ the set of all integers i such that $1 \leq i \leq k$. Throughout the article, unless stated otherwise, we let n denote the number of vertices of the input digraph of the problem under consideration.

3 Complexity dichotomy in terms of the length of the paths

In this section we focus on the two natural optimization versions of finding subdivisions of spindles mentioned in the introduction, namely Max $(k \times \bullet)$-Spindle Subdivision and Max $(\bullet \times \ell)$-Spindle Subdivision.

It is easy to prove that the first problem is NP-hard for any integer $k \geq 1$, by a simple reduction from Longest Path.

▶ **Theorem 5.** *Let $k \geq 1$ be a fixed integer. The* Max $(k \times \bullet)$-Spindle Subdivision *problem is* NP*-hard.*

Proof. We provide a polynomial reduction from the Longest Path problem on general digraphs, which is NP-hard as it generalizes Hamiltonian Path [15]. For $k = 1$, Max $(k \times \bullet)$-Spindle Subdivision is exactly the Longest Path problem, and the result follows. For $k > 1$, let G be an instance of Longest Path with n vertices, and we build an instance

G' of Max $(k \times \bullet)$-Spindle Subdivision as follows. We start with G and we add to it two vertices s and t together with $k-1$ internally vertex-disjoint paths of length $n+1$ between them; then we add an arc from s to every vertex of G, and an arc from every vertex of G to t. Let G' be the obtained graph. We claim that G has a path on ℓ vertices if and only if G' has a subdivision of a $(k, \ell+1)$-spindle. Indeed, if G has a (u,v)-path P on ℓ vertices, then the $k-1$ disjoint (s,t)-paths together with the (s,t)-path going through P form a subdivision of a $(k, \ell+1)$-spindle in G'. Conversely, suppose that S is a subdivision of a $(k, \ell+1)$-spindle in G' with tail u and head v. If $u = s$ and $v = t$, then the intermediate part of one of the k (s,t)-paths must be entirely contained in G, thus defining a path in G of length $\ell - 1$ (hence on ℓ vertices). Otherwise, because s is a source and t is a sink, one can verify that the intermediate part of each of the k (u,v)-paths is contained in G, yielding in particular a path on ℓ vertices in G. ◀

We now present the complexity dichotomy for the second problem, in order to prove Theorem 1. We start with the hardness result.

▶ **Theorem 6.** *Let $\ell \geq 4$ be a fixed integer. The* Max $(\bullet \times \ell)$-Spindle Subdivision *problem is* NP*-hard, even when restricted to DAGs.*

Proof. We provide a polynomial reduction from 3-Dimensional Matching, which is NP-hard [15]. In the 3-Dimensional Matching problem, we are given three sets A, B, C of the same size and a set of triples $\mathcal{T} \subseteq A \times B \times C$. The objective is to decide whether there exists a set $\mathcal{T}' \subseteq \mathcal{T}$ of pairwise disjoint triples with $|\mathcal{T}'| = |A|$. Given an instance $(A, B, C, \mathcal{T})$ of 3-Dimensional Matching, with $|A| = n$ and $|\mathcal{T}| = m$, we construct an instance G of Max $(\bullet \times \ell)$-Spindle Subdivision as follows. We first present the reduction for $\ell = 4$, and then we explain how to modify it for a general $\ell > 4$.

For every $i \in [n]$, we add to G three vertices a_i, b_i, c_i, corresponding to the elements in the sets A, B, C, respectively. Let H be the digraph with vertices $x_0, x_1, y_0, y_1, z_0, z_1, a, b, c$ and arcs $(x_0, x_1), (x_1, a), (x_1, y_0), (y_0, y_1), (y_1, b), (x_0, z_0), (z_0, z_1), (z_1, c)$ (see Figure 2(a)). For every triple $T \in \mathcal{T}$, with $T = (a_i, b_j, c_p)$, we add to G a copy of H and we identify vertex a with a_i, vertex b with b_j, and vertex c with c_p. Finally, we add a new vertex s that we connect to all other vertices introduced so far, and another vertex t to which we connect all other vertices introduced so far except s.

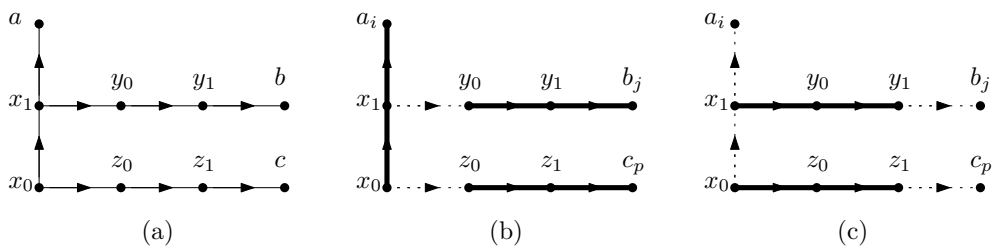

Figure 2 (a) Digraph H. (b) Selected paths when $T \in \mathcal{T}'$. (c) Selected paths when $T \in \mathcal{T} \setminus \mathcal{T}'$.

The constructed digraph G is easily seen to be a DAG. Indeed, we can define a topological ordering of $V(G)$ so that all arcs go from left to right as follows. We select s (resp. t) as the leftmost (resp. rightmost) vertex. We divide the remaining vertices of G into two blocks. On the right, we place all the vertices $\{a_i, b_i, c_i : i \in [n]\}$, and we order them arbitrarily. On the left, we place the remaining vertices of G, which we also order arbitrarily, except that for every triple $T \in \mathcal{T}$, we order the vertices in its copy of H, distinct from a, b, c, such that

$x_0 < x_1 < y_0 < y_1 < z_0 < z_1$ holds. One can check that, with respect to this ordering, all the arcs of G go from left to right.

Note that $|V(G)| = 3n + 6m + 2$, and since each path in a $(k \times 4)$-spindle has at least 3 internal vertices, we get that the largest possible integer k for which G contains a subdivision of a $(k \times 4)$-spindle is $k^* := n + 2m$. We claim that $(A, B, C, \mathcal{T})$ is a Yes-instance of 3-Dimensional Matching if and only if G contains a subdivision of a $(k^* \times 4)$-spindle.

Suppose first that $(A, B, C, \mathcal{T})$ is a Yes-instance, and let $\mathcal{T}' \subseteq \mathcal{T}$ be a solution. We proceed to define a set $\mathcal{P}$ of $n + 2m$ vertex-disjoint 2-paths in $G \setminus \{s, t\}$, which together with s and t yield the desired spindle. For every $T \in \mathcal{T}'$, with $T = (a_i, b_j, c_p)$, we add to $\mathcal{P}$ the three paths (x_0, x_1, a_i), (y_0, y_1, b_j), and (z_0, z_1, c_p) (see the thick arcs in Figure 2(b)). On the other hand, for every $T \in \mathcal{T} \setminus \mathcal{T}'$, with $T = (a_i, b_j, c_p)$, we add to $\mathcal{P}$ the two paths (x_1, y_0, y_1) and (x_0, z_0, z_1) (see the thick arcs in Figure 2(c)). Since $\mathcal{T}'$ is a solution of 3-Dimensional Matching, it holds that $|\mathcal{T}'| = n$, and thus $\mathcal{P} = 3n + 2(m - n) = n + 2m$, as required.

Conversely, suppose that G contains a subdivision of a $(k^* \times 4)$-spindle S. First, note that each $u \in V(G) \setminus (A \cup B \cup C \cup \{s, t\})$ has in-degree at most two and out-degree at most three. Also, each $d_i \in A \cup B \cup C$ has in-degree equal to one plus the number of triples it is involved in, which is at most $m + 1$, and out-degree at most one. Therefore, we get that s and t are the only vertices in G with in-degree and out-degree at least k^*, respectively, hence they must be the endpoints of S. Since $|V(G) \setminus \{s, t\}| = 3k^*$, it follows that $S \setminus \{s, t\}$ consists of a collection $\mathcal{P}$ of k^* vertex-disjoint 2-paths that covers all the vertices in $V(G) \setminus \{s, t\}$. Let H be the subdigraph in G associated with an arbitrary triple $T \in \mathcal{T}$, and consider $\mathcal{P} \cap H$. By construction of H, it follows that if $\mathcal{P} \cap H$ is not equal to one of the configurations corresponding to the thick arcs of Figure 2(b) or Figure 2(c), necessarily at least one vertex in $V(H)$ would not be covered by $\mathcal{P}$, a contradiction. Let $\mathcal{T}'$ be the set of triples in $\mathcal{T}$ such that the corresponding gadget H intersects $\mathcal{P}$ as in Figure 2(b). It follows that $3|\mathcal{T}'| + 2(m - |\mathcal{T}'|) = |\mathcal{P}| = k^* = n + 2m$, and therefore $|\mathcal{T}'| = n$. Since all the 2-paths in $\mathcal{P}$ associated with the triples in $\mathcal{T}'$ are vertex-disjoint, we have that $\mathcal{T}'$ is a collection of n pairwise disjoint triples, hence a solution of 3-Dimensional Matching.

For a general $\ell > 4$, we define the digraph G in the same way, except that we subdivide the arcs outgoing from s exactly $\ell - 4$ times. The rest of the proof is essentially the same, and the result follows. ◀

Note that, in the above hardness result, if we drop the hypothesis that the input digraph is a DAG, a simple NP-hardness reduction can be obtained directly from the problem of packing a maximum number of vertex-disjoint P_3's in a directed graph [7].

We now turn to the cases that can be solved in polynomial time. We first need some ingredients to deal with the case $\ell = 3$, which is the most interesting one. Let G be a digraph and let X and Y be two subsets of $V(G)$. We say that a (simple) path P is *directed from X to Y* if P is a directed path with first vertex x and last vertex y such that $x \in X$ and $y \in Y$. The path P is *nontrivial* if its endpoints are distinct.

The following proposition will be the key ingredient in the proof of Theorem 8. Its proof is inspired by similar constructions given by Schrijver [27] on undirected graphs and by Kriesell [22] on directed graphs, usually called *vertex splitting procedure* (see [2, Section 5.9]). In fact, the conclusion of Proposition 7 can be also derived as a corollary of the main result in [22], noting that a polynomial-time algorithm can be extracted from that proof. We present a simpler proof here for completeness.

▶ **Proposition 7.** *Let G be a digraph and let X and Y be two subsets of $V(G)$. The maximum number of vertex-disjoint directed nontrivial paths from X to Y can be computed in polynomial*

time.

Proof. Let G' be the undirected graph built from G as follows. The vertex set of G' is obtained from $V(G)$ by adding a copy v' of each vertex v not in $X \cup Y$. We build the edge set of G' starting from the empty set as follows. For every vertex v not in $X \cup Y$, add the edge $\{v, v'\}$. For each arc (u, v) in G, we add the edge $\{u, v\}$ if $v \in X \cup Y$ and the edge $\{u, v'\}$ otherwise. See Figure 3(a)-(b) for an example.

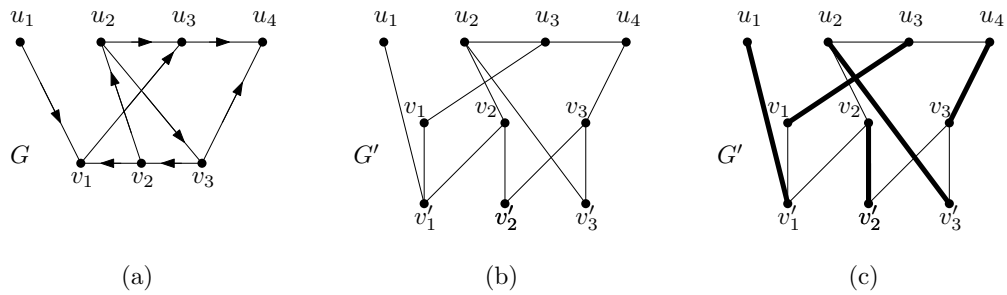

Figure 3 (a) Digraph G with $X = \{u_1, u_2, u_3\}$ and $Y = \{u_3, u_4\}$. (b) Graph G' associated with G. (c) The thick edges define a matching of size five in G', corresponding to the two vertex-disjoint directed nontrivial paths (u_1, v_1, u_3) and (u_2, v_3, u_4) from X to Y in G.

▶ **Claim 1.** *The digraph G contains a family of k vertex-disjoint directed nontrivial paths from X to Y if and only if G' has a matching of size $k + |V(G) \setminus (X \cup Y)|$.*

Proof of the claim. Let $\mathcal{P}$ be a family of k vertex-disjoint directed nontrivial paths from X to Y in G. If $P \in \mathcal{P}$ is a path between $x \in X$ and $y \in Y$ that contains an internal vertex w in X, then the set $\mathcal{P}'$ obtained from $\mathcal{P}$ by removing P and adding P', the subpath of P between w and y, has the same cardinality as $\mathcal{P}$ and also contains only vertex-disjoint directed nontrivial paths, since w is internal, i.e., $w \neq y$. A similar argument also holds when $w \in Y$, hence we can suppose that each path in $\mathcal{P}$ has no internal vertices in $X \cup Y$. Observe that we can therefore assume that G has no arcs to a vertex in $X \setminus Y$ or from a vertex in $Y \setminus X$. Let U be the subset of vertices of $V(G) \setminus (X \cup Y)$ that are not in a path in $\mathcal{P}$.

We build a matching M of G' starting with $M = \{\{u, u'\} : u \in U\}$ as follows. For every arc (u, v) used in some path of $\mathcal{P}$, we add to M either $\{u, v\}$ if $v \in X \cup Y$, or $\{u, v'\}$ otherwise (see Figure 3(c)). Note that M is indeed a matching, as vertices in $X \cup Y$ appear in at most one arc on a path in $\mathcal{P}$. For a vertex v not in $X \cup Y$, v appears at most once as an internal vertex in a path P of $\mathcal{P}$. Therefore, it appears in exactly two arcs of P and exactly once in an arc to v and once in an arc from v.

We now claim that the number of M-saturated vertices in G' is $2(k + |V(G) \setminus (X \cup Y)|)$. This claim implies that M has $k + |V(G) \setminus (X \cup Y)|$ edges. To prove this claim, first note that all vertices in $V(G) \setminus (X \cup Y)$ are saturated. Indeed, if v is in U, then both v and v' are initially saturated. Otherwise, v is an internal vertex of a path in $\mathcal{P}$ and is contained in two edges that saturate both v and v'. To conclude, note that every path in $\mathcal{P}$ contains exactly two vertices in $X \cup Y$, namely its endpoints, and, therefore, saturates exactly two vertices of $X \cup Y$ in G'.

Now, let M be a matching of G' of size $k + |V(G) \setminus (X \cup Y)|$. Let N be the matching $\{\{v, v'\} : v \in V(G) \setminus (X \cup Y)\}$ and $H = G[M \triangle N]$, where $\triangle$ denotes the symmetric difference. Since $|M| = k + |N|$, H contains at least k components with more edges in M than in N. We claim that from these components we can obtain k vertex-disjoint nontrivial paths in G.

To prove this claim, let C be a component of H with more edges in M than in N. Since C has more edges in M, then it is a path alternating between edges of M and N that starts and ends with an edge of M, its endpoints are N-unsaturated and its internal vertices are N-saturated. Thus, the endpoints of C are its only vertices in $X \cup Y$. Also note that if a vertex w of $V(C) \cap V(G)$ is not in $X \cup Y$, then both w and w' are in C and neither w nor w' appear in any other component of H.

Let u and v be the endpoints of C and the set W of internal vertices of C that are also in $V(G)$ be $\{w_1, \ldots, w_\ell\}$. Note that $u \neq v$ as C contains at least one edge in M. If $W = \emptyset$, then uv is an edge of G' and assume the edge $\{u, v\}$ is directed in G from u to v. If $W \neq \emptyset$, then assume the transversal of C from u to v visits the vertices in the order $u, w'_1, w_1, w'_2, w_2, \ldots, w'_\ell, w_\ell, v$. In both cases, note that $u, w_1, \ldots, w_\ell, v$ is the transversal of a directed path from u to v in G. Since G has no edge leaving a vertex of $Y \setminus X$ and no edge going into a vertex of $X \setminus Y$, then $u \in X$ and $v \in Y$. □

Claim 1 tells us that we can obtain a maximum number of vertex-disjoint nontrivial paths from X to Y in G by finding a maximum matching in the graph G', which can be done in polynomial time [11]. The proposition follows. ◀

We are now ready to prove the main algorithmic result of this section.

▶ **Theorem 8.** *Let $\ell \leq 3$ be a fixed positive integer. Then, the* Max $(\bullet \times \ell)$-Spindle Subdivision *problem can be solved in polynomial time.*

Proof. If $\ell = 1$, then the problem can be solved just by computing a maximum flow with unit capacities between every pair of vertices of the input digraph, which can be done in polynomial time [2]. If $\ell = 2$, we use the same algorithm, except that for every pair of vertices we first delete all the arcs between them before computing a maximum flow, as the paths of length one are the only forbidden ones in a subdivision of a $(k \times 2)$-spindle.

Let us now focus on the case $\ell = 3$. We first guess a pair of vertices s and t of $V(G)$ as candidates for being the tail and head of the desired spindle, respectively, and we delete the arcs between s and t, if any. The crucial observation is that the largest integer k such that G contains a $(k \times 3)$-spindle having s and t as tail and head, respectively, equals the maximum number of vertex-disjoint directed nontrivial paths from $N^+(s)$ to $N^-(t)$ in the digraph $G \setminus \{s, t\}$. Now the result follows directly by applying the polynomial-time algorithm given by Proposition 7 with input graph $G \setminus \{s, t\}$, $X = N^+(s)$, and $Y = N^-(t)$. ◀

We now prove a generalization of the result given in Theorem 8, using similar techniques.

▶ **Theorem 9.** *Given a digraph G and three non-negative integers k_1, k_2, k_3, let α be the sequence containing k_1 1's, followed by k_2 2's, followed by k_3 3's (for instance, for $k_1 = k_2 = k_3 = 2$, $\alpha = (1, 1, 2, 2, 3, 3)$). Then, deciding whether G contains a subdivision of an α-spindle can be solved in polynomial time.*

Proof. We iterate on pairs of vertices s and t in G to decide if the desired spindle exists with tail s and head t. From now on, we consider a fixed pair of vertices s and t. If an α-spindle subdivision exists with tail s and head t, let S be one such subdivision.

Let p be the number of arcs with tail s and head t. Note that if S exists, it can use at most $\min\{p, k_1\}$ arcs between s and t. In fact, we can assume S uses exactly $\min\{p, k_1\}$ arcs between s and t as, otherwise, there is a 1-path which was subdivided and can be changed to an unused arc from s to t. All other 1-paths of the spindle must have been subdivided and have length at least two in S. Therefore, S exists if and only if there is a subdivision of an

α'-spindle with tail s and head t, where α' consists of $k_2 + k_1 - \min\{p, k_1\}$ 2's, followed by k_3 3's. So, from now on assume $k_1 = 0$.

Let $X = N^+(s)$, $Y = N^-(t)$, and H be obtained from G by removing s and t. Observe that S exists if and only if there are $k_2 + k_3$ paths from X to Y in H such that at most k_2 of these paths are trivial. For $W \subseteq X \cap Y$, let G'_W be the graph built in Proposition 7 from $H - W$ to find paths from $X \setminus W$ to $Y \setminus W$. We build G'' from $G'_\emptyset$ by adding a set N of k_2 new vertices and adding an edge from every vertex in N to every vertex in $X \cap Y$. We claim that S exists if and only if G'' has a matching of size $k_2 + k_3 + |V(H) \setminus (X \cup Y)|$. Note that the theorem follows from the proof of this claim as we can build G'' and find a maximum matching in it in polynomial time [11].

Let $W \subseteq X \cap Y$. Note that $G'_W = G'_\emptyset - W$ by construction. Furthermore, Claim 1 states that a matching in G'_W of size $k + |V(H) \setminus (X \cup Y)|$ exists if and only if there are k vertex-disjoint directed nontrivial paths from $X \setminus W$ to $Y \setminus W$ in $H \setminus W$.

Let first M be a matching in G'' and let W be the set of vertices of $X \cap Y$ matched to vertices in N. Note that $|W| \leq |N| = k_2$. Now, let M' be the matching of size $|M| - |W|$ obtained from the edges of M in G'_W. From Claim 1, M' has size at least $k_2 + k_3 - |W| + |V(H) \setminus (X \cup Y)|$ if and only if there is a collection of $k_2 + k_3 - |W|$ vertex-disjoint directed nontrivial paths from $X \setminus W$ to $Y \setminus W$ in $H \setminus W$, and this happens if and only if M has size at least $k_2 + k_3 + |V(H) \setminus (X \cup Y)|$. We can find the desired number of paths by choosing $|W|$ trivial paths using the vertices of W.

Conversely, let $\mathcal{P}$ be a collection of $k_2 + k_3$ paths from X to Y in H such that at most k_2 of these paths are trivial. Let W be the set of vertices of $X \cap Y$ which are in trivial paths of $\mathcal{P}$ and note that $|W| \leq k_2$. Since $\mathcal{P}$ contains $k_2 + k_3 - |W|$ nontrivial paths from $X \setminus W$ to $Y \setminus W$ in $H \setminus W$, Claim 1 tells us that G'_W has a matching M of size $k_2 + k_3 - |W| + |V(H) \setminus (X \cup Y)|$. We find a matching in G'' of size $k_2 + k_3 + |V(H) \setminus (X \cup Y)|$ from M by matching the vertices of W arbitrarily to vertices of N. ◀

4 Finding subdivisions of 2-spindles

In this section we focus on finding subdivisions of 2-spindles, and we prove Theorem 2 and Theorem 3. We prove the negative and the positive results of both theorems separately. Namely, we provide the hardness results in Section 4.1 and we focus on the FPT algorithms in Section 4.2.

4.1 Hardness results

We start by proving the NP-hardness results.

▶ **Proposition 10.** *The* Max $(\bullet, \bullet)$-Spindle Subdivision *problem is* NP*-hard. For every fixed integer* $\ell_1 \geq 1$*, the* Max $(\ell_1, \bullet)$-Spindle Subdivision *problem is* NP*-hard.*

Proof. For both problems, we present a reduction from the Directed Hamiltonian (s,t)-Path problem, which consists in, given a digraph G and two vertices $s, t \in V(G)$, deciding whether G has an (s,t)-path that is Hamiltonian. This problem is easily seen to be NP-hard by a simple reduction from Directed Hamiltonian Cycle, which is known to be NP-hard [15]: given an instance G of Directed Hamiltonian Cycle, construct from G an instance G' of Directed Hamiltonian (s,t)-Path by choosing an arbitrary vertex $v \in V(G)$ and splitting it into two vertices s and t such that s (resp. t) is incident to exactly those arcs in G that were outgoing from (resp. incoming at) v.

We first prove the hardness of Max $(\bullet,\bullet)$-Spindle Subdivision. Given an instance G of Directed Hamiltonian (s,t)-Path, with $|V(G)| = n$, build an instance G' of Max $(\bullet,\bullet)$-Spindle Subdivision as follows. Start from G, and delete all the arcs incoming at s or outgoing from t, if any. Finally, add a new vertex v and arcs (s,v) and (v,t). We claim that G has a Hamiltonian (s,t)-path if and only if G' contains a subdivision of an (ℓ_1,ℓ_2)-spindle with $\min\{\ell_1,\ell_2\} \geq 1$ and $\ell_1+\ell_2 = n+1$. Assume first that G has a Hamiltonian (s,t)-path P. Then G' contains a $(2,n-1)$-spindle defined by the 2-path (s,v,t) together with the Hamiltonian (s,t)-path P. Conversely, assume that G' contains a subdivision S of an (ℓ_1,ℓ_2)-spindle with $\min\{\ell_1,\ell_2\} \geq 1$ and $\ell_1+\ell_2 = n+1$. Suppose that the newly added vertex $v \in V(G')$ does not belong to S, which implies that $|V(S)| \leq |V(G)| = n$. Since an (ℓ_1,ℓ_2)-spindle contains exactly $\ell_1+\ell_2$ vertices, it follows that $|V(S)| \geq \ell_1+\ell_2 = n+1$, a contradiction to the previous sentence. Therefore, $v \in V(S)$ and so (s,v,t) is one of the two paths of S. Thus, the remaining path of S is an (s,t)-path of length $n-1$ in G, that is, a Hamiltonian (s,t)-path in G.

We now prove the hardness of Max $(\ell_1,\bullet)$-Spindle Subdivision for every fixed integer $\ell_1 \geq 1$. Given an instance G of Directed Hamiltonian (s,t)-Path, with $|V(G)| = n$, build an instance G' of Max $(\ell_1,\bullet)$-Spindle Subdivision as follows. Start from G, and delete all the arcs incoming at s (resp. outgoing from t), if any, and the arc (s,t), if it exists. Finally, add an (s,t)-path with ℓ_1 arcs consisting of new vertices and arcs. One can easily check that G has a Hamiltonian (s,t)-path if and only if G' contains a subdivision of an $(\ell_1,n-1)$-spindle. ◀

Björklund et al. [5] showed that assuming the Exponential Time Hypothesis[1] (ETH) of Impagliazzo et al. [18], the Directed Hamiltonian Cycle problem cannot be solved in time $2^{o(n)}$. This result together with the proof of Proposition 10 directly imply the following two results assuming the ETH, claimed in Theorem 2 and Theorem 3, respectively. The first one is that, given a digraph G and a positive integer ℓ, the problem of deciding whether there exist two strictly positive integers ℓ_1,ℓ_2 with $\ell_1+\ell_2 = \ell$ such that G contains a subdivision of an (ℓ_1,ℓ_2)-spindle cannot be solved in time $2^{o(\ell)} \cdot n^{O(1)}$. The second one is that, given a digraph G and two integers ℓ_1,ℓ_2 with $\ell_2 \geq \ell_1 \geq 1$, the problem of deciding whether G contains a subdivision of an (ℓ_1,ℓ_2)-spindle cannot be solved in time $2^{o(\ell_2)} \cdot n^{O(\ell_1)}$.

Concerning the existence of polynomial kernels, it is easy to prove that none of the above problems admits polynomial kernels unless NP $\subseteq$ coNP/poly. Indeed, taking the disjoint union of t instances of any of these two problems defines a *cross-composition*, as defined by Bodlaender et al. [6], from the problem to itself, directly implying the desired results as both problems are NP-hard by Proposition 10. We refer to [6] for the missing definitions.

4.2 FPT algorithms

Our FPT algorithms for finding subdivisions of (ℓ_1,ℓ_2)-spindles are based on the technique of *representative families* introduced by Monien [24]. We use the improved version of this technique recently presented by Fomin et al. [14] and, more precisely, our algorithms and notation are inspired by the ones for Long Directed Cycle given in [14]. We start with some definitions introduced in [14] that can also be found in [10].

▶ **Definition 11.** Let $\mathcal{M}$ be a matroid. Two independent sets A,B of $\mathcal{M}$ *fit* if $A \cap B = \emptyset$ and $A \cup B$ is independent. Also, let $\mathcal{A}$ be a family of sets of size p in $\mathcal{M}$. A subfamily $\mathcal{A}' \subseteq \mathcal{A}$

[1] The ETH states that there is no algorithm solving 3-SAT on a formula with n variables in time $2^{o(n)}$.

is said to *q-represent* $\mathcal{A}$ if for every set B of size q such that there is an $A \in \mathcal{A}$ that fits B, there is an $A' \in \mathcal{A}'$ that also fits B. If $\mathcal{A}'$ q-represents $\mathcal{A}$, we write $\mathcal{A}' \subseteq^q_{\text{rep}} \mathcal{A}$.

4.2.1 Finding 2-spindles with large total size

We start with the algorithm to solve the problem of, given a digraph G and a positive integer ℓ, deciding whether there exist two strictly positive integers ℓ_1, ℓ_2 with $\ell_1 + \ell_2 = \ell$ such that G contains a subdivision of an (ℓ_1, ℓ_2)-spindle, running in time $2^{O(\ell)} \cdot n^{O(1)}$.

If a subdigraph S of G is a subdivision of an (ℓ_1, ℓ_2)-spindle, for some choice of ℓ_1, ℓ_2 such that $\min\{\ell_1, \ell_2\} \geq 1$ and $\ell_1 + \ell_2 = \ell$, we say that S is a *good spindle for* ℓ. Suppose that S is a good spindle for ℓ with tail u and head v, and let ℓ_1, ℓ_2 be as before. A subset $X \subseteq V(G)$ containing two internally vertex-disjoint subpaths of S starting at u on ℓ_1 and ℓ_2 vertices is called a *starting set for* S. Recall that ℓ_1, ℓ_2 are the minimum number of edges in the paths defining S. Therefore, in X, we are taking only the first $\ell_1 - 1, \ell_2 - 1$ edges in each path, which implies that $v \notin X$. Also, we may assume that $\max\{\ell_1, \ell_2\} \geq 2$, as otherwise the desired spindle is just an arc with multiplicity two, which can be detected in polynomial time by using a maximum flow algorithm.

Our algorithm is inspired by the following naive approach. One could compute, for every pair of integers $\ell_1, \ell_2 \geq 1$ with $\ell_1 + \ell_2 = \ell$, and for every triple of vertices u, u_1, u_2, the family $\mathcal{S}^{\ell_1,\ell_2}_{u,u_1,u_2}$ containing the possible candidates for starting sets of good spindles for ℓ with tail u. More formally, let

$$\begin{aligned}\mathcal{S}^{\ell_1,\ell_2}_{u,u_1,u_2} = \Big\{X :\ & X \subseteq V(G), |X| = \ell_1 + \ell_2 - 1, \text{ and } G[X] \text{ contains a} \\ & (u,u_1)\text{-path } P^u_1 \text{ on } \ell_1 \text{ vertices and a } (u,u_2)\text{-path } P^u_2 \\ & \text{on } \ell_2 \text{ vertices such that } V(P^u_1) \cap V(P^u_2) = \{u\}\Big\}.\end{aligned}$$

Then, for every $X \in \mathcal{S}^{\ell_1,\ell_2}_{u,u_1,u_2}$, in order to complete the desired spindle we can just run a flow algorithm in the graph $G - (X \setminus \{u_1, u_2\})$ to decide whether there exists some vertex v that is reachable from u_1 and u_2. Observe that, if such a vertex exists, this would give us the desired spindle and, conversely, if such a spindle exists, then some execution of the described algorithm finds it. The drawback of this approach is the size of the set $\mathcal{S}^{\ell_1,\ell_2}_{u,u_1,u_2}$: it can be roughly as large as $n^{\ell_1+\ell_2}$. The trick is to prove that only a "small" subset of $\mathcal{S}^{\ell_1,\ell_2}_{u,u_1,u_2}$ is needed, namely a subset that "represents" a good spindle.

To give some intuition before getting into the details, let S be a good spindle for ℓ with tail u and head v, and let P^u_1, P^u_2 be subpaths in S starting at u containing ℓ_1, ℓ_2 vertices, respectively, with $\min\{\ell_1, \ell_2\} \geq 1$ and $\ell_1 + \ell_2 = \ell$. Note that this means that $X = V(P_1) \cup V(P_2) \in \mathcal{S}^{\ell_1,\ell_2}_{u,u_1,u_2}$ and $|X| = \ell - 1$. Therefore, by letting $w \in S \setminus X$ (it exists since $v \notin X$), if we consider the uniform matroid with ground set $V(G)$ and rank ℓ, then X fits $\{w\}$, and any 1-representative family (see Definition 11) for $\mathcal{S}^{\ell_1,\ell_2}_{u,u_1,u_2}$ must contain some X' that also fits $\{w\}$. We shall prove that a good spindle for ℓ is found when iterating over X'. For the proof to work, we need to compute q-representative families for larger values of q.

More precisely, for every triple of vertices $u, u_1, u_2 \in V(G)$ and positive integers ℓ_1, ℓ_2, q with $1 \leq \ell_1, \ell_2 \leq \ell - 1$ and $1 \leq q \leq \ell - 1$, we compute in time $2^{O(\ell)} \cdot n^{O(1)}$ a q-representative family for $\mathcal{S}^{\ell_1,\ell_2}_{u,u_1,u_2}$ in the uniform matroid with ground set $V(G)$ and rank $\ell + q - 1$, denoted by

$$\widehat{\mathcal{S}}^{\ell_1,\ell_2,q}_{u,u_1,u_2} \subseteq^q_{\text{rep}} \mathcal{S}^{\ell_1,\ell_2}_{u,u_1,u_2}.$$

We postpone the computation of the above q-representative families in time $2^{O(\ell)} \cdot n^{O(1)}$ to Section 4.2.3, and assume now that we already have these families at hand. The following lemma states that they are enough to find the desired good spindle.

▶ **Lemma 12.** *G contains a good spindle S for ℓ with tail u if and only if G contains a good spindle for ℓ with starting set X', for some $X' \in \widehat{\mathcal{S}}^{\ell_1,\ell_2,q}_{u,u_1,u_2}$, $u_1, u_2 \in V(G)$, $\ell_1, \ell_2 \in [\ell-1]$ such that $\ell_1 + \ell_2 = \ell$, and some $q \in [\ell-1]$.*

Proof. Clearly, we only need to prove the sufficiency part. Let S be a good spindle for ℓ in G with minimum number of vertices, which exists by hypothesis, and let u and v be the tail and the head of S, respectively. Let $P_1^u = (u, \ldots, u_1)$ and $P_2^u = (u, \ldots, u_2)$ be two subpaths in S outgoing from u, on ℓ_1 and ℓ_2 vertices, respectively, with $\ell_1 + \ell_2 = \ell$. Let also $P_1^v = (u_1, \ldots, v)$ and $P_2^v = (u_2, \ldots, v)$ be the two subpaths in S from u_1 and u_2 to v, respectively; see Figure 4. Let $S_u = V(P_1^u) \cup V(P_2^u)$, and note that $|S_u| = \ell - 1$ and $S_u \in \mathcal{S}^{\ell_1,\ell_2}_{u,u_1,u_2}$.

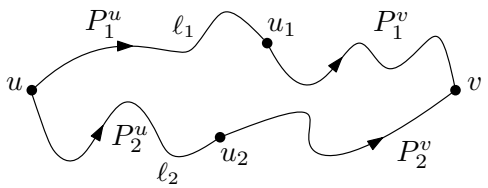

Figure 4 Illustration of the vertices and paths defined in the proof of Lemma 12.

If $|V(S) \setminus S_u| \leq \ell - 2$, then, by letting $B = V(S) \setminus S_u$ and $q = |B|$, there exists $\widehat{S}_u \in \widehat{\mathcal{S}}^{\ell_1,\ell_2,q}_{u,u_1,u_2}$, and $\widehat{S}_u \cup B$ clearly contains a good spindle with starting set $\widehat{S}_u$. Therefore, suppose $|V(S) \setminus S_u| \geq \ell - 1$, and let $q = \ell - 1$. Recall that $v \notin \{u_1, u_2\}$, and let B be the union of two subpaths $P_1^B = (v_1, \ldots, v)$ and $P_2^B = (v_2, \ldots, v)$ contained in $V(P_1^v) \cup V(P_2^v)$ with $V(P_1^B) \cap V(P_2^B) = \{v\}$ and $|V(P_1^B) \cup V(P_2^B)| = \ell - 1$. Note that there may be several choices for the lengths of P_1^B and P_2^B, as far as their joint number of vertices is equal to $\ell - 1$. Since $S_u \in \mathcal{S}^{\ell_1,\ell_2}_{u,u_1,u_2}$ and $S_u \cap B = \emptyset$, by definition of q-representative family there exists $\widehat{S}_u \in \widehat{\mathcal{S}}^{\ell_1,\ell_2,q}_{u,u_1,u_2}$ such that $\widehat{S}_u \cap B = \emptyset$. We claim that $\widehat{S}_u \cap (V(P_1^v) \cup V(P_2^v)) = \{u_1, u_2\}$, which concludes the proof of the lemma. Let $\widehat{P}_1^u$ and $\widehat{P}_2^u$ be the two paths in $G[\widehat{S}_u]$ with $V(\widehat{P}_1^u) \cap V(\widehat{P}_1^u) = \{u\}$. If $(\widehat{S}_u \cap (V(P_1^v) \cup V(P_2^v))) \setminus \{u_1, u_2\} = \emptyset$, then a good spindle can be easily obtained, so suppose otherwise; we distinguish two cases.

Suppose first that each of the paths $\widehat{P}_1^u$ and $\widehat{P}_2^u$ intersects exactly one of the paths P_1^v and P_2^v. By hypothesis, there exists a vertex $w \in (\widehat{S}_u \cap (V(P_1^v) \cup V(P_2^v))) \setminus \{u_1, u_2\}$, and suppose without loss of generality that $w \in V(\widehat{P}_1^u) \cap V(P_1^v)$; see Figure 5(a) for an illustration. We define a good spindle for ℓ, $\widehat{S}$, in G as follows. The tail and head of $\widehat{S}$ are vertices u and v, respectively. The first path of $\widehat{S}$ starts at u, follows $\widehat{P}_1^u$ until its first intersection with P_1^v (vertex w in Figure 5(a)), which is distinct from u_1 by hypothesis, and then follows P_1^v until v. The second path of $\widehat{S}$ starts at u, follows $\widehat{P}_2^u$ until its first intersection with P_2^v, which may be vertex u_2, and then follows P_2^v until v. Since $|B| = \ell - 1$ and each of $\widehat{P}_1^u$ and $\widehat{P}_2^u$ intersects exactly one of P_1^v and P_2^v, it follows that $\widehat{S}$ is indeed a good spindle for ℓ. On the other hand, since $|V(\widehat{P}_1^u) \cup V(\widehat{P}_2^u)| = |V(P_1^u) \cup V(P_2^u)|$ and vertex w comes strictly after u_1 in P_1^v, it follows that the first path of $\widehat{S}$ is strictly shorter than the corresponding path of S, while the second one is not longer. Therefore, $|V(\widehat{S})| < |V(S)|$, a contradiction to the choice of S.

Suppose now that one of the paths $\widehat{P}_1^u$ and $\widehat{P}_2^u$, say $\widehat{P}_1^u$, intersects both P_1^v and P_2^v. Without loss of generality, suppose that, starting from u, $\widehat{P}_1^u$ meets P_1^v before P_2^v. Let w_1

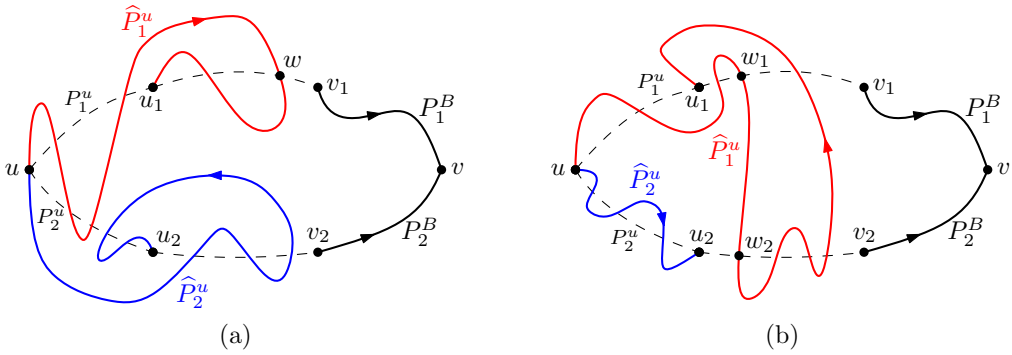

Figure 5 Illustration of the last two cases in the proof of Lemma 12.

and w_2 be vertices of $\widehat{P}_1^u$ such that $w_1 \in V(P_1^v)$, $w_2 \in V(P_2^v)$, and there is no vertex of $\widehat{P}_1^u$ between w_1 and w_2 that belongs to $V(P_1^v) \cup V(P_2^v)$; see Figure 5(b) for an illustration. We define a good spindle for ℓ, $\widehat{S}$, in G as follows. The tail and head of $\widehat{S}$ are vertices w_1 and v, respectively. The first path of $\widehat{S}$ starts at w_1 and follows P_1^v until v. The second path of $\widehat{S}$ starts at w_1, follows $\widehat{P}_1^u$ until w_2, and then follows P_2^v until v. By the choice of w_1 and w_2 and since $|B| = \ell - 1$, it follows that $\widehat{S}$ is indeed a good spindle for ℓ. On the other hand, by construction $|V(\widehat{S})| \leq |V(S)| - |V(\widehat{P}_2^u)| < |V(S)|$, contradicting again the choice of S. ◀

Wrapping up the algorithm. We finally have all the ingredients to describe our algorithm, which proceeds as follows. First, for every triple of vertices $u, u_1, u_2 \in V(G)$ and positive integers ℓ_1, ℓ_2, q with $1 \leq \ell_1, \ell_2 \leq \ell - 1$ and $1 \leq q \leq \ell - 1$, we compute, as explained in Section 4.2.3, a q-representative family $\widehat{\mathcal{S}}^{\ell_1,\ell_2,q}_{u,u_1,u_2} \subseteq^q_{\text{rep}} \mathcal{S}^{\ell_1,\ell_2}_{u,u_1,u_2}$ of size $2^{O(\ell)}$ in time $2^{O(\ell)} \cdot n^{O(1)}$. Then the algorithm checks, for every combination of values $u, u_1, u_2, \ell_1, \ell_2$, for every $S \in \widehat{\mathcal{S}}^{\ell_1,\ell_2,q}_{u,u_1,u_2}$ and every $v \in V(G - (S \setminus \{u_1, u_2\})))$, whether G contains a (u_1, v)-path P_1^v and a (u_2, v)-path P_2^v such that $V(P_1^v) \cap V(P_2^v) = \{v\}$ and $S \cap (V(P_1^v) \cup V(P_2^v)) = \{u_1, u_2\}$. Note that the latter check can be easily performed in polynomial time by a flow algorithm [2]. The correctness of the algorithm follows directly from Lemma 12, and its running time is $2^{O(\ell)} \cdot n^{O(1)}$, as claimed. In order to keep the exposition as simple as possible, we did not focus on optimizing either the constants involved in the algorithm or the degree of the polynomial factor. Nevertheless, explicit small constants can be derived by carefully following the details in Fomin et al. [14], or alternatively in [29].

4.2.2 Finding 2-spindles with two specified lengths

We now turn to the problem of finding 2-spindles with two specified lengths. Namely, given a digraph G and two integers ℓ_1, ℓ_2 with $\ell_2 \geq \ell_1 \geq 1$, our objective is to decide whether G contains a subdivision of an (ℓ_1, ℓ_2)-spindle in time $2^{O(\ell_2)} \cdot n^{O(\ell_1)}$. Note that this problem differs from the one considered in Section 4.2.1, as now we specify *both* lengths of the desired spindle, instead of just its total size. Our approach is similar to the one presented in Section 4.2.1, although some more technical ingredients are needed, and we need to look at the problem from a slightly different point of view.

In the proof of Lemma 12 in the previous section, we consider the existence of an (ℓ_1, ℓ_2)-spindle S, which is good for the total size $\ell = \ell_1 + \ell_2$, and with starting set S_u, and prove that there must exist a good spindle for ℓ with a starting set $\widehat{S}_u$ that is contained in the set of q-representatives. For this, we use the fact that the part of $S \setminus S_u$ is either so small that it must be disjoint from any q-representative, or that it is already big enough, i.e., has *total*

size at least ℓ. However, we cannot ensure that the "residual" paths in $S \setminus S_u$ have sizes at least ℓ_1 and ℓ_2, respectively. This is why in this section we need to use a different approach. Note that the running time we are aiming at allows us to guess the first ℓ_1 vertices in the "short" path.

Given integers ℓ_1, ℓ_2 with $\ell_2 \geq \ell_1 \geq 1$, we say that a subdigraph S of G is a *good spindle (for* (ℓ_1, ℓ_2)*)* if it is a subdivision of an (ℓ_1, ℓ_2)-spindle. We may again assume that $\max\{\ell_1, \ell_2\} \geq 2$. The main difference with respect to Section 4.2.1 is that now we will only represent the candidates for the first ℓ_2 vertices of the "long" path. To this end, we define, similarly to [14], the following set for every pair of vertices $u, u' \in V(G)$ and positive integer ℓ_2:

$$\mathcal{P}^{\ell_2}_{u,u'} = \Big\{X : \ X \subseteq V(G), |X| = \ell_2, \text{ and } G[X] \text{ contains a } (u,u')\text{-path on } \ell_2 \text{ vertices}\Big\}.$$

The above sets are exactly the same as those defined by Fomin et al. [14] to solve the Long Directed Cycle problem. Therefore, we can just apply [14, Lemma 5.2] and compute, for every pair of vertices $u, u' \in V(G)$ and positive integers ℓ_2, q with $q \leq \ell_1 + \ell_2 \leq 2\ell_2$, a q-representative family

$$\widehat{\mathcal{P}}^{\ell_2,q}_{u,u'} \subseteq^q_{\text{rep}} \mathcal{P}^{\ell_2}_{u,u'}$$

of size $2^{O(\ell_2)}$ in time $2^{O(\ell_2)} \cdot n^{O(1)}$.

Now we would like to state the equivalent of Lemma 12 adapted to the new representative families. However, it turns out that the families $\widehat{\mathcal{P}}^{\ell_2,q}_{u,u'}$ do not yet suffice in order to find the desired spindle. To circumvent this cul-de-sac, we use the following trick: we first try to find "short" spindles using the color-coding technique of Alon et al. [1], and if we do not succeed, we can guarantee that all good spindles have at least one "long" path. In this situation, we can prove that the families $\widehat{\mathcal{P}}^{\ell_2,q}_{u,u'}$ are indeed enough to find a good spindle. More precisely, a good spindle S is said to be *short* if both its paths have at most $2\ell_2$ vertices, and it is said to be *long* otherwise. Note that Lemma 13 only applies to digraphs without good short spindles.

Similarly as before, given a long spindle S defined by two internally vertex-disjoint paths P_1, P_2, where P_2 contains at least $2\ell_2$ vertices, we say that the subset $X \subseteq V(G)$ containing the first ℓ_2 vertices of P_2 is the *starting set for* S. Also, if u' is the extremity of X distinct from u, we say that X *ends in* u'.

▶ **Lemma 13.** *Let G be a digraph containing no good short spindles. If G contains a good long spindle with starting set that ends in a vertex u', then G contains a long spindle with starting set $X \subseteq V(\widehat{P}^u_2)$, for some $V(\widehat{P}^u_2) \in \widehat{\mathcal{P}}^{\ell_2,q}_{u,u'}$ with $q = \ell_1 + \ell_2 - 1$.*

Proof. Let S be a good long spindle in G with minimum number of vertices, which exists by hypothesis, and let u and v be the tail and the head of S, respectively. Let P_1 be the shortest of the two (u,v)-paths of S, and let u' be the vertex on the other path of S at distance exactly $\ell_2 - 1$ from u. Let P^u_2 and P^v_2 be the (u,u')-path and the (u',v)-path in S, respectively; see Figure 6. Note that $P^u_2 \in \mathcal{P}^{\ell_2}_{u,u'}$. Since by hypothesis S is a *long* spindle, it follows that $|V(P^v_2)| > \ell_2$.

Again, in order to apply the properties of q-representative families, we define a vertex set $B \subseteq V(S)$ as follows, crucially using the hypothesis that S is a good long spindle. Namely, B contains the last ℓ_1 vertices of the path P_1 together with the last ℓ_2 vertices of the path P^v_2, including v. Note that $|B| = \ell_1 + \ell_2 - 1$ and that, since $|V(P^v_2)| > \ell_2$, we have $V(P^u_2) \cap B = \emptyset$.

Let $q = |B|$. Since $P^u_2 \in \mathcal{P}^{\ell_2}_{u,u'}$ and $V(P^u_2) \cap B = \emptyset$, by definition of q-representative family there exists a set in $\widehat{\mathcal{P}}^{\ell_2,q}_{u,u'}$ corresponding to a (u,u')-path $\widehat{P}^u_2$ such that $V(\widehat{P}^u_2) \cap B = \emptyset$.

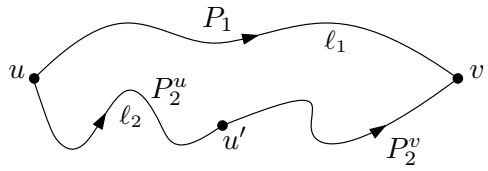

Figure 6 Illustration of the vertices and paths described in the proof of Lemma 13.

We claim that $V(\widehat{P}_2^u) \cap V(S) \subseteq V(P_2^u)$, which concludes the proof of the lemma. Assume for contradiction that $(V(\widehat{P}_2^u) \cap V(S)) \setminus V(P_2^u) \neq \emptyset$, and we again distinguish two cases.

Suppose first that $\widehat{P}_2^u$ is disjoint from P_1, except for vertex u. Let w be the first vertex of $\widehat{P}_2^u$ in $V(P_2^v) \setminus \{u'\}$; see Figure 7(a) for an illustration. We define a good spindle $\widehat{S}$ in G as follows. The tail and head of $\widehat{S}$ are vertices u and v, respectively. The first path of $\widehat{S}$ is equal to P_1. The second path of $\widehat{S}$ starts at u, follows $\widehat{P}_2^u$ until its first intersection with P_1^v (vertex w in Figure 7(a)), which is distinct from u_1 by hypothesis, and then follows P_1^v until v. By definition of B, it follows that $\widehat{S}$ is a good spindle, and by construction $|V(\widehat{S})| < |V(S)|$, a contradiction to the choice of S.

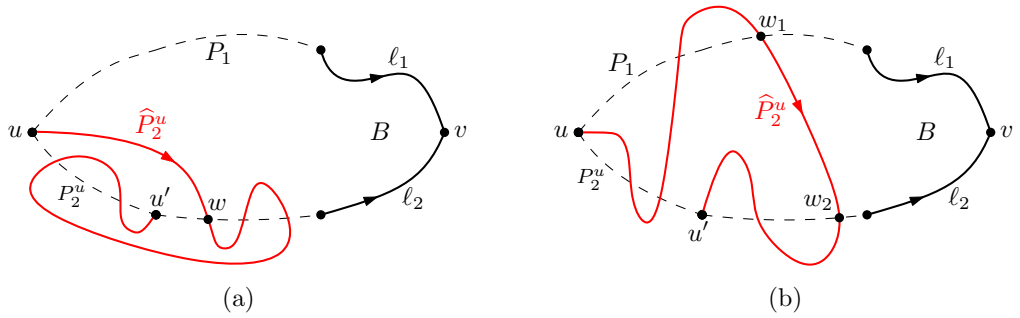

Figure 7 Illustration of the two cases in the proof of Lemma 13.

Suppose now that $\widehat{P}_2^u$ intersects P_1. Since $\widehat{P}_2^u$ ends at vertex $u' \notin V(P_1)$, there exist vertices w_1, w_2 such that $w_1 \in V(P_1)$, $w_2 \in V(\widehat{P}_2^u) \cup V(P_2^v)$, and there is no vertex of $\widehat{P}_1^u$ between w_1 and w_2 that belongs to $V(P_1) \cup V(\widehat{P}_2^u) \cup V(P_2^v)$; see Figure 7(b) for an illustration. We define a good spindle $\widehat{S}$ in G as follows. The tail and head of $\widehat{S}$ are vertices w_1 and v, respectively. The first path of $\widehat{S}$ starts at w_1 and follows P_1 until v. The second path of $\widehat{S}$ starts at w_1, follows $\widehat{P}_2^u$ until w_2, and then follows $\widehat{P}_2^u \cup P_2^v$ until v. By the choices of B, w_1, and w_2, it follows that $\widehat{S}$ is a good spindle with $|V(\widehat{S})| < |V(S)|$, contradicting again the choice of S. ◀

Wrapping up the algorithm. We start by trying to find good small spindles. Namely, for every pair of integers ℓ_1', ℓ_2' with $\ell_1 \leq \ell_1' \leq 2\ell_2$ and $\ell_2 \leq \ell_2' \leq 2\ell_2$, we test whether G contains an (ℓ_1', ℓ_2')-spindle as a *subgraph*, by using the color-coding technique of Alon et al. [1]. Since the treewidth of an undirected spindle is two, this procedure takes time $2^{O(\ell_2)} \cdot n^{O(1)}$.

If we succeed, the algorithm stops. Otherwise, we can guarantee that G does not contain any good short spindle, and therefore we are in position to apply Lemma 13. Before this, we first compute, for every pair of vertices $u, u' \in V(G)$ and positive integers ℓ_2, q with $q \leq \ell_1 + \ell_2 \leq 2\ell_2$, a q-representative family $\widehat{\mathcal{P}}_{u,u'}^{\ell_2,q} \subseteq_{\text{rep}}^{q} \mathcal{P}_{u,u'}^{\ell_2}$ of size $2^{O(\ell_2)}$ in time $2^{O(\ell_2)} \cdot n^{O(1)}$, using [14, Lemma 5.2].

Now, for each path $\widehat{P}_2^u$ such that $V(\widehat{P}_2^u) \in \widehat{\mathcal{P}}_{u,u'}^{\ell_2,q}$, with $q = \ell_1 + \ell_2 - 1$, we proceed as follows. By Lemma 13, it is enough to guess a vertex $v \in V(G)$ and check whether G contains

a (u,v)-path P_1 of length at least ℓ_1, and a (u',v)-path P_2^v such that $V(P_1) \cap V(\widehat{P}_2^u) = \{u\}$, $V(P_1) \cap V(P_2^v) = \{v\}$, and $V(\widehat{P}_2^u) \cap V(P_2^v) = \{u'\}$. In order to do so, we apply brute force and we guess the first ℓ_1 vertices of P_1 in time $n^{O(\ell_1)}$. Let these vertices be $u, u_2, \ldots, u_{\ell_1}$. All that remains is to test whether the graph $G \setminus \{u_2, \ldots, u_{\ell_1-1}\} \setminus (V(\widehat{P}_2^u) \setminus \{u'\})$ contains two internally vertex-disjoint paths from u_{ℓ_1} and u' to u, which can be done in polynomial time by using a flow algorithm [2]. The correctness of the algorithm follows by the above discussion, and its running time is $2^{O(\ell_2)} \cdot n^{O(\ell_1)}$, as claimed. Again, we did not focus on optimizing the constants involved in the algorithm.

4.2.3 Computing the representative families efficiently

We now explain how the representative families used in Sections 4.2.1 and 4.2.2 can be efficiently computed, by using the results of Fomin et al. [14]. As discussed in Section 4.2.2, the families $\widehat{\mathcal{P}}_{u,u'}^{\ell_2,q}$ are exactly the same as those used by Fomin et al. [14], so we can directly use [14, Lemma 5.2] and compute them in time $2^{O(\ell_2)} \cdot n^{O(1)}$. Let us now explain how the results of Fomin et al. [14] can be used to compute efficiently the families $\widehat{\mathcal{S}}_{u,u_1,u_2}^{\ell_1,\ell_2,q}$ used in Section 4.2.1. We need the following lemma.

▶ **Lemma 14** (Fomin et al. [14]). *Let $M = (E, \mathcal{I})$ be a matroid and $\mathcal{S}$ be a family of subsets of E. If $\mathcal{S} = \mathcal{S}_1 \cup \cdots \cup \mathcal{S}_k$ and $\widehat{\mathcal{S}}_i \subseteq_{rep}^q \mathcal{S}_i$ for $1 \le i \le k$, then $\cup_{i=1}^k \widehat{\mathcal{S}}_i \subseteq_{rep}^q \mathcal{S}$.*

The key observation is that the families $\mathcal{S}_{u,u_1,u_2}^{\ell_1,\ell_2}$ can be obtained by combining pairs of elements in the families $\mathcal{P}_{u,u'}^{\ell_2}$. More precisely, for every triple of vertices u, u_1, u_2 and positive integers ℓ_1, ℓ_2, it holds that

$$\mathcal{S}_{u,u_1,u_2}^{\ell_1,\ell_2} \subseteq \mathcal{P}_{u,u_1}^{\ell_1} \cup \mathcal{P}_{u,u_2}^{\ell_2}.$$

Note that in the above equation we do not have equality, as some pairs of paths in $\mathcal{P}_{u,u_1}^{\ell_1}$ and $\mathcal{P}_{u,u_2}^{\ell_2}$, respectively, may intersect at other vertices distinct from u.

By Lemma 14, if $\widehat{\mathcal{P}}_{u,u_1}^{\ell_1,q} \subseteq_{\text{rep}}^q \mathcal{P}_{u,u_1}^{\ell_1}$ and $\widehat{\mathcal{P}}_{u,u_2}^{\ell_2,q} \subseteq_{\text{rep}}^q \mathcal{P}_{u,u_2}^{\ell_2}$, then

$$\widehat{\mathcal{P}}_{u,u_1}^{\ell_1,q} \cup \widehat{\mathcal{P}}_{u,u_2}^{\ell_2,q} \subseteq_{\text{rep}}^q \mathcal{P}_{u,u_1}^{\ell_1} \cup \mathcal{P}_{u,u_2}^{\ell_2}.$$

To conclude, it just remains to observe that, by the definition of q-representative family, it holds that if $M = (E, \mathcal{I})$ is a matroid, $\mathcal{S}$ is a family of subsets of E, $\mathcal{S}' \subseteq \mathcal{S}$ and $\widehat{\mathcal{S}} \subseteq_{\text{rep}}^q \mathcal{S}$, then $\widehat{\mathcal{S}} \subseteq_{\text{rep}}^q \mathcal{S}'$ as well.

Therefore, for every triple of vertices u, u_1, u_2 and positive integers ℓ_1, ℓ_2 with $\ell_1, \ell_2 \le \ell$, in order to compute a q-representative family for $\mathcal{S}_{u,u_1,u_2}^{\ell_1,\ell_2}$, we can just take the union of q-representative families for $\mathcal{P}_{u,u_1}^{\ell_1}$ and $\mathcal{P}_{u,u_2}^{\ell_2}$, and these latter families can be computed in time $2^{O(\ell)} \cdot n^{O(1)}$ by [14, Lemma 5.2].

5 Finding spindles on directed acyclic graphs

In this section we focus on the case where the input digraph is acyclic. We start by proving Theorem 4. The proof uses classical dynamic programming along a topological ordering of the vertices of the input acyclic digraph.

Proof of Theorem 4. Given an acyclic digraph G and positive integers k, ℓ, recall that we want to prove that one can decide in time $O(2^k \cdot \ell^k \cdot n^{2k+1})$ whether G has a subdivision of a $(k \times \ell)$-spindle. For this, let H be obtained from the empty digraph by adding, for each

vertex $u \in V(G)$, vertices u^+, u^- and an arc (u^+, u^-) between them, and adding arc (u^-, v^+) for each arc $(u, v) \in A(G)$. Note that H is also acyclic, and fix an arbitrary topological ordering of $V(H)$.

▶ **Claim 2.** *There exists a subdivision of a $(k \times \ell)$-spindle in G if and only if there exist $x, y \in V(H)$ and k arc-disjoint (x, y)-paths in H, each of length at least $2\ell - 1$.*

Proof of the claim. On the one hand, each path of a $(k \times \ell)$-spindle gives rise to a path in H of length at least $2\ell - 1$, since each internal vertex of a path is split into two (these paths are actually vertex-disjoint). On the other hand, let $P_1, \ldots, P_k$ be arc-disjoint paths between $x, y \in V(H)$, each of length at least $2\ell - 1$. Since either $|N^+(z)| = 1$ or $|N^-(z)| = 1$ for every $z \in V(H)$, and since $P_1, \ldots, P_k$ are arc-disjoint, we get that $P_1, \ldots, P_k$ are actually internally vertex-disjoint. Now, to obtain the desired $(k \times \ell)$-spindle, it suffices to observe that if $u^+ \in V(P_i) \setminus \{y\}$, for some $u \in V(G)$ and some $i \in \{1, \ldots, k\}$, then $u^- \in V(P_i)$. □

We want to decide whether H has the desired paths. For each $x \in V(H)$, we define the table P_x with entries $(e_1, t_1, \ldots, e_k, t_k)$, for each choice of at most k distinct arcs $e_1, \ldots, e_k$ (some of these may not exist, in which case we represent it by 'null'), and for each choice of k values $t_1, \ldots, t_k$ from the set $\{0, 1, \ldots, 2\ell - 1\}$. Observe that P_x has size $(|A(H)| + 1)^k \cdot (2\ell)^k$, which, since we need to analyze the table of each vertex, gives us the claimed complexity of the algorithm. The meaning of an entry is given below:

> $P_x(e_1, t_1, \ldots, e_k, t_k) = \mathsf{true}$ if and only if there exist k arc-disjoint paths $P_1, \ldots, P_k$ starting at x and ending at $e_1, \ldots, e_k$ of length at least $t_1, \ldots, t_k$, respectively.

We compute these tables starting at small values of $\sum_{i=1}^{k} t_i$. Namely, for $t_1 = t_2 = \ldots = t_k = 0$, it holds that $P_x(e_1, t_1, \ldots, e_k, t_k) = \mathsf{true}$ if and only if $\{e_1, \ldots, e_k\} = \emptyset$.

Now, to compute $P_x(e_1, t_1, \ldots, e_k, t_k)$, let w be the greatest vertex in $\{z \in V(H) : (z', z) \in \{e_1, \ldots, e_k\}\}$, and let w' be the greatest vertex in $\{z \in V(H) : (z, w) \in \{e_1, \ldots, e_k\}\}$, according to the chosen topological ordering of $V(H)$. Also, let $e_i = (w', w)$. If $w = x$, then the entry is given above, so suppose otherwise.

▶ **Claim 3.** $P_x(e_1, t_1, \ldots, e_k, t_k) = \mathsf{true}$ *if and only if* $P_x(e_1, t_1, \ldots, e, t_i - 1, \ldots, e_k, t_k) = \mathsf{true}$, *for some arc $e \in A(H) \setminus \{e_1, \ldots, e_k\}$ incoming at w'.*

Proof of the claim. Suppose first that $P_x(e_1, t_1, \ldots, e_k, t_k) = \mathsf{true}$, and let $P_1, \ldots, P_k$ be arc-disjoint paths starting at x and ending at $e_1, \ldots, e_k$ of length at least $t_1, \ldots, t_k$, respectively. Let e be the arc preceding e_i in path P_i (e can denote the empty set when e_i is incident to x). Then, $P_1, \ldots, P_{i-1}, P_i - e_i, P_{i+1}, \ldots, P_k$ are arc-disjoint paths ending at $e_1, \ldots, e_{i-1}, e, e_{i+1}, \ldots, e_k$ of length at least $t_1, \ldots, t_{i-1}, t_i - 1, t_{i+1}, \ldots, t_k$, respectively.

Conversely, let $P_1, \ldots, P_k$ be arc-disjoint paths that certify entry $P_x(e_1, t_1, \ldots, e, t_i - 1, \ldots, e_k, t_k)$. If $e_i \notin A(P_j)$, for every $j \in \{1, \ldots, k\}$, then $P_1, \ldots, P_{i-1}, P_i + e_i, P_{i+1}, \ldots, P_k$ are the desired paths. So suppose that $e_i \in A(P_j)$. If $j = i$, then we get a cycle in H, a contradiction. Otherwise, because $e_i \neq e_j$ and P_j ends in e_j, we get that there is a path starting in w and ending in z', where $e_j = (z, z')$. This contradicts the choice of w. □

By Claim 3, the entry $P_x(e_1, t_1, \ldots, e_k, t_k)$ can be computed by verifying at most $|N^+(w')|$ smaller entries. By Claim 2, the desired spindle exists if and only if there exist $x, y \in V(H)$ and k arcs $e_1, \ldots, e_k$ incoming at y such that $P_x(e_1, 2\ell - 1, e_2, 2\ell - 1, \ldots, e_k, 2\ell - 1) = \mathsf{true}$. The theorem follows. □

Motivated by the fact that finding a subdivision of a general digraph F is in XP parameterized by $|V(F)|$ on acyclic digraphs [3, 23], we now present two hardness results about

finding subdivisions of disjoint spindles on acyclic digraphs. The first result holds even for *planar* acyclic digraphs.

▶ **Proposition 15.** *If F is the disjoint union of (2×1)-spindles, then deciding whether a planar acyclic digraph contains a subdivision of F is* NP*-complete.*

Proof. We reduce from the problem of deciding whether the edges of a tripartite graph can be partitioned into triangles, which is known to be NP-complete [15], even restricted to planar tripartite graphs [31]. Let G be an input planar tripartite (undirected) graph, and let $A \cup B \cup C$ be a tripartition of $V(G)$. We build from G a planar acyclic digraph G' by orienting all edges from A to B, from B to C, and from A to C. It is clear that $E(G)$ admits a partition into triangles if and only if G' contains as a subdivision (in fact, as a subdigraph) the digraph containing $|E(G)|/3$ disjoint copies of a (2×1)-spindle. ◀

Our next result shows that, for some choices of F, finding a subdivision of F is W[1]-hard on acyclic digraphs. We just present a sketch of proof, as the reduction is based on a minor modification of an existing reduction of Slivkins [30].

▶ **Proposition 16.** *If F is the disjoint union of a $(k_1 \times 1)$-spindle and a $(k_2 \times 1)$-spindle, then deciding whether an acyclic digraph contains a subdivision of F is* W[1]*-hard parameterized by $k_1 + k_2$.*

Sketch of proof. The proof is done by appropriately modifying the reduction for Edge-Disjoint Paths on acyclic digraphs given by Slivkins [30], which carries over to the vertex-disjoint version as well. The reduction is from k-Clique, and the sets of demands to be satisfied consist just of a multiarc with multiplicity $\binom{k}{2}$ and another one with multiplicity k between two given pairs of terminals. The idea is the following: since in our problem we do not have *fixed* terminals, we "simulate" them by leaving only four vertices of high degree, so that finding the desired subdivision will only be possible by using the prescribed four vertices as endpoints. To do so, we take the construction of Slivkins [30] and for each vertex, except for the four prescribed ones, we replace its outgoing (resp. incoming) arcs by an out-arborescence (resp. in-arborescence) of out-degree (resp. in-degree) at most two. Note that this operation may blow up the size of the subdivision, but it does not matter, as the parameter remains the same. By taking F to be the disjoint union of a $(\binom{k}{2} \times 1)$-spindle and a $(k \times 1)$-spindle, the result follows. □

It is worth noting that the problem considered in Proposition 16 is para-NP-hard on general digraphs, as the conditions of [3, Theorem 8] are easily seen to be fulfilled.

6 Conclusions

We studied the complexity of several problems consisting in finding subdivisions of spindles on digraphs. For a general spindle F, we do not know if finding a subdivision of F is FPT on general digraphs parameterized by $|V(F)|$, although we believe that it is indeed the case. As a partial result, one could try to prove that, for a *fixed* value of $\ell \geq 4$, finding a subdivision of a $(k \times \ell)$-spindle is FPT parameterized by k (the problem is NP-hard by Theorem 1).

The above question is open even if the input digraph is acyclic (note that Theorem 4 does *not* answer this question), or even if F is a 2-spindle. Concerning 2-spindles, one may try to use the technique we used to prove Theorems 2 and 3, based on representative families in matroids. However, the technique does not seem to be easily applicable when the parameter is the total size of a prescribed 2-spindle. Namely, using the terminology from

Section 4.2.2, the bottleneck is to find spindles that have one "short" and one "long" path. On the other hand, generalizing this technique to spindles with more than two paths seems pretty complicated.

It may be possible that the trick used by Zehavi [32] to avoid the use of representative families to solve Long Directed Cycle can be adapted to our setting as well. Another approach might be to use the *divide-and-color* technique [21].

Cai and Ye [8] recently studied the problem of finding two edge-disjoint paths on undirected graphs with length constraints between specified vertices. These length constraints can be an upper bound, a lower bound, or an equality on the lengths of each of the two desired paths, or no restriction at all, resulting in nine different problems. Interestingly, out of these nine problems, Cai and Ye [8] gave FPT algorithms for seven of them, and left open only the following two cases: when there is only one constraint of type 'at least', and when both constraints are of type 'at least'. Observe that this latter problem is closely related to finding a subdivision of a 2-spindle.

In general, very little is known about the complexity of finding subdivisions on digraphs. Bang-Jensen et al. [3] conjectured that, considering $|V(F)|$ as a constant, the problem of finding a subdivision of F is either polynomial-time solvable or NP-complete. This conjecture is wide open. Recently, Havet et al. [17] studied the cases where $|V(F)| = 4$, and managed to classify all of them up to five exceptions. Even less is known about the parameterized complexity of the cases that are polynomial-time solvable for fixed F, that is, the cases in XP. In this article we focused on spindles, but there are other potential candidates such as, using the terminology of [3], windmills, palms, or antipaths.

References

1. N. Alon, R. Yuster, and U. Zwick. Color-coding. *Journal of the ACM*, 42(4):844–856, 1995.
2. J. Bang-Jensen and G. Gutin. *Digraphs: Theory, Algorithms and Applications*. Springer-Verlag, 2nd edition, 2008.
3. J. Bang-Jensen, F. Havet, and A. K. Maia. Finding a subdivision of a digraph. *Theoretical Computer Science*, 562:283–303, 2015.
4. A. Benhocine and A. P. Wojda. On the existence of specified cycles in a tournament. *Journal of Graph Theory*, 7(4):469–473, 1983.
5. A. Björklund, T. Husfeldt, and S. Khanna. Approximating longest directed paths and cycles. In *Proc. of the 31st International Colloquium on Automata, Languages and Programming (ICALP)*, volume 3142 of *LNCS*, pages 222–233, 2004.
6. H. L. Bodlaender, B. M. P. Jansen, and S. Kratsch. Kernelization lower bounds by cross-composition. *SIAM Journal on Discrete Mathematics*, 28(1):277–305, 2014.
7. R. C. Brewster, P. Hell, S. H. Pantel, R. Rizzi, and A. Yeo. Packing paths in digraphs. *Journal of Graph Theory*, 44(2):81–94, 2003.
8. L. Cai and J. Ye. Finding two edge-disjoint paths with length constraints. In *Proc. of the 42nd International Workshop on Graph-Theoretic Concepts in Computer Science (WG)*, volume 9941 of *LNCS*, pages 62–73, 2016.
9. N. Cohen, F. Havet, W. Lochet, and N. Nisse. Subdivisions of oriented cycles in digraphs with large chromatic number. *Journal of Graph Theory*, 89(4):439–456, 2018.
10. M. Cygan, F. V. Fomin, L. Kowalik, D. Lokshtanov, D. Marx, M. Pilipczuk, M. Pilipczuk, and S. Saurabh. *Parameterized Algorithms*. Springer, 2015.
11. R. Diestel. *Graph Theory, 4th Edition*, volume 173 of *Graduate texts in mathematics*. Springer, 2012.
12. R. G. Downey and M. R. Fellows. *Fundamentals of Parameterized Complexity*. Texts in Computer Science. Springer, 2013.

13 J. Flum and M. Grohe. *Parameterized Complexity Theory.* Springer-Verlag, 2006.

14 F. V. Fomin, D. Lokshtanov, F. Panolan, and S. Saurabh. Efficient computation of representative families with applications in parameterized and exact algorithms. *Journal of the ACM*, 63(4):29:1–29:60, 2016.

15 M. R. Garey and D. S. Johnson. *Computers and intractability. A guide to the theory of NP-completeness.* W. H. Freeman and Co., 1979.

16 M. Grohe, K. Kawarabayashi, D. Marx, and P. Wollan. Finding topological subgraphs is fixed-parameter tractable. In *Proc. of the 43rd ACM Symposium on Theory of Computing (STOC)*, pages 479–488, 2011.

17 F. Havet, A. K. Maia, and B. Mohar. Finding a subdivision of a prescribed digraph of order 4. *Journal of Graph Theory*, 87(4):536–560, 2018.

18 R. Impagliazzo, R. Paturi, and F. Zane. Which problems have strongly exponential complexity? *Journal of Computer and System Sciences*, 63(4):512–530, 2001.

19 A. Itai, Y. Perl, and Y. Shiloach. The complexity of finding maximum disjoint paths with length constraints. *Networks*, 12(3):277–286, 1982.

20 R. Kim, S. Kim, J. Ma, and B. Park. Cycles with two blocks in k-chromatic digraphs. *Journal of Graph Theory*, 88(4):592–605, 2018.

21 J. Kneis, D. Mölle, S. Richter, and P. Rossmanith. Divide-and-color. In *Proc. of the 32nd International Workshop on Graph-Theoretic Concepts in Computer Science (WG)*, volume 4271 of *LNCS*, pages 58–67, 2006.

22 M. Kriesell. Disjoint A-paths in digraphs. *Journal of Combinatorial Theory, Series B*, 95(1):168–2005, 2005.

23 A. Metzlar. Disjoint paths in acyclic digraphs. *Journal of Combinatorial Theory, Series B*, 57(2):228–238, 1993.

24 B. Monien. How to find long paths efficiently. *Annals of Discrete Mathematics*, 25:239–254, 1985.

25 R. Niedermeier. *Invitation to Fixed-Parameter Algorithms.* Oxford University Press, 2006.

26 J. G. Oxley. *Matroid theory.* Oxford University Press, 1992.

27 A. Schrijver. A Short Proof of Mader's $\mathcal{S}$-Paths Theorem. *Journal of Combinatorial Theory, Series B*, 82(2):319–321, 2001.

28 R. Sedgewick and K. Wayne. *Algorithms (4th Edition).* Addison-Wesley, 2011.

29 H. Shachnai and M. Zehavi. Representative families: A unified tradeoff-based approach. *Journal of Computer and System Sciences*, 82(3):488–502, 2016.

30 A. Slivkins. Parameterized tractability of edge-disjoint paths on directed acyclic graphs. *SIAM Journal on Discrete Mathematics*, 24(1):146–157, 2010.

31 G. Venkatesan and C. Pandu Rangan. Approximate triclique coloring for register allocation. *Information Processing Letters*, 60:249–253, 1996.

32 M. Zehavi. A randomized algorithm for long directed cycle. *Information Processing Letters*, 116(6):419–422, 2016.