



International Association  
of Applied Mathematics and Mechanics  
– Archive for Students –



# A method for finding all connections in the Pantelides algorithm for delay differential-algebraic equations

Daniel Collin<sup>a,\*</sup> 

<sup>a</sup> Institute of Mathematics, Technische Universität Berlin, Berlin, Germany

received 23.01.2022, accepted 10.06.2022, published 11.07.2022

---

\* corresponding author: daniel.collin@web.de

supervisor: Ines Ahrens, Technische Universität Berlin, Berlin, Germany,

Benjamin Unger, Stuttgart Center for Simulation Science, Universität Stuttgart, Stuttgart, Germany, and

Volker Mehrmann, Technische Universität Berlin, Berlin, Germany,

---

**Abstract:** *The Pantelides algorithm for systems of delay differential-algebraic equations (DDAEs) is a method to structurally analyse such systems with the goal to detect which equations have to be differentiated or shifted to construct a solution. In this process, one has to detect implicit connections between equations in the shifting graph, making it necessary to check all possible connections. The problem of finding these efficiently remained unsolved so far. It is explored in further detail and a reformulation is introduced. Additionally, an algorithmic approach for its solution is presented.*

**Keywords:** delay differential-algebraic equation, Pantelides algorithm, structural analysis, enumeration algorithm, spanning tree

## 1 Introduction

Delay differential-algebraic equations (DDAEs) are a class of differential equations for a function  $x: \mathbb{R} \rightarrow \mathbb{R}^n$  on a time interval  $[0, T] \subseteq \mathbb{R}$ ,  $T > 0$ , that, in their simplest form, not only depend on the time derivative  $\dot{x}(t)$  but also on a previous time state  $\Delta_{-\tau}x(t) := x(t - \tau)$  with a delay  $\tau > 0$ . Additionally, the system may possess

algebraic constraints such that it can only be formulated in implicit form. Here, the DDAE is assumed to be a system of  $n \in \mathbb{N}$  equations  $F = (F_1, \dots, F_n)$  and variables  $x = (x_1, \dots, x_n)$ . Thus, consider a DDAE of the form

$$F(t, x(t), \dot{x}(t), \Delta_{-\tau}x(t)) = 0, \quad (1)$$

where

$$x: [-\tau, T] \rightarrow \mathbb{R}^n \quad \text{and} \quad F: [0, T] \times \mathbb{D}_x \times \mathbb{D}_{\dot{x}} \times \mathbb{D}_{\Delta x} \rightarrow \mathbb{R}^n,$$

with  $\mathbb{D}_x, \mathbb{D}_{\dot{x}}, \mathbb{D}_{\Delta x} \subseteq \mathbb{R}^n$  open. Here,  $\dot{x}$  denotes the derivative of  $x$  with respect to  $t$  from the right. To obtain an initial value problem, (1) has to be equipped with a consistent initial condition

$$x(t) = \phi(t) \quad \text{for} \quad t \in [-\tau, 0]. \quad (2)$$

Equations of this form arise in many applications, such as multibody control systems, electric circuits or fluid dynamics (see [7, 18]). They combine features of delay differential equations (DDEs) and differential-algebraic equations (DAEs), which makes them particularly difficult to solve. Solutions may depend on derivatives of  $F$  and on evaluations of  $F$  at future time points

(see [5, 9]). Therefore, the interplay of the differentiation operator  $\frac{d}{dt}$  and the shift operator  $\Delta_{-\tau}$  has to be treated carefully (cf. [11]) and solutions have to be constructed by differentiating and shifting equations (cf. [5, 9, 10, 11, 15, 16]). Even for linear DDAEs, general existence and uniqueness results can only be obtained using a distributional solution concept (see [15, 18]) or imposing further restrictions on the DDAE (see [9, 10, 16]) or its initial function in eq. (2) (see [8, 16]). For nonlinear DDAEs, the existence of solutions can be established for certain classes (cf. [2, 17]).

In most cases, the construction of a solution for a DDAE involves the method of steps. This means that the equation is successively integrated over the time intervals  $[i\tau, (i+1)\tau)$ ,  $i \in \mathbb{N}$ . By substituting the delayed variables with the already computed solution of the previous interval, the problem can be reduced to solve a DAE in each step (cf. [3, 4, 10]). However, this method does not always succeed and a reformulation of eq. (1) is required such that the DAE that has to be solved in each interval is regular and has a small index. Hereby, the index is, roughly speaking, a measure of how often parts of the DAE have to be differentiated to reformulate the DAE as an ordinary differential equation. This reformulation can be done by a compress-and-shift algorithm (see [5, 15]) or by a combined shift and derivative array (see [10]). The differentiation and shifting of certain equations is necessary in both cases.

Determining which equations one has to differentiate or shift is therefore a central aspect of the solution process of a DDAE. In [1], the Pantelides algorithm for DDAEs is presented as a tool to exploit the structure of (1), i.e., the information about which variable appears in which equation, to determine the number of differentiations and shifts necessary to solve the DDAE. It is based on the Pantelides algorithm for DAEs (see [14]) and will be simply referred to as Pantelides algorithm from now on. The approach consists of defining different bipartite graphs, where each equation and certain equivalence classes of variables are represented by the nodes. Edges exist between equation nodes and variable nodes if and only if one of the variables of the equivalence class appears in that equation. In other words, the graphs represent the structure of the DDAE. Then, matchings between equation nodes and variable nodes of highest shift and differentiation order are constructed in these graphs. This is achieved by following a specific pattern of shifting and differentiating equations and the variables belonging to it. At the end of the process, each equation can be resolved for a variable of highest shift and differentiation order. For more details on the whole procedure and the algorithm, see the original paper [1].

In this work, the focus is put on a specific subproblem that appears in the Pantelides algorithm. During the first part of the algorithm, called the shifting step, one has to shift equations that are connected to each other in a certain way through edges in a specific graph, called the shifting graph. Generally, such a connection is not unique and all possible connections have to be found to allow for correct shifting. The identification of all of these connections, however, may be computationally very expensive, and no efficient algorithmic solution is known so far.

The contribution of this paper is a deeper exploration of the problem of finding all connections. First, the problem is explained in further detail and all important preliminaries are given in Section 2. Then, a new solution approach is proposed based on the reformulation to a known enumeration problem from graph theory. The equivalence of both problems is proven (Section 3). Additionally, an algorithm from [6] for the solution of the enumeration problem is presented (Section 4). This algorithm is applied to the original problem, yielding a method for finding all connections in the Pantelides algorithm, and two detailed examples of its usage are given (Section 5). Finally, a numerical demonstration of the advantageous properties of the new method is shown (Section 6) and the paper is concluded with a summary and some final remarks (Section 7).

### Notation:

The natural numbers, the non-negative integers, and the reals are denoted by  $\mathbb{N}$ ,  $\mathbb{N}_0$  and  $\mathbb{R}$ , respectively. For a differentiable function  $f: \mathbb{I} \rightarrow \mathbb{R}^n$ , the notation  $\dot{f} := \frac{d}{dt}f$  is used to denote the derivative with respect to the (time) variable  $t$  and  $\ddot{f} := \frac{d}{dt}\dot{f}$  for the second derivative. For higher derivatives of order  $q \in \mathbb{N}_0$ , the abbreviation  $f^{(q)}$  is used. Similarly, the shift operator  $\Delta_\tau$  is defined as  $\Delta_\tau f(t) = f(t + \tau)$ . The union of two sets  $A$  and  $B$  is denoted by  $A \cup B$ . A disjoint union of sets is written as  $A \dot{\cup} B$ . The cardinality of the set  $A$  is denoted by  $|A|$ .

## 2 Problem description

This section describes the overall problem of the paper in detail and introduces the most important definitions to give all preliminaries needed to understand the solution approach. Since this paper can be seen as an extension of [1], most information of this section is based on that work and all derivations can be found there. In comparison, new concepts (which are not used in [1]) are introduced in Sections 3 and 4 and include Definition 5, Definition 6, Definition 9, and Definition 14, as well as results based on those.

The Pantelides algorithm translates the structural information of the DDAE into graphs. First, the *shifting graph*  $G^S$  is constructed by combining all variables of the same index  $k$  (for each  $k = 1, \dots, n$ ) and shift order  $p \in \mathbb{N}_0 \cup \{-1\}$  (but possibly different differentiation order) into the same equivalence class, i.e.,

$$[\Delta_{p\tau} x_k] := \{\Delta_{p\tau} x_k, \Delta_{p\tau} \dot{x}_k, \Delta_{p\tau} \ddot{x}_k, \dots\}.$$

Then, one can define the set of equation nodes, variable nodes, and edges as

$$V_E^S := \{F_1, \dots, F_n\},$$

$$V_V^S := \left\{ [\Delta_{p\tau} x_k] \mid \exists k, p \in \mathbb{N}_0 \cup \{-1\} \text{ s.t. } \Delta_{p\tau} x_k^{(q)} \text{ appears in DDAE for any } q \in \mathbb{N}_0 \right\},$$

$$E^S := \left\{ \{F_i, v_k\} \in V_E^S \times V_V^S \mid \exists \tilde{x} \in v_k \text{ that appears in } F_i \right\},$$

respectively, which yields the shifting graph defined as  $G^S := (V_E^S \cup V_V^S, E^S)$ . Note that although writing the equivalence classes in the form  $[\Delta_{p\tau} x_k]$  makes the notation more compact, the members of the equivalence classes will be given explicitly in set form to enhance comprehension for the reader in the examples throughout this paper.

**Example 1.** Consider the DDAE from [1, Example 3.11, p.17]:

$$\begin{aligned} \dot{x}_1 &= f_1, \\ \dot{x}_1 &= x_2 + f_2, \\ 0 &= x_1 + x_2 + \Delta_{-\tau} x_3 + f_3. \end{aligned} \quad (3)$$

To construct the shifting graph, one defines the equivalence classes

$$\begin{aligned} [x_1] &:= \{x_1, \dot{x}_1\}, \\ [x_2] &:= \{x_2\}, \\ [\Delta_{-\tau} x_3] &:= \{\Delta_{-\tau} x_3\}, \end{aligned}$$

and the sets of equation nodes, variable nodes, and edges as

$$\begin{aligned} V_E^S &:= \{F_1, F_2, F_3\}, \\ V_V^S &:= \{[x_1, \dot{x}_1], [x_2], [\Delta_{-\tau} x_3]\}, \\ E^S &:= \{\{F_1, [x_1, \dot{x}_1]\}, \{F_2, [x_1, \dot{x}_1]\}, \{F_2, [x_2]\}, \\ &\quad \{F_3, [x_1, \dot{x}_1]\}, \{F_3, [x_2]\}, \{F_3, [\Delta_{-\tau} x_3]\}\}. \end{aligned}$$

Then, the shifting graph is given as  $G^S := (V_E^S \cup V_V^S, E^S)$ . In Figure 1, an illustration of the graph can be found.

In the bipartite shifting graph, one successively assigns to each equation node  $F_i \in V_E^S$  an equivalence class  $v_k \in V_V^S$  of highest shift, i.e., if  $v_k = [\Delta_{p\tau} x_k]$  is of

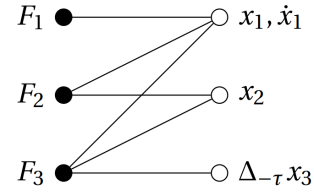


Figure 1 – The shifting graph of eq. (3).

highest shift and occurs in  $F_i$ , then  $[\Delta_{(p+\ell)\tau} x_k]$ , for  $\ell > 0$ , does not occur in any equation. By definition, a variable node  $[\Delta_{p\tau} x_k]$  with negative shift  $p = -1$  is never of highest shift and cannot be matched to an equation node. Like that, a matching  $\mathcal{M}$  is constructed, consisting of all assigned pairs  $\{F_i, v_k\}$ .

**Example 2.** Continuing with Example 1, one can assign the equivalence classes  $[x_1] = \{x_1, \dot{x}_1\}$  to  $F_1$  and  $[x_2] = \{x_2\}$  to  $F_2$  in the shifting step. Thus, this yields the matching

$$\mathcal{M} = \{\{F_1, [x_1, \dot{x}_1]\}, \{F_2, [x_2]\}\}.$$

Figure 2a illustrates  $\mathcal{M}$  by coloring the matching edges in blue.

If a particular  $F_j$  cannot be matched to a variable node that is not in  $\mathcal{M}$  yet, the node  $F_j$  is called *exposed with respect to  $\mathcal{M}$* . The corresponding equation is shifted, together with all other equations that  $F_j$  is connected to via alternating paths with respect to  $\mathcal{M}$  in  $G^S$ . An *alternating path with respect to  $\mathcal{M}$*  is a sequence of edges

$$\left( \{F_{i_1}, v_{k_1}\}, \{v_{k_1}, F_{i_2}\}, \{F_{i_2}, v_{k_2}\}, \dots, \{v_{k_{N-1}}, F_{i_N}\} \right)$$

in  $G^S$ , where all  $i_\ell$ ,  $\ell = 1, \dots, N$ , and all  $k_m$ ,  $m = 1, \dots, N-1$ , are distinct, respectively, and that has alternating non-matching and matching edges while starting with a non-matching edge.

**Example 3.** Consider again Example 1 and the matching

$$\mathcal{M} = \{\{F_1, [x_1, \dot{x}_1]\}, \{F_2, [x_2]\}\},$$

illustrated in Figure 2a. Equation  $F_3$  is exposed and cannot be matched directly to any equivalence class, but it is connected via alternating paths to the other equation nodes. There are several possibilities to connect the equation nodes via alternating paths, e.g., via the path

$$(\{F_3, [x_2]\}, \{[x_2], F_2\}, \{F_2, [x_1, \dot{x}_1]\}, \{[x_1, \dot{x}_1], F_1\}),$$

or via the paths

$$(\{F_3, [x_1, \dot{x}_1]\}, \{[x_1, \dot{x}_1], F_1\}) \quad \text{and} \quad (\{F_3, [x_2]\}, \{[x_2], F_2\}).$$

Here, all other equation nodes are connected to  $F_3$  via alternating paths, and thus, all equations are shifted. However, there are also scenarios where only a subset of equation nodes is connected.

Simply shifting all connected equations may not be sufficient, since the connection may be given only implicitly through the equivalence classes of the variable nodes. To see this, define  $G := (V_E \dot{\cup} V_V, E)$  as the *graph of the DDAE* with

$$\begin{aligned} V_E &:= \{F_1, \dots, F_n\}, \\ V_V &:= \{\Delta_{p\tau} x_k^{(q)} \mid \exists k, p, q \in \mathbb{N}_0 \cup \{-1\} \text{ s.t.} \\ &\quad \Delta_{p\tau} x_k^{(q)} \text{ appears in DDAE}\}, \\ E &:= \{\{F_i, v_k\} \in V_E \times V_V \mid v_k \text{ appears in } F_i\}. \end{aligned}$$

In other words, the graph of the DDAE contains all variables explicitly as distinct nodes without using equivalence classes. An implicit connection in the shifting graph means that the involved equations contain variables with the same shift but a different differentiation order. Thus, they belong to the same variable node in the shifting graph but not to the same node in the graph of the DDAE. There is an alternating path connecting the exposed equation  $F_j$  and the equation that has to be shifted in  $G^S$  but not in  $G$ . In this case, an explicit connection has to be established by differentiating the involved equations that do not depend on the highest derivative in the equivalence class. To ensure that all implicit connections are resolved, all possible connections have to be identified and checked.

**Example 4.** The graph of the DDAE of eq. (3) is given as  $G := (V_E \dot{\cup} V_V, E)$  with

$$\begin{aligned} V_E &:= \{F_1, F_2, F_3\}, \\ V_V &:= \{x_1, \dot{x}_1, x_2, \Delta_{-\tau} x_3\}, \\ E &:= \{\{F_1, \dot{x}_1\}, \{F_2, \dot{x}_1\}, \{F_2, x_2\}, \\ &\quad \{F_3, x_1\}, \{F_3, x_2\}, \{F_3, \Delta_{-\tau} x_3\}\}. \end{aligned}$$

A visualization is shown in Figure 2d.

Since  $F_3$  is exposed with respect to the matching

$$\mathcal{M} = \{\{F_1, \{x_1, \dot{x}_1\}\}, \{F_2, \{x_2\}\}\},$$

in the shifting graph (Figure 2a), all possible connections for  $F_3$  with respect to  $\mathcal{M}$  have to be found. These are given by the alternating paths presented in Example 3, i.e.,

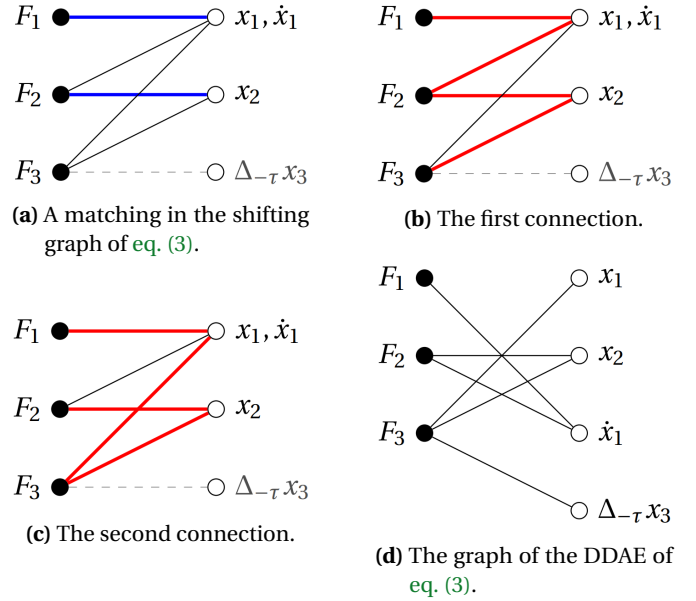
$$\mathcal{C}_1 = \{(\{F_3, \{x_2\}\}, \{\{x_2\}, F_2\}), (\{F_2, \{x_1, \dot{x}_1\}\}, \{\{x_1, \dot{x}_1\}, F_1\})\}$$

(pictured red in Figure 2b) and

$$\mathcal{C}_2 = \{(\{F_3, \{x_1, \dot{x}_1\}\}, \{\{x_1, \dot{x}_1\}, F_1\}), (\{F_3, \{x_2\}\}, \{\{x_2\}, F_2\})\}$$

(pictured red in Figure 2c). One can see that connection  $\mathcal{C}_1$  does also exist in  $G$  via the path

$$(\{F_3, x_2\}, \{x_2, F_2\}, \{F_2, \dot{x}_1\}, \{\dot{x}_1, F_1\})$$



**Figure 2** – Visualization of the problem of finding all connections for  $F_3$  with respect to  $\mathcal{M}$  in the shifting step of the DDAE eq. (3) (figure taken from [1, Figure 5, p.18]).

and hence, is an explicit connection. However,  $\mathcal{C}_2$  is implicit, as the alternating path  $(\{F_3, \{x_1, \dot{x}_1\}\}, \{\{x_1, \dot{x}_1\}, F_1\})$  does not connect  $F_3$  and  $F_1$  in  $G$ . One has to differentiate  $F_3$  to establish an explicit connection. It can be seen that checking one connection is not enough, all have to be identified to resolve possible implicit connections in the shifting graph.

Theoretically, one could identify all connections by just checking all possible combinations of edges of  $E^S$  that yield alternating paths. In practice, however, this approach is not feasible, because the amount of combinations increases rapidly with the number of nodes and edges of the graph  $G^S$ . By reformulating the problem, it can be solved much more efficiently.

### 3 Reformulation of problem

First, a connection has to be technically defined. To simplify the notation, let  $G = (V_E \dot{\cup} V_V, E)$  be a bipartite graph with equation nodes  $V_E$  and variable nodes  $V_V$ , where  $E$  contains only edges between  $V_E$  and  $V_V$ , not between nodes of one set. Further, let a matching

$$\mathcal{M} = \left\{ \{F_{i_1}, v_{k_1}\}, \dots, \{F_{i_M}, v_{k_M}\} \right\} \in E^M$$

be given in  $G$  with  $M < n$  and  $F_j \in V_E$  as an exposed equation node with respect to  $\mathcal{M}$ . Then,

$$C_{F_j} := \{ F_k \in V_E \mid \exists \text{ alternating path between } F_j \text{ and } F_k \text{ in } G \}$$



denotes the equation nodes that are connected to  $F_j$  via an alternating path. This set is automatically generated by the algorithm Augmentpath (see [1, Algorithm 1, p.10], [14, Algorithm 3.2, p.217]).

**Definition 5.** Let  $G = (V_E \dot{\cup} V_V, E)$  be a bipartite graph and  $\mathcal{M}$  a matching in  $G$ . Further, let  $F_j \in V_E$  be exposed with respect to  $\mathcal{M}$  and  $C_{F_j}$  as defined above. A *connection for  $F_j$  with respect to  $\mathcal{M}$*  is defined as a set of connected alternating paths  $(F_i, v_k, F_\ell) \in V_E \times V_V \times V_E$ , with  $\{F_i, v_k\} \in E \setminus \mathcal{M}$  and  $\{v_k, F_\ell\} \in \mathcal{M}$ . Additionally, it has to hold that for all  $F_\ell \in C_{F_j}$  the corresponding matching edge  $\{v_k, F_\ell\} \in \mathcal{M}$  occurs exactly once and there is at least one alternating path starting in  $F_j$ .

Note that the definition of a connection for  $F_j$  with respect to  $\mathcal{M}$  has been changed in comparison to the definition from [1, p.18]. That is reasonable because the previous definition is not sufficient and allows sets of alternating paths that contain cycles and not necessarily the exposed node  $F_j$ , two aspects that do not make sense in this context. Therefore, the definition was expanded to exclude these cases. In the forthcoming Corollary 13, it is shown that a connection in the sense of Definition 5 is indeed cycle-free. Also, a connection for  $F_j$  with respect to  $\mathcal{M}$  will simply be referred to as a connection when it is clear which node is exposed and which matching the connection is based on.

With the exact definition of a connection, one can further define the connection graph by interpreting the alternating paths from this definition as directed edges between the equation nodes.

**Definition 6.** Let  $G = (V_E \dot{\cup} V_V, E)$  be a bipartite graph and  $\mathcal{M}$  a matching in  $G$ . Further, let  $F_j \in V_E$  be exposed with respect to  $\mathcal{M}$  and  $C_{F_j}$  as defined above. Define the set of nodes  $V_H := C_{F_j} \dot{\cup} \{F_j\}$  and directed edges

$$E_H := \{ (F_i, F_\ell) \in V_H \times V_H \mid (F_i, v_k, F_\ell) \text{ is an alternating path with } \{F_i, v_k\} \in E \setminus \mathcal{M}, \{v_k, F_\ell\} \in \mathcal{M} \}.$$

Then, the directed graph  $H := (V_H, E_H)$  is called *connection graph for  $F_j$  with respect to  $\mathcal{M}$* .

**Remark 7.** Denoting the alternating path  $(F_i, v_k, F_\ell)$  as  $(F_i, F_\ell)$ , it might seem as if information is lost about which variable node  $v_k$  connects the equation nodes  $F_i$  and  $F_\ell$ . However, since each  $F_\ell$  is uniquely matched to one  $v_k$  in  $\mathcal{M}$ , the variable node can easily be reconstructed from the directed edge  $(F_i, F_\ell)$  using  $\mathcal{M}$ . A second approach to not lose information is to define edge weights  $w_{i\ell} = k$  for the edges  $(F_i, F_\ell)$ , i.e., if  $v_k$  is to be reconstructed from  $(F_i, F_\ell)$ , it holds that  $v_k = v_{w_{i\ell}}$ .

Similar to before, the connection graph for  $F_j$  with respect to  $\mathcal{M}$  will be referred to simply as connection graph when it is clear which node is exposed and which matching the connection graph is based on.

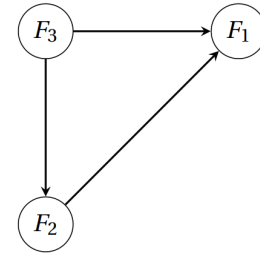
**Example 8.** Consider again the DDAE eq. (3) from Example 1 and the shifting graph from Figure 2a. Based on the matching

$$\mathcal{M} = \{\{F_1, \{x_1, \dot{x}_1\}\}, \{F_2, \{x_2\}\}\}$$

and  $C_{F_3} = \{F_1, F_2\}$ , the connection graph for  $F_3$  with respect to  $\mathcal{M}$  can be defined according to Definition 6 as  $H = (V_H, E_H)$  with

$$V_H = \{F_1, F_2, F_3\}, \\ E_H = \{(F_2, F_1), (F_3, F_1), (F_3, F_2)\}.$$

A picture of  $H$  can be seen in Figure 3.



**Figure 3** – The connection graph for  $F_3$  with respect to  $\mathcal{M}$  of the DDAE eq. (3).

The connection graph facilitates to reformulate the problem of finding all connections in the sense of Definition 5 by transferring it from the shifting graph to the connection graph. It translates to finding all spanning trees (defined below) with root  $F_j$  in  $H$ . To prove this, some definitions and lemmas from graph theory are needed (see [12, p.71-73] for reference and proofs; note that the term arborescence in [12] has been changed to spanning tree for consistency with the rest of the paper).

**Definition 9.** [12, Definition 6.15, p.71, Definition 6.20, 6.22, p.73]

- i) An undirected graph is called *forest* if it contains no cycles.
- ii) An undirected graph is called *tree* if it is a forest and connected.
- iii) Given a directed graph  $G$ , one can replace every directed edge by an undirected edge to get an undirected graph. The arising graph is called the *underlying undirected graph* of  $G$ .
- iv) A directed graph is called *branching* if its underlying undirected graph is a forest and every node has at most one edge ending in it.

- v) A directed graph is called *spanning tree* if it is a connected branching.

**Lemma 10.** [12, Theorem 6.18, p.72] *Let  $G$  be an undirected graph with  $n$  vertices. Then,  $G$  is a tree if and only if  $G$  has  $n - 1$  edges and is connected.*

The underlying undirected graph of a spanning tree has to be a connected forest, i.e., a tree. According to Lemma 10, a spanning tree with  $n$  nodes thus has  $n - 1$  edges and, according to Definition 9, every node has at most one edge ending in it. Therefore, there is exactly one node  $r$  with no incoming edge. Let  $\delta^-(v)$  be the set of incoming edges of a vertex  $v$  of  $G$ . Then, this condition can be formulated as  $\delta^-(r) = \emptyset$ . In this case,  $r$  is called *root* of the spanning tree. For any edge  $(u, v)$  of a spanning tree,  $v$  is called a *child* of  $u$  and  $u$  the *predecessor* of  $v$ . Vertices with no children are called *leaves* (see [12, p.73]).

**Lemma 11.** [12, Theorem 6.23, p.73] *Let  $G$  be a directed graph with  $n$  vertices and  $r$  a vertex of  $G$ . Then, the following statements are equivalent:*

- i)  $G$  is a spanning tree with root  $r$ .
- ii)  $G$  is a branching with  $n - 1$  edges and  $\delta^-(r) = \emptyset$ .
- iii)  $\delta^-(r) = \emptyset$  and there exists a uniquely determined directed path from  $r$  to every vertex in  $G$ .

**Theorem 12.** *Let  $G = (V_E \dot{\cup} V_V, E)$  be a bipartite graph,  $\mathcal{M}$  a matching in  $G$ ,  $F_j \in V_E$  exposed with respect to  $\mathcal{M}$ , and  $H = (V_H, E_H)$  the connection graph for  $F_j$  with respect to  $\mathcal{M}$ . Then, the following statements are equivalent:*

- i) The set

$$\mathcal{C} := \left\{ (F_{i_1}, v_{k_1}, F_{\ell_1}), \dots, (F_{i_M}, v_{k_M}, F_{\ell_M}) \right\}$$

*is a connection for  $F_j$  with respect to  $\mathcal{M}$ .*

- ii) The directed subgraph

$$\mathcal{H} := \left( V_H, \left\{ (F_{i_1}, F_{\ell_1}), \dots, (F_{i_M}, F_{\ell_M}) \right\} \right) \subseteq H$$

*is a spanning tree with root  $F_j$ .*

*Proof.* i)  $\Rightarrow$  ii): Let  $F_j \in V_E$  be exposed, denote a connection for  $F_j$  as  $\mathcal{C} = \{(F_{i_1}, v_{k_1}, F_{\ell_1}), \dots, (F_{i_M}, v_{k_M}, F_{\ell_M})\}$ , and let  $H$  be the connection graph for  $F_j$ , with respect to  $\mathcal{M}$ , respectively. Additionally, denote

$$\mathcal{H} = (V_H, E_{\mathcal{H}}), \quad E_{\mathcal{H}} := \left\{ (F_{i_1}, F_{\ell_1}), \dots, (F_{i_M}, F_{\ell_M}) \right\} \subseteq E_H,$$

as a subgraph of  $H$ , and

$$\mathcal{H}_u := (V_H, E_u), \quad E_u := \left\{ \{F_{i_1}, F_{\ell_1}\}, \dots, \{F_{i_M}, F_{\ell_M}\} \right\},$$

as the underlying undirected graph of  $\mathcal{H}$ . According to Definition 6, it holds that  $|V_H| = |C_{F_j}| + 1$  and therefore,

$$M = |\mathcal{C}| = |C_{F_j}| = |V_H| - 1.$$

It follows that  $\mathcal{H}_u$  has  $|V_H| - 1$  edges. According to Definition 5, the alternating paths  $(F_i, v_k, F_{\ell}) \in \mathcal{C}$  are connected and each  $F_{\ell} \in C_{F_j}$  occurs exactly once. The variable nodes  $v_k$  are uniquely determined by the equation nodes  $F_{\ell}$  via the matching  $\mathcal{M}$ . This implies that each  $v_{k_m}$ , for  $m = 1, \dots, M$ , also occurs only once in  $\mathcal{C}$  and the alternating paths must be connected via the equation nodes  $F_i$  and  $F_{\ell}$ . Thus, the edges in  $E_u$  are connected and thereby,  $\mathcal{H}_u$  is connected as well.

In summary,  $\mathcal{H}_u$  has  $|V_H| - 1$  edges and is connected. Therefore, according to Lemma 10, the underlying undirected graph of  $\mathcal{H}$  is a tree (and also a forest). Additionally,  $V_H = C_{F_j} \dot{\cup} \{F_j\}$  and it was already mentioned that each  $F_{\ell} \in C_{F_j}$  occurs exactly once as second equation node in the alternating paths  $(F_i, v_k, F_{\ell}) \in \mathcal{C}$ . The node  $F_j$  itself has no alternating path leading to it because it is exposed. Therefore, every node of  $\mathcal{H}$  has at most one edge ending in it. According to Definition 9, the graph  $\mathcal{H}$  is a branching and, as mentioned above, it has  $|V_H| - 1$  edges.

Finally, one can choose the exposed node  $F_j$  as the root of  $\mathcal{H}$  because  $\delta^-(F_j) = \emptyset$ . By Lemma 11, it follows that  $\mathcal{H}$  is a spanning tree with root  $F_j$ .

ii)  $\Rightarrow$  i): Let  $\mathcal{H} = (V_H, \{(F_{i_1}, F_{\ell_1}), \dots, (F_{i_M}, F_{\ell_M})\}) \subseteq H$  be a spanning tree with root  $F_j$ , and denote

$$\mathcal{C} = \left\{ (F_{i_1}, v_{k_1}, F_{\ell_1}), \dots, (F_{i_M}, v_{k_M}, F_{\ell_M}) \right\}.$$

Since  $\mathcal{H}$  is a subgraph of the connection graph for  $F_j$  with respect to  $\mathcal{M}$ , it follows that all  $(F_i, v_k, F_{\ell}) \in \mathcal{C}$  are alternating paths with  $\{F_i, v_k\} \in E \setminus \mathcal{M}$ ,  $\{v_k, F_{\ell}\} \in \mathcal{M}$ , where  $v_k$  is uniquely determined by the matching  $\mathcal{M}$  (see Definition 6 and Remark 7). From Lemma 11, it is known that the root  $F_j$  of the spanning tree is a vertex of  $\mathcal{H}$  with  $\delta^-(F_j) = \emptyset$  and, therefore, is also included as the starting node in an alternating path from  $\mathcal{C}$ .

Additionally, there exists a uniquely determined directed path from the root  $F_j$  to every vertex in  $\mathcal{H}$ . Thus, the same property applies to the alternating paths in  $\mathcal{C}$ , which yields that they are connected and each  $F_{\ell} \in C_{F_j}$  occurs exactly once as final node of an alternating path. Hence,  $\mathcal{C}$  fulfills all properties of a connection for  $F_j$  with respect to  $\mathcal{M}$ .  $\square$

**Corollary 13.** *All connections for  $F_j$  with respect to  $\mathcal{M}$  are cycle-free.*

*Proof.* The proof follows directly from Theorem 12, using the equivalence of a connection for  $F_j$  with respect to  $\mathcal{M}$

to a spanning tree in the connection graph for  $F_j$  with respect to  $\mathcal{M}$ . Spanning trees are by [Definition 9](#) cycle-free.  $\square$

In the literature (e.g., [6, 12]), a spanning tree of a directed graph is also referred to as an *arborescence*. The sort of problem where all possible solutions to a computational problem have to be computed and explicitly returned as an output is called *enumeration problem*. Methods for solving these problems are called *enumeration algorithms*.

With [Theorem 12](#), a reformulation of the initial problem (finding all connections) has been derived by showing that it is equivalent to the problem of enumerating all spanning trees in the corresponding connection graph. Each spanning tree can then be interpreted as a connection. There are efficient algorithms to solve this enumeration problem, one of which is discussed in the next section.

## 4 Enumeration of spanning trees

One enumeration algorithm for finding all spanning trees of a directed graph was published in [6]. It turns out to be an effective method for the purpose of this work and is therefore implemented in the (overall) PanTelides algorithm to solve the subproblem of finding all connections. In this section, a short summary is given of how this algorithm works. For further details of the implementation as well as theoretical results and their proofs, see [6].

First, the important concept of so-called bridges has to be introduced.

**Definition 14.** [6, p.280] Let  $G = (V, E)$  be a directed graph and  $r \in V$  a vertex.

- i)  $G$  is called *rooted at  $r$*  if there exists a spanning tree with root  $r$  in  $G$ .
- ii) An edge  $e \in E$  is called a *bridge for  $r$*  if  $G$  is rooted at  $r$  but  $G \setminus \{e\}$  is not rooted at  $r$ .
- iii) Equivalently, an edge  $e \in E$  is a *bridge for  $r$*  if it is part of every spanning tree rooted at  $r$  in  $G$ .

Assume a directed graph  $G = (V, E)$  and a root vertex  $r$  are given and all spanning trees of  $G$  rooted at  $r$  have to be computed. This goal is accomplished by finding all spanning trees containing different subtrees  $T \subseteq G$ , also rooted at  $r$ .

Given a subtree  $T$ , the approach consists of successively adding edges to  $T$  in the following way: A new edge  $e_i := (u, v) \in E$ , directed from a vertex  $u \in T$  to a vertex  $v \notin T$ , is added to  $T$ , and all spanning trees containing  $T \cup \{e_i\}$  are computed. When this is done, the edge  $e_i$

is deleted from  $G$  and  $T$  and another edge  $e_j \in E \setminus \{e_i\}$  (directed from  $T$  to a vertex not in  $T$ ), is added to  $T$ . Again, all spanning trees containing  $T \cup \{e_j\}$  are computed, then  $e_j$  is deleted from  $G$  and  $T$ . The same process continues with the next edge and is repeated until an edge is processed that is a bridge for  $r$  in the modified graph  $G \setminus \{e_i, e_j, \dots\}$ . Each spanning tree containing  $T$  has now been found exactly once.

A key point in this approach is to discover efficiently if an edge  $e$  is a bridge. Assume all spanning trees containing  $T \cup \{e\}$  have been computed and let  $L$  be the last found spanning tree. It has to be checked if  $e := (u, v)$  is a bridge.

There are several possibilities. The idea that is pursued in this algorithm is to consider the descendants and nondescendants of  $v$  in  $L$ . *Descendants of  $v$  in  $L$*  are vertices that can be reached following a directed path starting in  $v$  and using only edges of the spanning tree  $L$ . Contrary, *nondescendants of  $v$  in  $L$*  are vertices for which there cannot be constructed such a path using edges from  $L$ .

Clearly, if there is an edge in  $G \setminus \{e\}$  that goes from a nondescendant of  $v$  (in  $L$ ) to  $v$ , then  $e$  cannot be a bridge, since one could delete  $e$  and replace it with that edge to construct another spanning tree. Thus,  $G \setminus \{e\}$  is still rooted at  $r$  and  $e$  could not have been a bridge. On the other hand, if no edge that goes from a nondescendant of  $v$  in  $L$  to  $v$  can be found,  $e$  must be a bridge, because deleting  $e$  leads to a graph  $G \setminus \{e\}$  where there does not exist a path to vertex  $v$  anymore.

For this to hold true, the way edges are added plays an important role. Here, the algorithm adds edges depth-first. The depth of a vertex contained in a tree is the length of the path between the vertex and the root of the tree it is contained in. Thus, adding an edge depth-first means that it is added to the vertex of  $T \cup \{e\}$  that has the greatest depth possible, i.e., that has the greatest distance from the root node. Particularly, this ensures that the last computed spanning tree that contains  $T \cup \{e\}$  (namely the tree  $L$ ) has the fewest descendants of  $v$  amongst all spanning trees containing  $T \cup \{e\}$ . This fact can be used to prove that this bridge test works correctly (see [6, Lemma 2, p.284] for more details).

Thus, it is important for the implementation to grow  $T$  depth-first. To do so, a stack  $F$  is used, where edges are stored that are directed from vertices in  $T$  to vertices not in  $T$ . Note that the action of removing an element from the top of a stack is referred to as *popping*, whereas the action of adding an element to the top of a stack is referred to as *pushing*. An edge  $e := (u, v)$  is always popped from the top of  $F$  if it is added to  $T$  and then, outgoing edges from  $v$  (that are not directed to vertices

already contained in  $T \cup \{e\}$ ) are pushed onto the top of  $F$ . Also, some edges might be removed from the inner part of  $F$  while growing  $T$ . This is necessary for all edges in  $F$  that are directed to  $v$ , the newest leaf of  $T$ . To ensure the depth-first property, these edges have to be restored at the exact same place in  $F$  after all spanning trees containing  $T \cup \{e\}$  have been found.

A second stack  $FF$  is used to store already processed edges since they are temporarily deleted from  $G$  but have to be restored later.

The full algorithm is stated in [Algorithm 1](#). Note that the pseudo code uses MATLAB notation for indexing, i.e., array indexing begins at 1 and the index "end" of an array points to the last element of it or to the top of a stack. The symbol ">" indicates a comment in the code. [Algorithm 2](#) illustrates how to initialize the method.

To conclude this section, the complexity of the algorithm is stated. For a directed graph  $G = (V, E)$  that has  $N$  spanning trees, the time complexity is  $O(|E|N)$  and the space or memory complexity is  $O(|E|)$  (see [6, Lemma 4, p.285]). Next, it is shown how to use this method in the Pantelides algorithm.

## 5 Algorithm for computation of connections

The enumeration algorithm from the previous section has to be applied to the initial problem of finding all connections. This would replace the computation stated in line 1 of Algorithm 3 from [1, p.20]. Thus, transferring the notation, the shifting graph  $G^S = (V_E^S \cup V_V^S, E^S)$ , the exposed equation  $F_j \in V_E^S$  and the matching  $\mathcal{M}$  are given. It has been shown that finding all connections for  $F_j$  with respect to  $\mathcal{M}$  is equivalent to enumerating the spanning trees with root  $F_j$  in the connection graph for  $F_j$  with respect to  $\mathcal{M}$ .

Therefore, the connection graph  $H$  is constructed according to [Definition 6](#) and used as input to [Algorithm 2](#), together with  $F_j$  as the root  $r$ . All spanning trees in  $H$  are returned. Given a spanning tree, one can reconstruct the corresponding connection by taking its edges  $(F_i, F_\ell)$  and inserting into each directed edge the variable node that was assigned to the equation node  $F_\ell$  by  $\mathcal{M}$ . That yields a set of alternating paths  $(F_i, v_k, F_\ell)$  as desired. The method is summarized in [Algorithm 3](#).

To illustrate the new method, a simple and a slightly more complex example are given in the following.

**Example 15.** Consider again the DDAE [eq. \(3\)](#) from [Example 1](#) with the shifting graph [Figure 2a](#). In [Example 8](#), it has been shown how to construct the connec-

---

### Algorithm 1 GROW

---

**Input:** directed graph  $G = (V, E)$ , directed subgraph  $T = (V_T, E_T) \subseteq G$ , stack of edges  $F$ , set of spanning trees  $S$

**Output:** set of spanning trees  $S$ , last computed spanning tree  $L$

---

```

1: if  $|V_T| = |V|$  then
2:    $L \leftarrow T$                                  $\triangleright$  store spanning tree in  $L$  and  $S$ 
3:    $S \leftarrow S \cup T$ 
4: else
5:    $FF \leftarrow [\emptyset]$ 
6:   while  $b = 0$  do
7:      $e \leftarrow F(\text{end})$   $\triangleright$  pop an edge  $e$  from  $F$ , add it to  $T$ 
8:      $v \leftarrow e(2)$ 
9:      $F \leftarrow \text{pop}(\{e\})$ 
10:     $T \leftarrow (V_T \cup \{v\}, E_T \cup \{e\})$ 
11:     $F \leftarrow F \setminus \{(u, w) \in E \mid u \in T, w = v\}$   $\triangleright$  update  $F$ 
12:     $F \leftarrow \text{push}(\{(u, w) \in E \mid u = v, w \notin T\})$ 
13:     $(S, L) \leftarrow \text{GROW}(G, T, F, S)$   $\triangleright$  recurse
14:     $F \leftarrow \text{pop}(\{(u, w) \in E \mid u = v, w \notin T\})$   $\triangleright$  restore  $F$ 
15:     $F \leftarrow F \cup \{(u, w) \in E \mid u \in T, w = v\}$   $\triangleright$  restore in same place as before
16:     $T \leftarrow (V_T \setminus \{v\}, E_T \setminus \{e\})$   $\triangleright$  delete  $e$  from  $T$  and  $G$ , add it to  $FF$ 
17:     $G \leftarrow (V, E \setminus \{e\})$ 
18:     $FF \leftarrow \text{push}(\{e\})$ 
19:    if  $\{(u, w) \in E \mid w = v, u \text{ is a nondescendant of } v \text{ in } L\} \neq \emptyset$  then
20:       $b \leftarrow 0$   $\triangleright$  bridge test
21:    else
22:       $b \leftarrow 1$ 
23:    end if
24:  end while
25:  while  $FF(\text{end}) \neq \emptyset$  do
26:     $e \leftarrow FF(\text{end})$   $\triangleright$  reconstruct  $G$ 
27:     $F \leftarrow \text{push}(\{e\})$ 
28:     $FF \leftarrow \text{pop}(\{e\})$ 
29:     $G \leftarrow (V, E \cup \{e\})$ 
30:  end while
31: end if

```

---



---

### Algorithm 2 Enumeration of spanning trees of a directed graph

---

**Input:** directed graph  $G = (V, E)$ , root node  $r \in V$

**Output:** set of all spanning trees  $S$

---

```

1:  $T \leftarrow (\{r\}, \emptyset)$ 
2:  $F \leftarrow \text{push}(\{(u, v) \in E \mid u = r\})$ 
3:  $S \leftarrow \emptyset$ 
4:  $S \leftarrow \text{GROW}(G, T, F, S)$ 

```

---



**Algorithm 3** Find all connections for  $F_j$  with respect to  $\mathcal{M}$

**Input:** shifting graph  $G^S = (V_E^S \cup V_V^S, E^S)$ , exposed node  $F_j \in V_E^S$ , matching  $\mathcal{M}$  stored in assign, colorE, colorV

**Output:** set of all connections  $P$

```

1:  $C_{F_j} \leftarrow \{F_k \in V_E^S \mid \exists \text{ alternating path between } F_j \text{ and } F_k \text{ in } G^S\}$ 
2:  $V_H \leftarrow C_{F_j} \cup \{F_j\}$ 
3:  $E_H \leftarrow \{(F_i, F_\ell) \in V_H \times V_H \mid (F_i, v_k, F_\ell) \text{ is an alternating path with } (v_k, F_\ell) \in \mathcal{M}, (F_i, v_k) \in E^S \setminus \mathcal{M}\}$ 
4:  $H \leftarrow (V_H, E_H) \triangleright \text{construct connection graph}$ 
5:  $P \leftarrow \text{Algorithm2}(H, F_j) \triangleright \text{enumeration algorithm}$ 
6: for all  $T = (V_T, E_T) \in P$  do
7:    $T \leftarrow E_T \triangleright \text{replace spanning trees by connections}$ 
8:   for all  $e = (F_i, F_\ell) \in T$  do
9:      $e \leftarrow (F_i, v_k, F_\ell) \text{ such that } (v_k, F_\ell) \in \mathcal{M}$ 
10:  end for
11: end for

```

tion graph for  $F_3$  with respect to the matching

$$\mathcal{M} = \{\{F_1, \{x_1, \dot{x}_1\}\}, \{F_2, \{x_2\}\}\}.$$

It is given as  $H = (V_H, E_H)$  with

$$\begin{aligned} V_H &= \{F_1, F_2, F_3\}, \\ E_H &= \{(F_2, F_1), (F_3, F_1), (F_3, F_2)\}, \end{aligned}$$

and is visualized in [Figure 3](#).

Hence, one defines  $G := H$  and  $r := F_3$  as the input to [Algorithm 2](#) and enumerates all spanning trees of  $G$  rooted at  $r$ . To initialize the process, set

$$\begin{aligned} T &= (\{F_3\}, \emptyset), \\ F &= [(F_3, F_2), (F_3, F_1)], \end{aligned}$$

and execute [Algorithm 1](#).

The recursion process of [Algorithm 1](#) can be visualized by the tree structure in [Figure 4](#). Note that the nodes of the computation tree will be called *bisections* and the edges *arrows* to not confuse them with the nodes and edges of  $T$  or  $G$ . In general, the notion of the tree is as follows: each bisection represents the current subgraph  $T \subseteq G$ , indicated by its edges  $E_T$ . As described in [Section 4](#), one then adds an edge  $e \in G$  from the stack  $F$  to  $T$  and computes all spanning trees containing  $T \cup \{e\}$ . Adding an edge  $e$  is represented by an arrow pointing away from a bisection, i.e., if

$$E_T = \{(F_{i_1}, F_{\ell_1}), \dots, (F_{i_m}, F_{\ell_m})\}$$

and  $e = (F_{i_{m+1}}, F_{\ell_{m+1}})$  is added, then this is visualized in the computation tree by an arrow pointing from

$$\begin{aligned} &\{(F_{i_1}, F_{\ell_1}), \dots, (F_{i_m}, F_{\ell_m})\} \quad \text{to} \\ &\{(F_{i_1}, F_{\ell_1}), \dots, (F_{i_m}, F_{\ell_m}), (F_{i_{m+1}}, F_{\ell_{m+1}})\}. \end{aligned}$$

Thus, the computation of all spanning trees containing  $T$ , or  $T \cup \{e\}$ , is represented by the subtree (of the computation tree) rooted at the bisection representing  $T$ , or  $T \cup \{e\}$ , respectively. The arrows pointing away from  $T$  are, from left to right, all edges from the stack  $F$  that are added to  $T$ . Also, remember that after the computation of all spanning trees containing  $T \cup \{e\}$ , it has to be checked if  $e$  is a bridge. If it is not,  $e$  is deleted from  $T$  and  $G$ . This is depicted by the red arrows pointing to the bisection that represents the addition of the next edge, together with the corresponding label indicating which edge is deleted. If it is a bridge, then all spanning trees containing  $T$  have been found and that iteration comes to an end. This is similarly depicted by a red arrow pointing to "END". Finally, each leaf in the lowest level of the computation tree is a complete and unique spanning tree.

In the case of the present example and as stated above, one starts with  $T$  containing no edge ( $E_T = \emptyset$ ) and pops the last element from  $F$  to add it to  $T$ , yielding

$$\begin{aligned} e &= (F_3, F_1), \\ T &= (\{F_3, F_1\}, \{(F_3, F_1)\}), \text{ and} \\ F &= [(F_3, F_2)]. \end{aligned} \tag{4}$$

As all spanning trees containing  $T$  shall be computed, one pops the next edge from  $F$ , here  $(F_3, F_2)$ . This results in

$$\begin{aligned} e &= (F_3, F_2), \\ T &= (\{F_3, F_1, F_2\}, \{(F_3, F_1), (F_3, F_2)\}), \text{ and} \\ F &= [\emptyset]. \end{aligned}$$

The tree  $T$  is now a complete spanning tree as it has  $n-1$  (here,  $n=3$ ) edges. Thus, one sets  $L = T$  and tests if  $e = (F_3, F_2)$  is a bridge. The nondescendants of  $F_2$  in  $L$  are  $F_3$  and  $F_1$ , and there is no edge in  $G$  that goes from a nondescendant of  $F_2$  to  $F_2$  besides  $e$  itself. Consequently,  $e$  is categorized as a bridge and indeed, all spanning trees containing the subtree with  $E_T = \{(F_3, F_1)\}$  have been computed. The iteration ends and the algorithm returns to the setting of (4). Doing the bridge test here reveals that  $e = (F_3, F_1)$  is not a bridge, since  $L$  remains unchanged and there exists the edge  $(F_2, F_1)$  in  $G$  where  $F_2$  is a nondescendant of  $F_1$  in  $L$ . Therefore, the edge  $(F_3, F_1)$  is deleted from  $G$  and  $T$  and the next

iteration begins, meaning that the next edge from  $F$  is added to  $T$ :

$$\begin{aligned} e &= (F_3, F_2), \\ T &= (\{F_3, F_2\}, \{(F_3, F_2)\}), \text{ and} \\ F &= [(F_2, F_1)]. \end{aligned} \quad (5)$$

Again, all spanning trees containing  $T$  have to be computed and the next edge is popped from  $F$ , resulting in

$$\begin{aligned} e &= (F_2, F_1), \\ T &= (\{F_3, F_2, F_1\}, \{(F_3, F_2), (F_2, F_1)\}), \text{ and} \\ F &= [\emptyset]. \end{aligned}$$

The tree  $T$  is a new and distinct spanning tree. One sets  $L = T$  and a test reveals that  $e$  is a bridge:  $F_3$  is the only nondescendant of  $F_1$  in  $L$  that does not belong to  $e$  itself, and the edge  $(F_3, F_1)$  was just deleted from  $G$ , so it does not exist anymore in the current graph (although it will be restored later). All spanning trees containing the subtree with  $E_T = \{(F_3, F_2)\}$  have been found. The current iteration is terminated and the algorithm returns to the iteration with the setting (5). Here, one checks if  $e = (F_3, F_2)$  is a bridge and again, it is. The edge  $e$  is the only edge leading to  $F_2$  in  $G$ . Hence, this iteration ends as well, which means that all spanning trees containing the subtree with  $E_T = \emptyset$  have been computed successfully, or in other words, all spanning trees existing in  $G$ . The original graph  $G$  is restored (i.e., the edge that was deleted,  $(F_3, F_1)$ , is added once again to  $G$ ), the whole algorithm terminates and returns

$$S = \{(\{F_3, F_1, F_2\}, \{(F_3, F_1), (F_3, F_2)\}), (\{F_3, F_2, F_1\}, \{(F_3, F_2), (F_2, F_1)\})\}.$$

One can easily check by hand that these two spanning trees are the only ones existing in  $G$ .

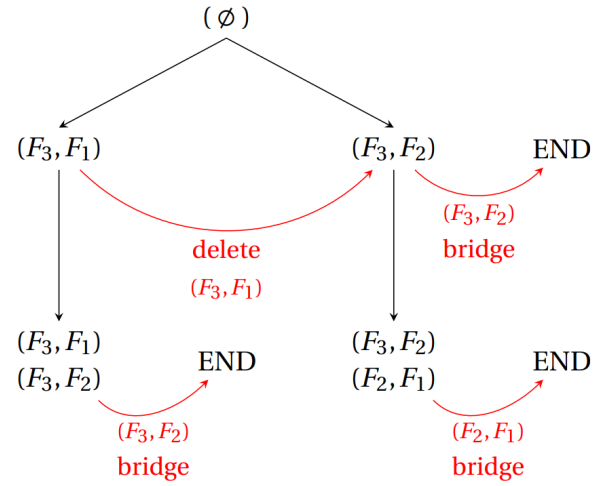
Finally, the result, still in tree structure, has to be converted back to a set of connections. Inserting the variable nodes stored in the matching  $\mathcal{M}$  gives

$$P = \{(\{F_3, \{x_1, \dot{x}_1\}, F_1\}, (F_3, \{x_2\}, F_2)\}, (\{F_3, \{x_2\}, F_2\}, (F_2, \{x_1, \dot{x}_1\}, F_1)\})\}.$$

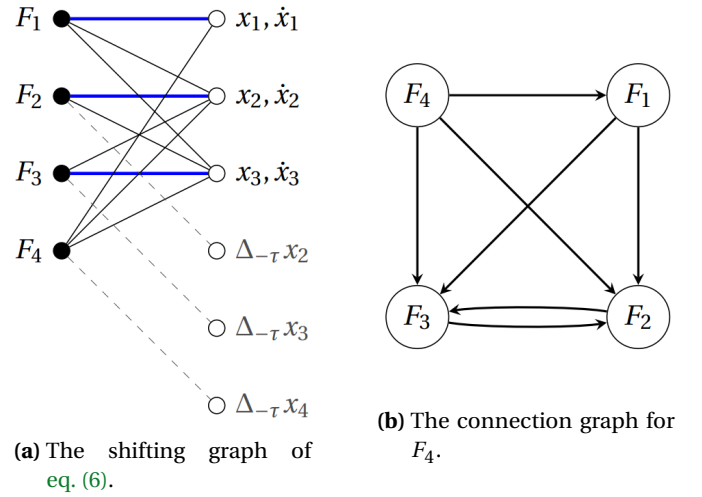
By comparing  $P$  to Figure 2b and Figure 2c, it can be seen that the algorithm successfully determined all desired connections for  $F_3$  with respect to  $\mathcal{M}$ .

**Example 16.** Consider the DDAE

$$\begin{aligned} \dot{x}_1 &= x_2 + x_3, \\ \dot{x}_2 &= x_3 + \Delta_{-\tau} x_2, \\ \dot{x}_3 &= x_2 + \Delta_{-\tau} x_3, \\ 0 &= x_1 + x_2 + x_3 + \Delta_{-\tau} x_4. \end{aligned} \quad (6)$$



**Figure 4** – Computation tree of Algorithm 1 for the construction of connections for  $F_3$  with respect to  $\mathcal{M}$  in the shifting step of the DDAE eq. (3).



**Figure 5** – Visualization of the construction of connections for  $F_4$  with respect to  $\mathcal{M}$  in the shifting step of the DDAE eq. (6).

The shifting graph, after assigning  $\{x_1, \dot{x}_1\}$  to  $F_1$ ,  $\{x_2, \dot{x}_2\}$  to  $F_2$  and  $\{x_3, \dot{x}_3\}$  to  $F_3$ , is shown in Figure 5a. The equation  $F_4$  is exposed and cannot be matched directly to any equivalence class, but it is connected via alternating paths to all other equation nodes. Therefore, it holds that  $C_{F_4} = \{F_1, F_2, F_3\}$  and all possible connections for  $F_4$  with respect to

$$\mathcal{M} = \{F_1, \{x_1, \dot{x}_1\}, F_2, \{x_2, \dot{x}_2\}, F_3, \{x_3, \dot{x}_3\}\}$$

have to be found. The connection graph  $H = (V_H, E_H)$  for  $F_4$  with respect to  $\mathcal{M}$  is given in Figure 5b with

$$\begin{aligned} V_H &= \{F_1, F_2, F_3, F_4\}, \\ E_H &= \{(F_1, F_2), (F_1, F_3), (F_2, F_3), (F_3, F_2), (F_4, F_1), (F_4, F_2), (F_4, F_3)\}. \end{aligned}$$

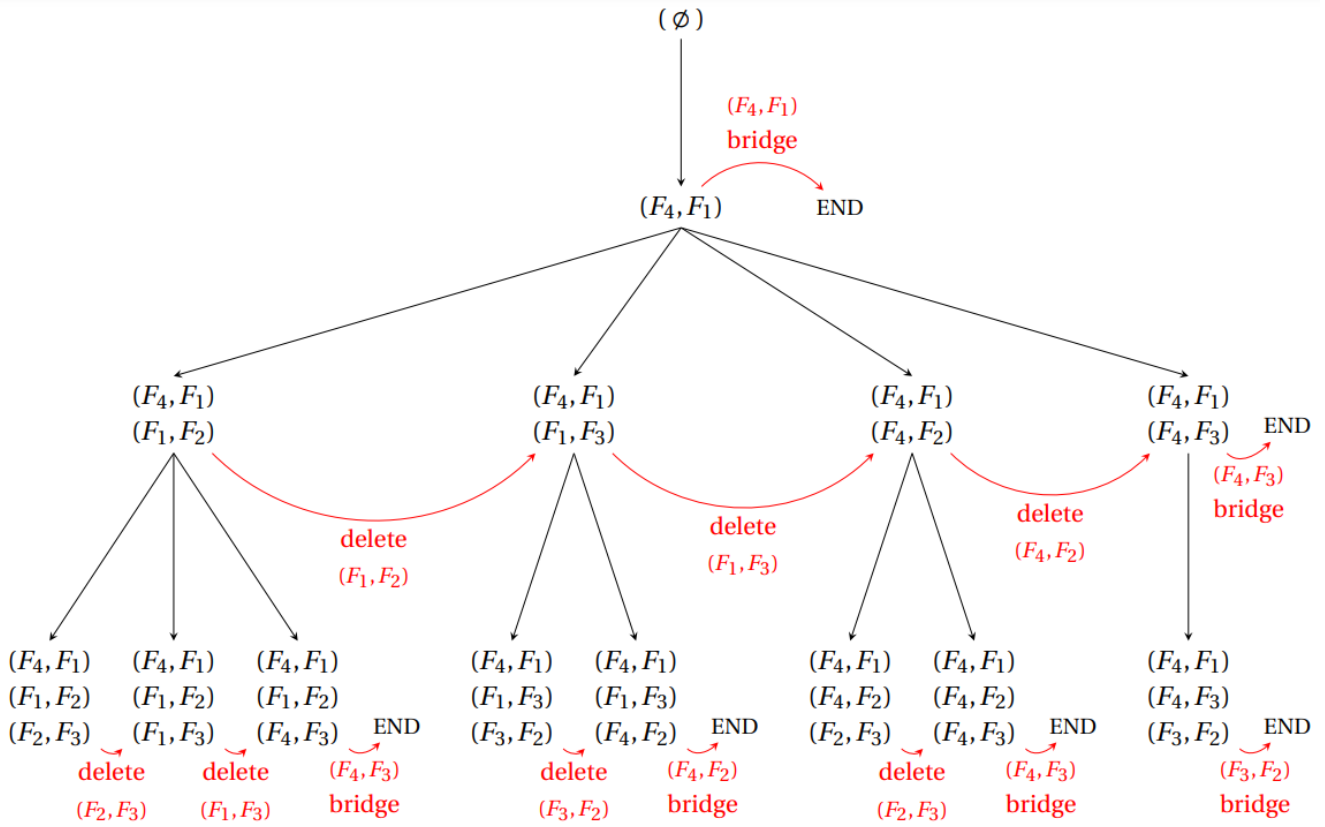


Figure 6 – Computation tree of Algorithm 1 for the DDAE eq. (6).

After defining  $G := H$  and  $r := F_4$  as the input to Algorithm 2 and initializing

$$T = (\{F_4\}, \emptyset), \text{ and } F = [(F_4, F_3), (F_4, F_2), (F_4, F_1)],$$

Algorithm 1 is executed to enumerate all spanning trees of  $G$  rooted at  $r$ .

The computation tree that represents the recursion structure of Algorithm 1 can be seen in Figure 6. Similarly to the last example, one can follow the different paths in the tree to retrace the construction of subtrees  $T$  by addition and deletion of edges  $e$ .

Note that even though  $F$  is initialized with all three edges outgoing from  $F_4$ , the algorithm already terminates after the first iteration where all spanning trees containing the subtree with  $E_T = \{(F_4, F_1)\}$  are computed. This is due to the fact that after deleting  $(F_4, F_1)$  from  $G$ , there is no edge leading to  $F_1$  anymore, and hence it is not possible to construct another spanning tree. The following eight spanning trees are returned:

$$S = \left\{ \begin{aligned} & \{ \{F_4, F_1, F_2, F_3\}, \{ (F_4, F_1), (F_1, F_2), (F_2, F_3) \} \}, \\ & \{ \{F_4, F_1, F_2, F_3\}, \{ (F_4, F_1), (F_1, F_2), (F_1, F_3) \} \}, \\ & \{ \{F_4, F_1, F_2, F_3\}, \{ (F_4, F_1), (F_1, F_2), (F_4, F_3) \} \}, \\ & \{ \{F_4, F_1, F_3, F_2\}, \{ (F_4, F_1), (F_1, F_3), (F_3, F_2) \} \}, \\ & \{ \{F_4, F_1, F_3, F_2\}, \{ (F_4, F_1), (F_1, F_3), (F_4, F_2) \} \}, \end{aligned} \right.$$

$$\begin{aligned} & \{ \{F_4, F_1, F_2, F_3\}, \{ (F_4, F_1), (F_4, F_2), (F_2, F_3) \} \}, \\ & \{ \{F_4, F_1, F_2, F_3\}, \{ (F_4, F_1), (F_4, F_2), (F_4, F_3) \} \}, \\ & \{ \{F_4, F_1, F_3, F_2\}, \{ (F_4, F_1), (F_4, F_3), (F_3, F_2) \} \}. \end{aligned}$$

After converting them into connections with respect to the matching  $\mathcal{M}$ , one finally obtains

$$P = \left\{ \begin{aligned} & \{ \{F_4, \{x_1, \dot{x}_1\}, F_1\}, (F_1, \{x_2, \dot{x}_2\}, F_2), (F_2, \{x_3, \dot{x}_3\}, F_3) \}, \\ & \{ \{F_4, \{x_1, \dot{x}_1\}, F_1\}, (F_1, \{x_2, \dot{x}_2\}, F_2), (F_1, \{x_3, \dot{x}_3\}, F_3) \}, \\ & \{ \{F_4, \{x_1, \dot{x}_1\}, F_1\}, (F_1, \{x_2, \dot{x}_2\}, F_2), (F_4, \{x_3, \dot{x}_3\}, F_3) \}, \\ & \{ \{F_4, \{x_1, \dot{x}_1\}, F_1\}, (F_1, \{x_3, \dot{x}_3\}, F_3), (F_3, \{x_2, \dot{x}_2\}, F_2) \}, \\ & \{ \{F_4, \{x_1, \dot{x}_1\}, F_1\}, (F_1, \{x_3, \dot{x}_3\}, F_3), (F_4, \{x_2, \dot{x}_2\}, F_2) \}, \\ & \{ \{F_4, \{x_1, \dot{x}_1\}, F_1\}, (F_4, \{x_2, \dot{x}_2\}, F_2), (F_2, \{x_3, \dot{x}_3\}, F_3) \}, \\ & \{ \{F_4, \{x_1, \dot{x}_1\}, F_1\}, (F_4, \{x_2, \dot{x}_2\}, F_2), (F_4, \{x_3, \dot{x}_3\}, F_3) \}, \\ & \{ \{F_4, \{x_1, \dot{x}_1\}, F_1\}, (F_4, \{x_3, \dot{x}_3\}, F_3), (F_3, \{x_2, \dot{x}_2\}, F_2) \} \}. \end{aligned} \right.$$

Indeed, all possible connections for  $F_4$  with respect to  $\mathcal{M}$  have been found.

## 6 Numerical demonstration

The developed algorithm presented in this paper has been implemented to empirically demonstrate its effectiveness. Also, a naive depth-first method is used to compute connections in order to estimate the efficiency

of the new algorithm in terms of computational complexity. The depth-first method works as described in [13, p.19]. However, it does not stop once a spanning tree is found but stores it and continues searching for other spanning trees by trying all possible combinations of edges. All computations are performed using MATLAB R2021a on a laptop with the processor Intel CORE i5-6267U CPU @2.90GHz (4 CPUs), ~2.8GHz.

For simplicity, a shifting graph  $G^S = (V_E^S \cup V_V^S, E^S)$  is assumed to be given where only the variable nodes  $v_k \in V_V^S$  of highest shift exist for  $k = 1, \dots, n-1$  (i.e., all other variable nodes have already been deleted) and each  $F_i \in V_E^S$  is matched to  $v_i$ , for  $i = 1, \dots, n-1$ . Thus,  $F_n$  is exposed with respect to the matching

$$\mathcal{M} = \{F_1, v_1\}, \dots, \{F_{n-1}, v_{n-1}\}.$$

Three different scenarios are tested. To illustrate the edge structures  $E^S$  of the corresponding shifting graphs, let  $A \in \mathbb{R}^{n \times (n-1)}$  be a matrix with entries

$$a_{ij} = \begin{cases} 1, & \text{if } \{F_i, v_j\} \in E^S, \\ 0, & \text{else.} \end{cases}$$

First, a shifting graph is constructed such that

$$A = \begin{bmatrix} 1 & 1 & & & & \\ 1 & \ddots & \ddots & & & \\ & \ddots & \ddots & 1 & & \\ & & & 1 & 1 & \\ 1 & \dots & \dots & 1 & & \end{bmatrix}, \quad (7)$$

i.e., each equation node  $F_i$ , for  $i = 1, \dots, n-1$ , is connected to at most three variable nodes and  $F_n$  is connected to each  $v_k$ , for  $k = 1, \dots, n-1$ . The computation times for shifting graphs of this structure for different  $n \in \mathbb{N}$  can be seen in Table 1. In all tables, "DFS" is the abbreviation for "depth-first search" and  $N$  denotes the number of possible connections. Some computations have been stopped after 10 minutes of computing time, which is indicated by "> 600". In these cases, computations for even higher  $n$  have not been executed. This is marked as "-" in the tables.

**Table 1** – Computation times in [s] for scenario (7).

$n$	5	6	7	8	9	10
$N$	21	55	144	377	987	2584
DFS	0.01	0.11	8.3	>600	-	-
Alg. 3	0.02	0.04	0.06	0.16	0.43	1.32

In a second test, a scenario is created such that

$$A = \begin{bmatrix} 1 & \dots & 1 \\ & \ddots & \vdots \\ & & 1 \\ 1 & \dots & 1 \end{bmatrix}, \quad (8)$$

i.e., each equation node  $F_i$ , for  $i = 1, \dots, n-1$ , is connected to  $n-i$  variable nodes and  $F_n$  is again connected to each  $v_k$ , for  $k = 1, \dots, n-1$ . The computation times for shifting graphs of this structure can be seen in Table 2.

**Table 2** – Computation times in [s] for scenario (8).

$n$	5	6	7	8	9	10
$N$	24	120	720	5040	40320	362880
DFS	0.01	0.22	31	>600	-	-
Alg. 3	0.04	0.09	0.39	2.2	14	126

For the third scenario, a complete graph is assumed, where each equation node is connected to all variable nodes, i.e.,

$$A = \begin{bmatrix} 1 & \dots & 1 \\ \vdots & & \vdots \\ 1 & \dots & 1 \end{bmatrix}. \quad (9)$$

The computation times are listed in Table 3.

**Table 3** – Computation times in [s] for scenario (9).

$n$	5	6	7	8	9
$N$	125	1296	16807	262144	4782969
DFS	0.03	0.41	318	>600	-
Alg. 3	0.05	0.48	6.3	88	2462

The results clearly show the advantage of Algorithm 3 as it is strongly superior in terms of computation time. For all scenarios, the depth-first search algorithm is only competitive for very small system sizes  $n$ , before its computation time suddenly explodes. This has a simple reason: by naively testing all possible combinations of edges, an extreme amount of possibilities arises. Even more, the majority of connections computed by the depth-first search algorithm are duplicates, meaning that they possess the same alternating paths in different order. All of these have to be identified and deleted after the algorithm terminates. Algorithm 3, however, does not have this problem, as only unique spanning trees (and thus, connections) are computed. Therefore, it scales well with the number of possible connections  $N$  and has a huge advantage in terms of computational complexity. Nevertheless, one can also see that the problem itself is very demanding, because  $N$



increases rapidly with the system size  $n$  and already for relatively small  $n$ , one cannot compute all connections in a reasonable time anymore. There are just too many in the case of dense graphs.

## 7 Conclusion

In this work, the problem of finding all connections in the shifting step of the Pantelides algorithm for DDAEs from [1] has been discussed. A new method, based on the reformulation of the problem into the problem of enumerating all spanning trees (or arborescences) in a directed graph, has been developed. This directed graph is constructed with the alternating paths of the shifting graph and is called connection graph. The equivalence of the solutions to these two problems has been proven in Theorem 12. That led to the possibility to exploit the fact that there already exist efficient methods to solve the enumeration problem. By introducing and implementing the method from [6], Algorithm 3 has been introduced to compute all connections in the shifting graph. Its effectiveness for the problem at hand has been shown by giving theoretical examples and its efficiency has been demonstrated by an implementation and numerical tests.

However, it was evident that the presented method is limited to small scale problems. There are several reasons for that. First, the enumeration algorithm uses a rather expensive system to add, delete and restore edges as well as to manage the different stacks of edges. The development of more efficient enumeration algorithms could improve the method in terms of computational complexity. Second, the problem of finding all connections itself scales really badly to larger system sizes as the pure number of existing connections grows exponentially with  $n$  for densely connected shifting graphs. Thus, if one restricts the DDAE to certain classes where each equation only has a limited number of variables in it, meaning that each node in the shifting graph has a limited number of edges, one could also limit the growth of the number of connections  $N$  and therefore extend the usage of the algorithm to larger system sizes. The effects of such restrictions could be further researched.

In summary, the lack of a satisfactory solution to the problem of finding all connections in the shifting step of the Pantelides algorithm for DDAEs has been overcome by this work for small problems. The new method now provides an efficient algorithm for its solution.

**Code Availability:** The MATLAB source code of the implementation used to compute the presented results is available as supplementary material and can be obtained under

DOI:10.14464/gammas.v4i1.503.

**Acknowledgements:** The author thanks his supervisors Ines Ahrens, Benjamin Unger and Volker Mehrmann for their help, valuable tips and the encouragement to publish this paper. His work is supported by the DFG Collaborative Research Center 910 *Control of self-organizing nonlinear systems: Theoretical methods and concepts of application*, project number 163436311.

## References

- [1] I. Ahrens and B. Unger. *The Pantelides algorithm for delay differential-algebraic equations*. *Trans. Math. Appl.*, 4(1):1–36, 2020.
- [2] U. Ascher and L. Petzold. *The numerical solution of delay-differential-algebraic equations of retarded and neutral type*. *SIAM J. Numer. Anal.*, 32(5):1635–1657, 1995.
- [3] A. Bellen and M. Zennaro. *Numerical methods for delay differential equations*. Numerical Mathematics and Scientific Computation. The Clarendon Press, Oxford University Press, New York, 2003.
- [4] S. Campbell. *Singular linear systems of differential equations with delays*. *Appl. Anal.*, 11(2):129–136, 1980.
- [5] S. Campbell. *Nonregular 2D descriptor delay systems*. *IMA J. Math. Control Inform.*, 12(1):57–67, 1995.
- [6] H. Gabow and E. Myers. *Finding all spanning trees of directed and undirected graphs*. *SIAM J. Comput.*, 7(3):280–287, 1978.
- [7] P. Ha. *Analysis and numerical solutions of delay differential-algebraic equations*. PhD thesis, Technische Universität Berlin. 2015.
- [8] P. Ha. *Spectral characterizations of solvability and stability for delay differential-algebraic equations*. *Acta Math. Vietnam.*, 43(4):715–735, 2018.
- [9] P. Ha and V. Mehrmann. *Analysis and reformulation of linear delay differential-algebraic equations*. *Electron. J. Linear Algebra*, 23:703–730, 2012.
- [10] P. Ha and V. Mehrmann. *Analysis and numerical solution of linear delay differential-algebraic equations*. *BIT*, 56(2):633–657, 2016.
- [11] P. Ha, V. Mehrmann, and A. Steinbrecher. *Analysis of linear variable coefficient delay differential-algebraic equations*. *J. Dynam. Differential Equations*, 26(4):889–914, 2014.
- [12] S. Hougardy and J. Vygen. *Algorithmic Mathematics*. Springer International Publishing, 2016.
- [13] D. Kozen and D. Gries. *The Design and Analysis of Algorithms*. Monographs in Computer Science. Springer, 1992.
- [14] C. Pantelides. *The consistent initialization of differential-algebraic systems*. *SIAM J. Sci. Statist. Comput.*, 9(2):213–231, 1988.
- [15] S. Trenn and B. Unger. *Delay regularity of differential-algebraic equations*. In *Proc. 58th IEEE Conf. Decision Control (CDC) 2019, Nice, France*, pages 989–994, 2019.
- [16] B. Unger. *Discontinuity propagation in delay differential-algebraic equations*. *Electron. J. Linear Algebra*, 34:582–601, 2018.
- [17] B. Unger. *Delay differential-algebraic equations in real-time dynamic substructuring*. *ArXiv e-print 2003.10195*, 2020.
- [18] B. Unger. *Well-Posedness and Realization Theory for Delay Differential-Algebraic Equations*. PhD thesis, Technische Universität Berlin. 2020.