Leaderless consensus of a hierarchical cyber-physical system

Shaoan Xiao\textsuperscript{a,b}, Xiao Chen\textsuperscript{c}, Yanjun Li\textsuperscript{a}, Arman Goudarzi\textsuperscript{c} and Ji Xiang\textsuperscript{c,d}

\textsuperscript{a}School of Information and Electrical Engineering, Zhejiang University City College, Hangzhou, Zhejiang, People’s Republic of China; \textsuperscript{b}College of Control Science and Engineering, Zhejiang University, Hangzhou, Zhejiang, People’s Republic of China; \textsuperscript{c}College of Electrical Engineering, Zhejiang University, Hangzhou, Zhejiang, People’s Republic of China; \textsuperscript{d}Zhejiang Awesome Intelligent Technology Co. Ltd., NingBo, People’s Republic of China.

ABSTRACT

This paper models a class of hierarchical cyber-physical systems and studies its consensus problem. The model has a pyramid structure, reflecting many realistic natural or human systems. By analysing the spectrum of the coupling matrix, it is shown that all nodes in the physical layer can reach consensus based on the proposed distributed protocols without interlayer delays. Then, the result is extended to the case with interlayer delays. A necessary and sufficient condition for consensus seeking is derived from the frequency domain perspective, which describes a permissible range of the delay. Finally, the application of the proposed model in the power-sharing problem is simulated to demonstrate the effectiveness and significance of the developed results.

1. Introduction

Advanced technologies in communication and computation techniques have been widely used in the information sensing, control and operation of physical systems such as power networks, medical systems and manufacturing (Gatouillat et al., 2018; Monostori et al., 2016; Yu & Xue, 2016). Such systems that connect the cyber world to the physical world are called cyber-physical systems (CPS), which are characterised by tightly coupling between computation, communications and physical processes (Antsaklis, 2014).

Intuitively, each CPS has a physical bottom layer with dynamic characteristics that is responsible for collecting data, sending state and output information, and multiple information layers that aggregate, process and make decisions on the data uploaded from the bottom layer. Thus, a CPS can naturally be modelled by a finite number of subsystems in a hierarchical pyramid structure. In a hierarchical pyramid structure, nodes are divided into groups which can be further subdivided into smaller groups. The nodes of the higher layers have aggregate information about their subordinate groups. This is completely different from multi-agent systems with peer-to-peer information exchange. This type of structure can be applied to analyse multifarious CPSs, such as systems described in Villegas et al. (2014), Xin et al. (2017), Wen et al. (2017) and Wang (2021).

The hierarchical pyramid structure is the most intuitive modelling structure for a certain number of CPSs. But how to model such a structure to achieve some coordination is still an open question. In this paper, a general model is built for this pyramid hierarchy system and its consistency conditions are discussed. We consider a power network with transmission system operators (TSOs) to illustrate the problem more vividly. A TSO manages local generators to supply customers without much dependence on neighbouring TSOs; meanwhile, the TSOs interconnect a broader region by sharing aggregate information for reliability and economic reasons (Wen et al., 2017). The above-mentioned example can be sketched into a hierarchical pyramid structure as shown in Figure 1.

The consensus problem, as a fundamental challenge in distributed control and coordination, has attracted extensive attention over the last decade and has encouraged a number of researchers to work on the consensus-based applications (Nedic, 2015; Olfati-Saber et al., 2007; Olfati-Saber & Murray, 2004; Xiang et al., 2017; Yang et al., 2016). There are also several works with the consensus of hierarchical network. In the field of leader-following consensus, Nguyen (2015) employs an LQR approach making the two-layer multi-agent system achieve consensus. The agents perform local actions in the lower layer and exchange information to achieve a cooperative purpose in the upper layer, such as tracking all agents to the reference system. And it is shown that the hierarchical network can achieve a fast convergence rate of consensus (Shao et al., 2016), which aligns with the phenomenon of pigeon flock or swarm intelligence (Nagy et al., 2010). Several studies have been conducted on the application of hierarchical structure in the area of leaderless consensus-based problems. Smith et al. introduce a hierarchical cyclic pursuit scheme where all the agents are placed in the cyclic pursuit within each group. At the same time, the centroid of each group follows the centroid of the next group in a sequential manner (Smith et al., 2005). Mukherjee and Ghose (2016) present a new method by taking the heterogeneous gains into account and generalised its convergence properties. Since the aforementioned hierarchical cyclic pursuit scheme failed to describe the weakness of intergroup...
couplings in the real world, Tsubakino and Hara (2012) proposes the concept of low-rank interactions to overcome this barrier. In Iqbal et al. (2018), a Cartesian product based hierarchical scheme is proposed, which does not necessarily exhibit circulant symmetry as required in the hierarchical cyclic pursuit method. Based on the Lyapunov function method, several researchers propose sufficient conditions for the consensus of a hierarchical multi-agent system with interlayer communication delay (Duan et al., 2015). Furthermore, in Pham et al. (2019), researchers achieve a robust control design for the two-layer hierarchical structure.

However, the discussed previous studies on consensus are often based on two-layer structure whether focus on leaderless or not. They cannot be simply migrated to the pyramid structure which is closer to the information transmission model in reality. Moreover, they are restricted to the two subsequent conditions: either the communication graphs of the subgroups which are located in the same layer must be identical, or a special circulant matrix is required. The main purpose of this paper is to propose a general mathematical model for a hierarchical pyramid CPS to break through the above two restrictions and investigate its related consensus problems. The first layer of the proposed model is the physical layer, where physical systems are restricted as first-order integrators. The other layers are hierarchical cyber layers for computation and communication. Interlayer communication delays are considered in the proposed model. The power-sharing problem in the power system is the direct application of the proposed method. The local generators can be regarded as the nodes in the physical layer, and superstructure composed of dispatch organisations (DOs) can be regarded as the hierarchical pyramid cyber layers. To be specific, the DOs in the second layer is responsible for inter-provincial power dispatch, while the DOs in the third layer perform power dispatch over a larger area such as different regions.

The major contributions of this study can be listed as follows:

1. Presenting a hierarchical model with distributed consensus protocols. It can be used to model some pyramidal CPS, and its subgroups are allowed to have different communication graphs.

2. Providing a necessary and sufficient condition for the consensus of the hierarchical CPS with interlayer delays.
Applying the proposed model to the power-sharing problem in power systems.

The remainder of this paper is organised as follows: Section 2 constructs the mathematical formulation of the proposed hierarchical CPS. Section 3 presents distributed protocols. Section 4 analyses the convergence properties of the hierarchy model without the interlayer delay, while Section 5 takes this interlayer delay into the account and presents a necessary and sufficient condition for the consensus. The simulation results of the proposed model are given in Section 6. Finally, Section 7 concludes the paper. All the proofs are placed in Appendix 1.

2. Mathematical formulation

2.1 Background in Laplacian matrix

Consider an undirected graph \( G = (V, E) \), where \( V \) is the set of nodes and \( E \) is the set of edges denoted by \((i, j)\). Two nodes \( u \) and \( v \) of \( G \) are neighbours if \( \{u, v\} \in E \). The binary adjacency matrix of \( G \) is the non-negative matrix \( A \), where \( a_{ij} = 1 \) if \((i, j) \in E \), and \( a_{ij} = 0 \) otherwise. The out-degree matrix \( D_{\text{out}} \) is a diagonal matrix defined as \( D_{\text{out}} = \text{diag}(\Delta) \), where \( \Delta \) is the column vector with compatible dimensions and all components being 1. Laplacian matrix of \( G \) is given by \( L = D_{\text{out}} - A \), which has the following properties: (i) its row-sums are zero, (ii) its diagonal entries are non-negative, and (iii) its non-diagonal entries are non-positive (Bullo, 2017).

2.2 Hierarchical structure

The rules that the proposed hierarchical pyramid structure should abide by are listed as follows:

\( \text{(r1)} \) The graph of each group should be connected and there is no communication link between the groups of the same layer.

\( \text{(r2)} \) The physical layer is at the bottom with the largest number of nodes, while the top layer has only one group with the smallest number of nodes.

\( \text{(r3)} \) Each group has one superior node in the upper layer, except for the group in the top layer; each node has a subordinate group in the lower layer, except for physical layer nodes.

The first and second rules show that the isolated groups are ultimately connected via the top layer nodes. The third rule implies that the number of nodes decreases as the number of layers increases, thereby forming a pyramid structure. The extreme case is that all the groups, except the top layer group, have only one node so that the number of nodes in all layers is the same.

Consider a hierarchical CPS with \( M \) layers, where the \( l \)-th layer contains \( N^{(l)} \) nodes, \( l = 1, 2, \ldots, M \). The bottom layer is the physical layer and the other layers are the cyber layers. The physical layer consists of \( N^{(1)} \) dynamic nodes, controlled by

\[
\dot{x}^{(1)}_i = u_i, \quad i = 1, 2, \ldots, N^{(1)},
\]

where \( x^{(1)}_i \in \mathcal{R} \) is the state value, \( u_i \) is the input and \( N^{(l)} \) denotes the number of nodes in the \( l \)-th layer, \( l = 1, \ldots, M \). The \( N^{(l)} \) nodes in the \( l \)-th layer form \( N^{(l+1)} \) groups, and each group is the subordinate group of one node in the \((l+1)\)-th layer. Consequently, all nodes in the top layer form only one group, namely \( N^{(M+1)} = 1 \).

The \( p \)-th group in the \( l \)-th layer contains \( k_p^{(l)} \) nodes, which has an undirected communication graph denoted by \( G_p^{(l)} \). All nodes in the same layer are numbered according to the ascending order of the respective group index. Let \( V_i^{(l)} \) denote the \( i \)-th node in the \( l \)-th layer. The set of neighbours of \( V_i^{(l)} \) in the \( l \)-th layer is denoted by \( N_i^{(l)} \).

Each group \( G_p^{(l)} \) in the \( l \)-th layer has the unique superior node \( V_p^{(l+1)} \), and \( G_p^{(l)} \) is called the subordinate group of node \( V_p^{(l+1)} \). Use \( V_p^{(l+1)} = G_p^{(l)} \) to represent that \( V_p^{(l+1)} \) is the superior node of \( G_p^{(l)} \). The superior node \( V_p^{(l+1)} \) collects and broadcasts the information from and to its subordinate group \( G_p^{(l)} \). This action represents the interlayer communication link, which is located between the superior node and the node of its subordinate group.

Similar to the definition of receptive field in convolutional neural networks, the receptive field of \( V_i^{(l)} \) in the hierarchical structure refers to the physical nodes that are visible to \( V_i^{(l)} \). The physical number of \( V_i^{(l)} \), denoted by \( N_i^{(l)} \), is the number of physical nodes in its receptive field. If each physical node has a nonnegative physical weight, we define the physical weight \( a_i^{(l)} \) of the cyber node \( V_i^{(l)} \) as the sum of the physical weights of the physical nodes in its receptive field.

Let \( \mu \) denote the information exchanged between layers. Given the communication delay \( \tau \), the available information is \( \mu(t - \tau) \) instead of \( \mu(t) \), which is in the Laplace domain is \( \mu(s) e^{-s\tau} \). In this sense, we can formulate the information exchanged between \( V_i^{(l)} \) and its superior node:

\[
\begin{align*}
\dot{z}_i^{(l)}(t) &= x_i^{(l)}(t - \tau), \\
\dot{y}_i^{(l)}(t) &= v_i^{(l)}(t - \tau),
\end{align*}
\]

where \( x_i^{(l)} \) is the information sent from \( V_i^{(l)} \), which is collected by its superior node \( V_p^{(l+1)} \) as \( z_i^{(l)} \). The received message \( y_i^{(l)} \) of \( V_i^{(l)} \) is broadcasted from its superior node as \( v_i^{(l)} \). The information collected by \( V_p^{(l+1)} \) from its subordinate group \( G_p^{(l)} \) can be written in a column vector \( z_p^{(l)} = [z_i^{(l)}]^T \in \mathcal{R}^{k_p^{(l)}} \) for \( V_i^{(l)} \in G_p^{(l)} \). Similarly, we use \( v_p^{(l)} = [v_i^{(l)}]^T \in \mathcal{R}^{k_p^{(l)}} \) to denote the information broadcasted by \( V_p^{(l+1)} \) to its subordinate group \( G_p^{(l)} \). All interlayer delays arising from the \( l \)-th layer to the \((l+1)\)-th layer are assumed to be the same, which is denoted by \( \tau_l \).

2.3 Three-layer example

In this section, we present a three-layer example as shown in Figure 1 to illustrate the parameters and contents introduced in Section 2.2. In the physical layer, there are \( V_1^{(1)} \), \( V_2^{(1)} \), \( V_3^{(1)} \), \( V_4^{(1)} \), \( V_5^{(1)} \) and \( V_6^{(1)} \) from top to bottom. These nodes are divided into three groups, and each group is the subordinate group of one node in the second layer. For instance, the third group in the
first layer, represented by $G_3^{(1)}$, is the subordinate group of the third node $V_3^{(2)}$ in the second layer. We call $V_3^{(2)}$ the superior node of $G_3^{(1)}$, denoted by $V_3^{(2)} = G_3^{(3)}$.

An intralayer edge connects $V_3^{(2)}$ and $V_6^{(1)}$ in $G_3^{(1)}$, but there is no intralayer edge between different groups in the same layer. The interlayer communication links are located between all nodes in every group and their superior nodes. The receptive field of $V_1^{(2)}$ is group $G_1^{(1)}$ and the physical number satisfies $n_1^{(2)} = 3$. All nodes in $G_3^{(1)}$ and $G_2^{(1)}$ are the receptive field of $V_1^{(3)}$, so the physical number of $V_1^{(3)}$ is $n_1^{(3)} = 4$.

The node $V_3^{(2)}$ has interlayer edges with its subordinate group $G_3^{(1)}$ and its superior node $V_2^{(3)}$. Let $x_3^{(2)}$ denote the information sent from $V_6^{(1)}$ to $V_3^{(2)}$. If we take the interlayer delay into account, then the information received by $V_3^{(2)}$ can be expressed by $y_3^{(2)}(t) = x_6^{(1)}(t - \tau_1)$, where $\tau_1$ is the delay time of interlayer communication between the first layer and the second layer. Similarly, we can obtain $y_6^{(1)}(t) = y_3^{(1)}(t - \tau_1)$, which describe the interdelay delay when the information is broadcasted from $V_3^{(2)}$ to $V_6^{(1)}$. Meanwhile, $x_3^{(2)}$ and $y_3^{(2)}$ are the information of $V_3^{(2)}$ which are exchanged with its superior node $V_2^{(3)}$.

### 2.4 Statement of the problem

Based on the presented communication structure, several assumptions are made as follows:

(A1) The available information of the physical node $V_j^{(1)}$ includes the neighbouring node state $x_j^{(1)}$ with $V_j^{(1)} \in \mathcal{N}_j^{(0)}$ and the information $y_j^{(1)}$ broadcasted from its superior node.

(A2) The available information of cyber layer node $V_p^{(l)}$ includes the neighbouring node state $x_p^{(l)}$ with $V_p^{(l)} \in \mathcal{N}_p^{(0)}$, the collected information $z_p^{(l-1)} \in \mathbb{R}^{k^{(l-1)}}$ from its subordinate group and the broadcasted information $y_p^{(l)}$ from its superior node. Here $y_p^{(M)}$ is null and $l = 2, \ldots, M$.

(A3) Every group has a connected graph.

Concerning the stated assumptions, the main goal of the study can be formulated as designing the following protocols:

(1) a control protocol $u_i = u(x_i^{(1)}, y_i^{(1)}, y_i^{(1)})$ with $V_j^{(1)} \in \mathcal{N}_i^{(n)}$,

(2) a collecting protocol $x_p^{(l)} = x(z_p^{(l-1)})$,

(3) a broadcasting protocol $y_p^{(l)} = y(x_p^{(l)}, x_p^{(l)}, y_p^{(l)})$ with $V_p^{(l-1)} \in G_p^{(l-1)}$ and $V_p^{(r)} \in \mathcal{N}_p^{(0)}$.

Let $u_i$, $x_i^{(1)}$, $y_i^{(1)}$ and $y_i^{(1)}$ be column vectors with entries $x_i^{(1)}$, $y_i^{(1)}$, $z_p^{(l-1)}$ and $y_p^{(l)}$, respectively $(i = 1, \ldots, N^{(l)})$. Then, the proposed hierarchical system can be expressed in a compact form.

### 3. Protocols

The proposed control protocol is

$$u_i = -\frac{1}{d_i^{(1)}} \sum_{j \in \mathcal{N}_i^{(n)}} (x_i^{(1)} - x_j^{(1)}) + y_i^{(1)}, \quad i = 1, \ldots, N^{(1)},$$

(3)

This protocol helps the nodes reach consensus within their group. By the received information $y_i^{(1)}$, the state values for isolated groups in the physical layer can converge to a common value.

For the superior node $V_p^{(l)}$ to obtain the weighted average of its subordinate group $G_p^{(l-1)}$, the collecting protocol is given by

$$x_p^{(l)} = C_p^{(l-1)} z_p^{(l-1)}, \quad l = 2, \ldots, M, \quad p = 1, \ldots, N^{(l)},$$

(4)

where $C_p^{(l-1)}$ is a $k^{(l-1)}$-dimensional row vector with non-negative entries that can add up to one.

The broadcasting protocol can be written as

$$y_p^{(l)} = \frac{1}{d_p} \sum_{r \in \mathcal{N}_p^{(0)}} (y_p^{(l)} - y_r^{(l)}) + y_p^{(l)}, \quad l = 2, \ldots, M, \quad p = 1, \ldots, N^{(l)},$$

(5)

where $V_p^{(l-1)} \in G_p^{(l-1)}$. It is clear that all nodes in $G_p^{(l-1)}$ receive the same message from their superior node $V_p^{(l)}$.

**Remark 3.1:** By modelling the collection and broadcast of information in human society through two proposed protocols, the proposed model can solve the consensus problem in the more practical pyramid structure. The model allows subgroups with the different communication graphs instead of with the same communication graphs as required in Mukherjee and Ghose (2016), Tsubakino and Hara (2012) and Iqbal et al. (2018). Moreover, the Kronecker product structure in Tsubakino and Hara (2012) and the circulant matrix in Mukherjee and Ghose (2016) are no longer required.

Equations (2)–(5) formulate the proposed hierarchical cyber-physical system, which can be illustrated by using an electrical power network. A DO in the cyber layer collects regional generation capacity and the required load demand information to control the power output of generators that are located in the physical layer. In power system operation, the power-sharing problem aims to force the output of generators to reach the same ratio concerning their maximum output power. Notation $x_i^{(1)}$ represents the output ratio of a generator, which can be affected by the state of the neighbouring node $x_j^{(1)}$ and the received information $y_i^{(1)}$ from its DO. These DOs share the aggregate information $x_p^{(2)}$ with their neighbours and coordinate power transfer between regions by sending $y_p$ to regional generators.

The issue of power balance in this model will be discussed in Section 6.

The proposed hierarchical system can be expressed in a compact form.
The node dynamics in the physical layer is given by

$$\dot{x}^{(1)} = -K^{(1)}L^{(1)}_D x^{(1)} + y^{(1)},$$

and the process in the cyber layer can be formulated by

$$\begin{align*}
z^{(l-1)}(t) &= x^{(l-1)}(t - \tau_{l-1}) \\
x^{(l)}(t) &= C^{(l-1)} \cdot z^{(l-1)}(t) \\
v^{(l-1)}(t) &= g^{(l-1)} \cdot [-K^{(l)}L^{(l)}_D x^{(l)}(t) + y^{(l)}(t)] \\
y^{(l-1)}(t) &= v^{(l-1)}(t - \tau_{l-1})
\end{align*}$$

(7)

where $L^{(l)}_D$ is the Laplacian matrix of the $l$th layer, which can be written by

$$L^{(l)}_D = \text{diag}(I^{(l)}_1, \ldots, I^{(l)}_{N^{(l)}_{(l+1)}}), \quad l = 1, \ldots, M,$$

(8)

where $I^{(l)}_p \in \mathbb{R}^{N^{(l+1)} \times 1}$ denote the Laplacian matrix of $G^{(l)}_p$. The matrix $K^{(l)}$ is a diagonal matrix represented by

$$K^{(l)} = \text{diag}(a^{(l)}_1, \ldots, a^{(l)}_{N^{(l+1)}_{(l+1)}}), \quad l = 1, \ldots, M.$$  

(9)

The matrix $B^{(l)}$ is a block diagonal matrix with $N^{(l+1)}$ blocks, describing the information broadcasted from the $(l+1)$th layer to the $l$th layer, and given by

$$B^{(l)} = \text{diag}(1^{(l)}_1, \ldots, 1^{(l)}_{N^{(l+1)}_{(l+1)}}), \quad l = 1, \ldots, M - 1,$$  

(10)

where $1^{(l)}_p$ is a $N^{(l+1)}$-dimensional column vector with all entries being 1. The matrix $C^{(l)}$ represents the information collection from the $l$th layer to the $(l+1)$th layer, which can be written by

$$C^{(l)} = \text{diag}(G^{(l)}_1, \ldots, G^{(l)}_{N^{(l+1)}_{(l+1)}}), \quad l = 1, \ldots, M - 1.$$  

(11)

If the interlayer delay is ignored ($\tau_l = 0$), then hierarchical system which has been defined by (6)–(7) can be simplified as

$$\dot{x}^{(1)} = -\sum_{l=1}^{M} I^{(l)}_p x^{(l)},$$

(12)

where

$$L^{(l)} = \begin{cases}
K^{(l)}L^{(l)}_D, & l = 1, \\
\prod_{i=1}^{l-1} B^{(i)} \cdot K^{(l)}L^{(l)}_D \cdot \prod_{i=l}^{M} C^{(l)}, & l = 2, \ldots, M,
\end{cases}$$

(13)

and

$$L = \sum_{l=1}^{M} L^{(l)}.$$  

(14)

**Remark 3.2:** The proposed model differs significantly from the consensus algorithm for two main reasons. First, instead of the traditional consensus protocol, specific collecting and broadcasting protocols are designed for cyber nodes to make full use of aggregate information in the hierarchical structure. According to the proposed model (12), there is no dynamic in the cyber nodes. This is in line with engineering practice. Unlike the peer-to-peer multi-agent system structure, the role of the upper node of CPS is often to monitor the data of the underlying physical nodes and issue control instructions to them, which generally does not have dynamics. Various CPSS can be applied this modelling method, such as Villegas et al. (2014), Xin et al. (2017), Voropai et al. (2018), Stennikov et al. (2022). Second, since $L$ given by (14) does not depict a Laplacian matrix, many well-known results cannot be directly used to analyse the proposed model.

### 4. Hierarchy model with no interlayer delay

In this section, we analyse the related consensus problems of the proposed hierarchical system (12) without delay. Throughout this paper, notation $\Lambda(X)$ denotes the eigenvalues set of $X$, and $1_x$ denotes the column vector with $x$-dimensions and all components being 1.

**Lemma 4.1:** Given two matrices $B \in \mathbb{R}^{m \times n}$ and $C \in \mathbb{R}^{n \times m}$ ($m \geq n$), satisfying $CB = I_{n \times n}$. Let $A \in \mathbb{R}^{n \times n}$ be a square matrix and $E \in \mathbb{R}^{n \times m}$, then $\Lambda(A) \subset \Lambda(E)$.

**Lemma 4.2:** Given two diagonal matrices $E = \text{diag}(E_1, \ldots, E_n)$ and $B = \text{diag}(B_1, \ldots, B_n)$, where $E_i \in \mathbb{R}^{m_i \times m_i}$ is a Laplacian matrix with $\sum_{i=1}^{n} m_i = m$, then for any matrix $C \in \mathbb{R}^{n \times m}$ satisfying $CB = I$, and any matrix $A \in \mathbb{R}^{n \times n}$, the eigenvalues set of matrix $F = E + BAC$ satisfies $[\Lambda(E) \cup \Lambda(A)] \subset \Lambda(F)$.

Let $L^{(l)}_c$ be a scaled matrix of $L^{(l)}_D$, given as

$$L^{(l)}_c = K^{(l)}L^{(l)}_D,$$  

(15)

then we can construct the following matrix sequence based on Equation (13),

$$
\begin{align*}
L^{(M)}_0 &= L^{(M)}_c \\
L^{(M)}_1 &= L^{(M-1)}_c + B^{(M-1)}I^{(M-1)}_D C^{(M-1)} \\
&\quad \ldots \\
L^{(M)}_k &= L^{(M-k)}_c + B^{(M-k)}I^{(M-k)}_D C^{(M-k)}
\end{align*}
$$

(16)

where $k < M$. It is clear that the matrix $L$ defined in Equation (14) can also be represented by $L = L^{(M)}_{M-1}$.

**Theorem 4.3:** Assume that in the top layer $M$, $N^{(M)}$ nodes form one connected graph, namely $N^{(M+1)} = 1$, then for any $k$, $\Lambda(L^{(M)}_k) = 0 \cup \Lambda(L^{(M-k)}_k) / 0 \cup \Lambda(L^{(M)}_{k-1}) / 0$,

(17)

where $\Lambda(L^{(M)}_{k-1}) / 0$ denotes the non-zero eigenvalues set of $L^{(M)}_{k-1}$. Furthermore, 0 is a simple eigenvalue of $L^{(M)}_k$, that is, the algebraic multiplicity of 0 is 1.

Theorem 4.3 implies that we can analyse $\Lambda(L^{(l)}_c)$ instead of studying $\Lambda(L)$ directly. Note that $K^{(l)}K^{(l)}_D = K^{(l)}_D K^{(l)} = 1$. $L^{(l)}_D K^{(l)}$ is a real symmetric matrix. One has that $\Lambda(L^{(l)}_c) = \Lambda(L^{(l)}_c) = \Lambda(1 \times 1) \subset \mathbb{R}$. Also notice that $L^{(l)}_c$ is a weighted Laplacian matrix, thus, the real part of eigenvalues of $L^{(l)}_c$ is not less than zero.
Furthermore, the real part of all other eigenvalues of $L$ is positive, except for the single zero eigenvalue.

Below is an immediate corollary based on Theorem 4.3 and its proof.

**Corollary 4.4:** Let $L_s = \sum_{l=1}^{M} s_l L^{(l)}$, where $s_l$ is a complex number except for zero, then

$$\Lambda(L_s) = 0 \cup \Lambda(s_1 L^{(1)}/0 \cup \cdots \cup \Lambda(s_M L^{(M)})/0. \tag{18}$$

**Theorem 4.5:** For the hierarchical system (12) without interlayer delay, all nodes in the physical layer asymptotically reach a consensus given by

$$\lim_{t \to \infty} x^{(1)}(t) = \frac{11^T K}{1^T K} x_0^{(1)}, \tag{19}$$

where $x_0^{(1)}$ is the initial value of $x^{(1)}$ and

$$K = \text{diag}(a_1^{(1)}, \ldots, a_{N^{(1)}}^{(1)}). \tag{20}$$

In particular, all physical layer nodes will reach an average consensus if $a_i^{(1)} = a_j^{(1)}$ for $i, j \in \{1, \ldots, N^{(1)}\}$.

### 5. The influence of interlayer delay

The main goal of this section is to analyse the impact of time constants on the convergence properties of the proposed system. First, the three-layer example from the perspective of the frequency domain is studied, and then the results are extended for the generalisation of the presented model.

#### 5.1 Three-layer example

Based on the proposed protocols in Section 3, the three-layer hierarchical system (as depicted in Figure 1) can be described by

$$\begin{cases}
\dot{x}^{(1)}(t) = -K^{(1)} L^{(1)} \cdot x^{(1)}(t) + y^{(1)}(t) \\
\dot{x}^{(2)}(t) = C^{(2)} \cdot x^{(1)}(t - \tau_1) \\
y^{(1)}(t) = B^{(1)} \cdot [-K^{(2)} L^{(2)} \cdot x^{(2)}(t - \tau_1) + y^{(2)}(t - \tau_1)] \\
\dot{x}^{(3)}(t) = C^{(3)} \cdot x^{(2)}(t - \tau_2) \\
y^{(2)}(t) = B^{(2)} \cdot [-K^{(3)} L^{(3)} \cdot x^{(3)}(t - \tau_2)]
\end{cases} \tag{21}$$

If we take the Laplace transform of Equation (21), then

$$sX^{(1)}(s) - x_0^{(1)} = -L^{(1)} X^{(1)}(s) - \frac{L^{(2)}}{e^{\tau_1 s}} X^{(2)}(s) - \frac{L^{(3)}}{e^{2(\tau_1 + \tau_2) s}} X^{(1)}(s), \tag{22}$$

where $X^{(1)}(s)$ is the Laplace transform of $x^{(1)}$, the initial value of $x^{(1)}$ is noted as $x_0^{(1)}$. Let $L_{\tau, s} = L^{(1)} + \frac{L^{(2)}}{e^{\tau_1 s}} + \frac{L^{(3)}}{e^{2(\tau_1 + \tau_2) s}}$, then

$$X^{(1)}(s) = [sI + L_{\tau, s}]^{-1} x_0^{(1)}, \tag{23}$$

and the characteristic function of Equation (21) is given by

$$|sI + L_{\tau, s}| = 0. \tag{24}$$

**Theorem 5.1:** For the three-layer hierarchical model (21), all nodes in the physical layer can asymptotically reach a consensus given by Equation (19) if and only if all other roots of Equation (24) are in the open left half-plane except for the single root at zero.

It is evident that the value of $s$ solved by Equation (24) is the eigenvalue of matrix $-L_{\tau, s}$. Together with the results given in Corollary 4.4 we can infer that

$$\Lambda(L_{\tau, s}) = 0 \cup \Lambda(L^{(1)})/0 \cup \frac{\Lambda(L^{(2)})/0}{e^{\tau_1 s}} \cup \frac{\Lambda(L^{(3)})/0}{e^{2(\tau_1 + \tau_2) s}},$$

and 0 is a simple eigenvalue of $L_{\tau, s}$. Thus, the solution set of Equation (24) is equal to the union of the solution sets of the following formulas

$$s = 0, \tag{25a}$$

$$s + \lambda^{(1)} = 0, \tag{25b}$$

$$s \cdot e^{2\tau_1 s} + \lambda^{(2)} = 0, \tag{25c}$$

$$s \cdot e^{2(\tau_1 + \tau_2) s} + \lambda^{(3)} = 0, \tag{25d}$$

where $\lambda^{(l)}$ denote the non-zero eigenvalue of $L^{(l)}$ for $l = \{1, 2, 3\}$.

Based on the analysis of Section 4, we know that $\lambda^{(l)} \in \mathbb{R}$ and the real part of eigenvalues of $L^{(l)}$ is not less than zero, so $\lambda^{(l)} \in \mathbb{R}^+$.  

**Theorem 5.2:** For a transcendental equation such as

$$s \cdot e^{Ts} + \lambda = 0, \tag{26}$$

where $s = \sigma + j\omega$ is a complex variable, $T, \lambda \in \mathbb{R}^+$. Then, it has all roots in the open left half-plane if and only if

$$T < T^* \simeq \frac{\pi}{2\lambda}. \tag{27}$$

**Remark 5.1:** A similar conclusion appears in Theorem 10 of Olfati-Saber and Murray (2004), however, in this paper, we have provided a new and explicit proof.

Here is an immediate Theorem based on Theorem 5.1 and 5.2.

**Theorem 5.3:** For the three-layer hierarchical model (21) with interlayer delay, all physical layer nodes converge to a consensus given by Equation (19) if and only if both of the following formulas hold

$$\tau_1 < \frac{\pi}{4\lambda_{\text{max}}^{(2)}}, \tag{28a}$$

$$\tau_1 + \tau_2 < \frac{\pi}{4\lambda_{\text{max}}^{(3)}}, \tag{28b}$$

where $\lambda_{\text{max}}^{(2)}$ and $\lambda_{\text{max}}^{(3)}$ are the maximum eigenvalues of $L^{(2)}$ and $L^{(3)}$, respectively.
5.2 Generalization

The proposed methods and conclusions in the three-layer example can be extended to establish a general model.

**Theorem 5.4:** The polynomial characteristics of the hierarchical system (6)–(7) is equivalent to a series of polynomials:

\[ s = 0, \quad (29a) \]
\[ s + \lambda^{(1)} = 0, \quad (29b) \]
\[ s \cdot e^{2\mu \sum_{i=1}^{l-1} \tau_i} + \lambda^{(l)} = 0, \quad l = 2, \ldots, M. \quad (29c) \]

All nodes in the physical layer asymptotically can reach a consensus given by Equation (19) if and only if

\[ \sum_{i=1}^{l-1} \tau_i < \frac{\pi}{4\lambda^{(1)}_{\max}}, \quad l = 2, \ldots, M. \quad (30) \]

**Remark 5.2:** Most of the previous studies, such as Mukherjee and Ghose (2016) and Iqbal et al. (2018), neglect the communication delays to simplify their analysis. Duan et al. (2013) takes the interlayer delays into account, but it is assumed that communication graphs of the subgroups which are located in the same layer must be identical. Here the effect of delay time has been studied from the perspective of the frequency domain. Corollary 4.4 splits the polynomial characteristic of the hierarchical system into a series of tractable polynomials. In addition, Theorem 5.2 shows the influence of interlayer delays on the distribution of characteristic roots.

**Remark 5.3:** According to the form of Theorem 5.4, we can find that when a new subsystem need to be plug into the \( p \)th layer of an existing hierarchical system, we just need to verify whether the conditions are met within the subsystem and updated \( p \)th layer. So the proposed pyramid hierarchical system is easy to expand.

6. Applications and simulation results

The proposed model can be applied to solve the power-sharing problem in the hierarchical power system, the purpose of which is to drive the output of the generators to converge to the same ratio with respect to their maximum power out. Taking the three-layer model as shown in Figure 1 for example, the six nodes in the physical layer represent six generators, and the nodes in the other two layers represent DOs. The three DOs in the second layer is responsible for inter-provincial power dispatch, while the two DOs in the third layer perform power dispatch in a larger area such as different regions. Notation \( x^{(1)}_i \) computed by \( \bar{p}^{(1)}_i / \bar{P}^{(1)}_i \) represents the power ration of a generator, where \( \bar{p}^{(1)}_i \) and \( \bar{P}^{(1)}_i \) denotes the output power and maximum output power of the generator, respectively. Let the physical weight of the generator node be equal to its maximum power output, that is \( a^{(1)}_i = \bar{P}^{(1)}_i \), then the vector form of the generator output power can be expressed as \( p^{(1)} = Kx^{(1)} \), where \( K \) is given by Equation (20). Suppose that the power output is regulated instantaneously and initially in a supply-demand equilibrium state, namely \( p^{(1)} = p^{(1)}_{\text{ref}} \) and \( 1^T \cdot p^{(1)}(0) = P_D \), where \( p^{(1)}_{\text{ref}} \) and \( P_D \) are the power out reference and the total power demand, respectively. It can be achieved \( 1^T \cdot \dot{p}^{(1)}(t) = 1^T \cdot Kx^{(1)}(t) = 0 \) from the proof of Theorem 5.1, thus, the supply-demand balance will not be violated in the transient process. Based on Theorem 5.3, we can obtain the permissible range for the inter-layer delay for the power-sharing in this hierarchical system.

In the following section, the effectiveness of the proposed model will be examined and verified by the simulation results.

![Figure 2. Power sharing test when \( \tau_1 = \frac{\pi}{7}, \tau_2 = \frac{\pi}{9} \) (Case 6.1).](image-url)
The aim of Case 6.1 is to demonstrate that all generators can achieve the power-sharing based on the proposed distributed protocols. The other cases (Cases 6.2–6.4) investigate the effect of time delay on the convergence properties of the proposed hierarchical system. Assuming that the maximum power output of the generator nodes are $P^{(1)} = [0.8, 0.7, 1.5, 1, 0.8, 1.2]^T$ MW. The initial power output are $p^{(1)}(0) = [0.24, 0.56, 0.9, 0.9, 0.56, 0.24]^T$ MW and the total demand is 3.4 MW. Therefore, the initial power ratio are $x^{(1)}(0) = [0.3, 0.8, 0.6, 0.9, 0.7, 0.2]^T$. Let all the edge weights of each graph $G^p(a)$ be equal to 1. Thus, the maximum eigenvalues of $L^{(2)}$ and $L^{(3)}$ can be listed as $\lambda^{(2)}_{\text{max}} = 4/3$ and $\lambda^{(3)}_{\text{max}} = 0.75$, respectively.

**Case 6.1:** $\tau_1 = \pi/7$ and $\tau_2 = \pi/9$: In this case, Equation (28a) holds, so the power ratio of all generators converges to a common value 0.5667 as shown in Figure 2(a), which is consistent with the result computed by Equation (19). It is clear that $\lim_{t \to \infty} p^{(1)}(t) = K \cdot \lim_{t \to \infty} x^{(1)}(t) = [0.4533, 0.3967, 0.85, 0.5667, 0.4533, 0.68]^T$ MW, which is shown in Figure 2(b). Figure 2(c) indicates that the power balance is maintained from beginning to end. It is worth noting that, the $u_i$ will change abruptly at $t = 2\tau_1$ and $t = 2(\tau_1 + \tau_2)$ due to the existence of the interlayer delays, so the non-derivable points appear at the corresponding moment.

**Case 6.2:** $\tau_1 = \pi/6$, $\tau_2 = \pi/6$: Figure 3(a) shows the power ratio trajectories of the generator nodes under this delay time. In this case, Equation (28a) holds but $\tau_1 + \tau_2 = \pi/(4\lambda^{(3)}_{\text{max}})$. It can be inferred from Theorem 5.2 that both Equations (25a) and (25b) have all the roots in the open left half-plane, but Equation (25c) has roots on the imaginary axis. Therefore, the system will be in a state of critical oscillation.

**Case 6.3:** $\tau_1 = 3\pi/16$, $\tau_2 = \pi/12$: In this case, Equation (28b) holds but $\tau_1 = \pi/(4\lambda^{(2)}_{\text{max}})$. So all the roots of Equations (25a) and (25c) are located in the open left half-plane, but Equation (25b) has roots on the imaginary axis. The three-layer example exhibits critical oscillation, as shown in Figure 3(b).

**Case 6.4:** $\tau_1 = 3\pi/16$, $\tau_2 = 7\pi/48$: Figure 3(c) shows the power ratio trajectories of the hierarchical example where $\tau_1 = \pi/(4\lambda^{(2)}_{\text{max}})$ and $\tau_1 + \tau_2 = \pi/(4\lambda^{(3)}_{\text{max}})$. Similarly, we can obtain all the roots of Equation (25a) that are located in the open left half-plane, but both Equations (25b) and (25c) have roots on the imaginary axis. Therefore, at the same time, the system will be in a state of critical oscillation.

7. Conclusions

In this paper, a hierarchical cyber-physical system has been introduced with distributed consensus protocols driving the nodes in the physical layer to reach a consensus. For the behaviour of the hierarchical model without interlayer delay, we have analysed its convergence properties. Also, an interlayer communication delay has been addressed with a necessary and sufficient condition that describes a permissible range for the delay time. The results of the simulation cases on the power sharing problem verify the practicality and the effectiveness of the proposed hierarchical model.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This work is supported by the Key Area Research and Development Program of Guangdong Province [grant number 2021B0101410005].
References

Antsaklis, P. (2014). Goals and challenges in cyber-physical systems research. IEEE Transactions on Automatic Control, 59(12), 3117–3119. https://doi.org/10.1109/TAC.2014.2363897

Bullo, F. (2017). Differential geometry of multi-agent systems. Communications in Information & Systems, 17(1), 71–88. https://doi.org/10.4310/CIS.2017.v17.n1.a5

Duan, Z., Zhai, G., & Xiang, Z. (2015). State consensus for hierarchical multi-agent dynamical systems with inter-layer communication time delay. Journal of the Franklin Institute, 352(3), 1235–1249. https://doi.org/10.1016/j.jfranklin.2014.12.013

Gatouillat, A., Badr, Y., Massot, B., & Sejdic, E. (2018). Internet of medical things: A review of recent contributions dealing with cyber-physical systems in medicine. Internet of Things Journal, 5(5), 3810–3822. https://doi.org/10.1016/j.iot.2018.06.005

Iqbal, M., Leth, J., & Ngo, T. D. (2018). Cartesian product-based hierarchical scheme for multi-agent systems. Automatica, 88, 70–75. https://doi.org/10.1016/j.automatica.2017.11.009

Monostori, L., Kádár, B., Bauerhansl, T., Kondoh, S., Kumara, S., Reinhart, G., Sauer, O., Schuh, G., Sihn, W., & Ueda, K. (2016). Cyber-physical systems in manufacturing. Circ Annals, 65(2), 621–641. https://doi.org/10.1016/j.cirp.2016.05.005

Nedich, A. (2015). Convergence rate of distributed averaging dynamics and optimization in networks. Foundations and Trends in Systems and Control, 2(1), 1–100. https://doi.org/10.1561/2000000004

Nguyen, D. H. (2015). A sub-optimal consensus design for linear multi-agent systems based on hierarchical LQR. Automatica, 55, 88–94. https://doi.org/10.1016/j.automatica.2015.02.037

Olfati-Saber, R., Fax, J. A., & Murray, R. M. (2007). Consensus and cooperation in networked multi-agent systems. Proceedings of the IEEE, 95(1), 215–233. https://doi.org/10.1109/JPROC.2006.868729

Olfati-Saber, R., & Murray, R. M. (2004). Consensus problems in networks of agents with switching topology and time-delays. IEEE Transactions on Automatic Control, 49(9), 1520–1533. https://doi.org/10.1109/TAC.2004.834113

Shao, J., Qin, J., Bishop, A. N., Huang, T.-Z., & Zheng, W. X. (2016). A novel analysis on the efficiency of hierarchy among leader-following systems. Automatica, 73, 215–222. https://doi.org/10.1016/j.automatica.2016.07.007

Smith, S. L., Broucke, M. E., & Francis, B. A. (2005). A hierarchical cyclic pursuit scheme for vehicle networks. Automatica, 41(6), 1045–1053. https://doi.org/10.1016/j.automatica.2005.01.001

Stennikov, V., Barakhtenko, E., Mayorov, G., Sokolov, D., & Zhou, B. (2022). Coordinated management of centralized and distributed generation in an integrated energy system using a multi-agent approach. Applied Energy, 309, 118487. https://doi.org/10.1016/j.apenergy.2021.118487

Tsubakino, D., & Hara, S. (2012). Eigenvector-based intergroup connection of low rank for hierarchical multi-agent dynamical systems. Systems & Control Letters, 61(2), 354–361. https://doi.org/10.1016/j.sysconle.2011.11.008

Pham, T. V., Nguyen, D. H., & Banjerdpoungchai, D. (2019). Design of robust hierarchical control for homogeneous multi-agent systems with parametric uncertainty and exogenous disturbance. In 2019 12th Asian Control Conference (ASC’19) (pp. 937–942). Kitakyushu, Japan.

and the National Natural Science Foundation of China [grant number 62073290, 62173295].

Appendix A. Proofs of Lemmas and Theorems

A.1 Proof of Lemma 4.1

Let $\lambda$ be any eigenvalue of $A$, that is, $AV = \lambda V$ for some non-zero vector $V$. Let $BV = W$, then

$$ BW = BACV = BACW = BA\lambda V = \lambda BV $$

Therefore, $A(A) \subset A(E)$.

A.2 Proof of Lemma 4.2

Let $AV = \lambda V$, due to $CB = I$, then $BACBV = \lambda BV$. Since $E_1m_0 = 0$, $EB = \text{diag}(E_1, \ldots , E_n)$ and $m_0$ is 0 and therefore $EBV = 0$. Thus,

$$ FBV = EBV + BACBV = \lambda BV $$

that is $A(A) \subset A(F)$. It is evident that 0 is an eigenvalue of $E$ and $F$, noting that $C_1m_0 = I_m$. The remainder is to study the relationship of non-zero eigenvalues between $E$ and $F$.

Let us be the left eigenvector of matrix $E$, that is, $w^TE = \lambda_2 w^T$, for some eigenvalue $\lambda_2$. Since $EB = 0$, one has $w^TB = 0$. Thus,

$$ w^TF = w^TE + w^TBAC = \lambda_2 w^T $$

that is $\Lambda(E) \subset \Lambda(F)$. To sum up, $[\Lambda(E) \cup \Lambda(A)] \subset \Lambda(F)$.

A.3 Proof of Theorem 4.3

According to Lemma 4.2, it follows that $[A(L^{(M-i)}_k) \cup A(L^{(M-i)}_{k+1})] \subset A(L^{(M-i)}_k)$. This, together with Equation (16), yields that any eigenvalue of $L^{(i)}_k$ is the eigenvalue of $L^{(i)}_{k-1}$ for $i = M - k, M - k + 1, \ldots , M$.

Since the number of non-zero eigenvalues of $L^{(i)}_k \in \mathbb{R}^{N^{(k)} \times N^{(k)}}$, the number of non-zero eigenvalues of $L^{(M-i)}_k \in \mathbb{R}^{N^{(M-i)} \times N^{(M-i)}}$ is not less than

$$ \sum_{i=M-k}^{M} N^{(i)} \geq N^{(M-k)} - N^{(M-i)} = N^{(M-k)} - 1. $$
This implies that 0 is a simple eigenvalue of $L_k^{(M)}$ and the algebraic multiplicity of 0 is 1.

### A.4 Proof of Theorem 4.5

This Theorem is a special case of Theorem 5.1 where $L_{r,s} = L$. For more details, please refer to the proof of Theorem 5.1.

### A.5 Proof of Theorem 5.1

Let $e(t) = x(t) - 1$, and our goal is to prove $\lim_{t \to \infty} e(t) = 0$. We use $E(s)$ to denote the Laplace transform of $e(t)$, then

$$E(s) = X^{(1)}(s) - \frac{1}{s} \cdot \frac{11_T K}{11_T K^s} X_0^{(1)}.$$  \hspace{1cm} (A2)

Substituting this equation into Equation (23), we can get

$$[sI + L_{r,s}] \cdot E(s) + \frac{1}{s} \cdot \frac{11_T K}{11_T K^s} X_0^{(1)} = X_0^{(1)},$$  \hspace{1cm} (A3)

Since $L_{r,s} \cdot 1 = 0$, then

$$[sI + L_{r,s}] \cdot E(s) = \left( I - \frac{11_T K}{11_T K^s} \right) X_0^{(1)} = e_0,$$  \hspace{1cm} (A4)

where $e_0$ is the initial value of $e(t)$. Since $11_T K L_{r,s} = 0$, then $11_T K x_0^{(1)} = 0$, so $11_T K x_0^{(1)} = 11_T K X_0^s$, and we can get

$$11_T K \cdot e(t) = 11_T K \cdot \left[ x^{(1)}(t) - \frac{11_T K}{11_T K^s} X_0^{(1)} \right]$$

$$= 11_T K X_0^{(1)} - 11_T K \frac{11_T K}{11_T K^s} X_0^{(1)} = 0.$$  \hspace{1cm} (A5)

Take the Laplace transform of the equation above, then $11_T K \cdot E(s) = 0$. Thus Equation (A4) is equivalent to

$$[sI + L_{r,s} + 11_T K] \cdot E(s) = e_0,$$  \hspace{1cm} (A6)

that is,

$$E(s) = [sI + L_{r,s} + 11_T K]^{-1} e_0,$$  \hspace{1cm} (A7)

so we can get the characteristic equation

$$[sI + L_{r,s} + 11_T K] = 0$$  \hspace{1cm} (A8)

If all roots of Equation (A8) are in the open left half-plane, then $\lim_{t \to \infty} e(t) = 0$ holds for any initial states. The value of $s$ solved by Equation (A8) is the eigenvalues of $-(L_{r,s} + 11_T K)$. Since

$$\lambda(L_{r,s} + 11_T K) = (11_T K^1) \cup \lambda(L_{r,s}) \setminus \{0\},$$  \hspace{1cm} (A9)

then $\lim_{t \to \infty} e(t) = 0$ if and only if all other roots of $[sI + L_{r,s}] = 0$ are in the open left half-plane except for one at zero.

### A.6 Proof of Theorem 5.2

Equation (26) can be written as

$$\sigma e^{\sigma T} \cdot e^{\omega_T} + \omega e^{\sigma T} \cdot e^{(\omega_T + \frac{\pi}{2})} + \lambda = 0,$$  \hspace{1cm} (A10)

which is equivalent to

$$\sigma e^{\sigma T} \cos(\omega_T) - \omega e^{\sigma T} \sin(\omega_T) = -\lambda,$$  \hspace{1cm} (A11a)

$$\sigma e^{\sigma T} \sin(\omega_T) + \omega e^{\sigma T} \cos(\omega_T) = 0.$$  \hspace{1cm} (A11b)

It is possible to prove that if $(\sigma, \omega)$ is a pair of roots of Equation (A11a), then it has a pair of roots $(\sigma, -\omega)$. Next, we first introduce how $T^*$ is derived, and then prove the validity of this theorem.

Let $\sigma = 0$, then $s = j\omega$, and Equation (26) is reduced to

$$\omega \cdot e^{(\omega_T + \frac{\pi}{2})} + \lambda = 0.$$  \hspace{1cm} (A12)

Without loss of generality, let $\omega \geq 0$ and we can get

$$\begin{align*}
\omega &= \frac{\lambda}{\omega_T + \frac{\pi}{2}} = \frac{\pi}{2} + 2k\pi,
\end{align*}$$  \hspace{1cm} (A13)

$$\begin{align*}
T &= \frac{\pi/2 + 2k\pi}{\lambda},
\end{align*}$$  \hspace{1cm} (A14)

Take $k = 0$, then $T^* = \frac{\pi}{2\lambda}$.

In the following proofs, we will discuss the cases where $\omega \neq 0$ and $\omega = 0$, respectively.

When $\omega \neq 0$, it can be inferred from Equation (A11b) that $\omega_T \neq k\pi$, $k = 0, \pm 1, \pm 2, \ldots.$ So Equation (A11b) can be expressed by

$$\sigma = -\frac{\omega}{\tan(\omega_T)}.$$  \hspace{1cm} (A15)

Without loss of generality, let $\sigma > 0$ when $\omega \neq 0$.

In the case of $T < T^*$, assume that $(\sigma, \omega)$ is a pair of roots of Equation (26), and $\sigma > 0$. Based on Equation (A11a) we can get that

$$e^{\sigma T} \sqrt{\sigma^2 + \omega^2} = \lambda.$$  \hspace{1cm} (A16)

It is clear that $\omega_0 < \lambda$ when $\sigma_0 > 0$. Then

$$0 < \omega_0 T < \lambda T^* = \frac{\pi}{2\lambda} = \frac{\pi}{2},$$  \hspace{1cm} (A17)

so $\sigma_0 < 0$ according to Equation (A15), which is contradict with the assumption above. To sum up, Equation (26) has all roots in the open left half plane when $T < T^*$.

Under the circumstance of $T > T^*$, we want to show that Equation (26) has at least one pair of roots $(\sigma_i, \omega_i)$ satisfying $\sigma_i > 0$. Substituting Equation (A15) into Equation (A11a), we can get the equivalent expression of Equation (26),

$$\frac{x}{\sin x} = \lambda T \cdot e^{-\pi x}$$  \hspace{1cm} (A18)

Let $f(x) = \frac{x}{\sin x} - \lambda T \cdot e^{-\pi x}$, it is obvious that $f(x)$ is continuous over $x \in (\frac{\pi}{2}, \pi)$ and

$$\lim_{x \to \frac{\pi}{2}^+} f(x) = \pi - \lambda T < \pi - \frac{\pi}{2} - \lambda T^* = \frac{\pi}{2} - \lambda \cdot \frac{\pi}{2\lambda} = 0,$$  \hspace{1cm} (A19)

$$\lim_{x \to \pi^-} f(x) = +\infty - \lambda T \cdot 0^+ = +\infty - 0 > 0.$$  \hspace{1cm} (A20)

So there must exist a real number $x \in (\frac{\pi}{2}, \pi)$ satisfying $f(x) = 0$ when $T > T^*$. That is to say, if $T > T^*$, then there exists a pair of roots $(\sigma, \omega)$ of Equation (26), which satisfying $\omega_0 T \in (\frac{\pi}{2}, \pi)$ and $\sigma_0 > 0$.

When $\omega = 0$, Equation (26) is equivalent to

$$\sigma \cdot e^{\sigma T} = -\lambda.$$  \hspace{1cm} (A21)

If $(\sigma, 0)$ is a pair of roots of Equation (26), it is clear that $\sigma_0 < 0$ based on Equation (A11a). Therefore, we only need to prove that Equation (A21) has no real roots when $T > T^*$. Let $g(\sigma) = \sigma \cdot e^{\sigma T}$, it is easy to prove that $g(\sigma)$ takes the minimum value when $\sigma = -\frac{1}{T}$, and the minimum value is $g_{\min} = g(-\frac{1}{T}) = -\frac{1}{Te}$. If $T > T^*$, then

$$g_{\min} = -\frac{1}{Te} > \frac{1}{Te} = \frac{2}{\pi e} \lambda > -\lambda,$$  \hspace{1cm} (A22)

so Equation (A21) has no real roots. This ends the proof.