Constructing Lyapunov functions for large-scale hierarchical networks of nonlinear differential equations

Michael Y. Li* and John Maidens†

Abstract

We present a graph theoretic framework for modeling complex hierarchical systems and a method of systematically constructing Lyapunov functions for these systems. The effectiveness of our method is illustrated by two applications: a new control protocol for the self-organization of swarms of mobile agents into a hierarchy of clusters, and a new result on the global stability of the endemic equilibrium for a multi-group, multi-stage epidemic model.

Keywords. Lyapunov functions, hierarchical graphs, nonlinear differential equations on networks, Kirchhoff’s Matrix Tree Theorem, Tree-Cycle Identity, global stability.

2010 AMS mathematics subject classification. 34D23, 93A13, 92D30

1 Introduction

Many complex systems arising in science and engineering can be modeled by differential equations on networks [16, 35]. The network is described by a graph where each vertex is assigned a set of state variables and a rule that describes its internal dynamics. A vertex can represent, for example, the concentration of chemicals in a solution [21], communities of species living in discrete spatial patches [19], the velocity and position of individual vehicles in a swarm [8] or single neurons in a neural network [5]. Directed edges link the vertices, corresponding to couplings between the systems on each vertex. In the above contexts, the couplings can model the production of new chemical compounds, dispersal of animals between patches, vehicle interactions or communication between neurons.

Often the structure of the system of differential equations induces a hierarchical structure on the network [27]. This can result from different dynamics on different scales such as inter- vs. intra-city movement in global transport networks, the flow of capital within vs. between financial institutions, or the dispersal of various species at different rates in a predator-prey system.

We are interested in using Lyapunov functions to study the asymptotic behaviour of these hierarchically coupled systems. A Lyapunov function for a system of differential equations \( \dot{x} = f(x) \) on a subset \( D \subset \mathbb{R}^n \) is a map \( V : D \to \mathbb{R} \) whose derivative along trajectories of the system \( \dot{V}(x) = \nabla V \cdot f(x) \) is non-positive everywhere in \( D \). The existence of a Lyapunov function can be used to demonstrate asymptotic properties of the system, such as the stability or dichotomy of invariant sets like equilibria.

We want to provide a general method of constructing Lyapunov functions for large-scale systems from simple Lyapunov functions \( V_i \) corresponding to each individual vertex \( i \). Li and Shuai [16] have recently developed a method of choosing coefficients \( c_i \) to construct global Lyapunov functions of the form

\[
V(x) = \sum_{i=1}^{n} c_i V_i(x_i).
\]

The goal of this paper is to develop a graph theoretic framework for modeling complex hierarchical systems and to extend the method of Li and Shuai to our new framework. To model hierarchy in networks, we label graph vertices with multiple indices and discuss how this labeling induces a multi-dimensional...
structure on the graph and its weight matrix. In many applications, the hierarchical information is essential to the understanding and analysis of the systems studied. Thus the methods described in this paper can be applied in a wider range of cases than the original work of Li and Shuai.

Our construction of Lyapunov functions for these systems relies on Kirchhoff’s Matrix Tree Theorem to choose coefficients. The method is inspired by the proof of Kirchhoff’s formula for calculating the effective conductance between two nodes of a linear electrical network. For each spanning tree in the network, adding a single “battery” edge between the source vertex and the sink vertex yields a graph with a unique cycle in which the conductance can easily be found [2, 34]. We use the same idea — creating unicyclic graphs from a spanning tree and a single edge — for the analysis of nonlinear networks. Noting that each spanning unicyclic graph can be formed in a number of different ways gives us a method of grouping terms in the Lyapunov function’s derivative. This new combinatorial result is summarized in the Tree Cycle Identity (Theorem 1). This then leads to a simple criterion for verifying that \( \dot{V} \) is non-positive (Theorem 2). Rather than considering the network as a whole, with its multiple coupled feedback loops, Theorem 2 allows us to show that \( \dot{V} \leq 0 \) by considering only individual feedback loops.

In the second half of the paper, we demonstrate the applicability of our results by using them to construct two Lyapunov functions of different forms in two vastly different contexts. Section 4 outlines a novel protocol for controlling swarms of mobile agents and gives a detailed convergence analysis based on our methods. Section 5 describes class of models of the transmission dynamics of multi-stage diseases in heterogeneous populations and uses the techniques we have developed to prove a new result on the uniqueness and global stability of the endemic equilibrium. In both cases, understanding the system’s hierarchical structure is essential in the stability analysis.

2 Preliminaries

2.1 Terminology from graph theory

A weighted directed graph, or weighted digraph, is a graph \( G = (V(G), E(G)) \) together with a weight function \( \omega : E(G) \to \mathbb{R} \) that assigns to each edge of \( G \) a real-valued weight. We can extend the domain of the weight function \( \omega \) to the set of subgraphs of \( G \) by defining the weight of a subgraph as the product of the weights of all edges in the subgraph.

A weighted digraph can be described by its weighted adjacency matrix \( A_G \), with entries \( a_{ij} \) representing the weight \( \omega(\vec{ij}) \) of the edge \( \vec{ij} \) directed from vertex \( i \) to vertex \( j \). The set of neighbours of vertex \( i \) is given by \( N_i = \{ j \in V(G) : \vec{ji} \in E(G) \} \) and consists of all vertices with incoming connections to \( i \).

A sequence of edges of \( G \) \( i \to j \to \cdots \to k \to i \) that begins and ends at the same vertex is called a directed cycle or feedback cycle if the vertices \( i, j, \ldots, k \) are all distinct. In the special case that the graph \( G \) is a signal flow graph or block diagram then the concept of a directed cycle coincides with the familiar notion of a feedback loop from linear control theory [7].

Given an \( n \)-vertex graph \( G \), a subgraph of \( G \) is called a directed spanning tree rooted at vertex \( i \) if it contains \( n - 1 \) edges and a directed path from the root vertex \( i \) to each other vertex of \( G \). A spanning unicyclic subgraph of \( G \) is a spanning directed subgraph consisting of a collection of disjoint rooted directed trees whose roots form a directed cycle. For a detailed discussion of these concepts, we refer the reader to [23].

2.2 Hierarchical graphs

**Definition 1.** A graph \( G = (V(G), E(G)) \) is said to be hierarchical if its vertices are indexed by ordered \( n \)-tuples of integers \((i_1, \ldots, i_n)\) where \( n \geq 2 \). We call \( n \) the dimension of the hierarchical graph.

Graphs whose vertices are labeled by multiple indices have been studied in limited contexts. Multi-dimensional Manhattan networks were indexed by vectors because of their inherent spatial structure [6], and graphs resulting from the Kronecker product of smaller graphs inherit the indexing of the original graphs [15, 36]. These graphs both have very regular structures. The generality of our definition allows the framework of hierarchical graphs to be used to model many real-world networks whose structures are often highly irregular. Examples of networks that lend themselves naturally to hierarchical modeling include community structures in sociology and ecology, spatially discrete systems that inherit a multi-dimensional structure from the ambient space, and distributed process control schemes.
To define certain familiar notions such as the adjacency matrix of a hierarchical graph on $M$ vertices, we need to adopt a system of ordering the vertices. Throughout this paper we will use the lexicographic ordering. We use the symbol $(i)$ to denote the $i$-th vertex in the lexicographic ordering. The graph’s weighted adjacency matrix $A_G$ is then the $M$ by $M$ matrix whose entry in the $j$-th row and $k$-th column is the weight $a_{jk}$ of the edge directed from vertex $(j)$ to vertex $(k)$.

We call graphs labeled by multiple indices hierarchical because a graph of dimension $n$ can naturally be organized into $n$ levels of nested clusters.

**Definition 2.** For each $l$ with $1 \leq l \leq n$ we define a level-$l$ cluster

$$[i_1, \ldots, i_l] = \{(j) = (j_1, \ldots, j_n) \in V(G) : j_i = i_1, \ldots, j_l = i_l\}.$$  

In the case $n = 2$ we will simply say cluster to mean level-1 cluster.

Each level of clustering induces a partitioning of the graph’s adjacency matrix into a block matrix. Conversely, we can always define a hierarchical graph based on a block matrix, where the system of blocks determines the clusters in the graph. The case $n = 2$ is illustrated in Figure 1 for a graph with $M = 13$. The higher-dimensional cases follow by partitioning the diagonal blocks in the same way.

![Figure 1](https://example.com/figure1.png)

(a) A 2-dimensional weighted hierarchical graph with 13 vertices and 4 clusters. (b) Vertex ordering (c) Weighted block adjacency matrix corresponding to the hierarchical graph in (a).

**Figure 1:** A hierarchical graph and the corresponding block adjacency matrix

The Laplacian matrix plays an important role in the study of dynamical systems on graphs. It is used, for example, in the study of consensus in networks of autonomous agents [24] and the study of electric [2] and hydraulic [20] flows. The Laplacian matrix $L_G$ is defined as

$$L_G = D_G - A_G$$

where $D_G$ is the in-degree matrix

$$D_G = \text{diag} \left( \sum_{(j) \in V(G)} a_{j1}^1, \ldots, \sum_{(j) \in V(G)} a_{jM}^M \right).$$

Since the columns of $L_G$ all sum to zero, the cofactors corresponding to the entries of each column of $L_G$ are all identical. Thus for each $i \in \{1, \ldots, M\}$ we can define the Laplacian cofactor $c_i$ as the cofactor corresponding to any entry of the $i$-th column of $L_G$. Kirchhoff’s Matrix Tree Theorem allows us to express the cofactor $c_i$ as the sum of the weights of directed spanning trees rooted at vertex $(i)$.

**Proposition 1** (Kirchhoff’s Matrix-Tree Theorem).

For each vertex $(i) \in V(G)$ we have

$$c_i = \sum_{T \in \mathcal{T}(i)} \omega(T)$$

where $\mathcal{T}(i)$ denotes the set of directed spanning trees in $G$ rooted at vertex $(i)$.

**Proof.** See [14, 23].
2.3 Differential equations on hierarchical networks

We can build a system of differential equations on an $n$-dimensional hierarchical graph $G$ by assigning to each vertex $(i) = (i_1, \ldots, i_n)$ a nonlinear equation of the form

$$\dot{x}_i = f_i(x_i)$$

where $x_i \in \mathbb{R}^{m_i}$. We then assign to each edge directed from vertex $(j)$ to vertex $(i)$ a coupling function $g_{ij}(x)$ that represents the influence of the variable $x_j$ on $x_i$. If there is not an edge directed from vertex $(j)$ to vertex $(i)$ we set $g_{ij} \equiv 0$. We can then write the coupled system as

$$\dot{x}_i = f_i(x_i) + \sum_{(j) \in N(i)} g_{ij}(x).$$

(1)

We assume that the functions $f_i$ and $g_{ij}$ are sufficiently well-behaved to guarantee the existence and uniqueness of solutions to (1).

Our goal is to systematically construct a global Lyapunov function for the system (1) as a linear combination of Lyapunov candidate functions $V(i)$ corresponding to each vertex $(i)$.

3 Main results

3.1 Tree cycle identity for hierarchical digraphs

Consider a weighted digraph $G$ with non-negative weights $a_{ij}$. Assign to each edge $\bar{ji}$ an arbitrary edge function $F_j : \mathbb{R}^M \to \mathbb{R}$. The vertex function for vertex $(i)$ is then given by

$$F_{(i)}(x) = \sum_{(j) \in N(i)} a_{ij} F_j(x).$$

We wish to assign weights $w_{(i)}$ to each vertex $(i)$ such that the weighted sum

$$H(x) = \sum_{(i) \in V(G)} w_{(i)} F_{(i)}(x)$$

can be expressed in terms of feedback cycles in $G$. This will allow us to determine properties of the overall sum $H$ based on properties of smaller sums around each feedback cycle, which are much easier to verify in practical applications. To develop a method of choosing these vertex weights, we consider unicyclic subgraphs of $G$. We denote the set of all spanning unicyclic subgraphs of $G$ by $\mathcal{P}$.

An element $\mathcal{P}$ of $\mathcal{P}$ with cycle length $l$ can be expressed in precisely $l$ distinct ways as the union of a directed spanning tree and a single directed edge (see Figure 2). Indeed, removing each edge $\bar{ji}$ from the cycle $C_P$ of $\mathcal{P}$ yields a directed spanning tree rooted at vertex $(i)$. 

Figure 2: Construction of spanning unicyclic graphs as the union of a directed spanning tree and a single edge

In the following theorem, we show that by choosing vertex weights \( w_{(i)} \) equal to the cofactors \( c_i \) of the weighted Laplacian matrix \( L_G \) we can re-organize the weighted sum \( H \) in terms of feedback cycles in \( G \). This result was first established for 1-dimensional networks by Li and Shuai [16]. We extend their result to the case of \( d \)-dimensional hierarchical networks.

**Theorem 1** (Tree Cycle Identity for hierarchical graphs).

Let \( c_i \) be given as in Proposition 1. Then

\[
H(x) = \sum_{(i) \in V(G)} c_i F_{(i)}(x) = \sum_{P \in P} \omega(P) \sum_{\vec{j} \in E(C_P)} F_{\vec{j}}^i(x)
\]

**Proof.** By Proposition 1

\[
c_i = \sum_{T \in T_{(i)}} \omega(T).
\]

Then

\[
\sum_{(i) \in V(G)} c_i F_{(i)}(x) = \sum_{(i) \in V(G)} \left( \sum_{T \in T_{(i)}} \omega(T) \right) \left( \sum_{(j) \in N_{(i)}} a_{ij}^i F_{\vec{j}}^i(x) \right) = \sum_{(i) \in V(G)} \sum_{T \in T_{(i)}} \sum_{(j) \in N_{(i)}} a_{ij}^i \omega(T) F_{\vec{j}}^i(x).
\]

Now, for each triple \( ((i), (j), T) \) with \( (i) \in V(G), (j) \in N_{(i)} \) and \( T \in T_{(i)} \) there is a unique spanning unicyclic graph \( P \in P \) formed from \( T \) and the edge \( \vec{j}i \). Conversely, for each pair \( (Q, \vec{j}i) \) with \( Q \in P \) and \( \vec{j}i \in E(C_P) \) there is a unique \( T \in T_{(i)} \) resulting from the removal of edge \( \vec{j}i \). Thus there is a direct correspondence between triples of the form \( ((i), (j), T) \) and pairs of the form \( (Q, \vec{j}i) \). Then noting that \( a_{ij}^i \omega(T) = \omega(P) \), we see that

\[
H(x) = \sum_{(i) \in V(G)} \sum_{T \in T_{(i)}} \sum_{(j) \in N_{(i)}} a_{ij}^i \omega(T) F_{\vec{j}}^i(x) = \sum_{P \in P} \omega(P) \sum_{\vec{j}i \in E(C_P)} F_{\vec{j}}^i(x).
\]

\[\square\]

### 3.2 Constructing Lyapunov functions via feedback cycles

Constructing Lyapunov functions \( V \) for large-scale networked systems is difficult due to the unwieldy number of terms in the Lyapunov derivative of \( V \) with respect to system (1). Even when each of the \( M \)
vertices is assigned only a single state variable, the number of terms in the expression

$$\dot{V}(x) = \sum_{(i)\in V(G)} \frac{\partial V}{\partial x_i} \left( f_i(x_i) + \sum_{(j)\in V(G)} g_{ji}^i(x) \right)$$

varies as the square of $M$.

Theorem 1 provides us with a method of grouping terms in the Lyapunov function’s derivative in terms of cycles in $G$. This gives us an easily-verifiable condition to guarantee that $\dot{V}(x) \leq 0$ for all $x$.

**Definition 3.** We say that the set of edge functions $\{F_{ji}^i : (i), (j) \in V(G)\}$ satisfies the cycle condition if, for any directed cycle $C$ in $G$, we have

$$\sum_{j \in E(C)} F_{ji}^i(x) \leq 0.$$

**Theorem 2.** Consider a Lyapunov candidate function of the form

$$V(x) = \sum_{(i)\in V(G)} c_i V_i(x)$$

where each $V_i : \mathbb{R}^M \to \mathbb{R}$ is a function corresponding to vertex $(i)$ and $c_i$ is given as in Proposition 1. Suppose that there exists a set of edge functions $\{F_{ji}^i : (i) \in V(G), (j) \in N_i\}$ such that

a) we have

$$\dot{V}_{(i)}(x) \leq F_{(i)}(x)$$

and

b) the set of edge functions satisfies the cycle condition given in Definition 3.

Then $V$ satisfies $\dot{V}(x) \leq 0$. Namely, $V$ is a Lyapunov function for the system (1).

**Proof.** By Theorem 1 we have

$$\dot{V}(x) \leq \sum_{(i)\in V(G)} c_i \sum_{(j)\in N_i} a_{ji}^i F_{ji}^i(x) = \sum_{\mathcal{P} \in \mathcal{P}} \omega(\mathcal{P}) \sum_{j \in E(C_P)} F_{ji}^i(x) \leq 0.$$

The simplicity of this proof illustrates the power of Theorem 1. We now move on to two applications from very different contexts to demonstrate the wide range of applicability of these techniques.

### 4 Application: Globally stable clustering for multi-agent swarms

#### 4.1 Background

The mathematical study of flocking behaviour in multi-agent systems has found applications in diverse areas, from the study of bird flocks and schools of fish [22, 26] to the behaviour of human crowds in emergency situation [18, 25] to engineering applications such as the control of vehicle formations [4]. To accurately model systems of autonomous agents we require that the motion of the agents be based on decentralized control schemes that use only information available locally to each member of the flock. The method of artificial potentials [32] provides one method of stabilizing flocks based on local information.

In [28, 29], Tanner, Jadbabaie and Pappas provide a method of stabilizing swarms of agents using artificial potentials. Their analysis uses a Lyapunov function to demonstrate that their protocol stabilizes the formation at a local minimum of the artificial potentials and leads to the alignment of all agents.

Communication networks among agents considered in [28, 29] are either complete (i.e. every agent communicates with every other agent) or dynamic (i.e. communication links appear and disappear based on proximity of agents).
In the design of control protocols for multi-agent systems, it is desirable to reduce the number of communication links among agents and make more efficient use of computational resources. We propose a hierarchical control scheme that is more efficient in this sense than the complete networks considered in [28, 29]. The key steps of our hierarchical potential clustering (HPC) protocol are:

a) run a clustering algorithm based on the initial positions of the agents
b) assign leaders to each cluster
c) implement the artificial potential scheme (with complete interaction graph) within each cluster and
d) implement a velocity consensus scheme between the cluster leaders.

4.2 Model
We model each agent by specifying its position \( r = (x, y) \in \mathbb{R}^2 \) and velocity \( v = \dot{r} \). The dynamics of each agent is governed by the 4-dimensional system

\[
\dot{r} = v, \\
\dot{v} = u,
\]

where the thrust \( u \in \mathbb{R}^2 \) is the control input.

Before designing the control inputs, agents are clustered based on their position using an algorithm seeking to minimize the sum-squared distance within each cluster. See [1, 12] for a review of clustering algorithms. If we label the resulting clusters with integers \( 1, \ldots, N \) then the clustering induces a labeling of the agents \((i, j)\) where \( i \) is the cluster index and \( j \in \{1, \ldots, n_i\} \) is used to distinguish agents within cluster \( i \). We use the index \((i, 1)\) to designate the leader of cluster \( i \). Clustering algorithms such as the one considered in [9] provide a natural choice of leader called an exemplar. For algorithms that do not provide a leader, one can be assigned arbitrarily. The choice of a leader is not important for our purposes since we will show that stability is independent of the choice of leader. With this labeling of vertices, we obtain a system that describes the collection of agents

\[
\dot{r}_{ij} = v_{ij} \quad \text{(2)} \\
\dot{v}_{ij} = u_{ij}. \quad \text{(3)}
\]

Now that we have clustered the agents, we need to design control inputs \( u \) to bring each cluster of agents into formation. Within each cluster \( i \), we implement the artificial potential control scheme. For each pair of agents \((i, j)\) and \((i, k)\) in cluster \( i \) we define an artificial potential function \( P_{ij}^{ik} : \mathbb{R}_+ \to \mathbb{R} \). The potentials \( P_{ij}^{ik} \) can be chosen to be any positive, continuous functions of the distance \( d = ||r_{ij} - r_{ik}|| \) between agents \((i, j)\) and \((i, k)\) that are:

- symmetric \( P_{ij}^{ik} = P_{ik}^{ij} \)
- unbounded near zero \( \lim_{d \to 0} P_{ij}^{ik}(d) = \infty \)
- and that have a unique minimum at the desired distance between agents.
Lyapunov functions for hierarchical networks

\[ P(d) = \frac{1}{d^2} + \log d^2. \]

Figure 3: Artificial potential function \( P(d) \).

An example of an appropriate potential function is shown in Figure 3. We can then define the total potential for vertex \((i, j)\) as

\[ P_{ij} = \sum_{k=1}^{n_i} P_{ik}^{ij}. \]

We introduce some notation that will be useful in defining a control protocol for this system of agents. For each pair of vertices \((i, j)\) and \((i, k)\) from the same cluster, define the function \( u_{ik}^{ij} : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R} \) as

\[ u_{ik}^{ij}(r_{ij}, r_{ik}) = ||r_{ij} - r_{ik}||. \]

The position derivative \( \nabla r_{ij} \) of the total potential function \( P_{ij} \) is then defined as the 2-dimensional vector

\[ \nabla r_{ij} P_{ij} = \sum_{k=1}^{n_i} \left( \frac{\partial (P_{ik}^{ij} \circ u_{ik}^{ij})}{\partial x_{ij}}, \frac{\partial (P_{ik}^{ij} \circ u_{ik}^{ij})}{\partial y_{ij}} \right). \]

This notation, used in [28, 29], emphasizes the fact that the control protocol given in the following paragraph is essentially a modified gradient descent, in which each agent seeks to minimize its total potential using the local information available to it. The artificial potential control protocol combines this gradient descent with a consensus protocol in the heading \( \theta_{ij} \) as follows:

\[ u_{ij} = -\nabla r_{ij} P_{ij} - \sum_{k} \frac{(\theta_{ij} - \theta_{ik})||v_{ij}||}{||r_{ij} - r_{ik}||} \hat{n}(ij), \quad j \neq 1, \quad (4) \]

where \( \hat{n}(ij) \) is the unit vector orthogonal to \( v_{ij} \), and \( \theta_{ij} = \arctan(\dot{y}_{ij}, \dot{x}_{ij}) \) is the heading of agent \((i, j)\). Among the leaders of each cluster we apply an additional thrust to coordinate their velocities

\[ u_{i1} = -\nabla r_{i1} P_{i1} - \sum_{k} \frac{(\theta_{i1} - \theta_{ik})||v_{i1}||}{||r_{i1} - r_{ik}||} \hat{n}(i1) + \sum_{h \in N_i} b_{ih}(v_{h1} - v_{i1}), \quad (5) \]

where \( b_{ih} \geq 0 \) are the entries of an irreducible matrix \( B \). This ensures that the leader communication graph \( \mathcal{G}_B \) that encodes the interactions among the cluster leaders is strongly connected [3].

4.3 Results

We wish to demonstrate that the system defined by Equations (2)-(5) will eventually reach a desirable equilibrium.

**Definition 4.** A control protocol \( u \) is said to solve the formation stabilization problem if solutions of (2)-(3) converge asymptotically to a state in which
a) the relative positions of each agent \((i, j)\) within a cluster are such that a local minimum of the total vertex potential \(P_{ij}\) is achieved, and

b) for any two agents \((i, j)\) and \((h, k)\), their headings satisfy \(\theta_{ij} = \theta_{hk}\).

The following lemma is needed in the proof of our main theorem. It’s proof is the same as that for equation (7) in [29] and is therefore omitted.

**Lemma 1.**

\[
\dot{\theta}_{ij} = -\sum_{k=1}^{n_i} \frac{\theta_{ij} - \theta_{ik}}{|r_{ij} - r_{ik}|}.
\]

**Theorem 3.** Given any clustering scheme, the HPC protocol solves the formation stabilization problem provided that the leader communication graph \(G_B\) is strongly connected.

**Proof.** For each agent indexed \((i, j)\), define a vertex Lyapunov function

\[
V_{ij} = \frac{1}{2} \left( \sum_{k=1}^{N_i} P_{ikj}^2(|r_{ij} - r_{ik}|) + v_{ij} \cdot \dot{v}_{ij} + \theta_{ij}^2 \right),
\]

and let \(c = (c_1, \ldots, c_N)\) denote the vector of cofactors of the columns of \(L_G\), as given by Kirchhoff’s Matrix Tree Theorem. Define

\[
V = \sum_{i=1}^{N} c_i \sum_{j=1}^{n_i} V_{ij}.
\]

We will show that \(V\) is a Lyapunov function for the multi-agent systems (2)-(5). Since \(G_B\) is strongly connected, there is a tree rooted at each vertex of \(G_B\). Thus all entries of \(c\) are strictly positive and hence \(V\) is positive definite. Taking the derivative of \(V\) along trajectories of (2)-(5) we get

\[
\dot{V} = \sum_{i=1}^{N} c_i \sum_{j=1}^{n_i} \dot{V}_{ij}
\]

\[
= \sum_{i=1}^{N} c_i \sum_{j=1}^{n_i} \left[ \nabla_{r_{ij}} P_{ij} \cdot \dot{r}_{ij} + v_{ij} \cdot \dot{v}_{ij} + \theta_{ij} \dot{\theta}_{ij} \right]
\]

\[
= \sum_{i=1}^{N} c_i \sum_{j=1}^{n_i} \left[ \nabla_{r_{ij}} P_{ij} \cdot v_{ij} + v_{ij} \cdot \left( -\nabla_{r_{ij}} P_{ij} - \sum_k \frac{(\theta_{ij} - \theta_{ik})|v_{ij}|}{|r_{ij} - r_{ik}|} \hat{n}(ij) \right) + \theta_{ij} \dot{\theta}_{ij} \right]
\]

\[
+ \sum_{i=1}^{N} c_i \sum_{j=1}^{n_i} b_{ih}(v_{h1} - v_{i1}) \cdot v_{i1}
\]

\[
= \sum_{i=1}^{N} c_i \sum_{j=1}^{n_i} \left[ -v_{ij} \cdot \sum_k \frac{(\theta_{ij} - \theta_{ik})|v_{ij}|}{|r_{ij} - r_{ik}|} \hat{n}(ij) + \theta_{ij} \dot{\theta}_{ij} \right] + \sum_{i=1}^{N} c_i \sum_{h \in \mathcal{N}_i} b_{ih}(v_{h1} - v_{i1}) \cdot v_{i1}.
\]

By definition, the vector \(\hat{n}(ij)\) is orthogonal to \(v_{ij}\) so we can further simplify our expression for \(\dot{V}\).
$$\dot{V} = \sum_{i=1}^{N} c_i \sum_{j=i}^{n_i} \theta_{ij} \dot{\theta}_{ij} + \sum_{i=1}^{N} c_i \sum_{h \in N_i} b_{ih}(v_{ih} - v_{i1}) \cdot v_{i1}. \quad (6)$$

For each cluster $i$, define the cluster distance graph $G_i$ as the complete graph on $n_i$ vertices (without self-loops) where the edge $j\bar{k}$ is assigned a weight $\frac{1}{||r_{ij} - r_{ik}\|}$ corresponding to the squared distance between agents $(i, j)$ and $(i, k)$. We can also define the vector $\theta_i = (\theta_{i1}, \ldots, \theta_{in_i})^T \in \mathbb{R}^{n_i}$ to be the concatenation of the headings of all agents in group $i$. Using Lemma 1 we can rewrite the expression for $V$ in terms of the weighted Laplacian matrix $L_i$ of the graph $G_i$ as follows

$$\dot{V} = -\sum_{i=1}^{N} c_i \theta_i^T L_i \theta_i + \sum_{i=1}^{N} c_i \sum_{h \in N_i} b_{ih}(v_{ih} - v_{i1}) \cdot v_{i1}.$$

For each group $i$, the Laplacian $L_i$ is symmetric and positive semidefinite. Since $G_i$ is strongly connected, the nullspace of $L_i$ is spanned by the consensus vector $1 = (1, \ldots, 1)^T$ [24]. Therefore we have

$$\dot{V} \leq \sum_{i=1}^{N} c_i \sum_{h \in N_i} b_{ih}(v_{ih} - v_{i1}) \cdot v_{i1}$$

with equality if and only if we have heading consensus $\theta_{ij} = \theta_{ik}$ for all agents within each group $i$. We now have

$$\dot{V} \leq \sum_{i=1}^{N} c_i \sum_{h \in N_i} b_{ih}(v_{ih} - v_{i1}) \cdot v_{i1}$$

$$= \sum_{i=1}^{N} c_i \sum_{h \in N_i} b_{ih}(\bar{x}_{ih} \bar{x}_{i1} - \bar{x}_{i1}^2 + \bar{y}_{ih} \bar{y}_{i1} - \bar{y}_{i1}^2)$$

$$\leq \sum_{i=1}^{N} c_i \sum_{h \in N_i} b_{ih}(\bar{x}_{ih}^2 + \bar{x}_{i1}^2 - \bar{x}_{i1}^2 + \bar{y}_{ih}^2 + \bar{y}_{i1}^2 - \bar{y}_{i1}^2)$$

$$= \sum_{i=1}^{N} c_i \sum_{h \in N_i} b_{ih} \frac{1}{2}(\dot{x}_{ih}^2 + \dot{x}_{i1}^2 + \dot{y}_{ih}^2 + \dot{y}_{i1}^2).$$

It is easy to verify that the set of edge functions $\{F_{ih} = \frac{1}{2}(\dot{x}_{ih}^2 - \dot{x}_{i1}^2 + \dot{y}_{ih}^2 - \dot{y}_{i1}^2) : \ i, h \in V(G_B)\}$ satisfies the cycle condition. Thus by Theorem 2 we have $\dot{V} \leq 0$. We also have $\dot{V} = 0$ if and only if $\theta_{ij} = \theta_{ik}$ for all agents within each group $i$ and the velocities of the leaders of each group are identical ($v_{i1} = v_{k1}$ for all group indices $i, k \in \{1, \ldots, N\}$). Thus the system's configuration will converge to the largest invariant set $M$ contained in $\{(r, v) : \theta_{11} = \cdots = \theta_{Nn,N} \text{ and } v_{11} = \cdots = v_{N1}\}$. So we see that part (b) of Definition 4 is satisfied.

To see that part (a) is satisfied, consider the derivative of $P_{ij}$ along trajectories of the system (2)-(5). In $M$, we have

$$\dot{P}_{ij} = \nabla_{r_{ij}} P_{ij} \cdot v_{ij} = \nabla_{r_{ij}} P_{ij} \cdot (-\nabla_{r_{ij}} P_{ij}) \leq 0.$$

Thus trajectories converge to a configuration corresponding to a local minimum of $P_{ij}$. □

### 4.4 Simulations

To demonstrate our results numerically, we begin by randomly assigning initial positions and velocities to a set of 13 agents. Positions are assigned within a 10 by 10 area using a uniform distribution, and vertical and horizontal velocities are assigned uniformly in the range $[-1, 1]$ (Figure 4a).

To group agents into clusters, we perform a $k$-means clustering [12] with the parameter $k$ set to 4 (Figure 4b). Each cluster is assigned a leader at random and any pair of distinct agents within a cluster is linked by the artificial potential function $P(x) = \frac{1}{x^2} + \log x^2$ that has a global minimum at $x = 1$. The
resulting formation is shown in Figure 4d. The distances between agents within each group converge to a value of approximately 1 which minimizes the potential functions. Although it cannot be seen from the diagrams, the agent headings also converge.

Figure 4

5 Application: Global stability of the endemic equilibrium for multi-stage diseases in heterogeneous populations

5.1 Background

Compartmental models have been used extensively to model the spread of disease in large populations [10, 33]. The Kermack-McKendrick SIR model [13] was introduced in 1927 and provided some of the first mathematical insights into how disease spreads through a population then disappears. To model infections in a heterogeneous population more accurately, the basic SIR models are extended by dividing the population into homogeneous groups based on factors including age, size, spatial position, social grouping and gender. This leads to a networked multi-group model coupled by the spread of disease between groups [17, 30, 31].

Many diseases such as HIV/AIDS progress through a long period of time and exhibit distinctive disease stages. At different stages, individuals may have varying levels of both infectivity and mortality. Thus we require a compartmental model structured based on both disease stage and population heterogeneity. In [11], Jacques et al. present a multi-stage multi-group model of HIV/AIDS. They demonstrate that when an endemic exists, it is unique and locally asymptotically stable, and they conjecture that
it is in fact globally stable. In this section, we present a model similar to the Jacques model and use Theorem 2 to construct a global Lyapunov function to show that the endemic equilibrium for this model is globally asymptotically stable whenever it exists. While most multi-stage models in the literature only considers forward progression of the disease through stages, the network topology of our model is more general than that of Jacquez et al. in that we allow the disease to progress or ameliorate between any two stages.

5.2 Model

The dynamics of an infection in a population structured into $N$ groups, in which an infected individual from the $i$-th group might progress through some number $n_i$ of stages can be described by a system of differential equations in double-indexed variables of the form $x_{ij}$, the index $i$ to denote the group and the index $j$ to denote the stage of infection. Such a system is described by the multi-group, multi-stage epidemic model given by the following differential equations:

$$\begin{align*}
\frac{dx_{i1}}{dt} &= \Lambda_i - d_{i1}x_{i1} - \sum_{h=1}^{N} \sum_{k=2}^{n_h} \beta_{ihk}x_{hk}x_{i1} \\
\frac{dx_{i2}}{dt} &= \sum_{h=1}^{N} \sum_{k=2}^{n_h} \beta_{ihk}x_{hk}x_{i1} + \sum_{l=2}^{n_i} \delta_{il}^i x_{il} - \left( d_{i2} + \sum_{l=2}^{n_i} \delta_{il}^i \right) x_{i2} \\
\frac{dx_{ij}}{dt} &= \sum_{l=1}^{n_i} \delta_{ij}^l x_{il} - \left( d_{ij} + \sum_{l=2}^{n_i} \delta_{ij}^l \right) x_{ij}, \quad \text{if } j \notin \{1, 2\},
\end{align*}$$

where

- $x_{i1}$ denotes the population of the susceptible compartment of group $i$
- $x_{ij}$ denotes the population of the $j$-th infectious stage of group $i$ where $j$ ranges from 2 to $n_i$
- $\Lambda_i$ denotes the birth rate into group $i$
- $d_{ij}$ denotes the rate of removal from $x_{ij}$ (the sum of the death rate and rate of progression to the terminal stage)
- $\delta_{ij}^k$ denotes the rate of movement from infectious stage $k$ to stage $j$ in group $i$
- $\beta_{ihk}$ is the transmission coefficient for infection of susceptible individuals in group $i$ by individuals in group $h$ at infectious stage $k$.

![Figure 5: Flow between disease stages corresponding to group $i$.](image-url)
To study the stability of an equilibrium
\[ x^* = (x_{11}^*, x_{12}^*, \ldots, x_{NN}^*) , \]
we define the equilibrium flow graph \( \mathcal{G} \) as the hierarchical graph with block adjacency matrix
\[
A_{\mathcal{G}} = [a_{ij}^{hk}] = \begin{bmatrix}
D_1 & B_{12} & B_{13} & \cdots & B_{1N} \\
B_{21} & D_2 & B_{23} & \cdots & B_{2N} \\
B_{31} & B_{32} & D_3 & \cdots & B_{3N} \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
B_{N1} & B_{N2} & B_{N3} & \cdots & D_N
\end{bmatrix},
\]
where the diagonal blocks \( D_i \) encode the progression and transmission of disease within each group once the system has reached equilibrium
\[
D_i = \begin{bmatrix}
0 & \sum_{h=1}^{N} \sum_{k=2}^{N} \beta_i^{hk} x_{hk}^* x_{i1}^* & 0 & \cdots & 0 \\
\beta_i^{12} x_{i2}^* x_{i1}^* & \delta_i^{12} x_{i2}^* & \delta_i^{12} x_{i2}^* & \cdots & \delta_i^{12} x_{i2}^* \\
\beta_i^{13} x_{i3}^* x_{i1}^* & \delta_i^{13} x_{i3}^* & 0 & \cdots & \delta_i^{13} x_{i3}^* \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
\beta_i^{1n} x_{in}^* x_{i1}^* & \delta_i^{1n} x_{in}^* & \delta_i^{1n} x_{in}^* & \cdots & 0
\end{bmatrix},
\]
and the other blocks each have a single nonzero column that encodes the transmission of disease between groups
\[
B_{hi} = \begin{bmatrix}
0 & 0 & \cdots & 0 \\
\beta_i^{h2} x_{h2}^* x_{i1}^* & 0 & \cdots & 0 \\
\beta_i^{h3} x_{h3}^* x_{i1}^* & 0 & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
\beta_i^{hn} x_{hn}^* x_{i1}^* & 0 & \cdots & 0
\end{bmatrix}.
\]
The equilibrium flow graph is shown in Figure 7.
5.3 Results

The following two lemmas are needed for the proof of the main theorem of this section. Lemma 2 establishes a property of the Laplacian cofactors of graph $G$, and Lemma 3 gives a general property of the Laplacian cofactors that will be useful in the proof of the main theorem.

**Lemma 2.** For each $i \in \{1, \ldots, N\}$ the Laplacian cofactors $c_{i1}$ and $c_{i2}$ are identical.

**Proof.** For each $i \in \{1, \ldots, N\}$ the $(i,1)$-th row of $L_G$ contains only two nonzero entries:

- the diagonal entry $\sum_{h=1}^{N} \sum_{k=2}^{n_h} \beta_{ih}^{hk} x_{hk}^* x_{i1}^*$ corresponding to the in-degree of vertex $(i,1)$ and

- the entry in the $((i, 1)+1)$-th column: $- \sum_{h=1}^{N} \sum_{k=2}^{n_h} \beta_{ih}^{hk} x_{hk}^* x_{i1}^*$

(The negative of the $(1,2)$ entry of $D_i$)

The Laplacian matrix is always singular. So performing a cofactor expansion along the $(i,1)$-th row we get

$$0 = \det(L_G) = c_{i1} \sum_{h=1}^{N} \sum_{k=2}^{n_h} \beta_{ih}^{hk} x_{hk}^* x_{i1}^* - c_{i2} \sum_{h=1}^{N} \sum_{k=2}^{n_h} \beta_{ih}^{hk} x_{hk}^* x_{i1}^*$$

and thus $c_{i1} = c_{i2}$. \qed

**Lemma 3.** The vector $c = (c_{11}, \ldots, c_{Nn_N})$ of Laplacian cofactors is a right eigenvector of $L_G$ corresponding to the eigenvalue $0$.

**Proof.** Let $(L_G)_i$ denote the $i$-th row of $L_G$. Performing a cofactor expansion along the $i$-th row we see that $0 = \det(L_G) = (L_G)_i \cdot c$ for all $i$ and hence $L_G c = 0$. \qed

Figure 7: The equilibrium flow graph $G$. Edge weights are not labeled to avoid clutter.
Let $K$ denote the closed positive orthant of the state space. An equilibrium of the epidemiological system (7) is called endemic if it lies in the interior $\hat{K}$ of $K$. Biologically, points in the interior of $K$ correspond to states in which disease persists throughout the population. The following theorem shows that whenever an endemic equilibrium exists it is necessarily unique. Furthermore, any disease introduced into the population will persist for all time, and the solutions will converge asymptotically to the endemic equilibrium.

**Theorem 4.** If the graph $G$ is strongly connected and there exists an endemic equilibrium $x^* = (x^*_1, x^*_2, \ldots, x^*_n)$ of the system (7) then $x^*$ is unique and globally stable in $\hat{K}$.

**Proof.** Consider the Lyapunov function

$$V(x) = \sum_{i=1}^{N} \sum_{j=1}^{n_i} c_{ij} \left[ x_{ij} - x^*_{ij} - x^*_{ij} \ln \left( \frac{x_{ij}}{x^*_{ij}} \right) \right],$$

where $c_{ij}$ is the cofactor corresponding to the $(i, j)$-th column of $L_G$. Since $G$ is strongly connected, all the cofactors are positive and therefore $V$ is globally positive definite and radially unbounded. Its derivative along trajectories of the systems is given by

$$\dot{V}(x) = \sum_{i=1}^{N} \sum_{j=1}^{n_i} c_{ij} \left( 1 - \frac{x^*_{ij}}{x_{ij}} \right) \frac{dx_{ij}}{dt}$$

$$= \sum_{i=1}^{N} \left[ c_{i1} \left( 1 - \frac{x^*_{i1}}{x_{i1}} \right) \left( \Lambda_i - d_{i1}x_{i1} - \sum_{h=1}^{n_h} \sum_{k=1}^{n_k} \beta^h_{ik} x_{hk}x_{i1} \right) + c_{i2} \left( 1 - \frac{x^*_{i2}}{x_{i2}} \right) \left( \sum_{h=1}^{n_h} \sum_{k=1}^{n_k} \beta^h_{ik} x_{hk}x_{i2} + \sum_{l=1}^{n_l} \delta^l_{i2}x_{il} - \left( d_{i2} + \sum_{l=1}^{n_l} \delta^l_{i2} \right) x_{i2} \right) + \sum_{j=3}^{n_i} c_{ij} \left( 1 - \frac{x^*_{ij}}{x_{ij}} \right) \left( \sum_{l=1}^{n_l} \delta^l_{ij}x_{il} - \left( d_{ij} + \sum_{l=1}^{n_l} \delta^l_{ij} \right) x_{ij} \right) \right].$$

We can then use the equilibrium equations

$$\Lambda_i = d_{i1}x^*_{i1} + \sum_{h=1}^{n_h} \sum_{k=1}^{n_k} \beta^h_{ik} x^*_{hk}x^*_{i1},$$

$$\left( d_{i2} + \sum_{l=1}^{n_l} \delta^l_{i2} \right) x^*_{i2} = \sum_{h=1}^{n_h} \sum_{k=1}^{n_k} \beta^h_{ik} x^*_{hk}x^*_{i2} + \sum_{l=1}^{n_l} \delta^l_{i2}x^*_{il},$$

$$\left( d_{ij} + \sum_{l=1}^{n_l} \delta^l_{ij} \right) x^*_{ij} = \sum_{l=1}^{n_l} \delta^l_{ij}x^*_{il}, \quad \text{if } j \notin \{1, 2\},$$

to rewrite our expression for $\dot{V}$:

$$\dot{V}(x) = \sum_{i=1}^{N} \left[ c_{i1} \left( 1 - \frac{x^*_{i1}}{x_{i1}} \right) \left( d_{i1}x^*_{i1} + \sum_{h=1}^{n_h} \sum_{k=1}^{n_k} \beta^h_{ik} x^*_{hk}x^*_{i1} - d_{i1}x_{i1} - \sum_{h=1}^{n_h} \sum_{k=1}^{n_k} \beta^h_{ik} x_{hk}x_{i1} \right) \right.$$

$$\left. + c_{i2} \left( 1 - \frac{x^*_{i2}}{x_{i2}} \right) \left( \sum_{h=1}^{n_h} \sum_{k=1}^{n_k} \beta^h_{ik} x_{hk}x_{i2} + \sum_{l=1}^{n_l} \delta^l_{i2}x_{il} - \left( d_{i2} + \sum_{l=1}^{n_l} \delta^l_{i2} \right) x_{i2} \right) + \sum_{j=3}^{n_i} c_{ij} \left( 1 - \frac{x^*_{ij}}{x_{ij}} \right) \left( \sum_{l=1}^{n_l} \delta^l_{ij}x_{il} - \left( d_{ij} + \sum_{l=1}^{n_l} \delta^l_{ij} \right) x_{ij} \right) \right].$$
We can now expand these expressions, and group terms with a single variable in the numerator separately.

\[
\dot{V}(x) = \sum_{i=1}^{N} \left[ c_{i1} \sum_{h=1}^{N} \sum_{k=2}^{n_{h}} \beta_{i}^{h} x_{hk} x_{i1}^{*} + (c_{i2} - c_{i1}) \sum_{h=1}^{N} \sum_{k=2}^{n_{h}} \beta_{i}^{h} x_{hk} x_{i1} + c_{i2} \sum_{l=1}^{N_{i}} \delta_{il} x_{il} \right]
- c_{i2} \left( \sum_{h=1}^{N} \sum_{k=2}^{n_{h}} \beta_{i}^{h} x_{hk} x_{i1}^{*} - \sum_{l=1}^{N_{i}} \delta_{il} x_{il} \right) \frac{x_{i2}}{x_{i1}} + \sum_{j=3}^{N} c_{ij} \left( \sum_{l=1}^{N_{i}} \delta_{ij} \frac{c_{il}}{x_{ij}} - \sum_{l=1}^{N_{i}} \delta_{ij} \frac{c_{il}}{x_{il}} \right)
+ \sum_{i=1}^{N} \left[ \sum_{h=1}^{N} \sum_{k=2}^{n_{h}} \beta_{i}^{h} x_{hk} x_{i1}^{*} - \sum_{l=1}^{N_{i}} \delta_{il} x_{il} \right] \frac{x_{i2}}{x_{i1}} + c_{i1} \sum_{h=1}^{N} \sum_{k=2}^{n_{h}} \beta_{i}^{h} x_{hk} x_{i1} \left( 1 - \frac{x_{i1}^{*}}{x_{i1}} \right)
+ \sum_{j=2}^{N} c_{ij} \left( \sum_{l=1}^{N_{i}} \delta_{ij} \frac{c_{il}}{x_{ij}} - \sum_{l=1}^{N_{i}} \delta_{ij} \frac{c_{il}}{x_{il}} \right) \left( 1 - \frac{x_{i1}^{*}}{x_{i1}} \right)
= \begin{bmatrix} x_{i1}^{*} & \cdots & x_{iN_{N_{i}}}^{*} \\ \vdots \\ x_{i1}^{*} \\ \vdots \\ x_{iN_{N_{i}}}^{*} \end{bmatrix} \begin{bmatrix} L \mathcal{G} \end{bmatrix} \begin{bmatrix} c_{i1} \\ \vdots \\ c_{iN_{N_{i}}} \end{bmatrix}
+ \sum_{i=1}^{N} \left[ \sum_{h=1}^{N} \sum_{k=2}^{n_{h}} \beta_{i}^{h} x_{hk} x_{i1}^{*} - \sum_{l=1}^{N_{i}} \delta_{il} x_{il} \right] \frac{x_{i2}}{x_{i1}} + c_{i1} \sum_{h=1}^{N} \sum_{k=2}^{n_{h}} \beta_{i}^{h} x_{hk} x_{i1} \left( 1 - \frac{x_{i1}^{*}}{x_{i1}} \right)
+ \sum_{j=2}^{N} c_{ij} \left( \sum_{l=1}^{N_{i}} \delta_{ij} \frac{c_{il}}{x_{ij}} - \sum_{l=1}^{N_{i}} \delta_{ij} \frac{c_{il}}{x_{il}} \right) \left( 1 - \frac{x_{i1}^{*}}{x_{i1}} \right).
\]

Using Lemma 3 we see that the term involving the Laplacian is always zero. We also have

\[
2 - \frac{x_{i1}^{*}}{x_{i1}} - \frac{x_{i1}^{*}}{x_{i1}} \leq 0
\]
for all positive \( x_{i1} \), with equality if and only if \( x_{i1} = x_{i1}^* \). So our equation becomes

\[
\dot{V}(x) \leq \sum_{i=1}^{N} \left[ c_{i2} \sum_{h=1}^{N} \sum_{k=2}^{n_h} \beta_{h,k} x_{h,k}^* x_{i1}^* \left( 1 - \frac{x_{i1}^*}{x_{i1}} \right) \right. \\
+ c_{i1} \sum_{k=2}^{n_h} \sum_{h=1}^{N} \beta_{h,k} x_{h,k}^* x_{i1}^* \left( 1 - \frac{x_{i1}^*}{x_{i2} x_{h,k} x_{i1}} \right) \\
\left. + \sum_{j=2}^{n_i} \sum_{l=1}^{n_j} \delta_{ij} x_{il}^* \left( 1 - \frac{x_{i1}^*}{x_{ij} x_{il}} \right) \right].
\]

Table 1

<table>
<thead>
<tr>
<th>Edge Type</th>
<th>Connects Vertices</th>
<th>Edge Function</th>
</tr>
</thead>
<tbody>
<tr>
<td>Infection edge</td>
<td>((h, k) \rightarrow (i, 1)) ((k \neq 1))</td>
<td>(F_{i1}^{hk}(x_{i1}, x_{i2}, x_{h,k}) = \left( 1 - \frac{x_{i2} x_{h,k} x_{i1}}{x_{h,k} x_{i1}} \right) )</td>
</tr>
<tr>
<td>Disease acquisition edge</td>
<td>((i, 1) \rightarrow (i, 2))</td>
<td>(F_{i1}^{i2}(x_{i2}) = \left( 1 - \frac{x_{i2} x_{i1}}{x_{i1}} \right) )</td>
</tr>
<tr>
<td>Disease progression edge</td>
<td>((i, l) \rightarrow (i, j)) ((j, l \neq 1))</td>
<td>(F_{ij}^{il}(x_{ij}, x_{il}) = \left( 1 - \frac{x_{ij} x_{il}}{x_{ij} x_{il}} \right) )</td>
</tr>
</tbody>
</table>

Defining edge functions \( F_{ij}^{hk} \) as in Table 1 and noting that

\[
a_{i2}^{i1} = \sum_{h=1}^{N} \sum_{k=2}^{n_h} \beta_{i2} x_{h,k}^* x_{i1}^* \\
a_{i1}^{hk} = \beta_{i1} x_{h,k}^* x_{i1}^* \quad (k \neq 1) \\
a_{ij}^{il} = \delta_{ij} x_{il}^* \quad (i, l \neq 1)
\]

we see that

\[
\dot{V}(x) \leq \sum_{i=1}^{N} \left[ c_{i2} a_{i2}^{i1} F_{i2}^{i1}(x_{i1}) \right. \\
+ c_{i1} \sum_{h=1}^{N} \sum_{k=2}^{n_h} a_{i1}^{hk} F_{i1}^{hk}(x_{i1}, x_{i2}, x_{h,k}) \\
\left. + \sum_{j=2}^{n_i} \sum_{l=1}^{n_j} a_{ij}^{il} F_{ij}^{il}(x_{ij}, x_{il}) \right] \\
= \sum_{(i,j) \in V(G)} c_{ij} \sum_{(h,k) \in V(G)} a_{ij}^{hk} F_{ij}^{hk}(x).
\]

Thus hypothesis (a) of Theorem 2 is satisfied. To verify hypothesis (b), consider a cycle \( C \) of length \( l \) in the graph \( G_A \) (shown in Figure 7). By construction, an infection edge \((h, k) \rightarrow (i, 1)\) appears in the cycle if and only if it is followed by the disease acquisition edge \((i, 1) \rightarrow (i, 2)\). Thus we have

\[
\left( 1 - F_{i1}^{hk}(x) \right) \left( 1 - F_{i2}^{i1}(x) \right) = \left( x_{i2} x_{h,k} x_{i1} \right) \left( x_{i1} \right) = x_{i2} x_{h,k} x_{i1} \frac{x_{i1} x_{h,k}}{x_{i1} x_{h,k}} = x_{i2} x_{h,k} x_{i1}.
\]
Since the geometric mean of $l$ positive numbers is bounded by their arithmetic mean, we now have

$$
\sum_{(h,k)\rightarrow(i,j)} F_{ij}^{hk}(x) = l - \sum_{(h,k)\rightarrow(i,j)} (1 - F_{ij}^{hk}(x)) \\
\leq l - l \left( \prod_{(h,k)\rightarrow(i,j)} (1 - F_{ij}^{hk}(x)) \right)^{1/l} \\
= l - l \left( \prod_{(h,k)\rightarrow(i,j)} \frac{x_{ij}^* x_{hk}}{x_{ij} x_{hk}^*} \right)^{1/l} \\
= 0.
$$

So all the hypotheses of Theorem 2 are satisfied. Thus we have $\dot{V}(x) \leq 0$ for all $x \in K$. To have the equality $\dot{V}(x) = 0$, we require that $x_{ij}^* = x_{ij}^*$ for all $i \in \{1, \ldots, N\}$ and that $(1 - F_{ij}^{hk}) = (1 - F_{ij}^{h'k'})$ for any pair of edges $(h,k) \rightarrow (i,j)$ and $(h', k') \rightarrow (i', j')$ in the graph. It is easy to see that this occurs if and only if for all $(i,j) \in V(G)$ we have $x_{ij} = x_{ij}^*$. So $\dot{V}(x) = 0$ if and only if $x = x^*$. Therefore the endemic equilibrium $x^* = (x_{11}^*, x_{12}^*, \ldots, x_{NN}^*)$ is globally stable in $K$. The uniqueness of $x^*$ follows from its global stability.

\section*{Acknowledgments}

This research is supported in part by grants from the Natural Sciences and Engineering Research Council of Canada (NSERC) and Canada Foundation for Innovation (CFI). MYL acknowledges the financial support from the NCE-MITACS. JM acknowledges the financial support of an NSERC USRA awarded at the University of Alberta.

\section*{References}


