2019-04-01

Exponential Stability of Intrinsically Stable Dynamical Networks and Switched Networks with Time-Varying Time Delays

David Patrick Reber
Brigham Young University

Exponential Stability of Intrinsically Stable Dynamical Networks and Switched Networks with Time-Varying Time Delays

David Patrick Reber

A thesis submitted to the faculty of
Brigham Young University
in partial fulfillment of the requirements for the degree of

Master of Science

Benjamin Webb, Chair
Jared Whitehead
Todd Fisher

Department of Mathematics

Brigham Young University

ABSTRACT

Exponential Stability of Intrinsically Stable Dynamical Networks and Switched Networks with Time-Varying Time Delays

David Patrick Reber
Department of Mathematics, BYU
Master of Science

Dynamic processes on real-world networks are time-delayed due to finite processing speeds and the need to transmit data over nonzero distances. These time-delays often destabilize the network's dynamics, but are difficult to analyze because they increase the dimension of the network.

We present results outlining an alternative means of analyzing these networks, by focusing analysis on the Lipschitz matrix of the relatively low-dimensional undelayed network.

The key criteria, intrinsic stability, is computationally efficient to verify by use of the power method. We demonstrate applications from control theory and neural networks.

Keywords: time-varying time-delays, neural network, switched system

Contents

Chapter 1. Introduction

Interactions between network elements are inherently time-delayed since they are seperated in some way, for example in space, and information can only be transmitted and processed at finite rates. For example, control systems often operate in the presence of delays, primarily due to the time it takes to acquire the information needed for decision-making, to create control decisions, and to execute these decisions. Consequently, systems with delays arise in engineering, biology, physics, operations research, and economics [1]. To assume that these delays remain constant in time is unrealistic, since these delays are often generated by stochastic processes.

These time delays make analyzing the dynamics of a network more complicated, which is perhaps a reason why the theory of time-delayed systems is less developed than the theory of undelayed systems. There are significant differences between the dynamics of delayed and undelayed systems; for example, the presense of delays may either stabilize or destabilize a dynamical system. For example, an undelayed, stable feedback system may become unstable for some delays; yet, delaying decision-making allows may allow supply-chain managers to better observe consumer trends and consequently make better purchasing and stocking decisions. Hence, understanding the dynamics of a time-delayed system cannot always be done by analyzing the system without its time-delays [1].

However, since the theory of time-varying time-delayed systems is less developed, we would like to be able to identify those situations in which analysis of a delayed system reduces to analysis of the undelayed system. Providing such a criteria justifies ignoring time-delays when modeling real-world networks.

Recent progress has been made in this regard, by the use of a criteria called *intrinsic stability*. It has been demonstrated that discrete-time systems which satisfy this stronger notion of global stability have the remarkable property that they remain stable even if time-delays are added or removed from the system, provided that these delays are constant in time [2]. Furthermore, for systems with a differentiable mapping rule, verification of intrinsic stability is performed using the power method for eigenvalues [18], making this criteria computationally feasible for analyzing large real-world dynamical networks.

However, as described above, the condition that the delays remain constant in time is unrealistic for stochastically generated delays. Consequently, we demonstrate that dynamical networks which are intrinsically stable retain stability not only in the presence of constant time-delays, but also in the presence of time-varying time-delays.

Chapter 2. Dynamical Networks

A network is composed of a set of *elements*, which are the individual units that make up the network, and a collection of interactions between these elements. An *interaction* between two network elements can be thought of as an element's ability to directly influence the behavior of the other network element. More generally, there is a *directed interaction* between the i^{th} and j^{th} elements of a network if the j^{th} network element can influence the state of the i^{th} network element (where there may be no influence of the i^{th} network element on the j^{th}). The dynamics of a network can be fomalized as follows:

Definition 2.1. (Dynamical Network) *Let (X_i, d_i) be a complete metric space for $1 \leq i \leq n$. Let (X, d_{max}) be the complete metric space formed by endowing the product space $X = \bigoplus_{i=1}^n X_i$ with the metric*

$$d_{max}(\mathbf{x}, \mathbf{y}) = \max_i d_i(x_i, y_i) \quad \textit{where } \mathbf{x}, \mathbf{y} \in X, \textit{ and } \quad x_i, y_i \in X_i.$$

Let $F : X \to X$ be a continuous map, with i^{th} component function $F_i : X \to X_i$ given by

$$F_i = F_i(x_1, x_2, \ldots, x_n) \quad \textit{in which} \quad x_j \in X_j \quad \textit{for } j = 1, 2, \ldots, n$$

where it is understood that there may be no actual dependance of F_i on x_j. The dynamical system (F, X) generated by iterating the function F on X is called a dynamical network. *If an initial condition $\mathbf{x}^0 \in X$ is given, we define the k^{th}* iterate *of $\mathbf{x}^0$ as $\mathbf{x}^k = F^k(\mathbf{x}^0)$, with the sequential orbit $\{F^k(\mathbf{x}^0)\}_{k=0}^{\infty} = \{\mathbf{x}^0, \mathbf{x}^1, \mathbf{x}^2, \ldots\}$ in which $\mathbf{x}^k$ is the state of the network at time $k \geq 0$.*

The component funtion $F_i = F_i(x_1, x_2, \ldots, x_n)$ describes the network elements that influence the i^{th} element of the network, where there is a directed interaction between the i^{th} and j^{th} elements if F_i actually depends on x_j. For the intial condition $\mathbf{x}^0 \in X$ the state of the i^{th} element at time $k \geq 0$ is $x_i^k = (F^k(\mathbf{x}^0))_i \in X_i$ and $\{(F^k(\mathbf{x}^0))_i\}_{k=0}^{\infty} = \{x_i^0, x_i^1, x_i^2, \ldots\}$.

Intuitively, the state space $X = \bigoplus_{i=1}^{n} X_i$ consists of n nodes, and the mapping $F : X \to X$ consists of the interactions between all nodes at each time step.

While the defintion of a dynamical network allows for the function F to be defined on general products of complete metric spaces, for the sake of intuition and direct applications of the theory we develop in this paper, our examples will focus on dynamical networks in which $X = \mathbb{R}^n$ with the infinity norm $\|\mathbf{x}\|_\infty = \max_i |x_i|$. To give a concrete example of a dynamical network and to illustrate results throughout this paper, we will use Cohen-Grossberg neural (CGN) networks.

Example 2.2. *(Cohen-Grossberg Neural Networks) For $W \in \mathbb{R}^{n \times n}$, $\sigma : \mathbb{R} \to \mathbb{R}$, and $c_i, \epsilon \in \mathbb{R}$ for $1 \leq i \leq n$ let $(C, \mathbb{R}^n)$ be the dynamical network with components*

$$C_i(\mathbf{x}) = (1 - \epsilon)x_i + \sum_{j=1}^{n} W_{ij}\sigma(x_j) + c_i \quad 1 \leq i \leq n, \tag{2.1}$$

which is a special case of a Cohen-Grossberg neural network in discrete-time [4]. The function σ is assumed to be bounded, differentiable, and monotonically increasing, with Lipschitz constant K, i.e. $|\sigma(x) - \sigma(y)| \leq K |x - y|$ for all $x, y \in \mathbb{R}$.

In a Cohen-Grossberg neural network the variable x_i represents the *activation* of the i^{th} neuron. The function σ is a bounded monotonically increasing function, which describes the i^{th} neuron's response to inputs. The matrix W gives the interaction strengths between each pair of neurons and describes how the neurons are connected within the network. The constants c_i indicate constant inputs from outside the network.

In a *globally stable* dynamical network (F, X), the state of the network tends toward an equilibruim irrespective of its initial condition. That is, the network has a *globally attracting fixed point* $\mathbf{y} \in X$ such that for any $\mathbf{x} \in X$, $F^k(\mathbf{x}) \to \mathbf{y}$ as $k \to \infty$.

Global stability is observed in a number of important systems including neural networks [4, 5, 6, 7, 8], epidemic models [9], and the study of congestion in computer networks [10]. In such systems the globally attracting equilibrium is typically a state in which the network can carry out a specific task. Whether or not this equilibrium stays stable depends on a number of factors including external influences but also internal processes such as the network's own growth, both of which can destabilize the network.

To give a sufficient condition under which a network (F, X) is stable, we define a *Lipschitz matrix* (called a *stability matrix* in [2]).

Definition 2.3. (Lipschitz Matrix) *For $F : X \to X$ suppose there are finite constants $a_{ij} \geq 0$ such that*

$$d_i(F_i(\mathbf{x}), F_i(\mathbf{y})) \leq \sum_{j=1}^{n} a_{ij} d_j(x_j, y_j) \quad \text{for all } \mathbf{x}, \mathbf{y} \in X.$$

Then we call $A = [a_{ij}] \in \mathbb{R}^{n \times n}$ a Lipschitz matrix *of the dynamical network (F, X).*

It is worth noting that if A is a Lipschitz matrix of a dynamical network then any matrix $B \preceq A$, where $\preceq$ denotes the element-wise inequality, is also a Lipschitz matrix of the network. However, if the function $F : X \to X$ is piecewise differentiable and each $X_i \subseteq \mathbb{R}$ then the matrix $A \in \mathbb{R}^{nxn}$ given by

$$A_{ij} = \sup_{x \in X} \left| \frac{\partial F_i}{\partial x_j}(\mathbf{x}) \right| \tag{2.2}$$

is the Lipschitz matrix of minimal spectral radius for (F, X) (see [2]). From a computational point of view, the Lipschitz matrix $A = [a_{ij}]$ of (F, X) can be more straightforward to find by use of Equation (2.2) if the function $F : X \to X$ is differentiable, compared to the more general formulation in Definition 2.3.

Using Equation (2.2) it follows that the Lipschitz matrix A of the Cohen-Grossberg Neural Network from Example 2.2 is given by

$$a_{ij} = \begin{cases} |1 - \epsilon| + K\,|W_{ii}| & \text{if } i = j \\ K\,|W_{ij}| & \text{otherwise.} \end{cases} \tag{2.3}$$

It is straightforward to verify that a Lipschitz matrix exists for a dynamical network (F, X) if and only if the mapping F is Lipschitz continuous. The idea is to use the Lipschitz matrix to simplify the stability analysis of nonlinear networks, using the following theorem of [2]. Here

$$\rho(A) = max_i\{|A| : A \in \sigma(A)\}$$

denotes the spectral radius of a matrix A.

Theorem 2.4. (Network Stability) *Let A be a Lipschitz matrix of a dynamical network (F, X). If $\rho(A) < 1$, then (F, X) is stable.*

It is worth noting that if we use the Lipschitz matrix A to define the dynamical network (G, X) by $G(\mathbf{x}) = A\mathbf{x}$ then (G, X) is stable if and only if $\rho(A) < 1$. Thus, a Lipschitz matrix of a dynamical network (F, X) can be thought of as the worst-case linear approximation to F. If this approximation has a globally attracting fixed point, then the original dynamical network (F, X) must also be stable. Note, however, that the condition that $\rho(A) < 1$ is sufficent but not neccesary for (F, X) to be stable. In fact, this stronger condition implies much more than network stability, so following the convention introduced in [2] we assign it the name of *intrinsic stability.*

Definition 2.5. (Intrinsic Stability) *Let $A \in \mathbb{R}^{nxn}$ be a Lipschitz matrix of a dynamical network (F, X). If $\rho(A) < 1$, then we say (F, X) is* intrinsically stable*.*

The Cohen-Grossberg neural network $(C, \mathbb{R}^{n\times n})$ has the stability matrix $A = |1-\epsilon|\, I + K\,|W|$ given by Equation (2.3) with spectral radius

$$\rho(A) = |1-\epsilon| + K\rho(|W|)$$

Here, $|W| \in \mathbb{R}^{n\times n}$ is the matrix W in which we take the absolute value of each entry. Thus, $(C, \mathbb{R}^{n\times n})$ is intrinsically stable if $|1-\epsilon| + K\rho(|W|) < 1$.

Chapter 3. Constant-Time-Delayed Dynamical Networks

As mentioned in the introduction, the dynamics of most real networks are *time-delayed*. That is, an interaction between two network elements will typically not happen instantaneously but will be delayed due to either the physical separation of these elements, their finite processing speeds, or be delayed due to other factors. We formalize this phenonenom by introducing a *delay distribution matrix* $D = [d_{ij}]$ into a dynamical network (F, X), where each d_{ij} is a nonnegative integer denoting the constant number of discrete time-steps by which the interaction from the j^{th} network element to the i^{th} network element is delayed.

Definition 3.1. (Constant Time-Delayed Dynamical Network)
Let (F, X) be a dynamical network and $D = [d_{ij}] \in \mathbb{N}^{n\times n}$ a delay distribution matrix with $\max_{i,j} d_{ij} \leq L$*, a bound on the delay length. Let X_L, the* extension *of X to delay-space, be defined as*

$$X_L = \bigoplus_{\ell=0}^{L} \bigoplus_{i=1}^{n} X_{i,\ell} \quad \text{where} \quad X_{i,\ell} = X_i \quad \text{for} \quad 1 \leq i \leq n \quad \text{and} \quad 0 \leq \ell \leq L.$$

Componentwise, define $F_D : X_L \to X_L$ *by*

$$(F_D)_{i,\ell+1} : X_{i,\ell} \to X_{i,\ell+1} \quad \textit{given by the identity} \quad (F_D)_{i,\ell+1}(x_{i,\ell}) = x_{i,\ell} \quad \textit{for} \quad 0 \le \ell \le L-1 \tag{3.1}$$

and

$$(F_D)_{i,0} : \bigoplus_{j=1}^{n} X_{j,d_{ij}} \to X_{i,0} \quad \textit{given by} \quad (F_D)_{i,0} = F_i(x_{1,d_{i1}}, x_{2,d_{i2}}, \ldots, x_{n,d_{in}}) \tag{3.2}$$

where $F_i : X \to X_i$ *is the* i^{th} *component function of* F *for* $i = 1, 2, \ldots, n$. *Then* (F_D, X_L) *is the the* delayed version *of* F *corresponding to the fixed-delay distribution* D *with* delay bound L.

We order the component spaces of X_L in the following way. If $\mathbf{x} \in X_L$ then

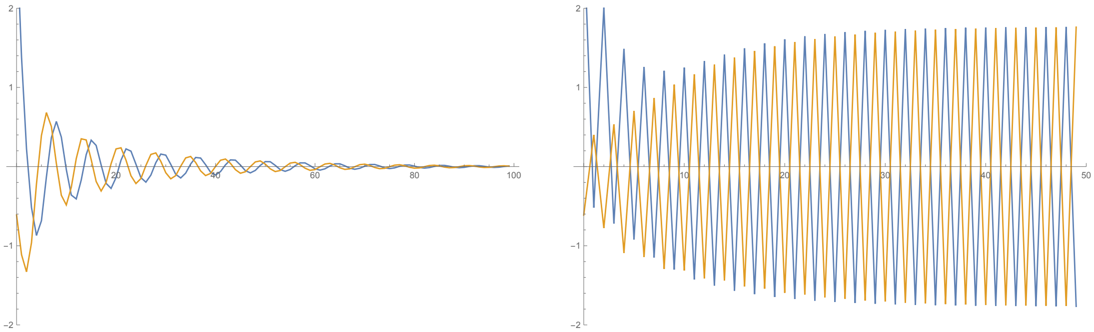
$$\mathbf{x} = [x_{1,0}, x_{2,0}, \ldots, x_{n,0}, x_{1,1}, x_{2,1}, \ldots, x_{n,L}]^T$$

where $x_{i,\ell} \in X_{i,\ell}$ for $i = 1, 2, \ldots, n$ and $\ell = 0, 1, \ldots, L$.

The formalization in Definition 3.1 captures the idea of adding a time-delay to an interaction: Each X_i is effectively copied L times, and past states of the i^{th} element are passed down this chain by the identity component functions $(F_D)_{i,\ell+1}$ for $0 \le \ell \le L-1$ in Equation (3.1). Finally, when a state of the i^{th} element has been passed through the chain d_{ij} times over d_{ij} time-steps it then influences the i^{th} network element, as captured by the entry-wise substitutions of $x_{j,d_{ij}}$ for x_j in $(F_D)_{i,0}$ in Equation (3.2).

Example 3.2. *Consider a simple 2-neuron version of a Cohen-Grossberg neural network* (C, X) *given by*

$$C\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = \begin{bmatrix} C_1(x_1, x_2) \\ C_2(x_1, x_2) \end{bmatrix} = \begin{bmatrix} (1-\epsilon)x_1 + W_{11}\phi(x_1) + W_{12}\phi(x_2) + c_1 \\ (1-\epsilon)x_2 + W_{21}\phi(x_1) + W_{22}\phi(x_2) + c_2 \end{bmatrix}, \tag{3.3}$$



Stable Network (C, X) Constant Time-Delayed Network (C_D, X_3)

Figure 3.1: Left: The stable dynamics of the two-neuron Cohen-Grossberg network (C, X) from Example 3.2 is shown. Right: The unstable dynamics of the constant time-delayed version of this network (C_D, X_3) is shown with the delay distribution given by the matrix D in Equation (3.4).

in which $X = \mathbb{R}^2$. For the delay distribution D given by

$$D = \begin{bmatrix} 1 & 2 \\ 1 & 3 \end{bmatrix}, \tag{3.4}$$

which has a maximum delay of $L = 3$ the time-delayed network (C_D, X_3) is given by

$$C_D\left(\begin{bmatrix} x_{1,0} \\ x_{2,0} \\ x_{1,1} \\ x_{2,1} \\ x_{1,2} \\ x_{2,2} \\ x_{1,3} \\ x_{2,3} \end{bmatrix}\right) = \begin{bmatrix} C_1(x_{1,1}, x_{2,2}) \\ C_2(x_{1,1}, x_{2,3}) \\ x_{1,1} \\ x_{2,1} \\ x_{1,2} \\ x_{2,2} \\ x_{1,3} \\ x_{2,3} \end{bmatrix} = \begin{bmatrix} (1-\epsilon)x_{1,1} + W_{11}\phi(x_{1,1}) + W_{12}\phi(x_{2,2}) + c_1 \\ (1-\epsilon)x_{2,3} + W_{21}\phi(x_{1,1}) + W_{22}\phi(x_{2,3}) + c_2 \\ x_{1,0} \\ x_{2,0} \\ x_{1,1} \\ x_{2,1} \\ x_{1,2} \\ x_{2,2} \end{bmatrix}$$

in which $X_3 = \mathbb{R}^8$. The time-delayed network (C_D, X_3) is the same as the original network (C, X) except that the state of $x_{1,0}$ gets passed through one identity mapping before it is input into F_1 and twice before it is input into F_2. Similarly, $x_{2,0}$ gets passed through one identity *mapping before it is input into F_1 and three identity mappings before it is input into F_2.*

A natural question is to ask is whether constant time-delays affect the stability of a network. We note that if

$$W = \begin{bmatrix} 0 & -\frac{3}{4} \\ \frac{3}{4} & 0 \end{bmatrix},$$

$c_1 = c_2 = 0$, $\sigma(x) = tanh(x)$, and $\epsilon = \frac{2}{5}$ then the dynamical network (C, X) given in Equation (3.3) is stable as can be seen in Figure (3.1) (left). However, the time-delayed version (C_D, X_3) of this network is unstable as is shown in the same figure (right). That is, the time-delays given by the delay distribution D have a destabilizing effect on this network.

An important fact about the network constructed in this example is that its spectral radius

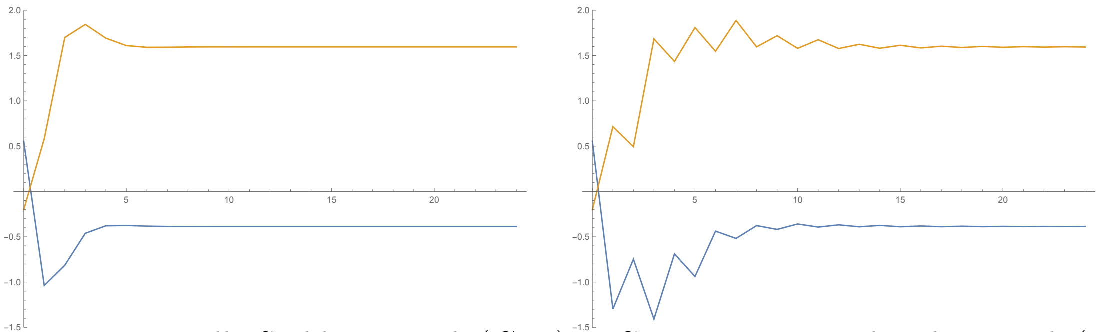
$$\rho(A) = |1 - \epsilon| + K\rho(|W|) = 1.35 > 1.$$

That is, although (C, X) is stable it is not intrinsically stable.

In [1], the authors demonstate that intrinsically stable systems are resilient to the addition of constant time-delays, as is stated in the following theorem.

Theorem 3.3. (Intrinsic Stability and Constant Delays) *Let (F, X) be a dynamical network and $D = [d_{ij}]$ a delay-distribution matrix. Let L satisfy $\max_{i,j} d_{ij} \leq L$. Then (F, X) is intrinsically stable if and only if (F_D, X_L) is intrinsically stable.*

Furthermore, we note that any fixed point(s) of an undelayed network (F, X) will also be fixed point(s) of any delayed version (F_D, X_L). This is formalized in the following proposition, and proven in the Appendix. Before stating this proposition, we require the following definition.



Intrinsically Stable Network (C, X) Constant Time-Delayed Network (C_D, X_3)

Figure 3.2: Left: The dynamics of the intrinsically stable network (C, X) from Example 3.6 is shown. Right: The stable dynamics of the constant time-delayed version of this network (C_D, X_3) is shown with the delay distribution given by the matrix D in Equation (3.4). Both systems are attracted to the fixed point $\mathbf{x}^* = (-.386, 1.595)$.

Definition 3.4. (Extension of a Point to Delay-Space) *Let $E_L(\mathbf{x}) \in X_L$ be equal to $L+1$ copies of $\mathbf{x} \in X$ stacked into a single vector, namely*

$$E_L(\mathbf{x}) = \begin{bmatrix} \mathbf{x}_0 \\ \mathbf{x}_1 \\ \vdots \\ \mathbf{x}_L \end{bmatrix} \quad \text{where} \quad \mathbf{x}_\ell = \mathbf{x} \quad \text{for} \quad 0 \leq \ell \leq L.$$

Proposition 3.5. (Fixed Points of Delayed Networks) *Let $\mathbf{x}^*$ be a fixed point of a dynamical network (F, X). Then for all delay distributions D with $\max_{ij} d_{ij} \leq L$, $E_L(\mathbf{x}^*)$ is a fixed point of (F_D, X_L).*

As an immediate consequence of Proposition 3.5 and Theorem 3.3, if an undelayed network (F, X) is intrinsically stable with a globally attracting fixed point $\mathbf{x}^* \in X$, then the delayed version (F_D, X_L) will have the "same" globally attracting fixed point $E_L(\mathbf{x}^*)$, in that $\mathbf{x}^*$ is the restriction of $\mathbf{y} = E_L(\mathbf{x}^*)$ to the first n component spaces of X_L. Thus, the asymptotic dynamics of an intrinsically stable network and any version of the network with constant time delays are essentially identical.

Example 3.6. *For example, consider the Cohen-Grossberg neural network* (C, X) *and delay matrix* D *given in Example 3.2 where* W *and* σ *are as before but* $\epsilon = \frac{4}{5}$, $c_1 = -1$ *and* $c_2 = 1$. *Since* $|1 - \epsilon| + \rho(|W|) = .95 < 1$ *then* (C, X) *is intrinsically stable with the globally attracting fixed point* $\mathbf{x}^* = (-.386, 1.595)$ *as seen in Figure 3.2 (left). Since* (C, X) *is intrinsically stable then Theorem 3.3 together with Proposition 3.5 imply that not only is* (C_D, X_3) *stable but its globally attracting fixed point is* $E_3(\mathbf{x}^*)$. *This is shown in Figure 3.2 (right).*

This result justifies the modeling choice of ignoring constant time-delays when analyzing intrinsically stable real-world neworks, but does not account for the potential time-dependance of delays arising from external or stochastic influences. Consequently, the main results of this paper presented in the next chapter focuses on strengthening the conclusion of Theorem 3.3.

Chapter 4. Time-Varying Time-Delayed Dynamical Networks

As the title of this chapter suggests, constant time-delays are not the only type of time delays that can occur in dynamical networks. More importantly, time-delays that vary with time occur in many real-world networks and in such systems are a significant source of instability [1]. It is worth noting that modeling such delays introduces even more complexity into models of dynamical networks that can already be quite complicated. This can hinder the tractability of analyzing such systems.

In order to define a network with time-varying time-delays, we first define the more general concept of a *switched network*.

Definition 4.1. (Switched Network) *Let* M *be a set of Lipschitz continuous mappings on* X*, such that for every* $F \in M$*,* (F, X) *is a dynamical network. Then we call* (M, X) *a* switched network *on* X*. Given some sequence* $\{F^{(k)}\}_{k=1}^{\infty} \subset M$*, we say that* $(\{F^{(k)}\}_{k=1}^{\infty}, X)$

is an instance of (M, X)*, with orbits determined at time* k *by the function*

$$\mathscr{F}^k(\mathbf{x}) = F^{(k)} \circ \ldots \circ F^{(2)} \circ F^{(1)}(\mathbf{x}) \quad \textit{for} \quad \mathbf{x} \in X.$$

For the switched network (M, X) *we construct a* Lipschitz set S *consisting of a set of* $n \times n$ *matrices as follows: For each* $F \in M$*, contribute exactly one Lipschitz matrix* A *of* (F, X) *to the set* S*.*

(M, X) is an ensemble of dynamical systems formed by taking all possible sequences of mappings $\{F^{(k)}\}_{k=1}^{\infty} \subset M$. For a switched network (M, X), the set S serves an analagous purpose to the Lipschitz matrix A of a dynamical network (F, X), as we soon demonstrate. Before describing this we consider the following example.

Example 4.2. *Let* $(P, \mathbb{R}^2)$ *and* $(Q, \mathbb{R}^2)$ *be the simple dynamical networks given by*

$$P(\mathbf{x}) = \begin{bmatrix} \epsilon & 1 \\ 0 & \epsilon \end{bmatrix} \mathbf{x} \quad \textit{and} \quad P(\mathbf{x}) = \begin{bmatrix} \epsilon & 0 \\ 1 & \epsilon \end{bmatrix} \mathbf{x} \quad \textit{for small} \quad \epsilon << 1.$$

For $M = \{P, Q\}$ *let* $\{F^{(k)}\}_{k=1}^{\infty}$ *be the sequence that alternates between* P *and* Q*, i.e.* $F^{(k)} = P$ *if* k *is odd and* $F^{(k)} = Q$ *if* k *is even. Note that if we let* $U = Q \circ P$ *then*

$$U(\mathbf{x}) = \begin{bmatrix} \epsilon^2 & \epsilon \\ \epsilon & 1 + \epsilon^2 \end{bmatrix} \quad \textit{with} \quad \rho(U) = \frac{1}{2}(1 + 2\epsilon^2 + \sqrt{1 + 4\epsilon^2}) > 1.$$

Since $\mathcal{F}^{(2k)}(\mathbf{x}) = U \circ \ldots \circ U(\mathbf{x})$ *for* $\mathbf{x} \in \mathbb{R}^2$ *then, as* $U(\mathbf{x})$ *is a linear system* $\lim_{k\to\infty} \mathcal{F}^{(2k)}(\mathbf{x}) = \infty$ *for any* $\mathbf{x} \neq 0$*. This is despite the fact that both* $(P, \mathbb{R}^2)$ *and* $(Q, \mathbb{R}^2)$ *are intrinsically stable both having the globally attracting fixed point* $\mathbf{0}$*.*

The issue is that although the spectral radius of both P and Q in this example are arbitrarily small, their joint spectral radius is not.

Definition 4.3. (Joint Spectral Radius) *Given some $\mathbf{z}^0 \in \mathbb{R}^n$ and some set of matrices $S \subset \mathbb{R}^{n\times n}$, let $\mathbf{z}^k = A_k \dots A_2 A_1 \mathbf{z}^0$ for some sequence $\{A_i\}_{i=1}^{\infty} \subset S$. The* joint spectral radius *$\overline{\rho}(S)$ of the set of matrices S is the smallest value $\overline{\rho} \geq 0$ such that for every $\mathbf{z}^0 \in \mathbb{R}^n$ there is some constant $C > 0$ for which*

$$\left\|\mathbf{z}^k\right\| \leq C(\overline{\rho})^k.$$

It is known that $\{\mathbf{z}^k\}_{k=1}^{\infty}$ converges to the origin for all $\mathbf{z}^0 \in \mathbb{R}^n$ if and only if $\overline{\rho}(S) < 1$ [11]. This allows us to state the following result regarding the asymptotic behavior of a nonlinear switched network (M, X) whose Lipschitz set S satisfies $\overline{\rho}(S) < 1$.

Main Result 1. (Independence of Initial Conditions for Switched Networks) *Let S be a Lipschitz set of a switched network (M, X) satisfying $\overline{\rho}(S) < 1$, and let $(\{F^{(k)}\}_{k=1}^{\infty}, X)$ be an instance of (M, X). Then for all initial conditions $\mathbf{x}^0, \mathbf{y}^0 \in X$, there exists some $C > 0$ such that*

$$d_{max}(\mathscr{F}^k(\mathbf{x}^0), \mathscr{F}^k(\mathbf{y}^0)) \leq C\overline{\rho}(S)^k.$$

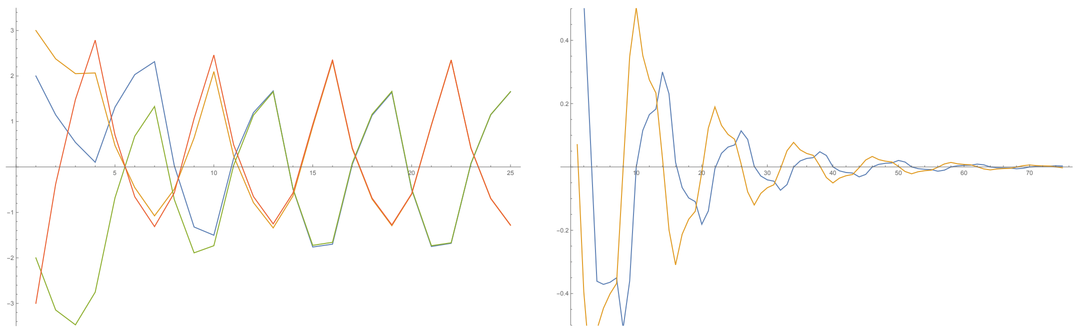
Additionally, if $\mathbf{x}^*$ is a shared fixed point of (F, X) for all $F \in M$, then $\lim_{k\to\infty} \mathscr{F}^k(\mathbf{x}^0) = \mathbf{x}^*$ for all initial conditions $\mathbf{x}^0 \in X$.

Hence, if the joint spectral radius of the Lipschitz set S of M is less than 1, all orbits in a switched network become asymptotically close to one another as time goes to infinity. Even if this limit-orbit is not convergent to any fixed point, it demonstrates an independence of asymptotic behavior to initial conditions. An example of this is the following.

Example 4.4. *Let (G, X) and (H, X) be the Cohen-Grossberg neural networks given by*

$$G\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = \begin{bmatrix} (1-\epsilon_1)x_1 - \frac{3}{4}\phi(x_2) + c_1 \\ (1-\epsilon_1)x_2 + \frac{3}{4}\phi(x_1) + c_2 \end{bmatrix} \quad \text{and} \quad H\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = \begin{bmatrix} (1-\epsilon_2)x_1 + \frac{1}{4}\phi(x_2) + d_1 \\ (1-\epsilon_2)x_2 + \frac{1}{4}\phi(x_1) + d_2 \end{bmatrix},$$

respectively, in which $\sigma(x) = \tanh(x)$, $\epsilon_1 = \frac{4}{5}$, $\epsilon_2 = \frac{3}{10}$, $c_1 = d_2 = -1$, and $c_2 = d_1 = 1$. Setting $M = \{G, H\}$ and letting S be the Lipschitz set of M given by



Independence to Initial Conditions   Dynamics with a Shared Fixed Point

Figure 4.1: Left: The dynamics of the switched network in Example 4.4 is shown for two different initial conditions $\mathbf{x}^0 = (2, 3)$ (shown in blue and yellow) and $\mathbf{y}^0 = (-2, -3)$ (shown in green and red). As the corresponding joint spectral radius $\overline{\rho}(S)$ of the network is less than 1, the orbits of these initial conditions converge to each other. Right: Modifying this switched network so that both $F, G \in M$ have the shared fixed point $\mathbf{0}$ results in a stable switched system with with the globally attracting fixed point $\mathbf{0}$.

$$S = \left\{ \begin{bmatrix} \frac{1}{5} & \frac{3}{4} \\ \frac{3}{4} & \frac{1}{5} \end{bmatrix}, \begin{bmatrix} \frac{7}{10} & \frac{1}{4} \\ \frac{1}{4} & \frac{7}{10} \end{bmatrix} \right\}.$$

It can be shown that the joint spectral radius $\overline{\rho}(S) = .95$. As this is less than 1, then for any instance $(\{F^{(k)}\}_{k=1}^{\infty}, X)$ of M and any initial conditions $\mathbf{x}^0$ and $\mathbf{y}^0$, we have

$$\lim_{k\to\infty} d_{max}(\mathcal{F}^k(\mathbf{x}^0), \mathcal{F}^k(\mathbf{y}^0)) = 0$$

by the Main Result 1. This can be seen in Figure 4.1 (left) where $(\{F^{(k)}\}_{k=1}^{\infty}, X)$ is the instance given by $\{F^{(k)}\}_{k=1}^{\infty} = \{G, G, G, H, H, H, \dots\}$.

If we set $c_1 = c_2 = d_1 = d_2 = 0$ in both (G, X) and (H, X) then both systems have the shared fixed point $\mathbf{0}$. In this case Main Result 1 indicated that any instance $(\{F^{(k)}\}_{k=1}^{\infty}, X)$ of M will be stable with the globally attracting fixed point $\mathbf{0}$. This is shown in Figure 4.1 (right) where again $\{F^{(k)}\}_{k=1}^{\infty} = \{G, G, G, H, H, H, \dots\}$.

Thus, as might be expected from the complicated nature of a switched system, the condition $\overline{\rho}(S) < 1$ alone is not able to match the strong implication of global stability, as $\rho(A) < 1$ does for a dynamical network (see Theorem [3.3]). Furthermore, it is often notoriously difficult to compute or approximate the joint spectral radius $\overline{\rho}(S)$ of a general set of matrices S [11].

Somewhat surprisingly, these issues are resolved when our switched system arises from a network experiencing time-varying time-delays. In this case, the computation of the joint spectral radius reduces to computing the spectral radius of the Lipschitz matrix of the original undelayed dynamical network (F, X), which for even large systems can be done efficiently using the power method [18]. This provides a general and computationally efficient method for verifying asymptotic stability despite time-varying time-delays.

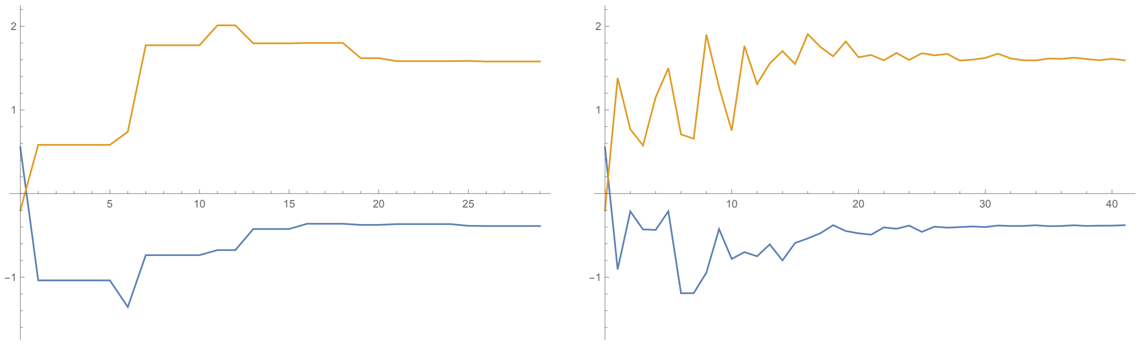
Main Result 2. (Intrinsic Stability and Time-Varying Time-Delayed Networks) *Suppose (F, X) is intrinsically stable with $\rho(A) < 1$, where $A \in \mathbb{R}^{n\times n}$ is a Lipschitz matrix of F and $\mathbf{x}^*$ is the network's globally attracting fixed point. Let $L > 0$ and*

$$M_d = \{F_D | D \in \mathbb{N}^{n\times n} \text{ with } \max_{ij} d_{ij} \leq L\}$$

and S_d be the Lipschitz set of M_d.
Then $E_L(\mathbf{x}^*)$ is a globally attracting fixed point of every instance $(\{F_{D^{(k)}}\}_{k=1}^{\infty}, X_L)$ of (M_d, X_L). Furthermore, $\overline{\rho}(S_d) = \rho(A_L) < 1$, where

$$A_L = \begin{bmatrix} \mathbf{0}_{n\times nL} & A \\ \mathbf{I}_{nL\times nL} & \mathbf{0}_{nL\times n} \end{bmatrix}.$$

Hence, any intrinsically stable dynamical network (F, X) retains convergence to the same equilibrium even when it experiences time-varying time-delays. This extends Theorem 2.3 of [2] to the much larger and more complicated class of switching-delay networks. Furthermore, note that intrinsic stability is a delay-independant result, which makes no assumption regarding the rate of growth of the time delays. In Main Result 2 we call $\overline{\rho}(M)$ the **convervence rate** of the system, since it provides the exponential rate to which all orbits converge to the fixed point $E_L(\mathbf{x}^*)$.



Periodically Delayed Network (C_P, X_5) Stochastically Delayed Network (C_U, X_{10})

Figure 4.2: Left: The dynamics of the intrinsically stable two-neuron Cohen-Grossberg network (C, X) from Example 3.6 is shown in which the network has periodic time-varying time-delays. Right: The dynamics of the same Cohen-Grossberg network is shown in which the network has stochastic time-varying time-delays. Both systems are attracted to the fixed point $\mathbf{x}^* = (-.386, 1.595)$ similar to the behavior in Figure 3.2.

It is worth emphasizing that as a consequence of this result, to determine the asymptotic behavior of a network (M, X_L) with $M = \{F_D | D \in \mathbb{N}^{n\times n}$ with $\max_{ij} d_{ij} \leq L\}$, in which the presence and magnitude of time delays is not exactly known, it suffices to study the dynamics of the much simpler undelayed system (F, X).

Example 4.5. (Periodic and Stochastic Time-Varying Delays) *Consider the intrinsically stable Cohen-Grossberg neural network (C, X) from Example 3.6 with the exception that we let $\{D^{(k)}\}_{k=1}^{\infty}$ be the sequence of delay distributions given by*

$$D^{(k)} = \begin{bmatrix} k \mod 5 & k \mod 6 \\ k \mod 6 & k \mod 5 \end{bmatrix}. \tag{4.1}$$

These time-varying delays are periodic with period 30 *and* $L = 5$. *The result of these delays on the network's dynamics can be seen in Figure 4.2 (left) where we let (C_P, X_5) denote the network (C, X) with time-varying time-delays given by 4.1. Note that although the trajectories are altered by these delays they still converge to the same fixed point $\mathbf{x}^* = (-.386, 1.595)$ as in the undelayed and constant time-delayed networks (cf. Figure 3.2) as guaranteed by Main Result 2.*

If instead we let

$$D^{(k)}_{U[0,10]} \in \mathbb{N}^{2\times 2}. \tag{4.2}$$

be the random matrix in which each entry is sampled uniformly from the integers $\{0, 1, \ldots, 10\}$ *the resulting switched network is still stable, as guaranteed by Main Result 2. Moreover, the network's trajectories still converge to the point* $\mathbf{x}^* = (-.386, 1.595)$ *as shown in Figure 4.2 (right) where we let* (C_U, X_{10}) *denote the network* (C, X) *with time-varying time-delays given by 4.2.*

4.1 Application: Linear Systems with Distict Delayed and Undelayed Interactions

A number of papers have published delay-dependant results regarding time-varying time-delayed systems whose delayed interactions differ from the undelayed interactions [13, 14, 15, 16, 17]. We demonstrate how to analyze these by use of Main Result 2.

Consider a linear system of the form

$$\mathbf{x}^{k+1} = A\mathbf{x}^k + B\mathbf{x}^{k-\tau(k)} \quad \text{where} \quad 1 \leq \tau(k) \leq L \tag{4.3}$$

where $A, B \in \mathbb{R}^{n\times n}$ are constant matrices, and $\tau(k)$ is a positive integer representing the magnitude of the time-varying time-delay, bounded by some $L > 0$. Here A and B represent distict weights of the delayed and undelayed interactions.

The minimally delayed version of system (4.3) is given by

$$\mathbf{x}^{k+1} = A\mathbf{x}^k + B\mathbf{x}^{k-1} \tag{4.4}$$

which we can express in terms of a single transition matrix $\widetilde{A}$ as

$$\widetilde{\mathbf{x}}^{k+1} = \widetilde{A}\widetilde{\mathbf{x}}^{k} \quad \text{where} \quad \widetilde{A} = \begin{bmatrix} A & B \\ I_{n\times n} & \mathbf{0}_{n\times n} \end{bmatrix} \quad \text{and} \quad \widetilde{\mathbf{x}}^{k} = \begin{bmatrix} \mathbf{x}^{k} \\ \mathbf{x}^{k-1} \end{bmatrix} \tag{4.5}$$

As in [17], we say that system (4.5) is the *lifted representation* of system (4.4).

Now observe that we may represent system (4.3) in the notation of Main Result 2 as the system $(F, \mathbb{R}^{2n})$ where

$$F(\widetilde{\mathbf{x}}^{k}) = \widetilde{A}\widetilde{\mathbf{x}}^{k}$$

where since F is linear, the Lipschitz matrix of F is given by $\left|\widetilde{A}\right|$ and the zero vector $\mathbf{0}_n$ is a fixed point. Then system (4.3) is the switched system instance $(\{F_{D^{(k)}}\}_{k=1}^{\infty}, X_L)$ obtained from the sequence of delay distributions

$$D^{(k)} = \tau(k) \begin{bmatrix} \mathbf{0}_{n\times n} & \mathbf{1}_{n\times n} \\ \mathbf{0}_{n\times n} & \mathbf{0}_{n\times n} \end{bmatrix}$$

ensuring that delays only occur to interactions corresponding to B, with magnitude $\tau(k)$. Immediately it follows by Main Result 2 that the system (4.3) is stable for arbitrary large delay bounds L when $(F, \mathbb{R}^{2n})$ is intrinsically stable, that is that $\rho(\widetilde{A}) < 1$ is satisfied.

Example 4.6. (Intrinsically Stable) *Consider system (4.3) with*

$$A = \begin{bmatrix} 0.6 & 0 \\ 0.35 & 0.7 \end{bmatrix}, \quad B = \begin{bmatrix} 0.1 & 0 \\ 0.2 & 0.1 \end{bmatrix}.$$

The transition matrix of the lifted representation is

$$\widetilde{A} = \begin{bmatrix} 0.6 & 0 & 0.1 & 0 \\ 0.35 & 0.7 & 0.2 & 0.1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}$$

which satisfies $\rho(\left|\widetilde{A}\right|) \approx 0.822 < 1$. *Hence this system is intrinsically stable, so is in fact stable for an arbitrarily large delay bound* $L > 0$.

However, there are systems of the form (4.3) which are not intrinsically stable.

Example 4.7. (Stable, not Intrinsically Stable) *Consider system (4.3) with*

$$A = \begin{bmatrix} 0.8 & 0 \\ 0.05 & 0.9 \end{bmatrix}, \quad B = \begin{bmatrix} -0.1 & 0 \\ -0.2 & -0.1 \end{bmatrix}.$$

The transition matrix of the lifted representation is

$$\widetilde{A} = \begin{bmatrix} 0.8 & 0 & -0.1 & 0 \\ 0.05 & 0.9 & -0.2 & -0.1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}$$

which satisfies $\rho(\left|\widetilde{A}\right|) = 1.0$. *Hence this system is not intrinsically stable, even though [13] showed that it is stable for all* $0 \leq L \leq 9.61 \times 10^8$.

However, even though some systems are not intrinsically stable, the fact that intrinsic stability can be verified much more efficiently than the results of [13, 14, 15, 16, 17] justifies checking anyways, in case the system under consideration is in fact intrinsically stable. For further analysis of the computational complexity of checking intrinsic stability, see chapter 6.

Chapter 5. Row-Independence Closure of Switched Networks

Using Main Result 1 and Main Result 2, we can extend our analysis of time-varying time-delays to a more general class of switched networks.

Definition 5.1. (Row-Independence Closure) *Let (M, X) be a switched network with Lipschitz set S. Then we denote $RI(S)$, the* row-independence closure *of S, by*

$$RI(S) = \{A^* | \ \mathbf{a}_i^* = \mathbf{a}_i^{(i)} \textit{ for some } A^{(1)}, \dots, A^{(n)} \in S\}$$

where $\mathbf{a}_i^{(i)}$ denotes the i^{th} row of the i^{th} matrix $A^{(i)}$.

Intuitively, row-independence indicates that there is no conditional relationship between rows of the matrices in $RI(S)$. Note that $S \subset RI(S)$.

The row-independence closure provides a computationally efficient sufficient condition for satisfying the hypothesis of Main Result 1. The following proposition follows directly from Definition 5.1 the results of [11] (Theorem B.4 in the Appendix).

Proposition 5.2. (Row-Independence Closure, Intrinsic Stability, and Switched Networks) *Let (M, X) be a switched network with Lipschitz set S. Then*

$$\overline{\rho}(S) \leq \overline{\rho}(RI(S)) = \max_{A \in RI(S)} \rho(A).$$

We then have the following extension of Main Result 2, which provides a sufficient condition ensuring that when time-varying time-delays are applied to a stable already-switched system, the resulting hyper-switched system retains stability.

Main Result 3. (Intrinsic Stability and Row-Independent Switched Networks) *Let (M, X) be a switched network with Lipschitz set S. Assume $\mathbf{x}^*$ is a shared fixed point of*

(F, X) for all $F \in M$. Assume $\rho(A) < 1$ for all $A \in RI(S)$. Let $L > 0$ and

$$M_d = \{F_D | F \in M, D \in \mathbb{N}^{n\times n} \text{ with } \max_{ij} d_{ij} \le L\}$$

and S_d be the Lipschitz set of M_d.

Then $E_L(\mathbf{x}^*)$ is a globally attracting fixed point of every instance $(\{F^{(k)}_{D^{(k)}}\}_{k=1}^{\infty}, X_L)$ of (M_d, X_L). Furthermore,

$$\overline{\rho}(S_d) \le \max_{A\in RI(S)} \rho(A_L) < 1$$

where given some $A \in \mathbb{R}^{n\times n}$, A_L is defined as

$$A_L = \begin{bmatrix} \mathbf{0}_{n\times nL} & A \\ \mathbf{I}_{nL\times nL} & \mathbf{0}_{n\times n} \end{bmatrix}.$$

Thus, a switched network with a fixed point satisfying $\rho(A) < 1$ for all $A \in RI(S)$ retains convergence to the same equilibrium, even when it experiences time-varying time-delays. This extends Main Result 2 further to the even more complicated class of switching-delay switched networks.

5.1 Application: Switched Linear Systems with Distict Delayed and Undelayed Interactions

We now extend our analysis of Section 4.1 to the case where system (4.3) is also a switched system:

$$\mathbf{x}^{k+1} = A_{\sigma(k)}\mathbf{x}^k + B_{\sigma(k)}\mathbf{x}^{k-\tau(k)} \quad \text{where} \quad 1 \le \tau(k) \le L \tag{5.1}$$

where as before, each $A, B \in \mathbb{R}^{n\times n}$ are constant matrices, and $\tau(k)$ is a positive integer representing the magnitude of the time-varying time-delay, bounded by some $L > 0$. The difference between system (5.1) and system (4.3) is that in system (5.1) there are multiple possibilities for the transition weights of the delayed and undelayed interactions.

The minimally delayed version of system (5.1) is given by

$$\mathbf{x}^{k+1} = A_{\sigma(k)}\mathbf{x}^k + B_{\sigma(k)}\mathbf{x}^{k-1} \tag{5.2}$$

so the lifted version of system (5.2) is

$$\widetilde{\mathbf{x}}^{k+1} = \widetilde{A}_{\sigma(k)}\widetilde{\mathbf{x}}^k \quad \text{where} \quad \widetilde{A}_{\sigma(k)} = \begin{bmatrix} A_{\sigma(k)} & B_{\sigma(k)} \\ I_{n\times n} & \mathbf{0}_{n\times n} \end{bmatrix} \quad \text{and} \quad \widetilde{\mathbf{x}}^k = \begin{bmatrix} \mathbf{x}^k \\ \mathbf{x}^{k-1} \end{bmatrix} \tag{5.3}$$

Now observe that we may represent system (5.1) in the notation of Main Result 3 as the system $(M_0, \mathbb{R}^{2n})$ where

$$M_0 = \{F|\ F(\widetilde{\mathbf{x}}^k) = \widetilde{A}_{\sigma(k)}\widetilde{\mathbf{x}}^k\}$$

where since each F is linear, the Lipschitz matrix of F is given by $\left|\widetilde{A}_{\sigma(k)}\right|$ and the zero vector $\mathbf{0}_n$ is a shared fixed point. Then system (5.1) is the switched system instance $(\{F^{(k)}_{D^{(k)}}\}_{k=1}^{\infty}, X_L)$ obtained from the sequence of transition matrices $\widetilde{A}_{\sigma(k)}$ and the sequence of delay distributions

$$D^{(k)} = \tau(k) \begin{bmatrix} \mathbf{0}_{n\times n} & \mathbf{1}_{n\times n} \\ \mathbf{0}_{n\times n} & \mathbf{0}_{n\times n} \end{bmatrix}$$

once the sequences $\sigma(k)$ and $\tau(k)$ are determined. ensuring that delays only occur to interactions corresponding to B, with magnitude $\tau(k)$. Immediately it follows by Main Result 3 that the system (5.1) is stable for arbitrary large delay bounds L when $(M_0, \mathbb{R}^{2n})$ satisfies $\rho(A) < 1$ for all $A \in RI(S)$ of the Lipschitz set S of M_0.

Example 5.3. (Switched Transitions) *Consider system (5.1) with*

$$A_1 = A_2 = \begin{bmatrix} 0 & 0.3 \\ -0.2 & 0.1 \end{bmatrix}, \quad A_3 = A_4 = \begin{bmatrix} 0 & 0.3 \\ -0.2 & -0.1 \end{bmatrix}$$

$$B_1 = B_3 = \begin{bmatrix} 0 & 0.1 \\ 0 & 0.2 \end{bmatrix}, \quad B_2 = B_4 = \begin{bmatrix} 0 & 0.1 \\ 0 & 0 \end{bmatrix}$$

The set M_0 of lifted transition matrices consists of the following four matrices:

$$\widetilde{A}_1 = \begin{bmatrix} 0 & 0.3 & 0 & 0.1 \\ -0.2 & 0.1 & 0 & 0.2 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} \quad \widetilde{A}_2 = \begin{bmatrix} 0 & 0.3 & 0 & 0.1 \\ -0.2 & 0.1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}$$

$$\widetilde{A}_3 = \begin{bmatrix} 0 & 0.3 & 0 & 0.1 \\ -0.2 & -0.1 & 0 & 0.2 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} \quad \widetilde{A}_4 = \begin{bmatrix} 0 & 0.3 & 0 & 0.1 \\ -0.2 & -0.1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}$$

Note that this satisfies $M_0 = RI(M_0)$, as well as $\rho(\left|\widetilde{A}_k\right|) < 1.0$ for all $\widetilde{A}_k \in M_0$. Hence this switched system is intrinsically stable, so is in fact stable for an arbitrarily large delay bound $L > 0$, no matter the sequences chosen.

Chapter 6. Analytical and Computational Considerations

We now summarize how these results are used to analyze real-world dynamical systems with network structure, that is, systems (F, X) of the form $X = \bigoplus_{i=1}^{n} X_i$ with a corresponding mapping F defined componentwise. Consider the simpler model of the network where all interactions occur instantaneously, excluding all delays of the form of Definition 3.1 (i.e. affecting only a single component by an identity mapping).

Find a Lipschitz matrix A of the network by equation 2.2 if the mapping F is piecewise differentiable and $X = \mathbb{R}^n$, otherwise directly determine the pairwise Lipschitz constants by use of Definition 2.3. Recall that there are infinitely many Lipschitz matrices for a given dynamical network; ideally, we would like to find one which minimizes the spectral radius, as this tightens the convergence rate of the system. However, in certain cases it may simplify analysis considerably to merely find upper bounds for each entry of the Lipschitz matrix.

The spectral radius $\rho(A)$ can be computed efficiently in $\mathcal{O}(mn)$ time by use of the power method, where m is the number of nonzero entries in A [18]. As soon as a single Lipschitz matrix A of (F, X) satisfies $\rho(A) < 1$, even if this A does not have minimal spectral radius, the results of Theorem 3.3 and Main Result 2 apply, guarenteeing that all delayed versions of (F, X) are stable, even if the delays are varying in time (so long as the magnitude of these delays are eventually bounded by some L). The convergence rate of the delayed system is given by $\rho(A_L)$, where A_L is defined in Main Result 2. Since A_L is sparse with $m+nL$ nonzero entries, $\rho(A_L)$ can also be computed in $\mathcal{O}(mnL + n^2L^2)$ time using the power method.

Chapter 7. Future Work

While Main Results 2 and 3 provide upper bounds on the convergence rate, further work is needed to find the expected convergence rate when the sequence of delay distributions $\{D^{(k)}\}_{k=1}^{\infty}$ is generated stochastically, as in Example 4.5. In addition, note that the addition of time-delays can be viewed as a form of network growth in which cycles are extended, but otherwise unaltered. Further research regarding the Lipschitz matrix of a dynamical nework may provide similar invariance of stability to network changes such as weight updates and cycle removal, as occurs with various machine learning techniques such as backpropagation and dropout. Lastly, the results presented here merit further investigation of applications, including high-dimensional, nonlinear switched systems with time-varying time-delays.

Chapter 8. Conclusion

Time-delays can destabilize a dynamical network, and increase the difficulty of anlayzing the asymptotic behavior of the network. An alternative method to typical Lyanpunov methods was presented in this thesis. This method, which consists of determining whether the lower-dimensional undelayed network is intrinsically stable, is computationally feasible to verify on large networks. Furthermore, this resilience to time-varying time-delays is extended to row-independent switched networks.

Appendix A. Proof of Proposition 3.5 and Main Result 1

A.1 Proof of Proposition 3.5

Proof. Let $\mathbf{x}^* = [x_1^*, x_2^*, \ldots, x_n^*]^T$ be a fixed point of a dynamical network (F, X). Then by definition

$$F_i(x_1^*, x_2^*, \ldots, x_n^*) = x_i^* \quad \textit{for all} \quad 1 \leq i \leq n.$$

Let D be a delay distribution with $\max_{ij} d_{ij} \leq L$ for some $L > 0$. With the usual ordering of the component spaces of $\mathbf{x} \in X_L$, we have

$$F_D(E_L(\mathbf{x}^*)) = \begin{bmatrix} (F_D)_{1,0}(x_1^*, x_2^*, \ldots, x_n^*) \\ (F_D)_{2,0}(x_1^*, x_2^*, \ldots, x_n^*) \\ \vdots \\ (F_D)_{n,0}(x_1^*, x_2^*, \ldots, x_n^*) \\ (F_D)_{1,1}(x_1^*) \\ (F_D)_{2,1}(x_2^*) \\ \vdots \\ (F_D)_{n,L}(x_n^*) \end{bmatrix} = \begin{bmatrix} F_1(x_1^*, x_2^*, \ldots, x_n^*) \\ F_2(x_1^*, x_2^*, \ldots, x_n^*) \\ \vdots \\ F_n(x_1^*, x_2^*, \ldots, x_n^*) \\ x_1^* \\ x_2^* \\ \vdots \\ x_n^* \end{bmatrix} = \begin{bmatrix} x_1^* \\ x_2^* \\ \vdots \\ x_n^* \\ x_1^* \\ x_2^* \\ \vdots \\ x_n^* \end{bmatrix} = E_L(\mathbf{x}^*). \quad \square$$

A.2 Proof of Main Result 1

Proof. Let $\mathbf{x}, \mathbf{y} \in X$ and $k > 0$ be arbitrary. Note that for any $F \in M$ with corresponding $A \in S$ we have, by definition of A being a Lipschitz matrix of F, that

$$\begin{bmatrix} d_1(F_1(\mathbf{x}), F_1(\mathbf{y})) \\ \vdots \\ d_n(F_n(\mathbf{x}), F_n(\mathbf{y})) \end{bmatrix} \preceq A \begin{bmatrix} d_1(x_1, y_1) \\ \vdots \\ d_n(x_n, y_n) \end{bmatrix}$$

where $\preceq$ denotes an element-wise inequality. Thus, for the specific instance $(\{F^{(k)}\}_{k=1}^{\infty}, X)$ given in the hypothesis, we have inductively

$$\begin{aligned} \begin{bmatrix} d_1(\mathscr{F}_1^k(\mathbf{x}), \mathscr{F}_1^k(\mathbf{y})) \\ \vdots \\ d_n(\mathscr{F}_n^k(\mathbf{x}), \mathscr{F}_n^k(\mathbf{y})) \end{bmatrix} &= \begin{bmatrix} d_1(F_1^{(k)} \circ \mathscr{F}^{k-1}(\mathbf{x}), F_1^{(k)} \circ \mathscr{F}^{k-1}(\mathbf{y})) \\ \vdots \\ d_n(F_n^{(k)} \circ \mathscr{F}^{k-1}(\mathbf{x}), F_n^{(k)} \circ \mathscr{F}^{k-1}(\mathbf{y})) \end{bmatrix} \\ &\preceq A^{(k)} \begin{bmatrix} d_1(\mathscr{F}_1^{k-1}(\mathbf{x}), \mathscr{F}_1^{k-1}(\mathbf{y})) \\ \vdots \\ d_n(\mathscr{F}_n^{k-1}(\mathbf{x}), \mathscr{F}_n^{k-1}(\mathbf{y})) \end{bmatrix} \\ &\preceq A^{(k)} A^{(k-1)} \ldots A^{(1)} \begin{bmatrix} d_1(x_1, y_1) \\ \vdots \\ d_n(x_n, y_n) \end{bmatrix}. \end{aligned}$$

Hence, by the definition of the joint spectral radius, there exists some positive constant C (possibly dependant on $\mathbf{x}$ and $\mathbf{y}$) such that

$$\begin{aligned} d_{max}(\mathscr{F}^k(\mathbf{x}), \mathscr{F}^k(\mathbf{y})) &= \left\| \begin{bmatrix} d_1(\mathscr{F}_1^k(\mathbf{x}), \mathscr{F}_1^k(\mathbf{y})) \\ \vdots \\ d_n(\mathscr{F}_n^k(\mathbf{x}), \mathscr{F}_n^k(\mathbf{y})) \end{bmatrix} \right\|_\infty \\ &\leq \left\| A^{(k)} A^{(k-1)} \dots A^{(1)} \begin{bmatrix} d_1(x_1, y_1) \\ \vdots \\ d_n(x_n, y_n) \end{bmatrix} \right\|_\infty \\ &\leq C(\overline{\rho}(S))^k \end{aligned}$$

where $\|\mathbf{x}\|_\infty = \max_i |x_i|$.

Now, assume $\mathbf{x}^*$ is a shared fixed point of (F, X) for all $F \in M$. Then $\mathscr{F}^1(\mathbf{x}^*) = F^{(1)}(\mathbf{x}^*) = \mathbf{x}^*$, and if it is assumed that $\mathscr{F}^{k-1}(\mathbf{x}^*) = \mathbf{x}^*$, then it follows immediately that

$$\mathscr{F}^k(\mathbf{x}^*) = F^{(k)} \circ \mathscr{F}^{k-1}(\mathbf{x}^*) = F^{(k)}(\mathbf{x}^*) = \mathbf{x}^*.$$

Hence by induction, $\mathbf{x}^*$ is a fixed point of $(\{F^{(k)}\}_{k=1}^\infty, X)$. Thus

$$d_{max}(\mathscr{F}^k(\mathbf{x}^0), \mathbf{x}^*) = d_{max}(\mathscr{F}^k(\mathbf{x}^0), \mathscr{F}^k(\mathbf{x}^*)) \leq C\overline{\rho}(S)^k$$

so $\lim_{k\to\infty} \mathscr{F}^k(\mathbf{x}^0) = \mathbf{x}^*$ for all initial conditions $\mathbf{x}^0 \in X$. $\square$

Appendix B. Proof of Main Results 2 and 3

We divide up the proof of Main Result 3 into several lemmata.

B.1 Preliminaries

First, we make explicit the effect of time delays on the Lipschitz matrix of a dynamical network.

Lemma B.1. (Structure of the Lipschitz Matrix of a Delayed Network) *Let (F, X) be a dynamical network with Lipschitz matrix $A = [a_{ij}] \in \mathbb{R}^{n\times n}$ and $D = [d_{ij}] \in \mathbb{N}^{n\times n}$ a delay distribution matrix with $\max_{i,j} d_{ij} \leq L$. Let A_D be defined in terms of A as*

$$A_D = \begin{bmatrix} A_0 & A_1 & \dots & A_{L-1} & A_L \\ I_n & 0 & \dots & 0 & 0 \\ 0 & I_n & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & I_n & 0 \end{bmatrix} \in \mathbb{R}^{n(L+1)\times n(L+1)}$$

where each $A_\ell \in \mathbb{R}^{n\times n}$ is defined element-wise as $A_\ell = \left[a_{ij}\mathbb{1}_{d_{ij}=\ell}\right]$, with the indicator function $\mathbb{1}_{d_{ij}=\ell}$ defined as

$$\mathbb{1}_{d_{ij}=\ell} = \begin{cases} 1 & \text{if } d_{ij} = \ell \\ 0 & \text{otherwise.} \end{cases} \qquad \text{for } 0 \leq \ell \leq L$$

Then A_D is a Lipschitz matrix of (F_D, X_L).

Proof. Recall that for $\mathbf{x} \in X_L$, we order the components $x_{i,\ell}$ of $\mathbf{x}$ as

$$\mathbf{x} = [x_{1,0}, x_{2,0}, \dots, x_{n,0}, x_{1,1}, x_{2,1}, \dots, x_{n,L}]^T$$

where $x_{i,\ell} \in X_{i,\ell}$ for $i = 1, 2, \ldots, n$ and $\ell = 0, 1, \ldots, L$. Let $\mathbf{x}, \mathbf{y} \in X$ be given. Then

$$\begin{aligned} d_{i,0}\left((F_D)_{i,0})(\mathbf{x}), (F_D)_{i,0})(\mathbf{y})\right) &= d_i\left(F_i(x_{1,d_{i1}}, x_{2,d_{i2}}, \ldots, x_{n,d_{in}}), F_i(y_{1,d_{i1}}, y_{2,d_{i2}}, \ldots, y_{n,d_{in}})\right) \\ &\le \sum_{j=1}^{n} a_{ij} d_j(x_{j,d_{ij}}, y_{j,d_{ij}}) = \sum_{\ell=0}^{L} \sum_{j=1}^{n} a_{ij} \mathbb{1}_{d_{ij}=\ell} d_{j,\ell}(x_{j,\ell}, y_{j,\ell}) \end{aligned}$$

which matches the first n rows of A_D. For $\ell \geq 1$,

$$d_{i,\ell}\left((F_D)_{i,\ell})(\mathbf{x}), (F_D)_{i,\ell})(\mathbf{y})\right) = d_{i,\ell-1}(x_{i,\ell-1}, y_{i,\ell-1})$$

yields the identity matrices I_n in A_D. $\square$

The following theorem follows as a direct corollary of [[3], Lemma 3.3], restated here for convenience and consistency of notation.

Theorem B.2. *Let (F, X) be a dynamical network with Lipschitz matrix A. Then for any delay distribution D the constant time-delayed dynamical network (F_D, X_D) has the Lipschitz matrix A_D with*
(i) $\rho(A) \leq \rho(A_D) < 1$ if $\rho(A) < 1$;
(ii) $\rho(A_D) = 1$ if $\rho(A) = 1$; and
(iii) $\rho(A) \geq \rho(A_D) > 1$ if $\rho(A) > 1$.

Suppose D and $\hat{D}$ are delay distribution matrices such that $D \preceq \hat{D}$, i.e. entries of D are less than or equal to the corresponding entries of $\hat{D}$. If (F, X), (F_D, X_D), and $(F_{\hat{D}}, X_{\hat{D}})$ have the corresponding Lipschitz matrices A, A_D, and $A_{\hat{D}}$, respectively, with $\rho(A) < 1$ then Theorem B.2 part (i) implies that

$$\rho(A) \leq \rho(A_D) \leq \rho(A_{\hat{D}}) < 1.$$

That is to say that the spectral radius of the network is monotonic with respect to the addition of delays if $\rho(A) < 1$.

B.1.1 Results Regarding the Joint Spectral Radius.

We require the following definition and theorem from [Equation (3.1), Theorem 2 in [11]] regarding sets of matrices with independent row uncertainties.

Definition B.3. (Independent Row Uncertainties) *We say that a set of matrices $S \subset \mathbb{R}^{n\times n}$ has independent row uncertainty if S can be expressed as*

$$S = \{(\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n)^T \mid \mathbf{a}_i \in Q_i, 1 \le i \le n\}$$

where the sets $Q_i \subset \mathbb{R}^n$, $1 \le i \le n$ are closed and bounded.

Theorem B.4. (Joint Spectral Radius of Nonnegative Matrices with Independent Row Uncertainty) *Let S be a set of nonnegative matrices with independent row uncertainty. Then*

$$\overline{\rho}(S) = \max_{A\in S} \rho(A).$$

Furthermore, we will use the equivalence of Definition 4.3 of the joint spectral radius with the following representation from (someone) and (someone) [cite me].

Theorem B.5. (Alternate Form of the Joint Spectral Radius) *Given a set of matrices $S \subset \mathbb{R}^{n\times n}$, the joint spectral radius $\rho(S)$ is given by*

$$\rho(S) = \limsup_{k\to\infty} \max\{\|A\|^{\frac{1}{k}} : A \textit{ is a product of length } k \textit{ of matrices in } S\}$$

It follows immediately by the form of Theorem B.5 that if $S_1 \subset S_2$, then $\rho(S_1) \le \rho(S_2)$.

B.1.2 Proof of Proposition 5.2.

Proof. Let (M, X) be a switched network with Lipschitz set S. By the form of B.5 we have $\overline{\rho}(S) \le \overline{\rho}(RI(S))$ since $S \subset RI(S)$. For $1 \le i \le n$, let $Q_i = \{\mathbf{a}_i | \ A \in \overline{S}\}$ be the set of i^{th}

rows of all $A \in \overline{S}$. Then $RI(S)$ may be expressed as

$$S = \{(\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n)^T \mid \mathbf{a}_i \in Q_i, 1 \leq i \leq n\}$$

so $RI(S)$ is row-independent. Thus

$$\overline{\rho}(S) \leq \overline{\rho}(RI(S)) = \max_{A \in RI(S)} \rho(A).$$

by Theorem B.4, as desired. $\square$

Lemma B.6. (Equality of A_D's) *Let $L > 0$, $1 \leq i \leq n$ be given. Let $A^{(1)}$, $A^{(2)}$ satisfy $\mathbf{a}_i^{(1)} = \mathbf{a}_i^{(2)}$, and $D^{(1)}$, $D^{(2)}$ satisfy $\mathbf{d}_i^{(1)} = \mathbf{d}_i^{(2)}$. Then $(A^{(1)}_{D^{(1)}})_i = (A^{(2)}_{D^{(2)}})_i$.*

Proof. It suffices to show that $(A_\ell^{(1)})_i = (A_\ell^{(2)})_i$, where given some A and D, A_ℓ is defined as in Lemma B.1. Let $0 \leq \ell \leq L$ be arbitrary. By Lemma B.1,

$$(A_\ell^{(1)})_{ij} = \mathbf{a}_{ij}^{(1)} \mathbb{1}_{d_{ij}^{(1)}=\ell} = \mathbf{a}_{ij}^{(2)} \mathbb{1}_{d_{ij}^{(2)}=\ell} = (A_\ell^{(2)})_{ij} \quad \text{for} \quad 1 \leq j \leq n.$$

Thus $(A_\ell^{(1)})_i = (A_\ell^{(2)})_i$, so by Lemma B.1, $(A^{(1)}_{D^{(1)}})_i = (A^{(2)}_{D^{(2)}})_i$. $\square$

Lemma B.7. (Equality of sets) *Let (M_0, X) be a switched network with Lipschitz set S. Let $L > 0$ and $\mathbb{D} = \{D \in \mathbb{N}^{n \times n} \mid \max_{ij} d_{ij} \leq L\}$.*

Then $RI(\{A_D \mid A \in S, D \in \mathbb{D}\}) = \{A_D \mid A \in RI(S), D \in \mathbb{D}\}$.

Proof. Let $(A_D)^* \in RI(\{A_D \mid A \in S, D \in \mathbb{D}\})$. Then there exist $(A_D)^{(1)}, \dots, (A_D)^{(n)} \in \{A_D \mid A \in S, D \in \mathbb{D}\}$ such that $(A_D)^*_i = ((A_D)^{(i)})_i$. Furthermore, there exist $A^{(1)}, \dots, A^{(n)} \in S$, $D^{(1)}, \dots, D^{(n)} \in \mathbb{D}$ such that $(A_D)^{(i)} = A^{(i)}_{D^{(i)}}$. Let A^* be constructed as $(A^*)_i = \mathbf{a}_i^{(i)}$, and D^* be constructed as $(D^*)_i = \mathbf{d}_i^{(i)}$. Then $A^* \in RI(S)$, and since each $D^{(i)}$ satisfies $d_{ij} \leq L$, we have $D^* \in \mathbb{D}$. Thus,

$$(A_D)^*_i = ((A_D)^{(i)})_i = (A^{(i)}_{D^{(i)}})_i = (A^*_{D^*})_i \quad \text{for} \quad 1 \leq i \leq n$$

where the last equality follows from Lemma B.6. Thus $(A_D)^* = A^*_{D^*} \in \{A_D|\ A \in RI(S), D \in \mathbb{D}\}$, so

$$RI(\{A_D|\ A \in S, D \in \mathbb{D}\}) \subset \{A_D|\ A \in RI(S), D \in \mathbb{D}\}.$$

Now let $A_D \in \{A_D|\ A \in RI(S), D \in \mathbb{D}\}$. Then there exist $A^{(1)}, \ldots, A^{(n)} \in S$ and $D \in \mathbb{D}$ such that $(A_D)_i = ((A^{(i)})_D)_i$. Let $(A_D)^{(i)} = (A^{(i)})_D$. Then $(A_D)^{(i)} \in \{A_D|\ A \in S, D \in \mathbb{D}\}$ so $A_D \in RI(\{A_D|\ A \in S, D \in \mathbb{D}\})$. Thus

$$\{A_D|\ A \in RI(S), D \in \mathbb{D}\} \subset RI(\{A_D|\ A \in S, D \in \mathbb{D}\})$$

□

B.2 Proof of Main Result 3

Proof. Let (M_0, X) be a switched network with Lipschitz set S. Assume $\mathbf{x}^*$ is a shared fixed point of (F, X) for all $F \in M_0$, and $\rho(A) < 1$ for all $A \in RI(S)$. Let $L > 0$, $\mathbb{D} = \{D \in \mathbb{N}^{n\times n}|\ \max_{ij} d_{ij} \leq L\}$, $M_d = \{F_D | F \in M_0, D \in \mathbb{D}\}$, and S_d be the Lipschitz set of M_d. We will show $\rho(A_D) < 1$ for all $A_D \in RI(S_d)$ and invoke Proposition 5.2.

By Lemma B.7, we have $RI(S_d) = \{A_D|\ A \in RI(S), D \in \mathbb{D}\}$. Then

$$\max_{A_D \in RI(S_d)} \rho(A_D) = \max_{A \in RI(S)} \max_{D \in \mathbb{D}} \rho(A_D) < 1$$

by the hypothesis and Theorem B.2. Now given some $A \in RI(S)$, by Lemma B.1 and repeated application of Theorem B.2 we have that

$$\max_{D \in \mathbb{D}} \rho(A_D) = \rho(A_L) \quad \text{where} \quad A_L = \begin{bmatrix} \mathbf{0}_{n\times nL} & A \\ \mathbf{I}_{nL\times nL} & \mathbf{0}_{n\times n} \end{bmatrix}.$$

Thus by Proposition 5.2,

$$\overline{\rho}(S_d) \leq \overline{\rho}(RI(S_d)) = \max_{A_D \in RI(S_d)} \rho(A_D) = \max_{A \in RI(S)} \rho(A_L) < 1.$$

Since $\mathbf{x}^*$ is a shared fixed point of (F, X) for all $F \in M_0$, by Proposition 3.5 $E_L(\mathbf{x}^*)$ is a shared fixed point of (F_D, X_L) for all $F_D \in M_d$. Thus by Main Result 1, $E_L(\mathbf{x}^*)$ is a globally attracting fixed point of every instance $(\{F_{D^{(k)}}^{(k)}\}_{k=1}^{\infty}, X_L)$ of (M_d, X_L). $\square$

B.3 Proof of Main Result 2

Proof. Let M_0 be the singleton set consisting of F. Then the Lipschitz set S of M_0 consists only of the matrix A, and so trivially satisfies $S = RI(S)$. Thus the hypothesis of Main Result 3 is trivially satisfied, and the result follows. $\square$

Bibliography

[1] Sipahi, R., S. Niculescu, Chaouki T. Abdallah, W. Michiels, and Keqin Gu. Stability and Stabilization of Systems with Time Delay. IEEE Control Systems, Feb, 2011. 38.

[2] Webb, Benjamin and Leonid Bunimovich. Intrinsic Stability, Time Delays and Transformations of Dynamical Networks. Advances in Dynamics, Patterns, Cognition. Springer International Publishing, 2017.

[3] Webb, Benjamin and Leonid Bunimovich. Isospectral Transformations: A New Approach to Analyzing Multidimensional Systems and Networks. New York, NY: Springer, 2014.

[4] Cohen,M.,Grossberg S. Absolute stability and global pattern formation and parallel memory storage by competitive neural networks. IEEE Trans.Syst.Man Cybern.SMC-13, 815–821 (1983)

[5] Cao,J. Global asymptotic stability of delayed bi-directional associative memory neural networks. Appl.Math.Comput.142(2–3),333–339 (2003)

[6] Chena,S.,Zhaoa,W.,Xub,Y. New criteria for globally exponential stability of delayed Cohen–Grossberg neural network. Math.Comput.Simul.79, 1527–1543 (2009)

[7] Cheng,C.-Y.,Lin,K.-H.,Shih,C.-W. Multistability in recurrent neural networks. SIAM J.Appl.Math.66(4),1301–1320 (2006)

[8] Tao,L.,Ting,W.,Shumin,F. Stability analysis on discrete-time Cohen–Grossberg neural networks with bounded distributed delay. In:Proceedings of the 30th Chinese Control Conference,Yantai,22–24 July (2011)

[9] Wang,L.,Dai,G.-Z. Global stability of virus spreading in complex heterogeneous networks. SIAMJ.Appl.Math.68(5),1495–1502 (2008)

[10] Alpcan,T.,Basar,T. A globally stable adaptive congestion control scheme for internet-style networks with delay. IEEE/ACMTrans.Netw.13, 6 (2005)

[11] Vincent D Blondel and Yurii Nesterov. Polynomial-Time Computation of the Joint Spectral Radius for some Sets of Nonnegative Matrices. SIAM Journal on Matrix Analysis and Applications 31, no. 3 (Sep 1, 2009): 865-876.

[12] Newman,M.E.J. Networks an Introduction. Oxford University Press, Oxford (2010)

[13] Stojanovic, Sreten B., Dragutin L. J. Debeljkovic, and Nebojsa Dimitrijevic. Stability of Discrete-Time Systems with Time-Varying Delay: Delay Decomposition Approach. International Journal of Computers Communications and Control 7, no. 4 (Sep 16, 2014): 776.

[14] Fridman, Emilia. Introduction to Time-Delay Systems : Analysis and Control, Chapter 6. Systems and Control: Foundations and Applications. 2014th ed. Cham: Birkhäuser Boston, 2014.

[15] Wan, Xiongbo, Min Wu, Yong He, and Jinhua She. Stability Analysis for Discrete Time-Delay Systems Based on New Finite-Sum Inequalities. Information Sciences 369, (Nov 10, 2016): 119-127, `https://www.sciencedirect.com/science/article/pii/S0020025516304467.`

[16] Li, Xu, Rui Wang, and Xudong Zhao. Stability of Discrete-Time Systems with Time-Varying Delay Based on Switching Technique. Journal of the Franklin Institute 355, no. 13 (Sep, 2018): 6026-6044, `https://www.sciencedirect.com/science/article/pii/S0016003218303910.`

[17] Zhang, Xian-Ming and Qing-Long Han. Abel Lemma-Based Finite-Sum Inequality and its Application to Stability Analysis for Linear Discrete Time-Delay Systems. Automatica 57, (Jul, 2015): 199-202, `https://www.sciencedirect.com/science/article/pii/S000510981500179X.`

[18] Trefethen, L.N., Bau, D., 1998. Numerical Linear Algebra. Philadelphia: Series in Applied Mathematics, Vol. 11. SIAM.