

Article

Human Brain Networks: Spiking Neuron Models, Multistability, Synchronization, Thermodynamics, Maximum Entropy Production, and Anesthetic Cascade Mechanisms

Wassim M. Haddad ^{1,*}, Qing Hui ² and James M. Bailey ³

¹ The School of Aerospace Engineering, Georgia Institute of Technology, Atlanta, GA 30332, USA

² Department of Mechanical Engineering, Texas Tech University, Lubbock, TX 79409-1021, USA;
E-Mail: qing.hui@ttu.edu

³ Department of Anesthesiology, Northeast Georgia Medical Center, Gainesville, GA 30503 USA;
E-Mail: james.bailey@nghs.com

* Author to whom correspondence should be addressed; E-Mail: wm.haddad@aerospace.gatech.edu;
Tel.: +1-404-894-1078; Fax: +1-404-894-2760.

Received: 6 May 2014; in revised form: 19 June 2014 / Accepted: 3 July 2014 /

Published: 17 July 2014

Abstract: Advances in neuroscience have been closely linked to mathematical modeling beginning with the integrate-and-fire model of Lapique and proceeding through the modeling of the action potential by Hodgkin and Huxley to the current era. The fundamental building block of the central nervous system, the neuron, may be thought of as a dynamic element that is “excitable”, and can generate a pulse or spike whenever the electrochemical potential across the cell membrane of the neuron exceeds a threshold. A key application of nonlinear dynamical systems theory to the neurosciences is to study phenomena of the central nervous system that exhibit nearly discontinuous transitions between macroscopic states. A very challenging and clinically important problem exhibiting this phenomenon is the induction of general anesthesia. In any specific patient, the transition from consciousness to unconsciousness as the concentration of anesthetic drugs increases is very sharp, resembling a thermodynamic phase transition. This paper focuses on multistability theory for continuous and discontinuous dynamical systems having a set of multiple isolated equilibria and/or a continuum of equilibria. Multistability is the property whereby the solutions of a dynamical system can alternate between two or more mutually exclusive Lyapunov stable and convergent equilibrium states under asymptotically slowly changing inputs or system parameters. In this paper, we extend the theory of multistability to continuous, discontinuous, and stochastic nonlinear dynamical systems. In particular, Lyapunov-based tests for

multistability and synchronization of dynamical systems with continuously differentiable and absolutely continuous flows are established. The results are then applied to excitatory and inhibitory biological neuronal networks to explain the underlying mechanism of action for anesthesia and consciousness from a multistable dynamical system perspective, thereby providing a theoretical foundation for general anesthesia using the network properties of the brain. Finally, we present some key emergent properties from the fields of thermodynamics and electromagnetic field theory to qualitatively explain the underlying neuronal mechanisms of action for anesthesia and consciousness.

Keywords: multistability; semistability; synchronization; biological networks; spiking neuron models; synaptic drive; discontinuous systems; thermodynamics; free energy; entropy; consciousness; arrow of time; excitatory and inhibitory neurons; Brownian motion; Wiener process; general anesthesia

1. Introduction

Advances in neuroscience have been closely linked to mathematical modeling beginning with the integrate-and-fire model of Lapicque [1] and proceeding through the modeling of the action potential by Hodgkin and Huxley [2] to the current era of mathematical neuroscience; see [3,4] and the numerous references therein. Neuroscience has always had models to interpret experimental results from a high-level complex systems perspective; however, expressing these models with dynamic equations rather than words fosters precision, completeness, and self-consistency. Nonlinear dynamical system theory, in particular, can provide a framework for a rigorous description of the behavior of large-scale networks of neurons. A particularly interesting application of nonlinear dynamical systems theory to the neurosciences is to study phenomena of the central nervous system that exhibit nearly discontinuous transitions between macroscopic states. One such example exhibiting this phenomenon is the induction of general anesthesia [5–9].

The rational, safe, and effective utilization of any drug in the practice of medicine is grounded in an understanding of the *pharmacodynamics* of the drug, loosely defined as what the drug does to the body [10]. A very important measure of the pharmacodynamics of any drug is the drug concentration parameter EC_{50} , which reflects the drug dose at which the therapeutic effect is achieved in 50% of the cases. This concept is certainly applicable for the administration of general inhalational anesthetics, where the potency of the drug is defined by the minimum alveolar concentration (MAC) of the drug needed to prevent a response to noxious stimuli in 50% of administrations [11].

The MAC concept is intrinsically embedded in a probabilistic framework [10]. It is the concentration at which the probability of a response to a noxious stimulus is 0.5. Typically the MAC of a particular anesthetic is determined by administering various doses of the agent to a population of patients and determining the dose at which there is a 0.5 chance of responding to a noxious stimulus. (Technically, we identify the concentration in the alveoli, the fundamental functional gas exchange units of the lung, at which the chance of response is 0.5.) It has been possible, however, to conduct studies of single subjects,

varying the anesthetic concentration and determining responsiveness. When this has been done, it has been noted that the transition from responsiveness to non-responsiveness in the individual patient is very sharp, almost an all-or-none transition [12]. This simply confirms the observations of generations of clinicians. And this raises the question of how to account for such a transition in terms of the known molecular properties of the anesthetic agent.

The mechanism of general anesthesia is still under considerable investigation. Theories range from a nonspecific perturbation of the lipid bilayer membrane of neurons, the cells responsible for the “information” function of the central nervous system, to the interaction of the anesthetic agent with specific protein receptors [13]. It is certainly possible that if the mechanism of general anesthesia is the binding of the anesthetic agent to a specific receptor protein, then the nearly all-or-none transition or bifurcation from the awake state to the anesthetized state could be explained by a highly cooperative binding of the anesthetic to the receptor. In fact, it has been common to mathematically model the probability of responsiveness to drug concentration using the Hill equation, a simplified static equation originally derived in 1909 to describe the cooperative binding of oxygen to the hemoglobin molecule [14]. However, an alternative explanation could be sought in the dynamic network properties of the brain.

The human central nervous system involves a complex large-scale interconnected neural network involving feedforward and feedback (or recurrent) networks, with the brain serving as the central element of this network system. The brain is interconnected to receptors that transmit sensory information to the brain, and in turn the brain delivers action commands to effectors. The neural network of the brain consists of approximately 10^{11} neurons (nerve cells) with each having 10^4 to 10^5 connections interconnected through subnetworks or nuclei. The nuclei in turn consist of clusters of neurons each of which performs a specific and defined function.

The most basic characteristic of the neurons that comprise the central nervous system is the electrochemical potential gradient across the cell membrane. All cells of the human body maintain an electrochemical potential gradient between the inside of the cell and the surrounding milieu. Neurons have the capacity of excitability. If stimulated beyond a threshold, then the neuron will “fire” and produce a large voltage spike (the action potential) before returning to the resting potential [3,4]. The neurons of the brain are connected in a complex network in which the firing of one neuron can be the stimulus for the firing of another neuron.

A major focus of theoretical neuroscience has been describing neuronal behavior in terms of this electrochemical potential, both at the single neuron level but more ambitiously, at the level of multi-neuron networks. In this type of analysis the specific properties of the single neuron that are most relevant are how the spike of a one neuron alters the electrochemical potential of another neuron, and how this change in the potential results in a neuronal spike. The physical connection between neurons occurs in the synapse, a small gap between the axon, the extension of the cell body of the transmitting neuron, and the dendrite, the extension of the receiving neuron. The signal is transmitted by the release of a neurotransmitter from the axon into the synapse. This neurotransmitter diffuses across the synapse, binds to a postsynaptic receptor membrane protein on the dendrite, and alters the electrochemical potential of the receiving neuron.

It is possible that the anesthetic bifurcation to unconsciousness or the nearly all-or-none characteristic induction of anesthesia is a type of phase transition of the neural network. This possibility was first considered by Steyn-Ross *et al.* (see [15] and the references therein). Their focus was on the mean voltage of the soma, or cell body, of neurons. Specifically, the authors in [15] show that the biological change of state to anesthetic unconsciousness is analogous to a thermodynamic phase change involving a liquid to solid phase transition. For certain ranges of anesthetic concentrations, their first-order model predicts the existence of multiple steady states for brain activity leading to a transition from normal levels of cerebral cortical activity to a quiescent, low-firing state.

In this paper, we develop an alternative approach to the possibility of neuronal network phase transition in terms of neuronal firing rates, using the concept of *multistability* for dynamical systems. Multistability is the property whereby the solutions of a dynamical system can alternate between two or more mutually exclusive Lyapunov stable and convergent states under asymptotically slowly changing inputs or system parameters. In particular, multistable systems give rise to the existence of multiple (isolated and/or a continuum of) Lyapunov stable equilibria involving a quasistatic-like behavior between these multiple *semistable* steady states [16–18]. Semistability is the property whereby the solutions of a dynamical system converge to Lyapunov stable equilibrium points determined by the system initial conditions [19,20].

Multistability is ubiquitous in biological systems ranging from biochemical networks to ecosystems to gene regulation and cell replication [21–23]. Since molecular studies suggest that one possible mechanism of action of anesthetics is the inhibition of synaptic transmission in cortical neurons [24,25], this suggests that general anesthesia is a phenomenon in which different equilibria can be attained with changing anesthetic agent concentrations. Hence, multistability theory can potentially provide a theoretical foundation for describing general anesthesia.

Although general anesthesia has been used in the clinical practice of medicine for over 150 years, the mechanism of action is still not fully understood [13] and is still under considerable investigation [5–9]. Early theories postulated that anesthesia is produced by disturbance of the physical properties of cell membranes. The work of Meyer and Overton [26,27] demonstrated that for some anesthetics there was a correlation between anesthetic potency and solubility in fat-like solvents. This led to a theory that anesthesia resulted from a nonspecific perturbation of the lipid bilayer membrane of neurons [9,28]. Subsequent research then found that membrane proteins performed functions of excitability and this led to a focus on anesthetic binding and perturbation of hydrophobic regions of membrane proteins [29]. Further research also revealed that some anesthetic gases follow the Meyer-Overton correlation but do not produce anesthesia and some Meyer-Overton gases are excitatory and can cause seizures [30,31]. These results led to the more common modern focus on the interaction of the anesthetic agent with specific protein receptors [13].

In particular, there has been extensive investigation of the influence of anesthetic agents on the binding of neurotransmitters to their postsynaptic receptors [7–9]. A plethora of receptors have been investigated, including receptors for glycine, serotonin type 2 and 3, N-methyl-d-aspartate (NMDA), α -2 adrenoreceptors, α -amino-3-hydroxy-5-methyl-4-isoxazolepropionic acid (AMPA), histamine, acetylcholine, and γ -aminobutyric acid (GABA). One attractive aspect of this focus on postsynaptic receptors is it facilitates mathematical analysis on the basis of the effect of receptor binding on the

postsynaptic potential. This is in marked contrast to the Meyer-Overton hypothesis, which failed to explicitly detail how a nonspecific perturbation of the lipid membrane would result in the anesthetic state.

In parallel with the investigation of the molecular interactions of general anesthetic agents, there has also been active investigation of the anatomic pathways involved in the transition from consciousness to anesthesia [5]. There is compelling evidence that the immobility created by some anesthetics is mediated at the level of the spinal cord. In contrast, functional imaging and electroencephalograph analysis has suggested that the site of suppression of consciousness is the thalamus, and thalamocortical tracts may play a critical role in the suppression of consciousness [9].

Despite these advances in our understanding of the molecular interactions of anesthetic agents and of specific anatomic loci for the action of anesthetic agents, there has been less development of a mathematical framework to understand this fascinating and clinically important phenomenon. It is certainly possible that if the mechanism of general anesthesia is the binding of the anesthetic agent to a specific receptor protein, then the nearly all-or-none transition from the awake state to the anesthetized state could be explained by a highly cooperative binding of the anesthetic to the receptor. In fact, as noted above, it has been common to mathematically model the probability of responsiveness to drug concentration using the Hill equation [14]. However, to date, no single unifying receptor mediating general anesthesia has been identified.

Rather, the most likely explanation for the mechanisms of action of anesthetics lies in the network properties of the brain. It is well established that there are two general types of neurons in the central nervous system—excitatory and inhibitory—interconnected in a complex dynamic network. The action potential of a spiking neuron is propagated along the axon to synapses where chemical neurotransmitters are released that generate a postsynaptic potential on the dendrites of connected neurons. Excitatory neurons generate a depolarizing postsynaptic potential on the dendrite of the connected neuron and if the depolarization is of sufficient magnitude, then a spike will be induced in the connected neuron. In contrast, inhibitory neurons generate a hyperpolarizing postsynaptic potential; an effect that acts to maintain a quiescent state.

There is considerable evidence that general anesthetics alter postsynaptic potentials [24,25]. An interesting example of how changes in the postsynaptic potential may be applied to the analysis of the induction of anesthesia is the view of anesthesia as a phase transition proposed by Steyn-Ross *et al.* (see [15] and the references therein). While their analysis was highly informative, in this paper we use a dynamical system theory framework in terms of neuronal firing rates, using the concepts of network thermodynamics [32] and multistability, for explaining the mechanisms of action for general anesthesia. This facilitates a focus on the network properties of the central nervous system. The firing rate models used for network analysis must have sufficient generality and include parameters that can account for such relevant physiological changes at the single neuron level. The synaptic drive firing model, introduced by Ermentrout and his collaborators [4,32], and the system thermodynamic framework, introduced in [32], is the underlying framework for this paper.

In this paper, the fundamental building block of the central nervous system, the neuron, is represented as a dynamic element that is “excitable”, and can generate a pulse or spike whenever the electrochemical potential across the cell membrane of the neuron exceeds a threshold value. More specifically, a nonlinear discontinuous system framework is developed in Sections 2 and 3 for describing the

relationship between the synaptic voltage and firing rates of excitatory and inhibitory neural networks. To establish convergence and semistability for discontinuous dynamical systems we introduce the notion of nontangency between a discontinuous vector field and a weakly invariant or weakly negatively invariant subset of level or sublevel sets of Lyapunov functions in Section 5. Specifically, to capture the notion of nontangency we introduce the direction cone of a discontinuous vector field. Then, using positive limit sets, restricted prolongations, and nontangency we develop Lyapunov analysis for convergence and semistability to establish multistability for discontinuous dynamical systems. Here, the restricted prolongation of a point is a subset of its positive prolongation as defined in [33]. In addition, using nontangency, we present Lyapunov results for convergence and semistability to develop sufficient conditions for multistability for discontinuous dynamical systems.

While previous treatments of nontangency-based Lyapunov tests for convergence and semistability for dynamical systems with continuous vector fields are given in [19], our results involve dynamical systems with discontinuous vector fields for capturing *plasticity* (*i.e.*, dynamic network connections) in our neural network model generating absolutely continuous solutions necessitating stability analysis via nonsmooth Lyapunov stability theory involving Clarke generalized gradients and set-valued Lie derivatives. Using the aforementioned dynamical system framework, we apply the results of Sections 4 and 5 to excitatory-inhibitory firing neural models in an attempt to understand the mechanisms of action of anesthetics. While there is ongoing debate as to whether information is encoded by the firing rates (*i.e.*, rate coding) of spiking neurons or by precise timing of single neuron spikes (*i.e.*, temporal coding) [34], it is evident that firing rates do characterize central nervous system activity. Firing rates are nonnegative entities and the nonnegativity constraint for neural network activity can be easily incorporated within nonlinear dynamical system theory using solutions of differential equations and differential inclusions evolving in cones [35].

There is extensive experimental verification that collections of neurons may function as oscillators and the synchronization of oscillators may play a key role in the transmission of information within the central nervous system. This may be particularly relevant to understand the mechanism of action for general anesthesia. In Section 7, we provide sufficient conditions for global asymptotic and exponential synchronization for our excitatory and inhibitory cortical neuronal network.

To avoid the complexity of large-scale and high connectivity models of the neural network of the brain, the scale and connectivity of the network can be simplified using mean field theories. Early mean field theories assumed that the brain is organized into a limited number of pools of identical spiking neurons [36]. However, more commonly mean field theories assume that the strength of connection between neurons is normally distributed around some mean value. Mean field theories can impose self-consistency on field variables; for example, if postsynaptic potentials are assumed to be a function of some mean firing rate, then those postsynaptic potentials should lead to a consistent predicted mean firing rate. The idea of applying mean field theories, drawn from the study of condensed matter, originated with [37]. Subsequently, Sompolinsky *et al.* [38] developed a mean field theory for neural networks analogous to the equations developed for spin glasses with randomly symmetric bonds [39]. The authors in [40] investigated the stability of system states for a network of integrate-and-fire neurons, whereas the authors in [41] extended this theoretical model to the analysis of oscillations. Gerstner *et al.* [42,43]

subsequently developed a mean field theory using a spike response model and demonstrated that the integrate-and-fire model was a special case of the spike response model.

In Section 8 of the paper, we extend our results further by demonstrating multistability in the mean when the coefficients of the neuronal connectivity matrix are random variables. Specifically, we use a stochastic multiplicative uncertainty model to include modeling of a priori uncertainty in the coefficients of the neuronal connectivity matrix by means of state-dependent noise. Our stochastic multiplicative uncertainty model uses state-dependent Gaussian white noise to represent parameter uncertainty by defining a measure of ignorance, in terms of an information-theoretic entropy, and then determining the probability distribution which maximizes this measure subjected to agreement with a given model. To account for time delay and memory effects in inhibitory and excitatory networks, in Sections 9 and 10 we extend the results of Section 8 to a large-scale excitatory and inhibitory synaptic drive firing rate model with time-varying delays and stochastic input uncertainty, and global mean-square synchronization of this model is investigated.

Finally, in Sections 11 and 12 we discuss key emergent properties from the fields of thermodynamics and electromagnetic field theory for developing plausible mechanisms of action for general anesthesia. Specifically, in Section 11, we highlight how the supreme law of nature—the second law of thermodynamics—can be used to arrive at mechanistic models for the anesthetic cascade using the principle of maximum entropy production and thermodynamics as applied to the human brain. In Section 12, we use the idea of anesthetics disrupting the inflow of free energy to the brain as electrical signals that generate electromagnetic fields causing a shielding effect leading to the emergence of unconsciousness.

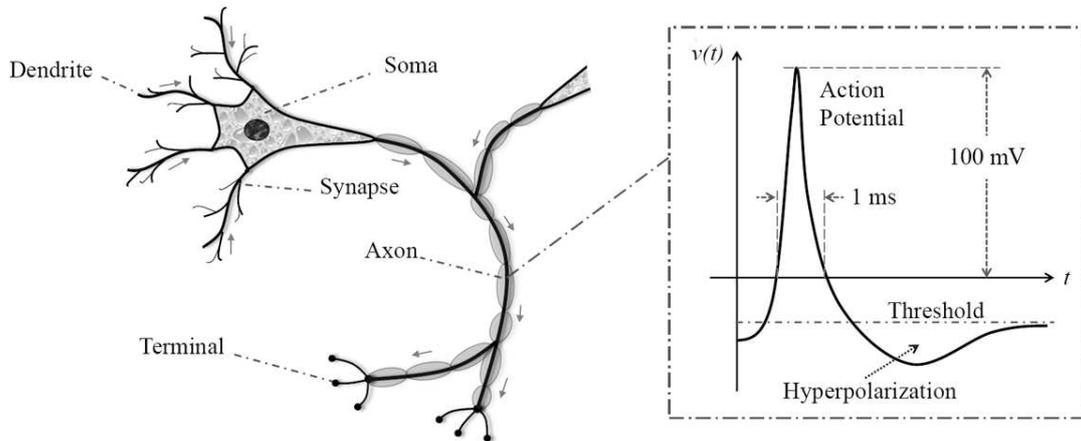
2. Biological Neural Networks: A Dynamical Systems Approach

The fundamental building block of the central nervous system, the *neuron*, can be divided into three functionally distinct parts, namely, the *dendrites*, *soma* (or cell body), and *axon* (see Figure 1). The dendrites play the role of input devices that collect signals from other neurons and transmit them to the soma; whereas the soma generates a signal that is transmitted to other neurons by the axon. The axons of other neurons connect to the dendrites and soma surfaces by means of connectors called *synapses*. The behavior of the neuron is best described in terms of the electrochemical potential gradient across the cell membrane. If the voltage gradient across the membrane increases to a critical threshold value, then there is a subsequent abrupt step-like increase in the potential gradient, the action potential. This action potential is transmitted from the soma along the axon to a dendrite of a receiving neuron. The action potential elicits the release of neurotransmitter molecules that diffuse to the dendrite of a “receiving” neuron. This alters the voltage gradient across the receiving neuron.

The electrochemical potential for a neuron can be described by a nonlinear four-state system [32]. Coupling these system equations for each neuron in a large neural population is computationally prohibitive. To simplify the mathematical modeling, it has been common to use phenomenological firing rate models for studying neural coding, memory, and network dynamics [4]. Firing rate models involve the averaged behavior of the spiking rates of groups of neurons rather than tracking the spike rate of each individual neuron cell. In such population models, the activity of a neuron, that is, the rate at which

the neuron generates an action potential (*i.e.*, “fires”) is modeled as a function of the voltage (across the membrane). The “firing” of a neuron evokes voltage changes, postsynaptic potentials on receiving neurons; that is, neurons electrically connected to the firing neurons via axon-dendrite connections.

Figure 1. Neuron anatomy.



In general, neurons are either excitatory or inhibitory depending on whether the postsynaptic potential increases or decreases the potential of the receiving neuron. In particular, excitatory neurotransmitters *depolarize* postsynaptic membranes by increasing membrane potentials and can collectively generate an action potential. Inhibitory neurotransmitters *hyperpolarize* the postsynaptic membrane by decreasing membrane potentials, thereby nullifying the actions of excitatory neurotransmitters and in certain cases prevent the generation of action potentials.

Biological neural network models predict a voltage in the receiving or postsynaptic neuron given by

$$v_i^X(t) = \sum_{j=1}^{n_E} A_{ij}^{XE} \sum_k \alpha_j^E(t - t_k) + \sum_{j'=1}^{n_I} A_{ij'}^{XI} \sum_{k'} \alpha_{j'}^I(t - t_{k'}), \tag{1}$$

where $v_i^X(\cdot)$, $X \in \{E, I\}$, $i = 1, \dots, n_E + n_I$, is the excitatory ($X = E$) and inhibitory ($X = I$) voltage in the i th receiving neuron, A_{ij}^{XY} , $X, Y \in \{E, I\}$, are constants representing the coupling strengths (in volts) of the j th neuron on the i th neuron, $k, k' = 1, \dots$, enumerate the action potential or firings of the excitatory and inhibitory transmitting (presynaptic) neurons at firing times t_k and $t_{k'}$, respectively, and $\alpha_j^E(\cdot)$ and $\alpha_{j'}^I(\cdot)$ are dimensionless functions describing the evolution of the excitatory and inhibitory postsynaptic potentials, respectively. Using a (possibly discontinuous) function $f_i(\cdot)$ to represent the firing rate (in Hz) of the i th neuron and assuming that the firing rate is a function of the voltage $v_i^E(\cdot)$ (resp., $v_i^I(\cdot)$) across the membrane of the i th neuron given by $f_i(v_i^E)$ (resp., $f_i(v_i^I)$), it follows that

$$v_i^E(t) = \sum_{j=1, j \neq i}^{n_E} A_{ij}^{EE} \int_{-\infty}^t \alpha_j^E(t - \tau) f_j(v_j^E(\tau)) d\tau + \sum_{j'=1}^{n_I} A_{ij'}^{EI} \int_{-\infty}^t \alpha_{j'}^I(t - \tau) f_{j'}(v_{j'}^I(\tau)) d\tau + v_{th_i}^E(t), \tag{2}$$

$i = 1, \dots, n_E,$

$$v_i^I(t) = \sum_{j=1}^{n_E} A_{ij}^{IE} \int_{-\infty}^t \alpha_j^E(t - \tau) f_j(v_j^E(\tau)) d\tau + \sum_{j'=1, j' \neq i}^{n_I} A_{ij'}^{II} \int_{-\infty}^t \alpha_{j'}^I(t - \tau) f_{j'}(v_{j'}^I(\tau)) d\tau + v_{th_i}^I(t), \tag{3}$$

$i = 1, \dots, n_I,$

where the *neuronal connectivity matrix* A^{XY} , with units of volts, contains entries A_{ij}^{XY} , $X, Y \in \{E, I\}$, representing the coupling strength of the j th neuron on the i th neuron such that $A_{ij}^{XE} > 0$ and $A_{ij}^{XI} < 0$, $X \in \{E, I\}$, if the j th neuron is connected (*i.e.*, contributes a postsynaptic potential) to the i th neuron, and $A_{ij}^{XY} = 0$, otherwise. Furthermore, $v_{thi}^E(\cdot)$ and $v_{thi}^I(\cdot)$ are continuous input voltages characterizing nerve impulses from sensory (pain) receptors, sensorimotor (temperature sensing) receptors, or proprioceptive (motion sensing) receptors. Alternatively, $v_{thi}^E(\cdot)$ and $v_{thi}^I(\cdot)$ can be thought of as inputs from the reticular activating system within the brainstem responsible for regulating arousal and sleep-wake transitions. Note that $A_{ii}^{EE} \triangleq A_{ii}^{II} \triangleq 0$ by definition.

Next, defining the *synaptic drive*—a dimensionless quantity—of each (excitatory or inhibitory) neuron by

$$S_i^{(E,I)}(t) \triangleq \int_{-\infty}^t \alpha_i^{(E,I)}(t - \tau) f_i(v_i^{(E,I)}(\tau)) d\tau, \tag{4}$$

and assuming an exponential decay of the synaptic voltages of the form ([4,44])

$$\alpha_i^{(E,I)}(t) = B_i^{(E,I)} e^{-\frac{t}{\lambda_i^{(E,I)}}}, \tag{5}$$

where the dimensionless gain $B_i^{(E,I)}$ is equal to B_i^E if the i th neuron is excitatory and B_i^I if the i th neuron is inhibitory, and similarly for $S_i^{(E,I)}$, $v_i^{(E,I)}$, $\alpha_i^{(E,I)}$, and $\lambda_i^{(E,I)}$, it follows from Equations (4) and (5) that

$$\frac{dS_i^{(E,I)}(t)}{dt} = -\frac{1}{\lambda_i^{(E,I)}} S_i^{(E,I)}(t) + B_i^{(E,I)} f_i(v_i^{(E,I)}(t)). \tag{6}$$

Now, using the expressions for the excitatory and inhibitory voltage given by Equations (2) and (3), respectively, it follows that

$$\frac{dS_i^E(t)}{dt} = -\frac{1}{\lambda_i^E} S_i^E(t) + B_i^E f_i \left(\sum_{j=1, j \neq i}^{n_E} A_{ij}^{EE} S_j^E(t) + \sum_{j'=1}^{n_I} A_{ij'}^{EI} S_{j'}^I(t) + v_{thi}^E(t) \right), \quad i = 1, \dots, n_E, \tag{7}$$

$$\frac{dS_i^I(t)}{dt} = -\frac{1}{\lambda_i^I} S_i^I(t) + B_i^I f_i \left(\sum_{j=1}^{n_E} A_{ij}^{IE} S_j^E(t) + \sum_{j'=1, j' \neq i}^{n_I} A_{ij'}^{II} S_{j'}^I(t) + v_{thi}^I(t) \right), \quad i = 1, \dots, n_I. \tag{8}$$

The above analysis reveals that a form for capturing the neuroelectric behavior of biological excitatory and inhibitory neuronal networks can be written as

$$\frac{dS_i(t)}{dt} = -\frac{1}{\tau_i} S_i(t) + B_i f_i \left(\sum_{j=1}^n A_{ij} S_j(t) + v_{thi}(t) \right), \quad S_i(0) = S_{i0}, \quad t \geq 0, \quad i = 1, \dots, n, \tag{9}$$

where $S_i(t) \in \mathcal{D} \subseteq \mathbb{R}$, $t \geq 0$, is the i th synaptic drive, $v_{thi}(t) \in \mathbb{R}$, $t \geq 0$, denotes the input voltage to the i th neuron, A_{ij} is a constant representing the coupling strength of the j th neuron on the i th neuron, τ_i is a time constant, B_i is a constant gain for the firing rate of the i th neuron, and $f_i(\cdot)$ is a nonlinear activation function describing the relationship between the synaptic drive and the firing rate of the i th neuron.

In this paper, we will explore different activation functions including discontinuous hard-limiter activation functions, and continuous half-wave rectification, saturation, and sigmoidal functions. Specifically, for a typical neuron ([3])

$$f_i(x) = [x]_+, \tag{10}$$

where $i \in \{1, \dots, n\}$ and $[x]_+ = x$ if $x \geq 0$, and $[x]_+ = 0$, otherwise. Alternatively, we can approximate $f_i(x)$ by the smooth (i.e., infinitely differentiable) half-wave rectification function

$$f_i(x) = \frac{x e^{\gamma x}}{1 + e^{\gamma x}}, \tag{11}$$

where $i \in \{1, \dots, n\}$ and $\gamma \gg 0$. Note that $f'_i(x) \approx 1$ for $x > 0$ and $f''_i(x) \approx 0$, $x \neq 0$. In addition, note that Equations (10) and (11) reflect the fact that as the voltage increases across the membrane of the i th neuron, the firing rate increases as well. Often, the membrane potential-firing rate curve exhibits a linear characteristic for a given range of voltages. At higher voltages, however, a saturation phenomenon appears, indicating that the full effect of the firing rate has been reached. To capture this effect, $f_i(\cdot)$ can be modeled as

$$f_i(x) = \frac{f_{\max}(e^{\gamma x} - 1)}{1 + e^{\gamma x}}, \tag{12}$$

where $i \in \{1, \dots, n\}$, $\gamma \gg 0$, and $f_{\max} = \lim_{\gamma \rightarrow \infty} f_i(x)$ denotes the maximum firing rate.

3. Connections to Mean Field Excitatory and Inhibitory Synaptic Drive Models

The excitatory and inhibitory neural network model given by Equations (7) and (8) can possess multiple equilibria. For certain values of the model parameters it can be shown that as the inhibitory time constants $\lambda_i^I, i = 1, \dots, n_I$, get larger, multiple stable and unstable state equilibria can appear [45]. Since molecular studies suggest that one possible mechanism of action of anesthetics is the prolongation of the time constants of inhibitory neurons [24,25], this suggests that general anesthesia is a phenomenon in which different equilibria can be attained with changing anesthetic agent concentrations. To explore this multistability phenomenon, in [45,46] we developed a simplified scale and connectivity neural network model using a mean field theory. As noted in the Introduction, mean field theories assume that the brain is organized into limited number of pools of identical spiking neurons [36]. However, more commonly, mean field theories assume that the strength of the connection between neurons is normally distributed around some mean value.

To see how our general excitatory and inhibitory synaptic drive model given by Equations (7) and (8) can be reduced to a mean excitatory and mean inhibitory model, consider Equations (7) and (8) with continuously differentiable $f_i(\cdot) = f(\cdot)$, $B_i^E = B_i^I = 1$, $\lambda_i^E = \lambda^E$, and $\lambda_i^I = \lambda^I$. In this case, Equations (7) and (8) become

$$\frac{dS_i^E(t)}{dt} = f \left(\sum_{j=1}^{n_E} A_{ij}^{EE} S_j^E(t) + \sum_{k=1}^{n_I} A_{ik}^{EI} S_k^I(t) + v_{th_i}^E(t) \right) - \frac{1}{\lambda^E} S_i^E(t), \quad i = 1, \dots, n_E, \tag{13}$$

$$\frac{dS_i^I(t)}{dt} = f \left(\sum_{j=1}^{n_E} A_{ij}^{IE} S_j^E(t) + \sum_{k=1}^{n_I} A_{ik}^{II} S_k^I(t) + v_{th_i}^I(t) \right) - \frac{1}{\lambda^I} S_i^I(t), \quad i = 1, \dots, n_I, \tag{14}$$

where $f(\cdot)$ is given by either Equation (11) or Equation (12) and $A_{ii}^{EE} = A_{ii}^{II} = 0$. Next, let $A_{ij}^{EE} = \bar{A}^{EE} + \Delta_{ij}^{EE}$, $A_{ij}^{EI} = \bar{A}^{EI} + \Delta_{ij}^{EI}$, $A_{ij}^{IE} = \bar{A}^{IE} + \Delta_{ij}^{IE}$, and $A_{ij}^{II} = \bar{A}^{II} + \Delta_{ij}^{II}$, where $\bar{A}^{XY} \triangleq \frac{1}{n_X n_Y} \sum_{i=1}^{n_X} \sum_{j=1}^{n_Y} A_{ij}^{XY}$, $X, Y \in \{E, I\}$, denote mean and $\Delta_{ij}^{XY}, X, Y \in \{E, I\}$, are deviations from the mean. In this case, it follows that

$$\sum_{i=1}^{n_E} \sum_{j=1}^{n_E} \Delta_{ij}^{EE} = \sum_{i=1}^{n_E} \sum_{j=1}^{n_I} \Delta_{ij}^{EI} = \sum_{i=1}^{n_I} \sum_{j=1}^{n_E} \Delta_{ij}^{IE} = \sum_{i=1}^{n_I} \sum_{j=1}^{n_I} \Delta_{ij}^{II} = 0. \tag{15}$$

Using the average and perturbed expression for A_{ij}^{XY} , $X, Y \in \{E, I\}$, Equations (13) and (14) can be rewritten as

$$\frac{dS_i^E(t)}{dt} = f \left(n_E \bar{A}^{EE} \bar{S}^E(t) + \sum_{j=1}^{n_E} \Delta_{ij}^{EE} S_j^E(t) + n_I \bar{A}^{EI} \bar{S}^I(t) + \sum_{k=1}^{n_I} \Delta_{ik}^{EI} S_k^I(t) + v_{thi}^E(t) \right) - \frac{1}{\lambda^E} S_i^E(t), \tag{16}$$

$i = 1, \dots, n_E,$

$$\frac{dS_i^I(t)}{dt} = f \left(n_E \bar{A}^{IE} \bar{S}^E(t) + \sum_{j=1}^{n_E} \Delta_{ij}^{IE} S_j^E(t) + n_I \bar{A}^{II} \bar{S}^I(t) + \sum_{k=1}^{n_I} \Delta_{ik}^{II} S_k^I(t) + v_{thi}^I(t) \right) - \frac{1}{\lambda^I} S_i^I(t), \tag{17}$$

$i = 1, \dots, n_I,$

where $\bar{S}^E(t) \triangleq \frac{1}{n_E} \sum_{j=1}^{n_E} S_j^E(t)$ and $\bar{S}^I(t) \triangleq \frac{1}{n_I} \sum_{j=1}^{n_I} S_j^I(t)$ denote the mean excitatory synaptic drive and mean inhibitory synaptic drive in dimensionless units, respectively. Now, defining $\delta_i^E(t) \triangleq S_i^E(t) - \bar{S}^E(t)$ and $\delta_i^I(t) \triangleq S_i^I(t) - \bar{S}^I(t)$, where $\delta_i^E(t)$ and $\delta_i^I(t)$ are deviations from the mean, Equations (16) and (17) become

$$\begin{aligned} \frac{dS_i^E(t)}{dt} = & f \left(n_E \bar{A}^{EE} \bar{S}^E(t) + \bar{S}^E(t) \sum_{j=1}^{n_E} \Delta_{ij}^{EE} + n_I \bar{A}^{EI} \bar{S}^I(t) + \bar{S}^I(t) \sum_{k=1}^{n_I} \Delta_{ik}^{EI} + \sum_{j=1}^{n_E} \Delta_{ij}^{EE} \delta_j^E(t) \right. \\ & \left. + \sum_{k=1}^{n_I} \Delta_{ik}^{EI} \delta_k^I(t) + v_{thi}^E(t) \right) - \frac{1}{\lambda^E} S_i^E(t), \quad i = 1, \dots, n_E, \end{aligned} \tag{18}$$

$$\begin{aligned} \frac{dS_i^I(t)}{dt} = & f \left(n_E \bar{A}^{IE} \bar{S}^E(t) + \bar{S}^E(t) \sum_{j=1}^{n_E} \Delta_{ij}^{IE} + n_I \bar{A}^{II} \bar{S}^I(t) + \bar{S}^I(t) \sum_{k=1}^{n_I} \Delta_{ik}^{II} + \sum_{j=1}^{n_E} \Delta_{ij}^{IE} \delta_j^E(t) \right. \\ & \left. + \sum_{k=1}^{n_I} \Delta_{ik}^{II} \delta_k^I(t) + v_{thi}^I(t) \right) - \frac{1}{\lambda^I} S_i^I(t), \quad i = 1, \dots, n_I. \end{aligned} \tag{19}$$

Next, assume that all terms with a factor Δ_{ij}^{XY} , $X, Y \in \{E, I\}$, $i = 1, \dots, n_X$ and $j = 1, \dots, n_Y$, in Equations (18) and (19) are small relative to the remaining terms in $f(\cdot)$. Then, a first-order expansion of Equations (18) and (19) gives

$$\begin{aligned} \frac{dS_i^E(t)}{dt} = & f \left(n_E \bar{A}^{EE} \bar{S}^E(t) + n_I \bar{A}^{EI} \bar{S}^I(t) + v_{thi}^E(t) \right) + f' \left(n_E \bar{A}^{EE} \bar{S}^E(t) + n_I \bar{A}^{EI} \bar{S}^I(t) + v_{thi}^E(t) \right) \\ & \times \left[\bar{S}^E(t) \sum_{j=1}^{n_E} \Delta_{ij}^{EE} + \bar{S}^I(t) \sum_{k=1}^{n_I} \Delta_{ik}^{EI} + \sum_{j=1}^{n_E} \Delta_{ij}^{EE} \delta_j^E(t) + \sum_{k=1}^{n_I} \Delta_{ik}^{EI} \delta_k^I(t) \right] - \frac{1}{\lambda^E} S_i^E(t), \end{aligned} \tag{20}$$

$i = 1, \dots, n_E,$

$$\begin{aligned} \frac{dS_i^I(t)}{dt} = & f \left(n_E \bar{A}^{IE} \bar{S}^E(t) + n_I \bar{A}^{II} \bar{S}^I(t) + v_{thi}^I(t) \right) + f' \left(n_E \bar{A}^{IE} \bar{S}^E(t) + n_I \bar{A}^{II} \bar{S}^I(t) + v_{thi}^I(t) \right) \\ & \times \left[\bar{S}^E(t) \sum_{j=1}^{n_E} \Delta_{ij}^{IE} + \bar{S}^I(t) \sum_{k=1}^{n_I} \Delta_{ik}^{II} + \sum_{j=1}^{n_E} \Delta_{ij}^{IE} \delta_j^E(t) + \sum_{k=1}^{n_I} \Delta_{ik}^{II} \delta_k^I(t) \right] - \frac{1}{\lambda^I} S_i^I(t), \end{aligned} \tag{21}$$

$i = 1, \dots, n_I.$

Now, assuming that the higher-order terms can be ignored, Equations (20) and (21) become

$$\begin{aligned} \frac{dS_i^E(t)}{dt} = & f \left(n_E \bar{A}^{EE} \bar{S}^E(t) + n_I \bar{A}^{EI} \bar{S}^I(t) + v_{th_i}^E(t) \right) + f' \left(n_E \bar{A}^{EE} \bar{S}^E(t) + n_I \bar{A}^{EI} \bar{S}^I(t) + v_{th_i}^E(t) \right) \\ & \times \left[\bar{S}^E(t) \sum_{j=1}^{n_E} \Delta_{ij}^{EE} + \bar{S}^I(t) \sum_{k=1}^{n_I} \Delta_{ik}^{EI} \right] - \frac{1}{\lambda^E} S_i^E(t), \quad i = 1, \dots, n_E, \end{aligned} \tag{22}$$

$$\begin{aligned} \frac{dS_i^I(t)}{dt} = & f \left(n_E \bar{A}^{IE} \bar{S}^E(t) + n_I \bar{A}^{II} \bar{S}^I(t) + v_{th_i}^I(t) \right) + f' \left(n_E \bar{A}^{IE} \bar{S}^E(t) + n_I \bar{A}^{II} \bar{S}^I(t) + v_{th_i}^I(t) \right) \\ & \times \left[\bar{S}^E(t) \sum_{j=1}^{n_E} \Delta_{ij}^{IE} + \bar{S}^I(t) \sum_{k=1}^{n_I} \Delta_{ik}^{II} \right] - \frac{1}{\lambda^I} S_i^I(t), \quad i = 1, \dots, n_I. \end{aligned} \tag{23}$$

Finally, summing Equations (22) and (23) over $i = 1, \dots, n_E$ and $i = 1, \dots, n_I$, dividing by n_E and n_I , respectively, using Equation (15), and assuming $v_{th_1}^E(t) = v_{th_2}^E(t) = \dots = v_{th_{n_E}}^E(t) = v_{th}^E$ and $v_{th_1}^I(t) = v_{th_2}^I(t) = \dots = v_{th_{n_I}}^I(t) = v_{th}^I, t \geq 0$, it follows that the average excitatory synaptic drive and the average inhibitory synaptic drive are given by

$$\frac{d\bar{S}^E(t)}{dt} = f \left(n_E \bar{A}^{EE} \bar{S}^E(t) + n_I \bar{A}^{EI} \bar{S}^I(t) + v_{th}^E \right) - \frac{1}{\lambda^E} \bar{S}^E(t), \quad \bar{S}^E(0) = \bar{S}_0^E, \quad t \geq 0, \tag{24}$$

$$\frac{d\bar{S}^I(t)}{dt} = f \left(n_E \bar{A}^{IE} \bar{S}^E(t) + n_I \bar{A}^{II} \bar{S}^I(t) + v_{th}^I \right) - \frac{1}{\lambda^I} \bar{S}^I(t), \quad \bar{S}^I(0) = \bar{S}_0^I. \tag{25}$$

Equations (24) and (25) represent the spatial average (mean) dynamics of the system given by Equations (13) and (14), and are predicated on a mean field assumption that reduces the complex (approximately $10^{11} \times 10^{11}$) neuronal connectivity matrix to a 2×2 excitatory–inhibitory system. This is a drastic assumption, but one which has been commonly used in theoretical neuroscience going back to the pioneering work of Wilson and Cowan [36]. Preliminary results using the simplified two-state mean excitatory and inhibitory synaptic drive model given by Equations (24) and (25) for connecting notions of general anesthesia to multistability and bifurcations are given in [45,46].

4. Multistability Theory and Discontinuous Spiking Neuron Models

Multistability is the property whereby the solutions of a dynamical system can alternate between two or more mutually exclusive semistable states under asymptotically slowly changing inputs or system parameters. In particular, the state of a multistable system converges to Lyapunov stable equilibria that belong to an equilibrium set that has a multivalued hybrid topological structure consisting of isolated points and closed sets homeomorphic to intervals on the real line. To develop the notion of multistability, consider the autonomous differential equation given by

$$\dot{x}(t) = f(x(t)), \quad x(0) = x_0, \quad \text{a.e. } t \geq 0, \tag{26}$$

where, for every $t \geq 0, x(t) \in \mathcal{D} \subseteq \mathbb{R}^n, f : \mathbb{R}^n \rightarrow \mathbb{R}^n$ is Lebesgue measurable and locally essentially bounded [47], that is, f is bounded on a bounded neighborhood of every point x , excluding sets of measure zero. Furthermore, let $\mathcal{E}_e \triangleq \{x \in \mathbb{R}^n : f(x) = 0\}$ denote the set of equilibria for Equation (26).

Definition 1 ([47,48]) An absolutely continuous function $x : [0, \tau] \rightarrow \mathbb{R}^n$ is said to be a Filippov solution of Equation (26) on the interval $[0, \tau]$ with initial condition $x(0) = x_0$, if $x(t)$ satisfies

$$\frac{d}{dt}x(t) \in \mathcal{K}[f](x(t)), \quad \text{a.e. } t \in [0, \tau], \tag{27}$$

where the Filippov set-valued map $\mathcal{K}[f] : \mathbb{R}^n \rightarrow 2^{\mathbb{R}^n}$ is defined by

$$\mathcal{K}[f](x) \triangleq \bigcap_{\delta > 0} \bigcap_{\mu(\mathcal{S})=0} \overline{\text{co}} \{f(\mathcal{B}_\delta(x) \setminus \mathcal{S})\}, \quad x \in \mathbb{R}^n,$$

and where $\mathcal{B}_\delta(x)$ denotes the open ball centered at x with radius δ , $2^{\mathbb{R}^n}$ denotes the collection of all subsets of \mathbb{R}^n , $\mu(\cdot)$ denotes the Lebesgue measure in \mathbb{R}^n , “ $\overline{\text{co}}$ ” denotes the convex closure, and $\bigcap_{\mu(\mathcal{S})=0}$ denotes the intersection over all sets \mathcal{S} of Lebesgue measure zero [49].

Note that since f is locally essentially bounded, $\mathcal{K}[f](\cdot)$ is upper semicontinuous and has nonempty, compact, and convex values. Thus, Filippov solutions are limits of solutions to Equation (26) with f averaged over progressively smaller neighborhoods around the solution point, and hence, allow solutions to be defined at points where f itself is not defined. Hence, the tangent vector to a Filippov solution, when it exists, lies in the convex closure of the limiting values of the system vector field $f(\cdot)$ in progressively smaller neighborhoods around the solution point. Note that $\mathcal{K}[f] : \mathbb{R}^n \rightarrow 2^{\mathbb{R}^n}$ is a map that assigns sets to points.

Dynamical systems of the form given by Equation (27) are called *differential inclusions* [50] and for each state $x \in \mathbb{R}^n$, they specify a *set* of possible evolutions rather than a single one. It follows from 1) of Theorem 1 of [51] that there exists a set $\mathcal{N}_f \subset \mathbb{R}^n$ of measure zero such that, for every set $\mathcal{W} \subset \mathbb{R}^n$ of measure zero,

$$\mathcal{K}[f](x) = \overline{\text{co}} \left\{ \lim_{i \rightarrow \infty} f(x_i) : x_i \rightarrow x, x_i \notin \mathcal{N}_f \cup \mathcal{W} \right\}, \tag{28}$$

where $\{x_i\}_{i \in \mathbb{Z}_+} \subset \mathbb{R}^n$ converges to $x \in \mathbb{R}^n$.

Differential inclusions include piecewise continuous dynamical systems as well as switched dynamical systems as special cases. For example, if $f(\cdot)$ is piecewise continuous, then Equation (26) can be represented as a differential inclusion involving the Filippov set-valued map of piecewise continuous vector fields given by $\mathcal{K}[f](x) = \overline{\text{co}} \{ \lim_{i \rightarrow \infty} f(x_i) : x_i \rightarrow x, x_i \notin \mathcal{S}_f \}$, where \mathcal{S}_f has measure zero and denotes the set of points where f is discontinuous [52]. Similarly, differential inclusions can include, as a special case, switched dynamical systems of the form

$$\dot{x}(t) = f_p(x(t)), \quad x(0) = x_0, \quad t \geq 0, \tag{29}$$

where $x(t) \in \mathcal{D} \subseteq \mathbb{R}^n$, $f_p : \mathbb{R}^n \rightarrow \mathbb{R}^n$ is locally Lipschitz continuous, and $p \in \mathcal{P} = \{1, \dots, d\}$ is a finite index set.

To see how the state-dependent switched dynamical system given by Equation (29) can be represented as a differential inclusion involving Filippov set-valued maps, define the switched system (29) with a piecewise linear partitioned state space as the triple $(\mathcal{D}, \mathcal{Q}, \mathcal{V})$, where $\mathcal{D} = \mathbb{R}^n$ or \mathcal{D} is a polytope in \mathbb{R}^n of dimension $\dim(\mathcal{D}) = n$, $\mathcal{Q} = \{\mathcal{D}_p\}_{p \in \mathcal{P}}$ is a piecewise linear partition of \mathcal{D} with index set \mathcal{P} , and $\mathcal{V} = \{f_p\}_{p \in \mathcal{P}^n}$, and where $f_p : \mathcal{U}_p \rightarrow \mathbb{R}^n$, \mathcal{U}_p is an open neighborhood of \mathcal{D}_p , and $\mathcal{P}^n = \{p \in \mathcal{P} : \dim(\mathcal{D}_p) = n\}$. Specifically, $\cup_{p=1}^d \mathcal{D}_p = \mathcal{D} \subseteq \mathbb{R}^n$, where \mathcal{D} is a polytope (*i.e.*, the convex hull of finitely many points, and hence, compact) and $\mathcal{D}_p, p = 1, \dots, d$, is a family of polyhedral sets in \mathbb{R}^n with nonempty interiors. Furthermore, for every $i, j \in \{1, \dots, d\}, i \neq j$, let $\overline{\mathcal{D}}_i \cap \overline{\mathcal{D}}_j = \emptyset$ or $\overline{\mathcal{D}}_i \cap \overline{\mathcal{D}}_j$ is a $(n - 1)$ -dimensional manifold included in the boundaries $\partial \mathcal{D}_i$ and $\partial \mathcal{D}_j$. Finally, since each

vector field f_i is Lipschitz continuous in the state x , it defines a continuously differentiable flow $\psi_i(t, x)$ within every open set $\mathcal{U}_i \supset \mathcal{D}_i$. In particular, each flow $\psi_i(t, x)$ is well defined on both sides of the boundary $\partial\mathcal{D}_j$. Thus, in the interior of each operating region \mathcal{D}_p , the global dynamics of Equation (29) is completely described by the local dynamics characterized by a particular vector field f_p , and hence, there exists a unique classical (i.e., continuously differentiable) solution to Equation (29). However, for a point $x \in \mathcal{H}_p \cap \mathcal{D}_p$, where \mathcal{H}_p is some supporting hyperplane of \mathcal{D}_p , nonuniqueness and nonexistence of solutions to Equation (29) can occur.

To address the problem of existence and extendability of solutions for Equation (29), let the global dynamics of Equation (29) be characterized by one of the following differential inclusions

$$\dot{x}(t) \in \mathcal{F}(x(t)) \triangleq \begin{cases} f_1(x(t)), & x(t) \in \mathcal{D}_1, \\ f_2(x(t)), & x(t) \in \mathcal{D}_2, \\ \vdots & \vdots \\ f_d(x(t)), & x(t) \in \mathcal{D}_d, \end{cases} \quad x(0) = x_0, \quad \text{a.e. } t \geq 0, \quad (30)$$

$$\dot{x}(t) \in \overline{\text{co}} \mathcal{F}(x(t)). \quad (31)$$

Note that $\mathcal{F} : \mathbb{R}^n \rightarrow 2^{\mathbb{R}^n}$ is nonempty and finite, and hence, compact. Moreover, since each $f_p : \mathcal{U}_p \rightarrow \mathbb{R}^n$ is continuous, the set-valued map $\mathcal{F}(\cdot)$ is upper semicontinuous, that is, for every $x \in \mathbb{R}^n$ and neighborhood \mathcal{N} of $\mathcal{F}(x)$, there exists a neighborhood \mathcal{M} of x such that $\mathcal{F}(\mathcal{M}) \subset \mathcal{N}$. However, it is important to note that $\mathcal{F}(x)$ is not convex. Alternatively, $\mathcal{F}^c(x) \triangleq \overline{\text{co}} \mathcal{F}(x)$ is convex, and hence, $\mathcal{F}^c : \mathbb{R}^n \rightarrow 2^{\mathbb{R}^n}$ is an upper semicontinuous set-valued map with nonempty, convex, and compact values. That is, for every $x \in \mathcal{D}$ and every $\varepsilon > 0$, there exists $\delta > 0$ such that, for all $z \in \mathbb{R}^n$ satisfying $\|z - x\| \leq \delta$, $\mathcal{F}^c(z) \subseteq \mathcal{F}^c(x) + \varepsilon \overline{\mathcal{B}}_1(0)$. In this case, it can be shown that there exists a Filippov solution to the differential inclusion given by Equation (30) at all interior points $x \in \mathcal{D} \subseteq \mathbb{R}^n$ [47,53]. In addition, if for each unbounded cell \mathcal{D}_p , $p \in \mathcal{P}^n$, $f_p(\mathcal{D}_p)$ is bounded and $\mathcal{F}^c(x) \cap T_x \mathcal{D} \neq \emptyset$, $x \in \mathcal{D}$, where $T_x \mathcal{D}$ denotes the tangent cone to \mathcal{D} at x (see Definition 5) [54], then there exists a Filippov solution $x : [0, \infty) \rightarrow \mathbb{R}^n$ to Equation (30) for every $x_0 \in \mathcal{D} \subseteq \mathbb{R}^n$ ([50], p. 180).

Switched dynamical systems are essential in modeling the plasticity of the central nervous system. In particular, recent neuroimaging findings have shown that the loss of top-down (feedback) processing in association with anesthetic-induced unconsciousness observed in electroencephalographic (EEG) and functional magnetic resonance imaging (fMRI) is associated with functional disconnections between anterior and posterior brain structures [55–57]. These studies show that topological rather than network connection strength of functional networks correlate with states of consciousness. Hence, changes in network topology are essential in capturing the mechanism of action for consciousness rather than network connectivity strength during general anesthesia. Dynamic neural network topologies capturing link separation or creation, which can be modeled by switched systems and differential inclusions, can provide a mechanism for the objective selective inhibition of feedback connectivity in association with anesthetic-induced unconsciousness.

Since the Filippov set-valued map $\mathcal{K}[f](x)$ is upper semicontinuous with nonempty, convex, and compact values, and $\mathcal{K}[f](x)$ is also locally bounded ([47](p. 85)), it follows that Filippov solutions to Equation (26) exist ([47](Theorem 1, p. 77)). Recall that the Filippov solution $t \mapsto x(t)$ to Equation (26) is a *right maximal solution* if it cannot be extended (either uniquely or nonuniquely) forward in time. We

assume that all right maximal Filippov solutions to Equation (26) exist on $[0, \infty)$, and hence, we assume that Equation (26) is *forward complete*. Recall that Equation (26) is forward complete if and only if the Filippov solutions to Equation (26) are uniformly globally sliding time stable ([58](Lemma 1, p. 182)).

We say that a set \mathcal{M} is *weakly positively invariant* (resp., *strongly positively invariant*) with respect to Equation (26) if, for every $x_0 \in \mathcal{M}$, \mathcal{M} contains a right maximal solution (resp., all right maximal solutions) of Equation (26) [48,59]. The set $\mathcal{M} \subseteq \mathbb{R}^n$ is *weakly negatively invariant* if, for every $x \in \mathcal{M}$ and $t \geq 0$, there exist $z \in \mathcal{M}$ and a Filippov solution $\psi(\cdot)$ to Equation (26) with $\psi(0) = z$ such that $\psi(t) = x$ and $\psi(\tau) \in \mathcal{M}$ for all $\tau \in [0, t]$. The set $\mathcal{M} \subseteq \mathbb{R}^n$ is *weakly invariant* if \mathcal{M} is weakly positively invariant as well as weakly negatively invariant. Finally, an equilibrium point of Equation (26) is a point $x_e \in \mathbb{R}^n$ such that $0 \in \mathcal{K}[f](x_e)$. It is easy to see that x_e is an equilibrium point of Equation (26) if and only if the constant function $x(\cdot) = x_e$ is a Filippov solution of Equation (26). We denote the set of equilibrium points of Equation (26) by \mathcal{E} . Since the set-valued map $\mathcal{K}[f]$ is upper semicontinuous, it follows that \mathcal{E} is closed.

To develop Lyapunov theory for nonsmooth dynamical systems of the form given by Equation (26), we need the notion of generalized derivatives and gradients. In this paper, we focus on Clarke generalized derivatives and gradients [52,60].

Definition 2 ([48,60]) Let $V : \mathbb{R}^n \rightarrow \mathbb{R}$ be a locally Lipschitz continuous function. The Clarke upper generalized derivative of $V(\cdot)$ at x in the direction of $v \in \mathbb{R}^n$ is defined by

$$V^\circ(x, v) \triangleq \limsup_{y \rightarrow x, h \rightarrow 0^+} \frac{V(y + hv) - V(y)}{h}. \tag{32}$$

The Clarke generalized gradient $\partial V : \mathbb{R}^n \rightarrow 2^{\mathbb{R}^{1 \times n}}$ of $V(\cdot)$ at x is the set

$$\partial V(x) \triangleq \text{co} \left\{ \lim_{i \rightarrow \infty} \nabla V(x_i) : x_i \rightarrow x, x_i \notin \mathcal{N} \cup \mathcal{S} \right\}, \tag{33}$$

where co denotes the convex hull, ∇ denotes the nabla operator, \mathcal{N} is the set of measure zero of points where ∇V does not exist, \mathcal{S} is any subset of \mathbb{R}^n of measure zero, and the increasing unbounded sequence $\{x_i\}_{i \in \mathbb{Z}_+} \subset \mathbb{R}^n$ converges to $x \in \mathbb{R}^n$.

Note that Equation (32) always exists. Furthermore, note that it follows from Definition 2 that the generalized gradient of V at x consists of all convex combinations of all the possible limits of the gradient at neighboring points where V is differentiable. In addition, note that since $V(\cdot)$ is Lipschitz continuous, it follows from Rademacher’s theorem ([61](Theorem 6, p. 281)) that the gradient $\nabla V(\cdot)$ of $V(\cdot)$ exists almost everywhere, and hence, $\nabla V(\cdot)$ is bounded. Specifically, for every $x \in \mathbb{R}^n$, every $\varepsilon > 0$, and every Lipschitz constant L for V on $\overline{B}_\varepsilon(x)$, $\partial V(x) \subseteq \overline{B}_L(0)$. Thus, since for every $x \in \mathbb{R}^n$, $\partial V(x)$ is convex, closed, and bounded, it follows that $\partial V(x)$ is compact.

In order to state the results of this paper, we need some additional notation and definitions. Given a locally Lipschitz continuous function $V : \mathbb{R}^n \rightarrow \mathbb{R}$, the *set-valued Lie derivative* $\mathcal{L}_f V : \mathbb{R}^n \rightarrow 2^{\mathbb{R}}$ of V with respect to f at x [48,62] is defined as

$$\begin{aligned} \mathcal{L}_f V(x) &\triangleq \left\{ a \in \mathbb{R} : \text{there exists } v \in \mathcal{K}[f](x) \text{ such that } p^T v = a \text{ for all } p^T \in \partial V(x) \right\} \\ &\subseteq \bigcap_{p^T \in \partial V(x)} p^T \mathcal{K}[f](x). \end{aligned} \tag{34}$$

If $\mathcal{K}[f](x)$ is convex with compact values, then $\mathcal{L}_f V(x)$, $x \in \mathbb{R}^n$, is a closed and bounded, possibly empty, interval in \mathbb{R} . If $V(\cdot)$ is continuously differentiable at x , then $\mathcal{L}_f V(x) = \{\nabla V(x) \cdot v : v \in \mathcal{K}[f](x)\}$. In the case where $\mathcal{L}_f V(x)$ is nonempty, we use the notation $\max \mathcal{L}_f V(x)$ (resp., $\min \mathcal{L}_f V(x)$) to denote the largest (resp., smallest) element of $\mathcal{L}_f V(x)$. Furthermore, we adopt the convention $\max \emptyset = -\infty$. Finally, recall that a function $V : \mathbb{R}^n \rightarrow \mathbb{R}$ is *regular* at $x \in \mathbb{R}^n$ ([60](Definition 2.3.4)) if, for all $v \in \mathbb{R}^n$, the right directional derivative $V'_+(x, v) \triangleq \lim_{h \rightarrow 0^+} \frac{1}{h}[V(x + hv) - V(x)]$ exists and $V'_+(x, v) = V^o(x, v)$. V is called *regular* on \mathbb{R}^n if it is regular at every $x \in \mathbb{R}^n$.

The next definition introduces the notion of semistability for discontinuous dynamical systems. This stability notion is necessary for systems having a continuum of equilibria. Specifically, since every neighborhood of a nonisolated equilibrium contains another equilibrium, a nonisolated equilibrium cannot be asymptotically stable. Hence, asymptotic stability is not the appropriate notion of stability for systems having a continuum of equilibria. Two notions that are of particular relevance to such systems are convergence and semistability. Convergence is the property whereby every system solution converges to a limit point that may depend on the system initial condition (*i.e.*, initial anesthetic concentrations). Semistability is the additional requirement that all solutions converge to limit points that are Lyapunov stable. Semistability for an equilibrium thus implies Lyapunov stability, and is implied by asymptotic stability. Thus, semistability guarantees that small perturbations from the limiting state of unconsciousness will lead to only small transient excursions from that state of unconsciousness. It is important to note that semistability is not merely equivalent to asymptotic stability of the set of equilibria. Indeed, it is possible for a trajectory to converge to the set of equilibria without converging to any one equilibrium point as examples in [19] show. For further details, see [20].

Definition 3 Let $\mathcal{D} \subseteq \mathbb{R}^n$ be an open strongly positively invariant set with respect to Equation (26). An equilibrium point $z \in \mathcal{D}$ of Equation (26) is Lyapunov stable if, for every $\varepsilon > 0$, there exists $\delta = \delta(\varepsilon) > 0$ such that, for every initial condition $x_0 \in \mathcal{B}_\delta(z)$ and every Filippov solution $x(t)$ with the initial condition $x(0) = x_0$, $x(t) \in \mathcal{B}_\varepsilon(z)$ for all $t \geq 0$. An equilibrium point $z \in \mathcal{D}$ of Equation (26) is semistable if z is Lyapunov stable and there exists an open subset \mathcal{D}_0 of \mathcal{D} containing z such that, for all initial conditions in \mathcal{D}_0 , the Filippov solutions of Equation (26) converge to a Lyapunov stable equilibrium point. The system given by Equation (26) is semistable with respect to \mathcal{D} if every Filippov solution with initial condition in \mathcal{D} converges to a Lyapunov stable equilibrium. Finally, the system given by Equation (26) is said to be globally semistable if the system given by Equation (26) is semistable with respect to \mathbb{R}^n .

Finally, we introduce the definition of multistability of the dynamical system given by Equation (26).

Definition 4 Consider the nonlinear dynamical system given by Equation (26). We say that the dynamical system given by Equation (26) is multistable if (i) there exists more than one equilibrium point of Equation (26) in \mathbb{R}^n ; (ii) all Filippov solutions to Equation (26) converge to one of these equilibrium points; and (iii) almost all Filippov solutions to Equations (26) converge to Lyapunov stable equilibria; that is, the set of initial conditions driving the Filippov solutions of Equation (26) to unstable equilibria has Lebesgue measure zero.

It is important to note that our definition of multistability is different from the definition given in [18]. Specifically, pertaining to condition (iii), the definition of multistability given in [18] requires

that almost all Filippov solutions to Equation (26) converge to asymptotically stable equilibria. This key difference between the two definitions allows for the dynamical system given by Equation (26) to possess a continuum of equilibria, rather than merely isolated equilibria. As we see later, if f_i is of the form given by Equation (10), then Equation (9) has a continuum of equilibria under certain conditions, and hence, Equation (26) is semistable in the sense of (iii) [63]. Hence, in this case, it is appropriate to use Definition 4 to characterize multistability.

Almost all of the existing results on multistability theory rely on linearization techniques based on the Hartman-Grobman theorem [64,65] involving the fact that the linearized system has the same topological property as the original system around a hyperbolic fixed point. When the system fixed point is not hyperbolic, however, these techniques fail to predict multistability. In this case, checking multistability becomes a daunting task. Rather than checking the transversality condition for hyperbolicity, we propose a new approach for guaranteeing multistability using equilibria-independent, semidefinite Lyapunov function methods. In particular, using the geometric structure of the vector field f for a given dynamical system, we develop nontangency-based Lyapunov tests for verifying conditions (ii) and (iii) in Definition 4 involving convergence and Lyapunov stability almost everywhere.

5. Direction Cones, Nontangency, Restricted Prolongations, and Nonsmooth Multistability Theory

To develop multistability theory for discontinuous dynamical systems of the form given by Equation (26) we use the notions of direction cones, nontangency, and restricted prolongations. In particular, to show condition (ii) in Definition 4 holds for dynamical systems of the form given by Equation (26), we adopt the notion of nontangency [19,63] to develop nontangency-based Lyapunov tests for convergence. Specifically, the authors in [19] develop a general framework for nontangency-based Lyapunov tests for the convergence of dynamical systems described by ordinary differential equations with continuous vector fields. In [63], the authors extend some of the results of [19] to nonsmooth dynamical systems, that is, systems described by ordinary differential equations with the discontinuous right-hand sides. Since the vector field f characterizing biological neural networks can involve either continuous (e.g., half-wave rectification functions) or discontinuous (e.g., hard-limiter activation functions) vector fields, and, more importantly, the fact that anesthetics reconfigure the topological structure of functional brain networks [55–57] by suppressing midbrain/pontine areas involved with regulatory arousal leading to a dynamic network topology, we use the more general definition for nontangency presented in [63].

Intuitively, a vector field is nontangent to a set at a point if the vector field at the point is not contained in the tangent space to the set at that point. We use this intuitive idea for the case where the vector field describing the system dynamics is discontinuous and the set is the set of equilibria of the system. However, this notion presents two key difficulties when the set under consideration is the set of singular points of the vector field, that is, the set of equilibria of the system. In particular, the vector field at an equilibrium point is zero, and hence, it is always contained in the tangent space to the set of equilibria. In this case, in order to capture the notion of nontangency, we introduce the *direction cone* of a vector field.

Alternatively, the set of equilibria may not be sufficiently regular to possess a tangent space at the equilibrium point under consideration and may have corners or self-intersections. For example, in firing

rate population models appearing in neuroscience, the firing rate is a nonnegative quantity representing the probability of the firing action potential by the neuron and can be interpreted as a measure of the neuron's activity. Since the firing rate of the excitatory-inhibitory network is nonnegative, all solutions of physical interest always take values in the nonnegative orthant $\overline{\mathbb{R}}_+^n$ of the state space for nonnegative initial conditions. For such systems, which evolve on possibly closed positively invariant subsets of \mathbb{R}^n , it is natural to consider the nonnegative orthant as their state space. Hence, the dynamical system evolves on the nonnegative orthant and can have the boundary of the orthant as its set of equilibria. In this case, the set of equilibria has a corner at the origin. We overcome this difficulty by considering the *tangent cone* [66,67], which extends the notion of a tangent space to a nonsmooth setting.

To introduce the notions of direction cone and tangent cone, some notation and definitions are required. A set $\mathcal{E} \subseteq \mathbb{R}^n$ is *connected* if and only if every pair of open sets $\mathcal{U}_i \subseteq \mathbb{R}^n$, $i = 1, 2$, satisfying $\mathcal{E} \subseteq \mathcal{U}_1 \cup \mathcal{U}_2$ and $\mathcal{U}_i \cap \mathcal{E} \neq \emptyset$, $i = 1, 2$, has a nonempty intersection. A *connected component* of the set $\mathcal{E} \subseteq \mathbb{R}^n$ is a connected subset of \mathcal{E} that is not properly contained in any connected subset of \mathcal{E} . Given a set $\mathcal{E} \subseteq \mathbb{R}^n$, let $\text{coco } \mathcal{E}$ denote the convex cone generated by \mathcal{E} .

Definition 5 Given $x \in \mathbb{R}^n$, the direction cone \mathcal{F}_x of the vector field f at x is the intersection of closed convex cones of the form $\overline{\bigcap_{\mu(\mathcal{S})=0} \text{coco}\{f(\mathcal{U} \setminus \mathcal{S})\}}$, where $\mathcal{U} \subseteq \mathbb{R}^n$ is an open neighborhood of x and $\overline{\mathcal{Q}}$ denotes the closure of the set \mathcal{Q} . Let $\mathcal{E} \subseteq \mathbb{R}^n$. A vector $v \in \mathbb{R}^n$ is tangent to \mathcal{E} at $z \in \mathcal{E}$ if there exist a sequence $\{z_i\}_{i=1}^\infty$ in \mathcal{E} converging to z and a sequence $\{h_i\}_{i=1}^\infty$ of positive real numbers converging to zero such that $\lim_{i \rightarrow \infty} \frac{1}{h_i}(z_i - z) = v$. The tangent cone to \mathcal{E} at z is the closed cone $T_z \mathcal{E}$ of all vectors tangent to \mathcal{E} at z . Finally, the vector field f is nontangent to the set \mathcal{E} at the point $z \in \mathcal{E}$ if $T_z \mathcal{E} \cap \mathcal{F}_z \subseteq \{0\}$.

The notion of nontangency introduced in Definition 5 is different from the well-known notion of transversality [68]. Transversality between a vector field and a set is possible only at a point in the set where the vector field is not zero and the set is locally a differentiable submanifold of codimension one. Alternatively, nontangency is possible even if the vector field is zero and the set is not a differentiable submanifold of codimension one.

Definition 5 formalizes the notion of nontangency by defining nontangency of a discontinuous vector field to a set at a point to be the condition that the tangent cone to the set at the point and the direction cone of the vector field at that point have no nonzero vector in common. Using the notion of nontangency, in [63] we developed necessary and sufficient conditions for convergence of discontinuous dynamical systems. Specifically, convergence of system trajectories are guaranteed if and only if the vector field is nontangent to the positive limit set of the point at some positive limit point. However, this result cannot be applied directly in practice since it is not generally possible to find the positive limit set of system solution trajectories. Since nontangency to any outer estimate of the positive limit set implies nontangency to the positive limit set itself, we use nontangency-based Lyapunov tests for convergence. In particular, if the vector field f is nontangent to the largest invariant subset of the zero-level set of the derivative of a Lyapunov function that is nonincreasing along the solutions of Equation (26), then every bounded solution converges to a limit.

Since the application of the convergence results discussed above depends on verifying the boundedness of trajectories, the well-known results for boundedness involving proper (that is, radially

unbounded in the case where the state space is \mathbb{R}^n) Lyapunov functions [69,70] by introducing the notion of a weakly proper function need to be extended. Specifically, in [63] we consider Lyapunov functions whose connected components of their sublevel sets are compact. In this case, the existence of a weakly proper Lyapunov function that is nonincreasing along the system trajectories implies that the trajectories are bounded.

Using the notion of nontangency we then developed Lyapunov measures for almost everywhere semistability to arrive at multistability theory for discontinuous dynamical systems [63]. Here, prolongations [33,71] play a role analogous to that played by positive limit sets in the aforementioned discussion. In particular, the notion of a restricted prolongation of a point is used to show that an equilibrium point of Equation (26) is Lyapunov stable if and only if the discontinuous vector field is nontangent at the equilibrium to its restricted prolongation. The restricted prolongation of a point is a subset of its positive prolongation [33,71] and is defined as follows.

Definition 6 Given a point $x \in \mathbb{R}^n$ and a bounded open neighborhood $\mathcal{U} \subset \mathbb{R}^n$ of x , the restricted prolongation of x with respect to \mathcal{U} is the set $\mathcal{R}_x^{\mathcal{U}} \subseteq \overline{\mathcal{U}}$ of all subsequential limits of sequences of the form $\{\psi_i(t_i)\}_{i=1}^{\infty}$, where $\{t_i\}_{i=1}^{\infty}$ is a sequence in $[0, \infty)$, $\psi_i(\cdot)$ is a solution to Equation (26) with $\psi_i(0) = x_i$, $i = 1, 2, \dots$, and $\{x_i\}_{i=1}^{\infty}$ is a sequence in \mathcal{U} converging to x such that the set $\{z \in \mathbb{R}^n : z = \psi_i(t), t \in [0, t_i]\}$ is contained in $\overline{\mathcal{U}}$ for every $i = 1, 2, \dots$

The utility of prolongations in stability analysis follows from the fact that an equilibrium point is Lyapunov stable if and only if the positive prolongation of the equilibrium consists only of the equilibrium point. See, ([71], Proposition 7.3) and ([33](Theorem V.1.12)). For systems with continuous vector fields this was first shown in [19]. Since the restricted prolongation of a point is a subset of the positive prolongation of the point, such a result provides a sharper version of the results ([71](Proposition 7.3)) and ([33](Theorem V.1.12)).

As in the case for positive limit sets discussed above, since it is not generally possible to find the restricted prolongation of an equilibrium point in practice and since nontangency to any outer estimate of the restricted prolongation implies nontangency to the restricted prolongation itself, we use outer estimates of restricted prolongations in terms of connected components of invariant and negatively invariant subsets of level and sublevel sets of Lyapunov functions and their derivatives. By assuming nontangency of the vector field to invariant or negatively invariant subsets of the level set of the Lyapunov function containing the equilibrium we can trap the restricted prolongation and the positive limit set, respectively, in the level sets of the Lyapunov function and its derivative.

The following theorem establishes sufficient conditions for convergence and Lyapunov stability almost everywhere for the system given by Equation (26). This result follows from the fact that $\mathcal{R}_x^{\mathcal{U}}$ is connected and the fact that if \mathcal{N} is composed of isolated equilibria of Equation (26) or f is nontangent to \mathcal{N} at every point in \mathcal{N} , then the solution to Equation (26) is convergent. For details of these facts; see [45]. For the statement of the next result, \dot{V} denotes the set-valued Lie derivative for the Filippov solutions to Equation (26) and $V^{-1}(0) \triangleq \{x \in \mathbb{R}^n : V(x) = 0\}$.

Theorem 1 ([45]) Assume there exists a locally Lipschitz continuous and regular function $V : \mathbb{R}^n \rightarrow \mathbb{R}$ such that \dot{V} is defined almost everywhere on \mathbb{R}^n and satisfies $\dot{V}(x) \leq 0$ for almost all $x \in \mathbb{R}^n$. Let $x \in \mathbb{R}^n$ be such that the solution of Equation (26) is bounded and let \mathcal{N} denote the largest weakly

invariant set contained in $\dot{V}^{-1}(0)$. If either every point in \mathcal{N} is Lyapunov stable or f is nontangent to \mathcal{N} at every point in \mathcal{N} , then almost all solutions of Equation (26) converge to Lyapunov stable equilibria.

Example 1 Consider the two-class mean excitatory and mean inhibitory synaptic drive network characterized by a discontinuous vector field for modeling plasticity in the network given by

$$\dot{S}^E(t) = -0.5S^E(t) + f_1(S^E(t), S^I(t)), \quad S^E(0) = S_0^E, \quad t \geq 0, \tag{35}$$

$$\dot{S}^I(t) = -S^I(t) + f_2(S^E(t), S^I(t)), \quad S^I(0) = S_0^I, \tag{36}$$

where $S_0^E \geq 0, S_0^I \geq 0$, and $f_i(\cdot), i = 1, 2$, are given by

$$f_1(S^E, S^I) = \left[\text{step}(S^I - S^E)(S^E - 0.5S^I) \right]_+, \tag{37}$$

$$f_2(S^E, S^I) = \left[\text{step}(S^I - S^E)(2S^E - S^I) \right]_+, \tag{38}$$

where $\text{step}(y) = 1$ for $y \geq 0$ and $\text{step}(y) = 0$, otherwise. For this system, the set of equilibria in $\overline{\mathbb{R}}_+^2$ are given by $\mathcal{E} \triangleq \{(S^E, S^I) \in \overline{\mathbb{R}}_+^2 : S^E = S^I\}$.

Next, we show that for almost all the initial conditions $(S_0^E, S_0^I) \in \overline{\mathbb{R}}_+^2$, the equilibrium set \mathcal{E} is attractive. To see this, consider the function $V : \mathbb{R}^2 \rightarrow \mathbb{R}$ given by $V(S^E, S^I) = \frac{1}{2}[(S^E)^2 + (S^I)^2]$. Now, it follows that the set-valued Lie derivative $\dot{V}(S^E, S^I)$ satisfies

$$\dot{V}(S^E, S^I) = \begin{cases} \left\{ -\frac{1}{2}(S^E)^2 - (S^I)^2 \right\}, & S^E > S^I \geq 0 \text{ or } S^I > 2S^E \geq 0, \\ \left\{ \frac{1}{2}(S^E - S^I)(S^E + 4S^I) \right\}, & 0 \leq S^E < S^I < 2S^E, \\ \overline{\text{co}} \left\{ -\frac{1}{2}(S^E)^2 - (S^I)^2, \frac{1}{2}(S^E - S^I)(S^E + 4S^I) \right\}, & \text{otherwise, } (S^E, S^I) \in \overline{\mathbb{R}}_+^2. \end{cases}$$

It can be verified that $\dot{V}(S^E, S^I) \leq 0$ for almost all $(S^E, S^I) \in \overline{\mathbb{R}}_+^2$ and $\dot{V}^{-1}(0) = \mathcal{E}$.

Next, we show that the vector field f of the system given by Equations (35) and (36) is nontangent to \mathcal{E} . Let $(S^E, S^I) \in \mathcal{E}$ and note that it follows from the expression of f that the direction cone $\mathcal{F}_{(S^E, S^I)}$ of the vector field f at $(S^E, S^I) \in \mathcal{E}$ is given by

$$\mathcal{F}_{(S^E, S^I)} = \{k[1, 2]^T : k \in \mathbb{R}\}. \tag{39}$$

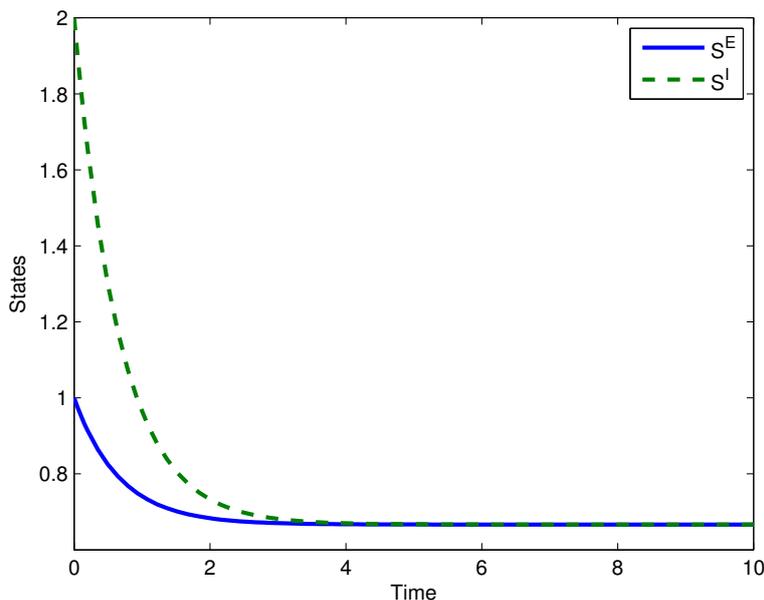
In addition, the tangent cone to \mathcal{E} at $(S^E, S^I) \in \mathcal{E}$ is given by

$$T_{(S^E, S^I)}\mathcal{E} = \{k[1, 1]^T : k \in \mathbb{R}\}. \tag{40}$$

Now, it follows from Equations (39) and (40) that $T_{(S^E, S^I)}\mathcal{E} \cap \mathcal{F}_{(S^E, S^I)} = \{0\}$, and hence, for every $(S^E, S^I) \in \mathcal{E}$, f is nontangent to \mathcal{E} at (S^E, S^I) .

Finally, it follows from Theorem 5.1 that almost all solutions of the system converge to Lyapunov stable equilibria in \mathcal{E} , and hence, by definition, the system given by Equations (35) and (36) is multistable. Figure 2 shows the state trajectories versus time for the initial condition $[S_0^E, S_0^I] = [1, 2]$. This system is additionally synchronized; a notion that we discuss in Section 7.

Figure 2. State trajectories versus time for Example 1 with initial condition $[S_0^E, S_0^I] = [1, 2]$.



6. Multistability of Excitatory-Inhibitory Biological Networks

Having developed multistability theory for discontinuous dynamical systems, in this section we apply the results of Section 5 to the excitatory-inhibitory neural firing rate model given by Equation (9). The form of biological neural network models given by Equation (9) represents a wide range of firing rate population models appearing in neuroscience [3,4]. The firing rate is a nonnegative quantity representing the probability of the firing action potential by the neuron and can be interpreted as a measure of the neuron’s activity. Since the firing rate of the excitatory-inhibitory network is nonnegative, all solutions of physical interest always take values in the nonnegative orthant of the state space for nonnegative initial conditions. For such systems, which evolve on possibly closed positively invariant subsets of \mathbb{R}^n , it is natural to consider the nonnegative orthant $\overline{\mathbb{R}}_+^n$ as their state space, and hence, these systems are nonnegative dynamical systems [35]. In this case, the stability and convergence result developed in Section 5 holds with respect to $\overline{\mathbb{R}}_+^n$ by replacing \mathbb{R}^n with $\overline{\mathbb{R}}_+^n$. For related details, see [35].

The following result, which follows from Proposition 2.1 of [35], gives necessary and sufficient conditions for the firing rates $S_i(t), i = 1, \dots, n$, to remain in the nonnegative orthant of the state space. For the statement of the next result recall that f is *nonnegative* [35] if and only if $f(x) \geq \geq 0, x \in \overline{\mathbb{R}}_+^n$, where “ $\geq \geq$ ” denotes a component-wise inequality.

Proposition 1 Consider the excitatory-inhibitory network given by Equation (9). The firing rate vector $S(t) \triangleq [S_1(t), \dots, S_n(t)]^T \in \mathbb{R}^n$ remains in the nonnegative orthant of the state space $\overline{\mathbb{R}}_+^n$ for all $t \geq 0$ if and only if, for every $S_i \geq 0$ and $v_{thi} \geq 0, i = 1, \dots, n$, the function $\tilde{f} = [f_1(\sum_{j=1}^n A_{1j}S_j + v_{th1}), \dots, f_n(\sum_{j=1}^n A_{nj}S_j + v_{thn})]^T : \mathbb{R}^n \rightarrow \mathbb{R}^n$ is nonnegative.

Consider the excitatory-inhibitory network given by Equation (9) where $f_i(\cdot)$ is given by Equation (10), and note that $\tilde{f} : \mathbb{R}^n \rightarrow \mathbb{R}^n$ is nonnegative. Thus, it follows from Proposition 1 that

if $S_0 \in \overline{\mathbb{R}}_+^n$, then $S(t) \in \overline{\mathbb{R}}_+^n$ for all $t \geq 0$. Next, assume $v_{thi}(t) \equiv 0$, so that the vector-matrix form of Equation (9) can be written as

$$\dot{S}(t) = -LS(t) + \tilde{f}(AS(t)), \quad S(0) = S_0, \quad t \geq 0, \tag{41}$$

where $L \triangleq \text{diag}[\frac{1}{\tau_1}, \dots, \frac{1}{\tau_n}] \in \mathbb{R}^{n \times n}$ is a time constant matrix, $A \triangleq [A_{ij}] \in \mathbb{R}^{n \times n}$ is a matrix representing the strength of the synaptic interconnections, and $\tilde{f}(S) \triangleq [f_1(\sum_{j=1}^m A_{1j}S_j), \dots, f_n(\sum_{j=1}^m A_{nj}S_j)]^T : \mathbb{R}^n \rightarrow \mathbb{R}^n$ is a vector activation function describing the relationship between the synaptic drives and the firing rates of the neurons, where $S = [S_1, \dots, S_n]^T$ and $f_i(\cdot)$ is defined in Equation (10). Finally, assume that the set $\mathcal{E}_e = \{S \in \overline{\mathbb{R}}_+^n : \tilde{f}(AS) - LS = 0\} \subset \overline{\mathbb{R}}_+^n$ has a nonzero element, that is, \mathcal{E}_e has a nonzero solution for $S \in \overline{\mathbb{R}}_+^n$. For the statement of the next theorem, $\mathcal{N}(X)$ denotes the nullspace of the matrix X .

Theorem 2 Consider the excitatory-inhibitory network given by Equation (41) with $f_i(\cdot)$, $i = 1, \dots, n$, given by Equation (10) and $S_0 \in \overline{\mathbb{R}}_+^n$. Let $\Omega_1 \geq 0$ and $\Omega_2 \geq 0$, where

$$\Omega_1 \triangleq \frac{1}{2}[(HL - \tilde{H}\tilde{A}) + (HL - \tilde{H}\tilde{A})^T], \tag{42}$$

$$\Omega_2 \triangleq (I - H)L + (\tilde{H} - I)\tilde{A}, \tag{43}$$

$H = H^T$, $\text{rank } H = n$, and \tilde{H} and \tilde{A} are $n \times n$ matrices whose entries are given by $\tilde{H}_{ij} = [H_{ij}]_+$ and $\tilde{A}_{ij} = [A_{ij}]_+$, and satisfy $(HL - \tilde{H}\tilde{A}) \geq 0$. Furthermore, assume that every point $S \in \mathcal{N}(\Omega_1) \cap \overline{\mathbb{R}}_+^n$ is Lyapunov stable with respect to Equation (41). Then Equation (41) is multistable.

Proof. Consider the function $V(S) = \frac{1}{2}S^T H S$, $S \in \overline{\mathbb{R}}_+^n$, and note that the derivative of $V(S)$ along the trajectories of Equation (41) with $f_i(\cdot)$, $i = 1 \dots, n$, given by Equation (10) is given by

$$\begin{aligned} \dot{V}(S) &= S^T H (-LS + \tilde{f}(AS)) \\ &= -\sum_{i=1}^n \sum_{j=1}^n \frac{1}{\tau_i} H_{ij} S_i S_j + \sum_{i=1}^n \sum_{j=1}^n H_{ij} S_j \left[\sum_{k=1, k \neq i}^n A_{ik} S_k \right]_+, \end{aligned} \tag{44}$$

where $[\cdot]_+$ is defined as in Equation (10). Now, since $S \in \overline{\mathbb{R}}_+^n$, $[x]_+ \geq 0$ and $[x + y]_+ \leq [x]_+ + [y]_+$ for all $x, y \in \mathbb{R}$, it follows that

$$\begin{aligned} \dot{V}(S) &\leq -\sum_{i=1}^n \sum_{j=1}^n \frac{1}{\tau_i} H_{ij} S_i S_j + \sum_{i=1}^n \sum_{j=1}^n [H_{ij}]_+ S_j \left[\sum_{k=1, k \neq i}^n A_{ik} S_k \right]_+ \\ &\leq -\sum_{i=1}^n \sum_{j=1}^n \frac{1}{\tau_i} H_{ij} S_i S_j + \sum_{i=1}^n \sum_{j=1}^n \sum_{k=1, k \neq i}^n [H_{ij}]_+ [A_{ik}]_+ S_j S_k \\ &= -S^T H L S + S^T \tilde{H} \tilde{A} S \\ &= -S^T \Omega_1 S, \quad S \in \overline{\mathbb{R}}_+^n. \end{aligned} \tag{45}$$

Next, since H is such that $(HL - \tilde{H}\tilde{A}) \geq 0$, it follows from Equation (45) that $\dot{V}(S) \leq 0$ for all $S \in \overline{\mathbb{R}}_+^n$, and hence, $\dot{V}^{-1}(0) \subseteq \mathcal{N}(\Omega_1)$.

Since $\Omega_1 \geq 0$, the dynamical system

$$\dot{x}(t) = -(HL - \tilde{H}\tilde{A})x(t), \quad x(0) = S_0, \quad t \geq 0, \tag{46}$$

where $S_0 \in \overline{\mathbb{R}}_+^n$, is Lyapunov stable. Next, for every $S \in \overline{\mathbb{R}}_+^n$, it follows that $\dot{S} = -LS + \tilde{f}(AS) \leq -LS + \tilde{A}S$. Now, since H is chosen such that all the entries of Ω_2 are nonnegative, it follows that

$$\dot{S} \leq -LS + \tilde{A}S \leq -LS + \tilde{A}S + \Omega_2 S = -(HL - \tilde{H}\tilde{A})S, \quad S \in \overline{\mathbb{R}}_+^n. \tag{47}$$

In this case, $S(t) \leq x(t)$, $t \geq 0$, where $x(t)$, $t \geq 0$, is the solution to Equation (46). In addition, since $S(t) \in \overline{\mathbb{R}}_+^n$, $t \geq 0$, $S_0 \in \overline{\mathbb{R}}_+^n$, and Equation (46) is Lyapunov stable, it follows that $S(t) \in \overline{\mathbb{R}}_+^n$ is bounded for all $t \geq 0$.

Finally, let \mathcal{W} be the largest weakly invariant set contained in $\dot{V}^{-1}(0)$ and note that since $S(t) \in \mathbb{R}_+^n$, $t \geq 0$, $\mathcal{W} \subseteq \mathcal{N}(\Omega_1) \cap \overline{\mathbb{R}}_+^n$. Now, since $S \in \mathcal{N}(\Omega_1) \cap \overline{\mathbb{R}}_+^n$ is Lyapunov stable with respect to Equation (41), it follows from Corollary 4.2 of [45] that all the solutions of the excitatory-inhibitory network given by Equation (41) converge to one of the Lyapunov stable equilibria in $\mathcal{N}(\Omega_1) \cap \overline{\mathbb{R}}_+^n$ for Equation (41) with $S_0 \in \overline{\mathbb{R}}_+^n$. Hence, it follows from Theorem 1 that Equation (41) is multistable. \square

Example 2 Consider the excitatory-inhibitory network characterized by the dynamics

$$\dot{S}_1(t) = -S_1(t) + f_1(2S_2(t)), \quad S_1(0) = S_{01}, \quad t \geq 0, \tag{48}$$

$$\dot{S}_2(t) = -4S_2(t) + f_2(2S_1(t)), \quad S_2(0) = S_{02}, \tag{49}$$

where $S_{01} \geq 0$, $S_{02} \geq 0$, and, $f_i(\cdot)$, $i = 1, 2$, are defined by Equation (10), and note that Equations (48) and (49) can be written in the form of Equation (41) with

$$L = \begin{bmatrix} 1 & 0 \\ 0 & 4 \end{bmatrix}, \quad A = \begin{bmatrix} 0 & 2 \\ 2 & 0 \end{bmatrix}.$$

Let $H = I_2$ and define \tilde{H} and \tilde{A} by $\tilde{H}_{ij} = [H_{ij}]_+$ and $\tilde{A}_{ij} = [A_{ij}]_+$ so that

$$\Omega_1 = \frac{1}{2}((HL - \tilde{H}\tilde{A}) + (HL - \tilde{H}\tilde{A})^T) = \begin{bmatrix} 1 & -2 \\ -2 & 4 \end{bmatrix} \geq 0$$

and

$$\Omega_2 = (I - H)L + (\tilde{H} - I)\tilde{A} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix},$$

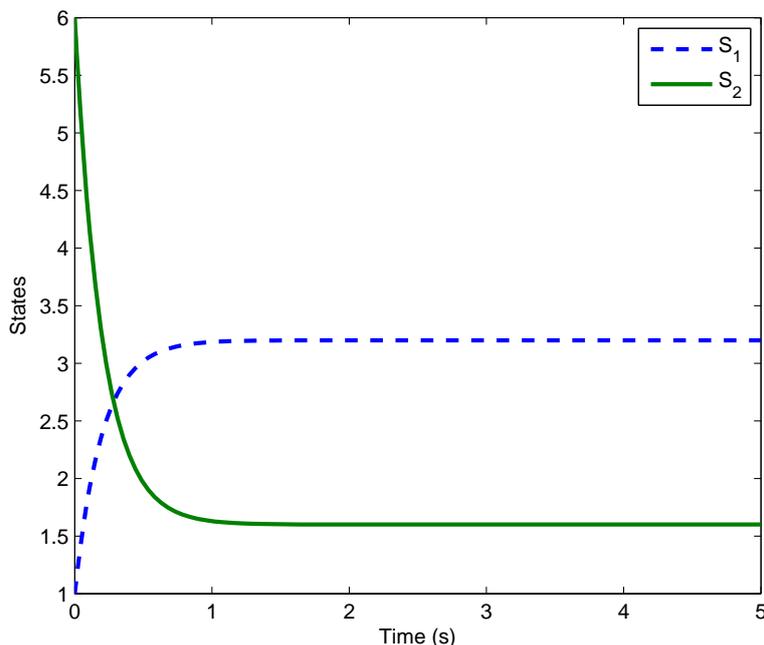
which contains nonnegative entries.

Next, it follows from Proposition 1 that $S(t) \geq 0$, $t \geq 0$, for all $(S_{10}, S_{20}) \in \overline{\mathbb{R}}_+^2$, and hence, Equations (48) and (49) collapse to the linear model given by

$$\dot{S}_1(t) = -S_1(t) + 2S_2(t), \quad S_1(0) = S_{01}, \quad t \geq 0, \tag{50}$$

$$\dot{S}_2(t) = -4S_2(t) + 2S_1(t), \quad S_2(0) = S_{02}. \tag{51}$$

Clearly, the system given by Equations (50) and (51) is Lyapunov stable, and hence, every $S \in \mathcal{N}(\Omega_1) \cap \overline{\mathbb{R}}_+^n$ is Lyapunov stable with respect to Equations (48) and (49). Now, it follows from Theorem 2 that the system given by Equations (48) and (49) is multistable. For the initial conditions $S_{01} = 3$ and $S_{02} = 1$ the trajectories of the state variables with respect to time is shown in Figure 3.

Figure 3. State trajectories versus time for Example 2.

7. Synchronization of Biological Neural Networks

Numerous complex large-scale dynamical networks often demonstrate a degree of synchronization. System synchronization typically involves coordination of events that allows a dynamical system to operate in unison resulting in system self-organization. The onset of synchronization in populations of coupled dynamical networks have been studied for various complex networks including network models for mathematical biology, statistical physics, kinetic theory, bifurcation theory, as well as plasma physics [72]. Synchronization of firing neural oscillator populations using probabilistic analysis has also been addressed in the neuroscience literature [73]. One of the most important questions in neuroscience is how do neurons, or collections of neurons, communicate. In other words, what is the neural code? There is extensive experimental verification that collections of neurons may function as oscillators [74–76] and the synchronization of oscillators may play a key role in the transmission of information within the central nervous system. This may be particularly relevant to understanding the mechanism of action for general anesthesia [45].

It has been known for a long time that general anesthesia has profound effects on the spectrum of oscillations in the electroencephalograph [77,78]. More recently, the authors in [79] have suggested that thalamocortical circuits function as neural pacemakers and that alterations in the thalamic oscillations are associated with the induction of general anesthesia. Furthermore, it is well known that anesthetic drugs frequently induce epileptiform activity as part of the progression to the state of unconsciousness [31].

Multiple lines of evidence indicate that anesthetic agents impact neural oscillators. In addition, epileptiform activity implies synchronization of oscillators. This leads to the possibility that synchronization of these oscillators is involved in the transition to the anesthetic state. To develop global synchronization properties for the biological neural network system given by Equation (41) we introduce the notions of asymptotic synchronization and exponential synchronization.

Definition 7 The biological neural network given by Equation (41) is said to be globally asymptotically synchronized if

$$\lim_{t \rightarrow \infty} |S_i(t) - S_j(t)| = 0 \tag{52}$$

for all $S_0 \in \overline{\mathbb{R}}_+^n$ and $i, j = 1, 2, \dots, n, i \neq j$.

Definition 8 The biological neural network given by Equation (41) is said to be globally exponentially synchronized if there exist constants $\rho > 0$ and $p > 0$ such that

$$|S_i(t) - S_j(t)| \leq \rho e^{-pt} |S_{0i} - S_{0j}|, \quad t \geq 0, \tag{53}$$

for all $S_0 = [S_{01}, \dots, S_{0n}]^T \in \overline{\mathbb{R}}_+^n$ and $i, j = 1, 2, \dots, n, i \neq j$.

The following theorems provide sufficient conditions for global asymptotic synchronization and global exponentially synchronization of the biological neural network system given by Equation (41). For the statement of the theorems we define the ones vector of order n by $e_n \triangleq [1, \dots, 1]^T$.

Theorem 3 Consider the biological neural network given by Equation (41) with $f_i(\cdot), i = 1, 2, \dots, n$, given by Equation (10). If there exist positive definite matrices $P, Q \in \mathbb{R}^{n \times n}$ and a diagonal positive-definite matrix $R \in \mathbb{R}^{n \times n}$ such that

$$\begin{bmatrix} Q & -P \\ -P & R \end{bmatrix} \geq 0, \tag{54}$$

and either $\Omega_3 < 0$ or both $\Omega_3 \leq 0$ and $\mathcal{N}(\Omega_3) = \text{span}(e_n)$ hold, where

$$\Omega_3 \triangleq -PL - LP + Q + A^T R A, \tag{55}$$

then Equation (41) is globally asymptotically synchronized.

Proof. Consider the Lyapunov function candidate $V : \overline{\mathbb{R}}_+^n \rightarrow \mathbb{R}$ given by $V(S) = S^T P S$. It follows that the derivative of $V(S)$ along the trajectories of Equation (41) is given by

$$\dot{V}(S) = 2S^T P (-LS + \tilde{f}(AS)) = -2S^T P L S + 2S^T P \tilde{f}(AS).$$

Next, it follows from Equation (54) that $\dot{V}(S)$ satisfies

$$\begin{aligned} \dot{V}(S) &\leq -2S^T P L S + S^T Q S + \tilde{f}^T(AS) R \tilde{f}(AS) \\ &= S^T (-PL - LP + Q) S + \tilde{f}^T(AS) R \tilde{f}(AS), \quad S \in \overline{\mathbb{R}}_+^n. \end{aligned} \tag{56}$$

Since $f_i^2(x) \leq x^2, x \in \mathbb{R}$, for $f_i(\cdot), i = 1, \dots, n$, given by Equation (10) and $R \in \mathbb{R}^{n \times n}$ is a positive-definite diagonal matrix, it follows that

$$\tilde{f}^T(AS) R \tilde{f}(AS) \leq S^T A^T R A S, \quad S \in \overline{\mathbb{R}}_+^n. \tag{57}$$

Now, it follows from Equations (56) and (57) that

$$\dot{V}(S) \leq S^T \Omega_3 S, \quad S \in \overline{\mathbb{R}}_+^n. \tag{58}$$

If $\Omega_3 < 0$, then $\dot{V}(S) \leq 0$, $S \in \overline{\mathbb{R}}_+^n$, and $\dot{V}(S) = 0$ if and only if $S = 0$. Hence, the zero solution $S(t) \equiv 0$ to Equation (41) is asymptotically stable, which implies that $\lim_{t \rightarrow \infty} |S_i(t) - S_j(t)| = 0$ for all $S_0 \in \mathbb{R}^n$ and $i, j = 1, 2, \dots, n, i \neq j$. Hence, Equation (41) is globally asymptotically synchronized. Alternatively, if $\Omega_3 \leq 0$ and $\mathcal{N}(\Omega_3) = \text{span}(\mathbf{e}_n)$ holds, then $\dot{V}(S(t)) \leq 0, t \geq 0$, and hence, $V(S(t)) \leq V(S_0)$ for all $t \geq 0$. Next, since P is positive definite and $\dot{V}(S(t))$ is a nonincreasing function of time, it follows that $V(S(t))$ is bounded for all $t \geq 0$, and hence, $S(t)$ is bounded for all $t \geq 0$, which further implies that $\ddot{V}(S(t))$ is bounded for all $t \geq 0$. Thus, $\dot{V}(S(t))$ is uniformly continuous in t . Now, it follows from Barbalat’s lemma ([20](p. 221)) that $\dot{V}(S(t)) \rightarrow 0$ as $t \rightarrow \infty$, which, since $\mathcal{N}(\Omega_3) = \text{span}(\mathbf{e}_n)$, implies that Equation (41) is globally asymptotically synchronized. \square

Theorem 4 Consider the biological neural network given by Equation (41) with $f_i(\cdot), i = 1, 2, \dots, n$, given by Equation (10). If there exist positive definite matrices $P, Q \in \mathbb{R}^{n \times n}$, a diagonal positive-definite matrix $R \in \mathbb{R}^{n \times n}$ and a scalar $\varepsilon > 0$ such that Equation (54) holds, and either $\Omega_4 < 0$ or both $\Omega_4 \leq 0$ and $\mathcal{N}(\Omega_4) = \text{span}(\mathbf{e}_n)$ hold, where

$$\Omega_4 \triangleq 2\varepsilon P - PL - LP + Q + A^T R A, \tag{59}$$

then Equation (41) is globally exponentially synchronized.

Proof. The proof is similar to the proof of Theorem 3 using the function $V : [0, \infty) \times \overline{\mathbb{R}}_+^n \rightarrow \mathbb{R}$ given by $V(t, S) = e^{2\varepsilon t} S^T P S$ and, hence, is omitted. \square

Definition 9 ([80]) Let $(\mathcal{R}, +, \cdot)$ be a ring with the two binary operations of addition (+) and multiplication (\cdot) connected by distributive laws, and let $\mathcal{T}(\mathcal{R}, K)$ be the set of matrices with entries in \mathcal{R} such that the sum of the entries in each row is equal to K for some $K \in \mathcal{R}$.

Lemma 1 ([81]) Let G be an $n \times n$ matrix such that $G \in \mathcal{T}(\mathcal{R}, K)$. Then there exists a matrix $\overline{G} \in \mathbb{R}^{(n-1) \times (n-1)}$ given by $\overline{G} = MGJ$ such that $MG = \overline{G}M$, where M and J are given by

$$M = \begin{bmatrix} 1 & -1 & 0 & \cdots & 0 \\ 0 & 1 & -1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & 1 & -1 \end{bmatrix} \in \mathbb{R}^{(n-1) \times n}, \quad J = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ 0 & 1 & \ddots & \vdots \\ 0 & \ddots & \ddots & 1 \\ 0 & \cdots & 0 & 1 \\ 0 & \cdots & 0 & 0 \end{bmatrix} \in \mathbb{R}^{n \times (n-1)}. \tag{60}$$

Next, we analyze the synchronization properties of a special class of biological neural network given by Equation (41) where $L, A \in \mathcal{T}(\mathcal{R}, K)$ and $f_i(\cdot), i = 1, \dots, n$, satisfies

$$f_i(x) = \begin{cases} 0, & x < 0, \\ x, & 0 \leq x \leq f_{\max}, \\ f_{\max}, & x > f_{\max}, \end{cases} \tag{61}$$

where f_{\max} denotes the maximum firing rate.

Lemma 2 Consider the biological neural network given by Equation (41) and assume that $f_i(\cdot)$, $i = 1, \dots, n$, is given by Equation (61). Then $S(t)$ is bounded for all $t \geq 0$.

Proof. Consider the function $V(S) = S^T S$ and note that the derivative of $V(S)$ along the trajectories of Equation (41) satisfies

$$\dot{V}(S) = 2S^T(-LS + \tilde{f}(AS)) = -2S^T L S + 2S^T \tilde{f}(AS). \tag{62}$$

Note that since $2x^T y \leq r x^T x + \frac{1}{r} y^T y$ for all $x, y \in \mathbb{R}^n$ and $r > 0$, it follows that

$$\dot{V}(S) \leq -2S^T L S + \lambda_{\min}(L) S^T S + \frac{1}{\lambda_{\min}(L)} \tilde{f}^T(AS) \tilde{f}(AS), \quad S \in \overline{\mathbb{R}}_+^n. \tag{63}$$

Now, since $0 \leq f_i(x) \leq f_{\max}$ for all $x \in \mathbb{R}$ and $S^T L S \geq \lambda_{\min}(L) S^T S$, it follows that

$$\begin{aligned} \dot{V}(S) &\leq -2\lambda_{\min}(L) S^T S + \lambda_{\min}(L) S^T S + \frac{n}{\lambda_{\min}(L)} f_{\max}^2 \\ &= -\lambda_{\min}(L) S^T S + \frac{n}{\lambda_{\min}(L)} f_{\max}^2, \quad S \in \overline{\mathbb{R}}_+^n. \end{aligned} \tag{64}$$

Next, we show that if $\|S_0\|^2 = c < m$, where $m = \frac{n}{\lambda_{\min}^2(L)} f_{\max}^2$, then $V(S(t)) \leq m$ for all $t \geq 0$. To see this, assume, *ad absurdum*, that $V(S(t)) > m$ for some $t > 0$, which holds since $V(S(t))$ is continuous in t , and note that there exists $\tau \in (0, t)$ such that $V(S(\tau)) = m$ and $\dot{V}(S(\tau)) > 0$. Now, using Equation (64) it follows that $\dot{V}(S(\tau)) \leq 0$ for all $\tau > 0$ such that $V(S(\tau)) = m$, which contradicts $\dot{V}(S(\tau)) > 0$. Alternately, if $\|S_0\|^2 = c \geq m$, then using a similar argument it can be shown that $V(S(t)) \leq c$ for all $t \geq 0$. Hence, $\|S(t)\|^2$ is bounded for all $t \geq 0$, and hence, the solution $S(t)$ is bounded for all $t \geq 0$. \square

Theorem 5 Consider the biological neural network given by Equation (41) where $L, A \in \mathcal{T}(\mathbb{R}, K)$ and $f_i(\cdot), i = 1, 2, \dots, n$, is given by Equation (61). If there exist positive definite matrices $P, Q \in \mathbb{R}^{(n-1) \times (n-1)}$, a diagonal positive-definite matrix $R \in \mathbb{R}^{(n-1) \times (n-1)}$ such that Equation (54) holds, and

$$\Omega_5 \triangleq -P\bar{L} - \bar{L}P + Q + \bar{A}^T R \bar{A} < 0, \tag{65}$$

where \bar{L} and \bar{A} are generated from L and A using Lemma 1, then Equation (41) is globally asymptotically synchronized.

Proof. Consider the function $V : \overline{\mathbb{R}}_+^n \rightarrow \mathbb{R}$ given by $V(S) = S^T M^T P M S$, where M is given by Equation (60). It follows that the derivative of $V(S)$ along the trajectories of Equation (41) is given by

$$\dot{V}(S) = 2S^T M^T P M (-LS + \tilde{f}(AS)) = -2S^T M^T P M L S + 2S^T M^T P M \tilde{f}(AS). \tag{66}$$

Next, it follows from Equation (54) that

$$\dot{V}(S) \leq -2S^T M^T P M L S + S^T M^T Q M S + \tilde{f}^T(AS) M^T R M \tilde{f}(AS), \quad S \in \overline{\mathbb{R}}_+^n. \tag{67}$$

Note that since $R \in \mathbb{R}^{(n-1) \times (n-1)}$ is a diagonal positive-definite matrix and $f_i(\cdot), i = 1, 2, \dots, n$, given by Equation (61) satisfies $(f_i(x) - f_j(y))^2 \leq (x - y)^2$ for all $x, y \in \mathbb{R}$, it follows that $\tilde{f}^T(AS) M^T R M \tilde{f}(AS) \leq S^T A^T M^T R M A S, S \in \overline{\mathbb{R}}_+^n$. Hence,

$$\begin{aligned} \dot{V}(S) &\leq -2S^T M^T P M L S + S^T M^T Q M S + S^T A^T M^T R M A S \\ &= S^T M^T (-P\bar{L} - \bar{L}P + Q) S + S^T M^T \bar{A}^T R \bar{A} M S \\ &= S^T M^T \Omega_5 M S, \quad S \in \overline{\mathbb{R}}_+^n. \end{aligned} \tag{68}$$

Since $\Omega_5 < 0$, $\dot{V}(S) \leq 0$, $S \in \overline{\mathbb{R}}_+^n$, and hence, $V(S(t)) \leq V(S_0)$ for all $t \geq 0$. Next, since P is positive definite and $\dot{V}(S(t))$ is a non-increasing function of time, it follows that $V(S(t))$ is bounded for all $t \geq 0$. Since, by Lemma 2, $S(t)$ is bounded for all t , it follows that $\ddot{V}(S(t))$ is bounded for all $t \geq 0$, and hence, $\dot{V}(S(t))$ is uniformly continuous in t . Now, it follows from Barbalat’s lemma ([20](p. 221)) that $\dot{V}(S(t)) \rightarrow 0$ as $t \rightarrow \infty$, which implies that $\lim_{t \rightarrow \infty} MS(t) = 0$. Hence, Equation (41) is globally asymptotically synchronized. \square

Theorem 6 Consider the biological neural network given by Equation (41) where $L, A \in \mathcal{T}(\mathcal{R}, K)$ and $f_i(\cdot), i = 1, 2, \dots, n$, is given by Equation (61). If there exist positive definite matrices $P, Q \in \mathbb{R}^{(n-1) \times (n-1)}$, a diagonal positive-definite matrix $R \in \mathbb{R}^{(n-1) \times (n-1)}$, a scalar $\varepsilon > 0$ such that Equation (54) holds, and

$$\Omega_6 \triangleq 2\varepsilon P - P\bar{L} - \bar{L}P + Q + \bar{A}^T R \bar{A} < 0, \tag{69}$$

where \bar{L} and \bar{A} are generated from L and A using Lemma 1, then Equation (41) is globally exponentially synchronized.

Proof. The proof is similar to the proof of Theorem 5 using the function $V : [0, \infty) \times \overline{\mathbb{R}}_+^n \rightarrow \mathbb{R}$ given by

$$V(t, S) = e^{2\varepsilon t} S^T M^T P M S$$

and, hence, is omitted. \square

Example 3 Consider the biological neural network given by Equation (41) consisting of a network of six excitatory and three inhibitory neurons shown in Figure 4. The neural connectivity matrix A for this network is given by

$$A = \left[\begin{array}{cccccc|ccc} 0 & 0 & 0 & 0 & 2 & 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 & 0 & 2 & 0 & 0 & -1 \\ 2 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 \\ 0 & 2 & 0 & 0 & 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 2 & 0 & 0 & 0 & -1 & 0 \\ \hline 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 1 & 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 0 & -1 & 0 \end{array} \right].$$

Here, we assume that the time constant of all nine neurons are the same, and hence, $L = \frac{1}{\tau} I_9$. Furthermore, we assume the functions $f_i(\cdot), i = 1, 2, \dots, 9$, are given by Equation (61) with the same maximum firing rate $f_{\max} = 0.5$.

Next, we construct \bar{A} and \bar{L} using Lemma 1 and solve the Linear Matrix Inequalities (LMIs) given by Equations (54) and (65) for the neuron time constants $\tau = 10, 1$, and 0.1 . We use the MATLAB toolbox YALMIP for solving the LMIs. For $\tau = 10$, there is no feasible solution to Equations (54) and (65), and, as shown in Figure 5, the synaptic drive of the inhibitory neurons oscillate whereas the synaptic drive of the excitatory neurons converge to a steady-state. For $\tau = 1$, Equations (54) and (65) are also not

satisfied and as shown in Figure 6, the synaptic drive of all the neurons converge to different steady state values, which implies that the network is not synchronized. Finally, for $\tau = 0.1$, Equations (54) and (65) are satisfied with P , Q , and R , given by

$$P = \begin{bmatrix} 25.85 & 11.25 & 9.479 & 4.939 & 4.977 & 0 & -1.008 & 0.7254 \\ 11.25 & 32.11 & 14.01 & 5.302 & 5.223 & 0 & -3.626 & 3.375 \\ 9.479 & 14.01 & 30.6 & 8.948 & 7.178 & 0 & -3.037 & 1.77 \\ 4.939 & 5.302 & 8.948 & 18.92 & 8.76 & 0 & -6.062 & -2.118 \\ 4.977 & 5.223 & 7.178 & 8.76 & 24.36 & 0 & -4.475 & -2.172 \\ 0 & 0 & 0 & 0 & 0 & 7.104 & 0 & 0 \\ -1.008 & -3.626 & -3.037 & -6.062 & -4.475 & 0 & 13.32 & 2.425 \\ 0.7254 & 3.375 & 1.77 & -2.118 & -2.172 & 0 & 2.425 & 10.53 \end{bmatrix},$$

$$Q = \begin{bmatrix} 65.66 & 19.39 & 15.22 & 7.273 & 7.601 & 0 & -5.527 & 2.403 \\ 19.39 & 92.92 & 25.48 & 13.29 & 12.26 & 0 & -5.246 & 1.426 \\ 15.22 & 25.48 & 74.49 & 9.073 & 11.16 & 0 & -6.843 & 0.72 \\ 7.273 & 13.29 & 9.073 & 67.94 & 5.833 & 0 & -0.2344 & 1.604 \\ 7.601 & 12.26 & 11.16 & 5.833 & 62.91 & 0 & -4.261 & 1.016 \\ 0 & 0 & 0 & 0 & 0 & 72.06 & 0 & 0 \\ -5.527 & -5.246 & -6.843 & -0.2344 & -4.261 & 0 & 65.52 & -3.424 \\ 2.403 & 1.426 & 0.72 & 1.604 & 1.016 & 0 & -3.424 & 66.32 \end{bmatrix},$$

$$R = \text{diag} \left[60.65 \quad 23.27 \quad 63.18 \quad 45.2 \quad 55.94 \quad 24.87 \quad 33.43 \quad 40.61 \right].$$

Hence, the biological neural network is asymptotically synchronized. See Figure 7.

Figure 4. A population of six excitatory and three inhibitory neurons. Neurons E_1, \dots, E_6 are excitatory and neurons I_1, \dots, I_3 are inhibitory. The synaptic weights shown on the connecting arcs represent the coupling strength of the neural interconnections.

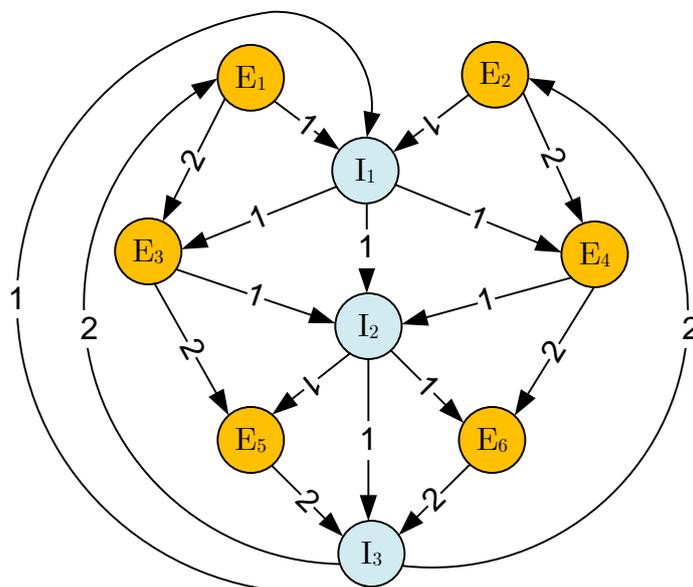


Figure 5. Solutions to Equation (41) with initial conditions $S(0) = [0.1, 0.2, 0.3, 0.5, 0.3, 0.7, 0.4, 0.8, 0.6]^T$ for $\tau = 10$. The synaptic drive of the inhibitory neurons oscillate, whereas the synaptic drive of excitatory neurons converge to zero.

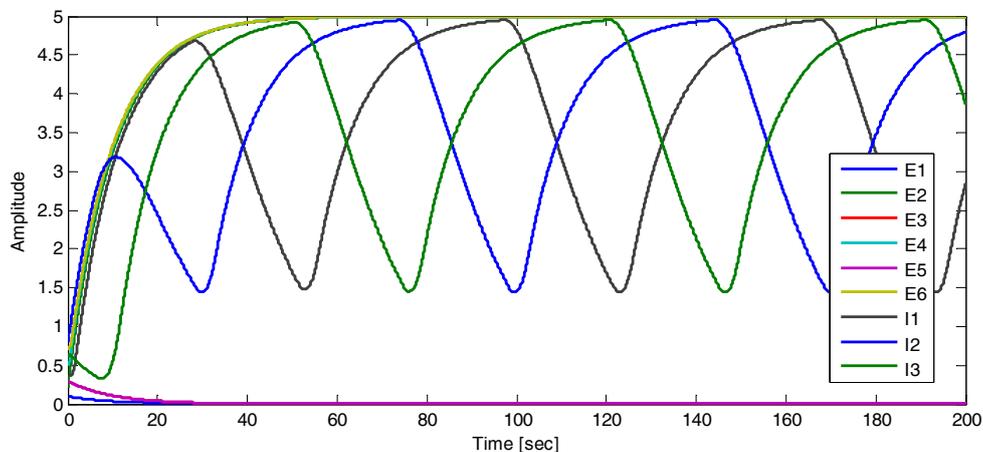


Figure 6. Solutions to Equation (41) with initial conditions $S(0) = [0.1, 0.2, 0.3, 0.5, 0.3, 0.7, 0.4, 0.8, 0.6]^T$ for $\tau = 1$. The synaptic drive of all the neurons converge to different values.

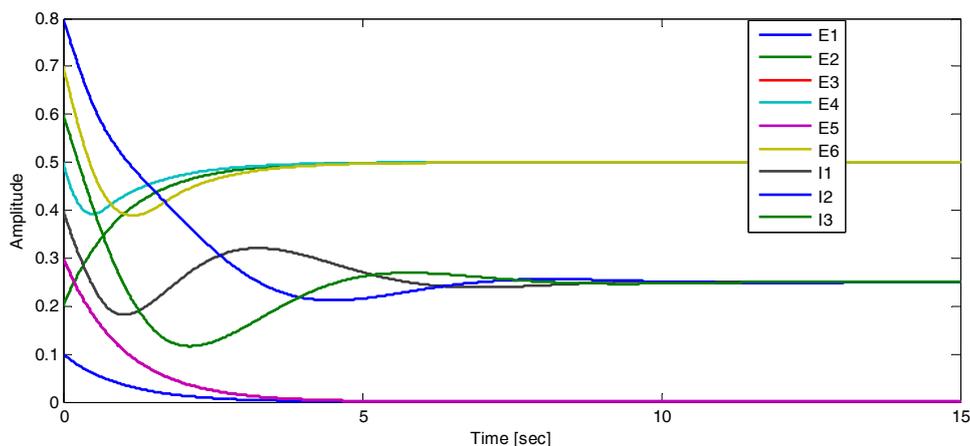
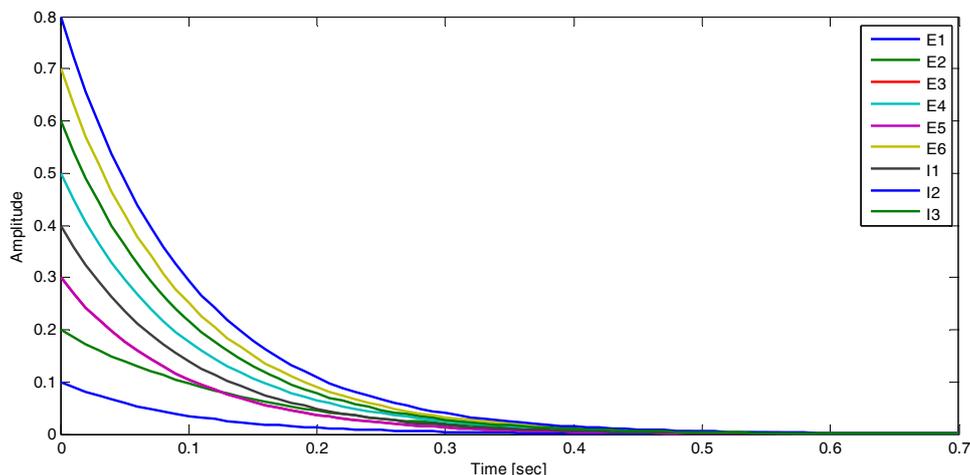


Figure 7. Solutions to Equation (41) with initial conditions $S(0) = [0.1, 0.2, 0.3, 0.5, 0.3, 0.7, 0.4, 0.8, 0.6]^T$ for $\tau = 0.1$. All neurons are synchronized.



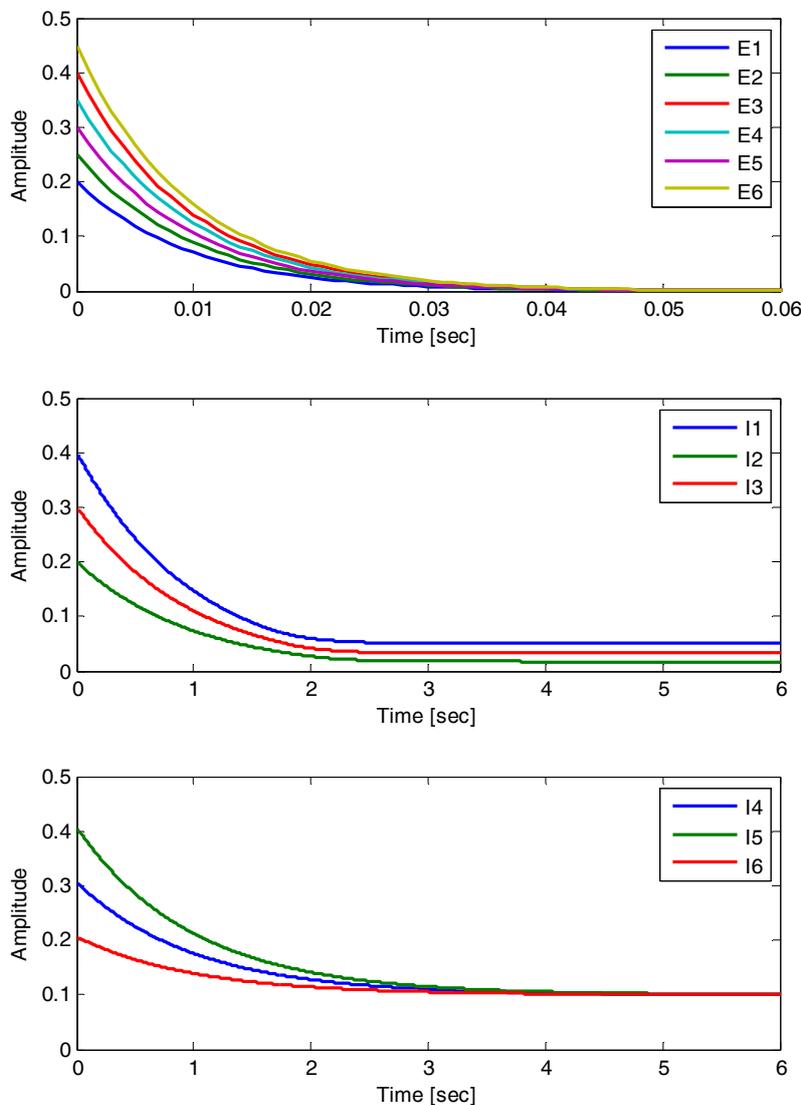
Example 4 When patients lose consciousness some parts of the brain are still functional (e.g., cardiac function) whereas other parts are suppressed. This can be captured by biological neural network models that exhibit partial synchronization wherein part of the system’s state is synchronized and the other parts fire at normal levels. In addition, the administration of increasing anesthetic doses can lead to a paradoxical state of excitement in the patient prior to decreases in the level of consciousness. This paradoxical boost in brain activity prior to hypnotic induction is known as drug biphasic response [82]. There is also a second biphasic surge in the EEG power as the patient emerges from unconsciousness. Models that predict the aforementioned characteristics are of great clinical importance in providing the phenomenological trends of the anesthetic cascade.

To demonstrate partial synchronization of the model presented in Section 2, consider the biological neural network given by Equation (41) consisting of a population of neurons with six excitatory neurons $E_1 - E_6$ and six inhibitory neurons $I_1 - I_6$. The neural connectivity matrix A for this network is given by

$$A = \begin{bmatrix} 0 & 1 & 1 & 1 & 1 & 1 & -1 & 0 & -1 & 0 & -1 & 0 \\ 1 & 0 & 1 & 1 & 1 & 1 & 0 & -1 & -1 & -1 & 0 & 0 \\ 1 & 1 & 0 & 1 & 1 & 1 & -1 & 0 & 0 & 0 & -1 & -1 \\ 1 & 1 & 1 & 0 & 1 & 1 & 0 & -1 & 0 & -1 & 0 & -1 \\ 1 & 1 & 1 & 1 & 0 & 1 & -1 & 0 & -1 & 0 & -1 & 0 \\ 1 & 1 & 1 & 1 & 1 & 0 & 0 & -1 & 0 & -1 & 0 & -1 \\ 1 & 0 & 1 & 0 & 1 & 0 & 0 & -1 & -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 & -1 & 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 1 & 0 & -1 & -1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix},$$

which implies that the three inhibitory neurons $I_4, I_5,$ and I_6 do not receive any inhibitory inputs. Here we assume that the excitatory neurons have a time constant $\lambda^E = 0.01$ s and the inhibitory neurons have a prolonged time constant $\lambda^I = 1$ s. Furthermore, we assume $v_{th}^E = v_{th}^I = 0.1V$ for all neurons, and the functions $f_i(\cdot), i = 1, 2, \dots, 12,$ are given by Equation (61) with the same maximum firing rate $f_{max} = 0.5$. Figure 8 shows that as λ^I increases, the excitatory neurons that are coupled to inhibitory neurons all go to a zero synaptic drive, whereas the inhibitory neurons that themselves are not coupled to inhibitory neurons synchronize to some finite value.

Figure 8. Solutions to Equation (41) with initial conditions $S(0) = [0.2, 0.25, 0.4, 0.35, 0.3, 0.45, 0.4, 0.2, 0.3, 0.3, 0.4, 0.2]^T$ for $\lambda^E = 0.01$ s and $\lambda^I = 1$ s. The synaptic drive of the excitatory neurons converges to zero, whereas the synaptic drive of the inhibitory neurons converge to nonzero values. Synchronization is observed in the synaptic drive of the inhibitory neurons $I_4, I_5,$ and I_6 that themselves do not receive inhibitory inputs.



8. Stochastic Multistability for a Mean Field Synaptic Drive Firing Neuronal Model

Since the neurocortex contains on the order of 10^{11} neurons, each supporting up to 10^5 synaptic contacts, we extend our dynamical system framework to a stochastic setting. Specifically, a large population of spiking neurons can be reduced to a distribution function describing their probabilistic evolution; that is, a function that captures the distribution of neuronal states at a given time [83]. In this section, we develop a stochastic field theory as in [84] for capturing neural activity in order to analyze system multistability. In the next section, we extend the results of this section to additionally address time delay functional models [85] in order to account for time delay and memory effects in inhibitory and excitatory networks.

In Sections 4 and 5 we developed deterministic multistability theory to explain the underlying mechanism of action for anesthesia and consciousness using a synaptic drive firing model framework [3]. In this section, we extend these results further by demonstrating multistability in the mean when the coefficients of the neuronal connectivity matrix are random variables. Specifically, we use a stochastic multiplicative uncertainty model to include modeling of a priori uncertainty in the coefficients of the neuronal connectivity matrix by means of state-dependent noise. The philosophy of representing uncertain parameters by means of multiplicative white noise is motivated by the Maximum Entropy Principle of Jaynes [86,87] and statistical analysis [88].

Maximum entropy modeling is a form of stochastic modeling wherein stochastic integration is interpreted in the sense of Itô to provide a model for system parameter uncertainty. The use of stochastic theory to model system parameter uncertainty has been used within a modern information-theoretic interpretation of probability theory [86,87,89]. In particular, rather than regarding the probability of an event as an *objective* quantity such as the limiting frequency of outcomes of numerous repetitions, maximum entropy modeling adopts the view that the probability of an event is a *subjective* quantity which reflects the observer’s certainty to a particular event occurring. This quantity corresponds to a measure of information. The validity of a stochastic model for a biological neural network does not rely on the existence of an ensemble model but rather in the interpretation that it expresses modeling certainty or uncertainty regarding the coefficients of the neuronal connectivity matrix. Hence, a stochastic multiplicative uncertainty model utilizes state-dependent Gaussian white noise to represent parameter uncertainty by defining a measure of ignorance, in terms of an information-theoretic entropy, and then determining the probability distribution which *maximizes* this measure subject to agreement with a given model.

To develop stochastic multistability synaptic drive model, consider for simplicity of exposition the simplified mean field synaptic drive model where the coefficients of Equations (24) and (25), with $f_i(\cdot)$ given by Equation (10), are randomly disturbed. Specifically, we assume that the initial value $S(0) \triangleq [\bar{S}^E(0), \bar{S}^I(0)]^T$ is deterministic and contained in the nonnegative orthant of the state space, and consider the stochastic differential mean field synaptic drive model given by

$$dS(t) = -LS(t)(dt + \nu dw(t)) + [AS(t)(dt + \nu dw(t))]_+, \quad S(0) = S_0, \quad t \geq 0, \quad (70)$$

where

$$L = \begin{bmatrix} \frac{1}{\lambda^E} & 0 \\ 0 & \frac{1}{\lambda^I} \end{bmatrix}, \quad A = \begin{bmatrix} n_E \bar{A}^{EE} & n_I \bar{A}^{EI} \\ n_E \bar{A}^{IE} & n_I \bar{A}^{II} \end{bmatrix},$$

$w(t)$ represents Brownian motion, that is, a Wiener process, $\nu \in \mathbb{R}$ indicates the intensity of the Gaussian white noise $dw(t)$, and $[x]_+ \triangleq [[x_1]_+, [x_2]_+]^T$ for $x = [x_1, x_2]^T \in \mathbb{R}^2$. Here, we assume that every entry of the matrices A and L of the *mean* dynamics given by Equations (24) and (25) (with $v_{th}^E = v_{th}^I = 0$) is synchronously perturbed.

For the statement of the results in this section, we require some additional notation and definitions. Specifically, let $(\Omega, \mathcal{F}, \mathbb{P})$ be the probability space associated with Equation (70), where Ω denotes the sample space, \mathcal{F} denotes a σ -algebra, and \mathbb{P} defines a probability measure on the σ -algebra \mathcal{F} , that is, \mathbb{P} is a nonnegative countably additive set function on \mathcal{F} such that $\mathbb{P}(\Omega) = 1$ [90]. Note that Equation (70)

is a Markov process, and hence, there exists a filtration $\{\mathcal{F}_t\}$ satisfying $\mathcal{F}_\tau \subset \mathcal{F}_t \subset \mathcal{F}$, $0 \leq \tau < t$, such that $\{\omega \in \Omega : S(t) \in B\} \in \mathcal{F}_t$, $t \geq 0$, for all Borel sets $B \subset \mathbb{R}^2$ contained in the Borel σ -algebra \mathcal{B} . Finally, $\text{spec}(X)$ denotes the spectrum of the square matrix X including multiplicity and $\mathcal{R}(Y)$ denotes the range space of the matrix Y .

When every component of the vector $AS(t)(dt + \nu dw(t))$, $t \geq 0$, is nonnegative, the stochastic dynamical system given by Equation (70) can be written as

$$dS(t) = \tilde{A}S(t)dt + \tilde{A}_s S(t)dw(t), \quad S(0) = S_0, \quad t \geq 0, \tag{71}$$

where $\tilde{A} \triangleq A - L$ and $\tilde{A}_s \triangleq \nu(A - L)$. The multiplicative white noise model given by Equation (71) can be regarded as a parameter uncertainty model where $dw(t)$ corresponds to an uncertain parameter whose *pattern* and *magnitude* are given by $\tilde{A}_s/\|\tilde{A}\|$ and $\|\tilde{A}_s\|$, respectively. Note that if $\text{rank}(A - L) < 2$, then every point $\alpha \in \mathcal{N}(\tilde{A})$ is an equilibrium point of Equation (71). With a slight abuse of notation, we use \mathcal{E} to denote the equilibrium set of Equation (70) or Equation (71). First, motivated by the definition of stochastic Lyapunov stability in [90], we have the following definition of stochastic semistability. For the statement of the next result, define $\text{dist}(x, \mathcal{E}) \triangleq \inf_{y \in \mathcal{E}} \|x - y\|$. For a similar definition of stochastic semistability, see [91].

Definition 10 *An equilibrium solution $S(t) \equiv \alpha \in \mathcal{E}$ of Equation (70) is stochastically semistable if the following statements hold.*

- (i) For every $\varepsilon > 0$, $\lim_{S(0) \rightarrow \alpha} \mathbb{P} \left[\sup_{0 \leq t < \infty} \|S(t) - \alpha\| \geq \varepsilon \right] = 0$.
- (ii) $\lim_{S(0) \rightarrow \mathcal{E}} \mathbb{P} [\lim_{t \rightarrow \infty} \text{dist}(S(t), \mathcal{E}) = 0] = 1$.

The dynamical system given by Equation (70) is stochastically semistable if every equilibrium solution of Equation (70) is stochastically semistable. Finally, the system given by Equation (70) is globally stochastically semistable if it is stochastically semistable and $\mathbb{P} \left[\lim_{t \rightarrow \infty} \text{dist}(S(t), \mathcal{E}) = 0 \right] = 1$ for every initial condition $S(0) \in \mathbb{R}^2$. If, alternatively, $S(t) \equiv \alpha \in \mathcal{E}$ only satisfies i), then the equilibrium solution $S(t) \equiv \alpha \in \mathcal{E}$ of Equation (70) is stochastically Lyapunov stable.

Definition 10 is a stability notion for the stochastic dynamical system given by Equation (70) having a continuum of equilibria and is a generalization of the notion of semistability from deterministic dynamical systems [63,92] to stochastic dynamical systems. It is noted in [92] that existing methods for analyzing the stability of deterministic dynamical systems with isolated equilibria cannot be used for deterministic dynamical systems with nonisolated equilibria due to the connectedness property of equilibrium sets. Hence, Definition 10 is essential for analyzing the stability of stochastic dynamical systems with nonisolated equilibria. Note that (i) in Definition 10 implies stochastic Lyapunov stability of an equilibrium, whereas (ii) implies almost sure convergence of trajectories to the equilibrium manifold.

Next, we extend the notion of multistability for deterministic dynamical systems defined in Section 4 to that of stochastic multistability for stochastic dynamical systems.

Definition 11 *Consider the dynamical system given by Equation (70) and let $\mu(\cdot)$ denote the Lebesgue measure in \mathbb{R}^2 . We say that the system given by Equation (70) is stochastically multistable if the following statements hold.*

- (i) $\mathcal{E} \setminus \{(0, 0)\} \neq \emptyset$.
- (ii) For every $S(0) \in \mathbb{R}^2$, there exists $\alpha(\omega) \in \mathcal{E}$, $\omega \in \Omega$, such that $\mathbb{P}[\lim_{t \rightarrow \infty} S(t) = \alpha(\omega)] = 1$.
- (iii) There exists a subset $\mathcal{M} \subset \mathbb{R}^2$ satisfying $\mu(\mathcal{M}) = 0$ such that, for every $S(0) \in \mathbb{R}^2 \setminus \mathcal{M}$, $\mathbb{P}[\lim_{t \rightarrow \infty} \text{dist}(S(t), \mathcal{E}) = 0] = 1$.

Stochastic multistability is a global stability notion for the stochastic dynamical system given by Equation (70) having isolated equilibria and/or a continuum of equilibria, whereas stochastic semistability is a local stability notion for the stochastic dynamical system given by Equation (70) having a continuum of equilibria. Hence, stochastic multistability is a stronger notion than stochastic semistability. The next result states a relationship between stochastic multistability and global stochastic semistability.

Proposition 2 Consider the dynamical system given by Equation (70). If Equation (70) is globally stochastically semistable, then Equation (70) is stochastically multistable.

Proof. Suppose that the dynamical system given by Equation (70) is globally stochastically semistable. Then, by definition, $\mathbb{P}[\lim_{t \rightarrow \infty} \text{dist}(S(t), \mathcal{E}) = 0] = 1$ for every initial condition $S(0) \in \mathbb{R}^2$. Next, we show that for every $S(0) \in \mathbb{R}^2$, there exists $\alpha = \alpha(\omega) \in \mathcal{E}$, $\omega \in \Omega$, such that $\mathbb{P}[\lim_{t \rightarrow \infty} S(t) = \alpha(\omega)] = 1$. Let

$$\Gamma(S) \triangleq \left\{ x \in \mathbb{R}^2 : \text{there exists a divergent sequence } \{t_i\}_{i=1}^\infty \text{ such that } \mathbb{P}\left[\lim_{i \rightarrow \infty} S(t_i) = x\right] = 1 \right\}$$

and $\mathcal{B}_\delta(z) \triangleq \{x \in \mathbb{R}^2 : \|x - z\| < \delta\}$.

Suppose $z \in \Gamma(S)$ is stochastically Lyapunov stable and let $\varepsilon_1, \varepsilon_2 > 0$. Since z is stochastically Lyapunov stable there exists an open neighborhood $\mathcal{B}_\delta(z)$, where $\delta = \delta(\varepsilon_1, \varepsilon_2) > 0$, such that, for every $S(0) \in \mathcal{B}_\delta(z)$, $\mathbb{P}[\sup_{t \geq 0} \|S(t) - z\| \geq \varepsilon_1] < \varepsilon_2$, and hence, $\mathbb{P}[\sup_{t \geq 0} \|S(t) - z\| < \varepsilon_1] \geq 1 - \varepsilon_2$. Now, since $z \in \Gamma(S)$, it follows that there exists a divergent sequence $\{t_i\}_{i=1}^\infty$ in $[0, \infty)$ such that $\mathbb{P}[\lim_{i \rightarrow \infty} S(t_i) = z] = 1$, and hence, for every $\varepsilon_3, \varepsilon_4 > 0$, there exists $k = k(\varepsilon_3) \geq 1$ such that $\mathbb{P}[\sup_{i \geq k} \|S(t_i) - z\| > \varepsilon_3] < \varepsilon_4$ or, equivalently, $\mathbb{P}[\sup_{i \geq k} \|S(t_i) - z\| < \varepsilon_3] \geq 1 - \varepsilon_4$.

Next, note that $\mathbb{P}[\sup_{t \geq t_k} \|S(t) - z\| < \varepsilon_1] \geq \mathbb{P}[\sup_{t \geq 0} \|S(t) - z\| < \varepsilon_1]$. It now follows that

$$\begin{aligned} \mathbb{P}\left[\sup_{t \geq t_k} \|S(t) - z\| < \varepsilon_1\right] &\geq \mathbb{P}\left[\sup_{t \geq t_k} \|S(t) - z\| < \varepsilon_1 \mid \sup_{i \geq k} \|S(t_i) - z\| < \varepsilon_3\right] \\ &\quad \times \mathbb{P}\left[\sup_{i \geq k} \|S(t_i) - z\| < \varepsilon_3\right] \\ &\geq (1 - \varepsilon_2)(1 - \varepsilon_4), \end{aligned}$$

where $\mathbb{P}[\cdot|\cdot]$ denotes conditional probability. Since $\varepsilon_1, \varepsilon_2$, and ε_4 were chosen arbitrarily, it follows that $\mathbb{P}[z = \lim_{t \rightarrow \infty} S(t)] = 1$. Thus, $\mathbb{P}[\lim_{n \rightarrow \infty} S(t_n) = z] = 1$ for every divergent sequence $\{t_n\}_{n=1}^\infty$, and hence, $\Gamma(S) = \{z\}$; that is, for every $S(0) \in \mathbb{R}^2$, there exists $\alpha = \alpha(\omega) \in \mathcal{E}$, $\omega \in \Omega$, such that $\mathbb{P}[\lim_{t \rightarrow \infty} S(t) = \alpha(\omega)] = 1$. □

Next, recall from [93] that a matrix $\tilde{A} \in \mathbb{R}^{n \times n}$ is *semistable* if and only if $\lim_{t \rightarrow \infty} e^{\tilde{A}t}$ exists. In other words, \tilde{A} is semistable if and only if for every $\lambda \in \text{spec}(\tilde{A})$, $\lambda = 0$ or $\text{Re } \lambda < 0$ and if $\lambda = 0$, then 0 is semisimple. Furthermore, if \tilde{A} is semistable, then the index of \tilde{A} is zero or one, and hence, \tilde{A} is group invertible. The *group inverse* $\tilde{A}^\#$ of \tilde{A} is a special case of the Drazin inverse \tilde{A}^D in the case where \tilde{A} has index zero or one [93]. In this case, $\lim_{t \rightarrow \infty} e^{\tilde{A}t} = I_q - \tilde{A}\tilde{A}^\#$ [93].

Proposition 3 *If \tilde{A} is semistable, then, for sufficiently small $|\nu|$, the dynamical system given by Equation (71) is globally stochastically semistable.*

Proof. First, note that the solution to Equation (71) is given by

$$S(t) = e^{\tilde{A}t}S(0) + \int_0^t e^{\tilde{A}(t-s)}\tilde{A}_s S(s)dw(s), \quad t \geq 0. \tag{72}$$

Since \tilde{A} is semistable, it follows that $\lim_{t \rightarrow \infty} e^{\tilde{A}t}S(0)$ exists. In this case, let $S_\infty = \lim_{t \rightarrow \infty} e^{\tilde{A}t}S(0)$. Furthermore, note that $S_\infty = (I_2 - \tilde{A}\tilde{A}^\#)S(0) \in \mathcal{N}(\tilde{A})$ [93], where $\tilde{A}^\#$ denotes the group inverse of \tilde{A} . Next, note that $\int_0^t e^{\tilde{A}(t-s)}\tilde{A}_s S(s)dw(s)$ is an Itô integral and let $\|\cdot\|$ denote the Euclidean norm on \mathbb{R}^2 . Then, it follows from Property e) of Theorem 4.4.14 of ([90](p. 73)) that

$$\begin{aligned} \mathbb{E} \left[\left\| \int_0^t e^{\tilde{A}(t-s)}\tilde{A}_s S(s)dw(s) \right\|^2 \right] &= \int_0^t \mathbb{E} \left[\left\| e^{\tilde{A}(t-s)}\tilde{A}_s S(s) \right\|^2 \right] ds \\ &= \nu^2 \int_0^t \mathbb{E} \left[\left\| e^{\tilde{A}(t-s)}\tilde{A}S(s) \right\|^2 \right] ds \\ &= \nu^2 \int_0^t \mathbb{E} \left[\left\| (e^{\tilde{A}(t-s)} - (I_2 - \tilde{A}\tilde{A}^\#))\tilde{A}S(s) \right\|^2 \right] ds \\ &= \nu^2 \int_0^t \mathbb{E} \left[\left\| (e^{\tilde{A}(t-s)} - (I_2 - \tilde{A}\tilde{A}^\#))\tilde{A}(S(s) - S_\infty) \right\|^2 \right] ds, \end{aligned} \tag{73}$$

where $\mathbb{E}[\cdot]$ denotes expectation with respect to the probability space $(\Omega, \mathcal{F}, \mathbb{P})$.

Next, define $e(t) \triangleq e^{\tilde{A}t}S(0) - (I_2 - \tilde{A}\tilde{A}^\#)S(0) = e^{\tilde{A}t}S(0) - S_\infty$. Then it follows from the semistability of \tilde{A} that $\lim_{t \rightarrow \infty} e(t) = 0$. Since $\dot{e}(t) = \tilde{A}e(t)$ for every $t \geq 0$, it follows from the equivalence of (uniform) asymptotic stability and (uniform) exponential stability for linear time-invariant systems [94] that there exist real scalars $\sigma, r > 0$ such that $\|e(t)\| \leq \sigma e^{-rt}\|e(0)\|, t \geq 0$, or, equivalently, $\| [e^{\tilde{A}t} - (I_2 - \tilde{A}\tilde{A}^\#)]S(0) \| \leq \sigma e^{-rt}\|\tilde{A}\tilde{A}^\#S(0)\|, t \geq 0$. Hence,

$$\begin{aligned} \| [e^{\tilde{A}t} - (I_2 - \tilde{A}\tilde{A}^\#)] \|' &= \max_{S(0) \in \mathbb{R}^2 \setminus \{0\}} \frac{\| [e^{\tilde{A}t} - (I_2 - \tilde{A}\tilde{A}^\#)]S(0) \|}{\|S(0)\|} \\ &\leq \sigma e^{-rt} \max_{S(0) \in \mathbb{R}^2 \setminus \{0\}} \frac{\|\tilde{A}\tilde{A}^\#S(0)\|}{\|S(0)\|} \\ &= \sigma e^{-rt}\|\tilde{A}\tilde{A}^\#\|', \quad t \geq 0, \end{aligned} \tag{74}$$

where $\|\cdot\|' = \sigma_{\max}(\cdot)$ and $\sigma_{\max}(\cdot)$ denotes the maximum singular value. Thus, Equation (74) implies

$$\| [e^{\tilde{A}t} - (I_2 - \tilde{A}\tilde{A}^\#)] \|' \leq \rho e^{-rt}, \quad t \geq 0, \tag{75}$$

where $\rho \triangleq \sigma\|\tilde{A}\tilde{A}^\#\|'$.

Next, it follows from Equations (73) and (75) that

$$\int_0^t \mathbb{E} \left[\left\| e^{\tilde{A}(t-s)}\tilde{A}_s S(s) \right\|^2 \right] ds \leq \nu^2 \rho^2 \|\tilde{A}\|^2 \int_0^t e^{-2r(t-s)} \mathbb{E} [\|S(s) - S_\infty\|^2] ds, \quad t \geq 0. \tag{76}$$

Now, it follows from Equations (72), (76), and the triangle inequality that

$$\begin{aligned} \mathbb{E} \left[\|S(t) - S_\infty\|^2 \right] &\leq \|e^{\tilde{A}t}S(0) - S_\infty\|^2 + \nu^2 \rho^2 \|\tilde{A}\|'^2 \int_0^t e^{-2r(t-s)} \mathbb{E} \left[\|S(s) - S_\infty\|^2 \right] ds \\ &\leq \|e^{\tilde{A}t}S(0) - S_\infty\|^2 + \nu^2 \rho^2 \|\tilde{A}\|'^2 e^{-2rt} \int_0^t e^{2rs} \mathbb{E} \left[\|S(s) - S_\infty\|^2 \right] ds, \quad t \geq 0, \end{aligned}$$

and hence,

$$\begin{aligned} e^{2rt} \mathbb{E} \left[\|S(t) - S_\infty\|^2 \right] &\leq e^{2rt} \|e^{\tilde{A}t}S(0) - S_\infty\|^2 + \nu^2 \rho^2 \|\tilde{A}\|'^2 \\ &\quad \times \int_0^t e^{2rs} \mathbb{E} \left[\|S(s) - S_\infty\|^2 \right] ds, \quad t \geq 0. \end{aligned}$$

Hence, it follows from the Gronwall-Bellman lemma ([20](p. 125)) that

$$\begin{aligned} e^{2rt} \mathbb{E} \left[\|S(t) - S_\infty\|^2 \right] &\leq e^{2rt} \|e^{\tilde{A}t}S(0) - S_\infty\|^2 + \nu^2 \rho^2 \|\tilde{A}\|'^2 \int_0^t e^{2rs} \\ &\quad \times \|e^{\tilde{A}s}S(0) - S_\infty\|^2 e^{\nu^2 \rho^2 \|\tilde{A}\|'^2 (t-s)} ds, \quad t \geq 0, \end{aligned}$$

or, equivalently, for $\nu \neq 0$,

$$\begin{aligned} \mathbb{E} \left[\|S(t) - S_\infty\|^2 \right] &\leq \|e^{\tilde{A}t}S(0) - S_\infty\|^2 + \nu^2 \rho^2 \|\tilde{A}\|'^2 \int_0^t e^{-2r(t-s)} \|e^{\tilde{A}s}S(0) - S_\infty\|^2 e^{\nu^2 \rho^2 \|\tilde{A}\|'^2 (t-s)} ds \\ &\leq \|e^{\tilde{A}t}S(0) - S_\infty\|^2 + \nu^2 \rho^4 \|\tilde{A}\|'^2 \|S(0)\|^2 \int_0^t e^{-2rt} e^{\nu^2 \rho^2 \|\tilde{A}\|'^2 (t-s)} ds \\ &= \|e^{\tilde{A}t}S(0) - S_\infty\|^2 + \nu^2 \rho^4 \|\tilde{A}\|'^2 \|S(0)\|^2 e^{-(2r-\nu^2 \rho^2 \|\tilde{A}\|'^2)t} \int_0^t e^{-\nu^2 \rho^2 \|\tilde{A}\|'^2 s} ds \\ &= \|e^{\tilde{A}t}S(0) - S_\infty\|^2 + \rho^2 \|S(0)\|^2 e^{-(2r-\nu^2 \rho^2 \|\tilde{A}\|'^2)t} \left(1 - e^{-\nu^2 \rho^2 \|\tilde{A}\|'^2 t} \right), \quad t \geq 0. \end{aligned}$$

Taking $|\nu|$ to be such that

$$\nu^2 \rho^2 \|\tilde{A}\|'^2 < 2r, \tag{77}$$

it follows that $\lim_{t \rightarrow \infty} e^{-(2r-\nu^2 \rho^2 \|\tilde{A}\|'^2)t} = 0$. In this case, $\lim_{t \rightarrow \infty} \mathbb{E}[\|S(t) - S_\infty\|^2] = 0$, that is, $S(t)$, $t \geq 0$, converges to S_∞ in the mean square.

Finally, by Theorem 7.6.10 of [95] or ([90](p. 187)) (Khasminskiy’s theorem), for every initial condition $S(0) \in \mathbb{R}^2$ and every $\varepsilon > 0$, we have

$$\mathbb{P} \left[\sup_{0 \leq t < \infty} \|S(t) - S_\infty\| \geq \varepsilon \right] \leq \frac{1}{\varepsilon^2} \mathbb{E}[\|S(0) - S_\infty\|^2]$$

and $\mathbb{P} \left[\lim_{t \rightarrow \infty} S(t) \text{ exists} \right] = 1$. Thus, the dynamical system given by Equation (71) is globally stochastically semistable. □

Remark 1 If \tilde{A} is semistable, then there exists an invertible transformation matrix $T \in \mathbb{R}^{2 \times 2}$ such that $T\tilde{A}T^{-1} = \text{diag}[-\lambda, 0]$, where $\lambda \in \text{spec}(\tilde{A})$ and $\lambda > 0$. In this case, defining the new coordinates

$[\hat{S}_1(t), \hat{S}_2(t)]^T \triangleq TS(t)$, Equation (71) yields the two decoupled stochastic differential equations given by

$$d\hat{S}_1(t) = -\lambda\hat{S}_1(t)dt - \nu\lambda\hat{S}_1(t)dw(t), \quad \hat{S}_1(0) = \hat{S}_{10}, \quad t \geq 0, \tag{78}$$

$$d\hat{S}_2(t) = 0, \quad \hat{S}_2(0) = \hat{S}_{20}(0). \tag{79}$$

Since the analytical solution to Equation (78) is given by $\hat{S}_1(t) = \hat{S}_1(0)e^{-\lambda(1+\frac{1}{2}\lambda\nu^2)t-\nu\lambda w(t)}$, it follows that

$$S(t) = T^{-1}\hat{S}(t) = T^{-1} \begin{bmatrix} \hat{S}_1(0)e^{-\lambda(1+\frac{1}{2}\lambda\nu^2)t-\nu\lambda w(t)} \\ \hat{S}_2(0) \end{bmatrix}.$$

Finally, we provide a sufficient condition for stochastic multistability for the dynamical system given by Equation (71). For this result, the following lemma is first needed.

Lemma 3 Let $\tilde{A} \in \mathbb{R}^{n \times n}$. If there exist $n \times n$ matrices $P = P^T \geq 0$ and $R = R^T \geq 0$, and a nonnegative integer k such that

$$0 = (\tilde{A}^k)^T(\tilde{A}^T P + P\tilde{A} + R)\tilde{A}^k, \tag{80}$$

$$k = \min \left\{ l \in \mathbb{Z}_+ : \bigcap_{i=1}^n \mathcal{N}(R\tilde{A}^{i+l-1}) = \mathcal{N}(\tilde{A}) \right\}, \tag{81}$$

then (i) $\mathcal{N}(P\tilde{A}^k) \subseteq \mathcal{N}(\tilde{A}) \subseteq \mathcal{N}(R\tilde{A}^k)$ and (ii) $\mathcal{N}(\tilde{A}) \cap \mathcal{R}(\tilde{A}) = \{0\}$.

Proof. The proof is similar to that of Lemma 4.5 of [96] and, hence, is omitted. □

Theorem 7 Consider the dynamical system given by Equation(71). Suppose there exist 2×2 matrices $P = P^T \geq 0$ and $R = R^T \geq 0$, and a nonnegative integer k such that Equations (80) and (81) hold with $n = 2$. If $\mathcal{N}(A - L) \setminus \{(0, 0)\} \neq \emptyset$ and $|\nu|$ is sufficiently small, then the dynamical system given by Equation (71) is stochastically multistable.

Proof. By Proposition 3 it suffices to show that $\tilde{A} \triangleq A - L$ is semistable. Consider the deterministic dynamical system given by

$$\dot{x}(t) = \tilde{A}x(t), \quad x(0) = x_0, \quad t \geq 0, \tag{82}$$

where $x(t) \in \mathbb{R}^2$. Note that \tilde{A} is semistable if and only if Equation (82) is semistable [93], and hence, it suffices to show that Equation (82) is semistable. Since, by Lemma 3, $\mathcal{N}(\tilde{A}) \cap \mathcal{R}(\tilde{A}) = \{0\}$, it follows from ([97](p. 119)) that \tilde{A} is group invertible. Thus, let $L \triangleq I_2 - \tilde{A}\tilde{A}^\#$ and note that $L^2 = L$. Hence, L is the unique 2×2 matrix satisfying $\mathcal{N}(L) = \mathcal{R}(\tilde{A})$, $\mathcal{R}(L) = \mathcal{N}(\tilde{A})$, and $Lx = x$ for all $x \in \mathcal{N}(\tilde{A})$.

Next, consider the nonnegative function

$$\mathbb{V}(x) = x^T(\tilde{A}^k)^T P \tilde{A}^k x + x^T L^T L x.$$

If $\mathbb{V}(x) = 0$ for some $x \in \mathbb{R}^2$, then $P\tilde{A}^k x = 0$ and $Lx = 0$. Now, it follows from Lemma 3 that $x \in \mathcal{N}(\tilde{A})$, whereas $Lx = 0$ implies $x \in \mathcal{R}(\tilde{A})$, and hence, $\mathbb{V}(x) = 0$ only if $x = 0$. Hence, $\mathbb{V}(\cdot)$

is positive definite. Next, since $L\tilde{A} = \tilde{A} - \tilde{A}\tilde{A}^\# \tilde{A} = 0$, it follows that the time derivative along the trajectories of Equation (82) is given by

$$\begin{aligned} \dot{\mathbb{V}}(x(t)) &= -x^T(t)(\tilde{A}^k)^T R \tilde{A}^k x(t) + x^T(t) \tilde{A}^T L^T L x(t) + x^T(t) L^T L \tilde{A} x(t) \\ &= -x^T(t)(\tilde{A}^k)^T R \tilde{A}^k x(t) \\ &\leq 0, \quad t \geq 0. \end{aligned}$$

Note that $\dot{\mathbb{V}}^{-1}(0) = \mathcal{N}(R\tilde{A}^k)$.

To find the largest invariant set \mathcal{M} contained in $\mathcal{N}(R\tilde{A}^k)$, consider a solution $x(\cdot)$ of Equation (82) such that $R\tilde{A}^k x(t) = 0$ for all $t \geq 0$. Then, $R\tilde{A}^k \frac{d^{i-1}}{dt^{i-1}} x(t) = 0$ for every $i \in \{1, 2, \dots\}$ and $t \geq 0$, that is, $R\tilde{A}^k \tilde{A}^{i-1} x(t) = R\tilde{A}^{k+i-1} x(t) = 0$ for every $i \in \{1, 2, \dots\}$ and $t \geq 0$. Equation (81) now implies that $x(t) \in \mathcal{N}(\tilde{A})$ for all $t \geq 0$. Thus, $\mathcal{M} \subseteq \mathcal{N}(\tilde{A})$. However, $\mathcal{N}(\tilde{A})$ consists of only equilibrium points, and hence, is invariant. Hence, $\mathcal{M} = \mathcal{N}(\tilde{A})$.

Finally, let $x_e \in \mathcal{N}(\tilde{A})$ be an equilibrium point of Equation (82) and consider the Lyapunov function candidate $\mathbb{U}(x) = \mathbb{V}(x - x_e)$, which is positive definite with respect to x_e . Then it follows that the Lyapunov derivative along the trajectories of Equation (82) is given by

$$\dot{\mathbb{U}}(x(t)) = -(x(t) - x_e)^T (\tilde{A}^k)^T R \tilde{A}^k (x(t) - x_e) \leq 0, \quad t \geq 0.$$

Thus, it follows that x_e is Lyapunov stable. Now, it follows from Theorem 3.1 of [63] that Equation (82) is semistable, that is, A is semistable. Finally, it follows from Proposition 3 that Equation (71) is stochastically multistable. \square

Example 5 In this example, we illustrate the stochastic multistability properties of the two-state nonlinear synaptic drive neuronal firing model given by Equation (70). Specifically, consider the mean field synaptic drive model given by Equation (70) with $n_E \bar{A}^{EE} = 1 \text{ V}$, $n_I \bar{A}^{EI} = -1 \text{ V}$, $n_E \bar{A}^{IE} = 1 \text{ V}$, $n_I \bar{A}^{II} = 0 \text{ V}$, and $\lambda^E = 10 \text{ ms}$, and let λ^I vary. In this case, the system matrices in Equation (71) are given by

$$A = \begin{bmatrix} 1 & -1 \\ 1 & 0 \end{bmatrix}, \quad L = \begin{bmatrix} 0.1 & 0 \\ 0 & \frac{1}{\lambda^I} \end{bmatrix}, \quad A - L = \begin{bmatrix} 0.9 & -1 \\ 1 & -\frac{1}{\lambda^I} \end{bmatrix}.$$

Figure 9 shows the eigenvalues of $A - L$ as a function of λ^I .

Note that for $\lambda^I < 0.9 \text{ ms}$ or $\lambda^I > 0.93 \text{ ms}$, $(A - L)$ is unstable, whereas for $0.9 \text{ ms} < \lambda^I < 0.93 \text{ ms}$, $(A - L)$ is asymptotically stable. Clearly, $\text{rank}(A - L) < 2$ for $\lambda^I = 0.9 \text{ ms}$. Hence, it follows from Theorem 7 that the stochastic dynamical system given by Equation (71) exhibits multistability for $\lambda^I = 0.9 \text{ ms}$. In this case, $A - L$ is semistable and the $\mathcal{N}(A - L)$ is characterized by the direction vector $[1, 0.9]^T$.

For our simulation, we use the initial condition $S(0) = [0.1, 0.5]^T$. Figures 10 and 11 show the time response for the average excitatory and inhibitory synaptic drives, and the phase portrait for $\lambda^I = 0.9 \text{ ms}$ with $\nu = 0.2$. Furthermore, for $\lambda^I = 0.9 \text{ ms}$ with $\nu = 0.2$, Figure 12 shows a histogram of the limit points of $\log_e S^E$ (or, equivalently, $\log_e 0.9S^I$) over 10,000 samples. Note that the mean and variance of $\log_e S^E$ is -1.6135 and 0.3152 , respectively. Similar plots shown in Figures 10 and 11 are shown for $\lambda^I = 0.78 \text{ ms}$ and $\lambda^I = 1.20 \text{ ms}$ with $\nu = 0.2$ in Figures 13–16. Finally, Figures 17 and 18 show similar simulations for the case where $\lambda^I = 0.9 \text{ ms}$ (i.e., $(A - L)$ is semistable) and $\nu = 1$. However, note that

in this case the condition given by Equation (77) of Proposition 3 is not satisfied, and hence, the model exhibits instability.

Figure 9. Eigenvalues of $A - L$ as a function of λ^I . Arrows indicate increasing values of λ^I .

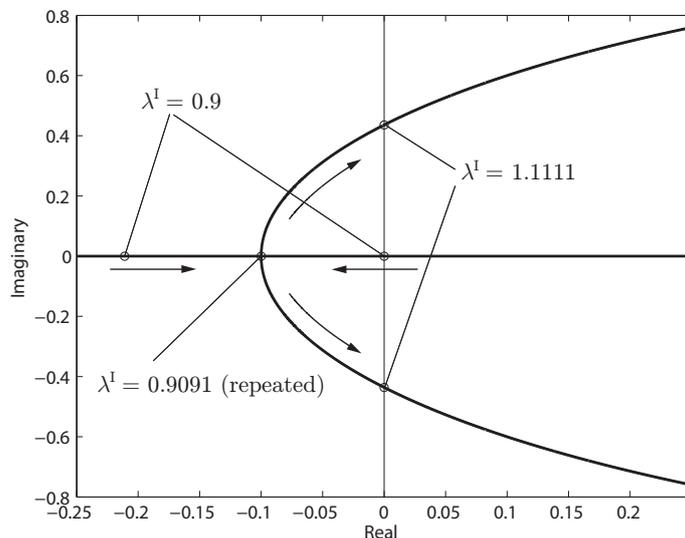


Figure 10. State trajectories of the sample trajectories of Equation (70) for $\lambda^I = 0.9$ with $\nu = 0.2$.

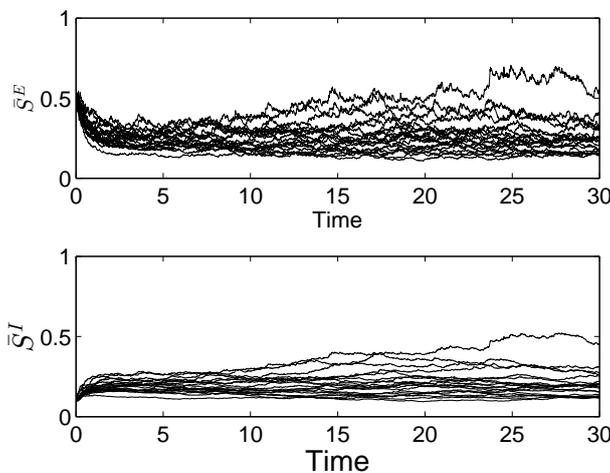


Figure 11. Phase portrait of the sample trajectories of Equation (70) for $\lambda^I = 0.9$ with $\nu = 0.2$.

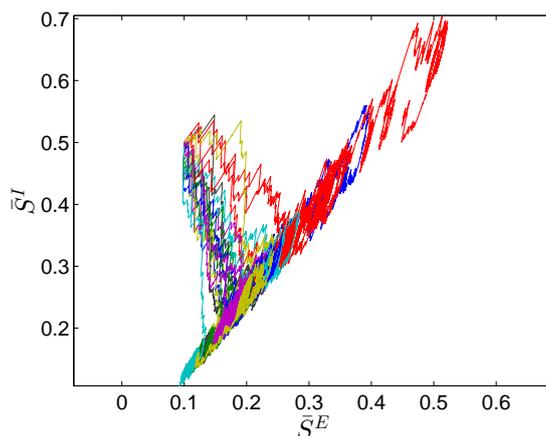


Figure 12. Histogram showing the limit points of $\log_e S^E$ over 10,000 samples.

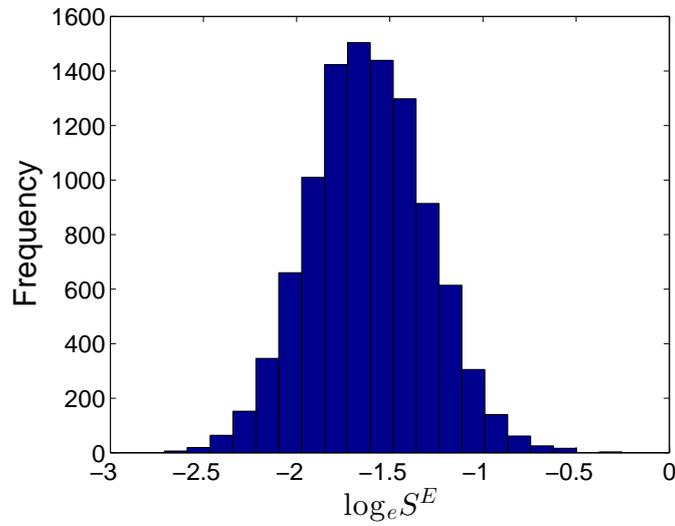


Figure 13. State trajectories of the sample trajectories of Equation (70) for $\lambda^I = 0.78$ with $\nu = 0.2$.

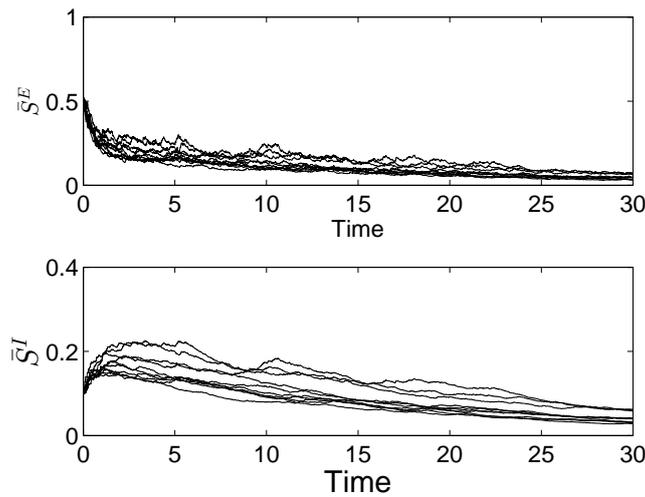


Figure 14. Phase portrait of the sample trajectories of Equation (70) for $\lambda^I = 0.78$ with $\nu = 0.2$.

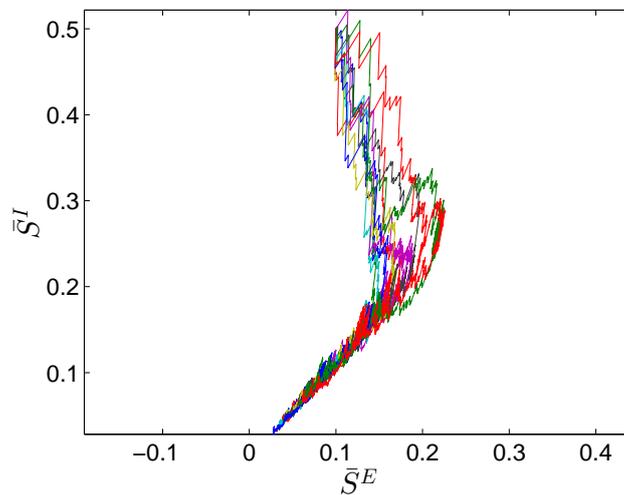


Figure 15. State trajectories of the sample trajectories of Equation (70) for $\lambda^I = 1.20$ with $\nu = 0.2$.

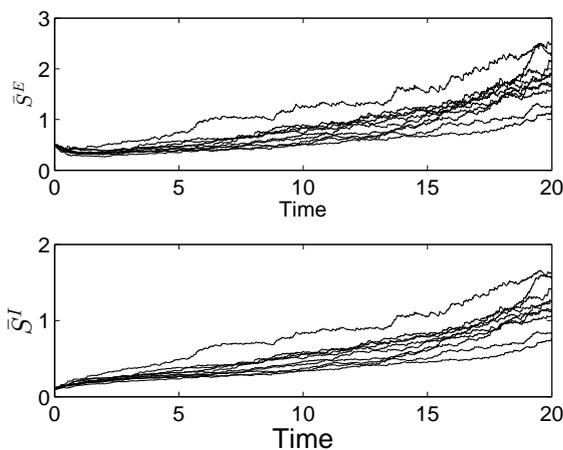


Figure 16. Phase portrait of the sample trajectories of Equation (70) for $\lambda^I = 1.20$ with $\nu = 0.2$.

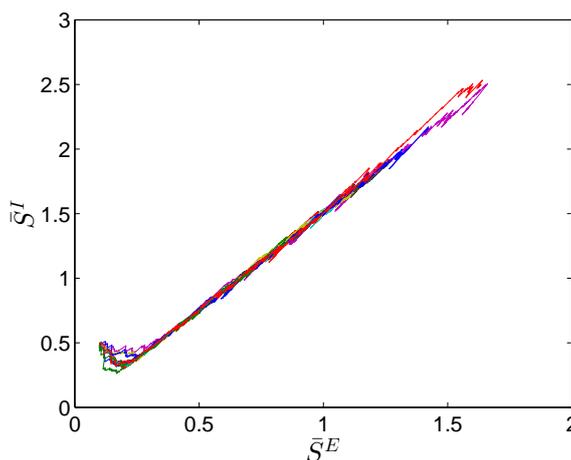
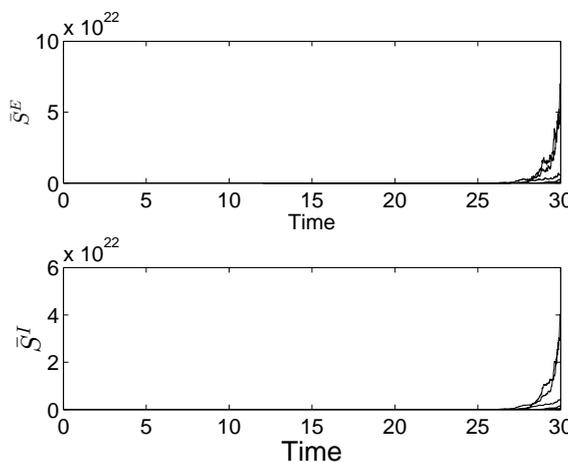


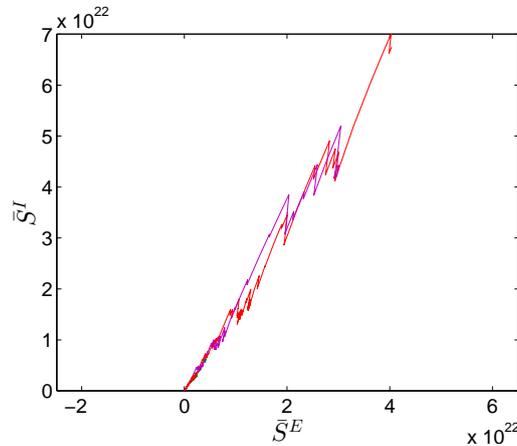
Figure 17. State trajectories of the sample trajectories of Equation (70) for $\lambda^I = 0.9$ with $\nu = 1$.



The trajectories of Equations (24) and (25) can exhibit unstable and multistable behaviors for different values of the parameters, which are similar to the simulation results for Equation (70).

Moreover, its averaging dynamics will be analogous to the results of the deterministic model given in [98].

Figure 18. Phase portrait of the sample trajectories of Equation (70) for $\lambda^1 = 0.9$ with $\nu = 1$.



9. A Synaptic Drive Firing Model with Time-Varying Delays and Stochastic Multiplicative Uncertainty

In this section, we extend the synaptic drive model developed in Section 2 to investigate the conditions that would lead to synchronization or neutral oscillators. In particular, we extend the biological neural network model of Section 2 to include time-varying delays and stochastic input uncertainty. The system uncertainty model involves a Markov process wherein stochastic integration is interpreted in the sense of Itô.

For the statement of the results of this section, we require some additional notation and definitions. Specifically, $\mathcal{C}([-\tau, 0], \mathbb{R}^n)$ with $\tau > 0$ denotes a Banach space of continuous vector-valued functions mapping the interval $[-\tau, 0]$ into \mathbb{R}^n with topology of uniform convergence and designated operator norm given by $\|\psi\| = \sup_{-\tau \leq \theta \leq 0} \|\psi(\theta)\|$ for $\psi \in \mathcal{C}([-\tau, 0], \mathbb{R}^n)$. Furthermore, let $S_t \in \mathcal{C}((-\infty, +\infty), \mathbb{R}^n)$ defined by $S_t(\theta) \triangleq S(t + \theta), \theta \in (-\infty, 0], t \geq 0$, denote an (infinite dimensional) state at time t corresponding to the *piece of trajectories* S between $-\infty$ and t , and assume $v_{thi}(t) \equiv 0$. To capture communication delays in our biological neural network model given by Equation (9), define $S(t) \triangleq [S_1(t), S_2(t), \dots, S_n(t)]^T, f(S) \triangleq [f_1(S_1), f_2(S_2), \dots, f_n(S_n)]^T$, where $f_i(\cdot)$ is defined by Equation (11) or Equation (12), $L \triangleq \text{diag}[\frac{1}{\tau_1}, \frac{1}{\tau_2}, \dots, \frac{1}{\tau_n}]$, and $B \triangleq \text{diag}[B_1, B_2, \dots, B_n]$. Furthermore, define

$$\hat{S}(t) \triangleq \begin{bmatrix} \sum_{j=2}^n A_{1j} S_j(t - \delta_{1j}(t)) \\ 0 \\ \vdots \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ \sum_{j=1, j \neq 2}^n A_{2j} S_j(t - \delta_{2j}(t)) \\ \vdots \\ 0 \end{bmatrix} + \dots + \begin{bmatrix} 0 \\ \vdots \\ 0 \\ \sum_{j=1}^{n-1} A_{nj} S_j(t - \delta_{nj}(t)) \end{bmatrix}, \tag{83}$$

where $\delta_{ij}(t)$ denotes the continuous, time-varying time delay of the transmission signal from the j th neuron to the i th neuron at time t , $\delta_{ij}(t) \geq 0$, $t \geq 0$, and $S_j(t)$ denotes the j th component of $S(t)$. The system delays $\delta_{ij}(t)$ correspond to the times of the spike hitting the synapse and t is the time after the spike, and hence, these delays account for the distance traveled by the voltage spikes down the axon.

We modify the biological neural network system given by Equation (9) to include the effects of stochastic perturbations as well as time delays. Specifically, we consider the model

$$dS(t) = (-LS(t) + Bf(\hat{S}(t)))dt + \sigma(S(t))dw(t), \quad S(\theta) = \phi(\theta), \quad -\infty < \theta \leq 0, \quad t \geq 0, \quad (84)$$

where $\phi(\cdot) \in \mathcal{C} \triangleq \mathcal{C}((-\infty, 0], \mathbb{R}^n)$ is a continuous vector-valued function specifying the initial state of the system given by Equation (84), $w(t) = [w_1(t), w_2(t), \dots, w_n(t)]^T$ captures noise in the input voltage and is represented by Brownian motion, that is, an n -dimensional mutually independent standard Wiener process, and $\sigma(S) = \text{diag}[\sigma_1(S), \sigma_2(S), \dots, \sigma_n(S)]$ represents the state-dependent noise intensity matrix for the Gaussian white noise process $dw(t)$. Henceforth, we consider Equation (84) as the model of the perturbed biological neural network.

Next, since $\hat{S}(t)$ defined by Equation (83) contains $n(n-1)$ terms with different time delays, each term can be written as the product of an $n \times n$ -dimensional matrix and an n -dimensional vector. Specifically, for $i' = 1, 2, \dots, n, j = 1, 2, \dots, n, i' \neq j$, define $i \triangleq i'(n-1)+j, i' > j$, and $i \triangleq i'(n-1)+j-1, i' < j$, where $i = 1, 2, \dots, n(n-1)$, define $\delta_i(t) \triangleq \delta_{i'j}(t)$, and define the matrix $A_i \in \mathbb{R}^{n \times n}$ whose (i', j) th entry is $A_{i'j}$ and all the other entries are 0. Thus, the i th term in Equation (83) can be replaced by $A_i S(t - \delta_i(t)), i \in \{1, 2, \dots, n(n-1)\}$. Hence, setting $N = n(n-1), \hat{S}(t)$ can be written as

$$\hat{S}(t) = \sum_{i=1}^N A_i S(t - \delta_i(t)). \quad (85)$$

For the statement of the results in this section, we define the infinitesimal operator $\mathcal{L} : [0, \infty) \times \mathcal{C}((-\infty, 0], \mathbb{R}^n) \rightarrow \mathbb{R}$ associated with the stochastic process given by Equation (84), acting on the functional $V : \mathbb{R} \times \mathcal{C} \rightarrow \mathbb{R}$, by

$$\mathcal{L}V(t, S_t) \triangleq \limsup_{h \rightarrow 0^+} \frac{\mathbb{E}[V(t+h, S_{t+h})|S_t] - V(t, S_t)}{h}. \quad (86)$$

For a two-times continuously differentiable function $V : [0, \infty) \times \mathbb{R}^n \rightarrow \mathbb{R}$ of the random variable S , the infinitesimal operator $\mathcal{L}V(t, S)$ is defined as [90]

$$\begin{aligned} \mathcal{L}V(t, S) &\triangleq \lim_{h \rightarrow 0^+} \frac{\mathbb{E}[V(t+h, S(t+h))] - V(t, S)}{h} \\ &= \frac{\partial V(t, S)}{\partial t} + V'(t, S)(-LS + Bf(\hat{S})) + \frac{1}{2}\sigma^T(S)V''(t, S)\sigma(S), \end{aligned} \quad (87)$$

where $V'(t, S)$ denotes the Fréchet derivative of V and $V''(t, S)$ denotes the Hessian matrix of V with respect to S at (t, S) . The following lemma provides an explicit formula for the infinitesimal operator on two kinds of functionals using the ideas from Lemma 3.1 of [99].

Lemma 4 Consider the biological neural network given by Equation (84) and let

$$V_1(t, \psi) = \int_{-d(t)}^0 \psi^T(\theta)H\psi(\theta)d\theta, \tag{88}$$

$$V_2(t, \psi) = \int_{-d(t)}^0 e^{\varepsilon(t+\theta)}\psi^T(\theta)H\psi(\theta)d\theta, \tag{89}$$

where $t \geq 0$, $\psi \in \mathcal{C}((-\infty, 0], \mathbb{R}^n)$, $\varepsilon > 0$, $H \in \mathbb{R}^{n \times n}$, $d : \mathbb{R} \rightarrow \mathbb{R}$ is differentiable, and $d(t) \geq 0$, $t \geq 0$. Then, the infinitesimal operator acting on $V_1 : [0, \infty) \times \mathcal{C} \rightarrow \mathbb{R}$ and $V_2 : [0, \infty) \times \mathcal{C} \rightarrow \mathbb{R}$ are given by

$$\mathcal{L}V_1(t, S_t) = S^T(t)HS(t) - (1 - \dot{d}(t))S^T(t - d(t))HS(t - d(t)), \tag{90}$$

$$\mathcal{L}V_2(t, S_t) = e^{\varepsilon t}S^T(t)HS(t) - e^{\varepsilon(t-d(t))}(1 - \dot{d}(t))S^T(t - d(t))HS(t - d(t)). \tag{91}$$

Proof. For sufficiently small $h > 0$,

$$\begin{aligned} \mathbb{E}[V_1(t + h, S_{t+h})|S_t] &= \mathbb{E} \left[\int_{-d(t+h)}^0 S^T(t + h + \theta)HS(t + h + \theta)d\theta \middle| S_t \right] \\ &= \mathbb{E} \left[\int_{h-d(t+h)}^h S^T(t + \theta)HS(t + \theta)d\theta \middle| S_t \right] \\ &= \mathbb{E} \left[\int_0^h S^T(t + \theta)HS(t + \theta)d\theta \middle| S_t \right] \\ &\quad + \int_{-d(t)}^0 S^T(t + \theta)HS(t + \theta)d\theta \\ &\quad + \mathbb{E} \left[\int_{-d(t+h)}^{-d(t)} S^T(t + \theta)HS(t + \theta)d\theta \middle| S_t \right] \\ &\quad + \mathbb{E} \left[\int_{h-d(t+h)}^{-d(t+h)} S^T(t + \theta)HS(t + \theta)d\theta \middle| S_t \right] \\ &= \mathbb{E} \left[\int_0^h S^T(t + \theta)HS(t + \theta)d\theta \middle| S_t \right] \\ &\quad + \int_{-d(t)}^0 S^T(t + \theta)HS(t + \theta)d\theta \\ &\quad - \mathbb{E} \left[\int_{-d(t)}^{-d(t)+h\dot{d}(t)+O(h)} S^T(t + \theta)HS(t + \theta)d\theta \middle| S_t \right] \\ &\quad - \mathbb{E} \left[\int_{-d(t+h)}^{h-d(t+h)} S^T(t + \theta)HS(t + \theta)d\theta \middle| S_t \right] \\ &= hS^T(t)HS(t) + V_1(t, S_t) \\ &\quad - h(1 - \dot{d}(t))S^T(t - d(t))HS(t - d(t)) + O(h), \quad t \geq 0, \tag{92} \end{aligned}$$

where $O(h)$ denotes higher-order terms in h . Substituting Equation (92) into Equation (86) yields Equation (90). The proof of Equation (91) is similar to the proof of Equation (90) and, hence, is omitted. \square

To develop a global synchronization property for the biological neural network system given by Equation (84), we introduce the notion of stochastic synchronization. Here, we focus on mean-square synchronization.

Definition 12 The biological neural network given by Equation (84) is said to be globally asymptotically mean-square synchronized if

$$\lim_{t \rightarrow \infty} \mathbb{E}[\|S_{it} - S_{jt}\|^2] = 0 \tag{93}$$

for all $\phi(\cdot) \in \mathcal{C}((-\infty, 0], \mathbb{R}^n)$ and $i, j = 1, 2, \dots, n, i \neq j$, where $S_{it} \triangleq S_i(t + \theta), \theta \in (-\infty, 0], t \geq 0$, and $\|S_{it} - S_{jt}\| = \sup_{-\tau \leq \theta \leq 0} |S_i(t + \theta) - S_j(t + \theta)|, \tau > 0$.

Definition 13 The biological neural network given by Equation (84) is said to be globally exponentially mean-square synchronized if there exist constants $\rho > 0$ and $p > 0$ such that

$$\mathbb{E}[\|S_{it} - S_{jt}\|^2] \leq \rho e^{-pt} \int_{-\infty}^0 |\phi_i(\theta) - \phi_j(\theta)|^2 d\theta, \quad t \geq 0, \quad p > 0, \tag{94}$$

for all $\phi(\cdot) = [\phi_1(\cdot), \dots, \phi_n(\cdot)]^T \in \mathcal{C}((-\infty, 0], \mathbb{R}^n)$ and $i, j = 1, 2, \dots, n, i \neq j$.

10. Synchronization of Stochastic Biological Neural Networks

In this section, we develop sufficient conditions for global mean-square synchronization for the biological neural network given by Equation (84) with differentiable time delays using Barbalat’s lemma and linear matrix inequalities (LMIs). Here, we assume that the noise intensity matrix function $\sigma(S)$ has a linear growth rate; that is, there exists $r > 0$ such that

$$\text{tr}[\sigma^2(S)] \leq r S^T M^T M S, \quad S \in \mathbb{R}^n, \tag{95}$$

where M is defined in Equation (60).

The following theorem provides sufficient conditions for global mean-square asymptotic synchronization of the biological neural network system given by Equation (84).

Theorem 8 Consider the biological neural network given by Equation (84) with $f_i(\cdot), i = 1, 2, \dots, n$, given by either Equation (11) or Equation (12), and assume that $\delta_i(t) \leq h_1 < 1$, and $\delta_i(t) \geq 0, t \geq 0, i = 1, 2, \dots, N$, hold. If there exist a positive-definite matrix $P \in \mathbb{R}^{n \times n}$, nonnegative definite matrices $Q_i \in \mathbb{R}^{n \times n}, i = 1, 2, \dots, N$, and $R \in \mathbb{R}^{n \times n}$, and a nonnegative-definite diagonal matrix $\Lambda \in \mathbb{R}^{n \times n}$ such that

$$\begin{bmatrix} R & -PB \\ -BP & \Lambda \end{bmatrix} \geq 0, \tag{96}$$

$$\Omega_7 \triangleq \begin{bmatrix} -(1 - h_1)Q_1 + A_1^T \Lambda A_1 & A_1^T \Lambda A_2 & \dots & A_1^T \Lambda A_N \\ A_2^T \Lambda A_1 & -(1 - h_1)Q_2 + A_2^T \Lambda A_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & A_{N-1}^T \Lambda A_N \\ A_N^T \Lambda A_1 & \dots & A_N^T \Lambda A_{N-1} & -(1 - h_1)Q_N + A_N^T \Lambda A_N \end{bmatrix} \leq 0, \tag{97}$$

and either $\Omega_8 < 0$ or both $\Omega_8 \leq 0$ and $\mathcal{N}(\Omega_8) = \text{span}(\mathbf{e}_n)$ hold, where

$$\Omega_8 \triangleq -PL - LP + k_1 r M^T M + R + \sum_{i=1}^N Q_i, \tag{98}$$

$k_1 \triangleq \lambda_{\max}(P)$, r is such that Equation (95) holds, M is given by Equation (60), and $A_i, i = 1, \dots, N$, is defined in Equation (85), then Equation (84) is globally asymptotically mean-square synchronized.

Proof. Consider the functional $V : [0, \infty) \times \mathcal{C} \rightarrow \mathbb{R}$ given by $V(t, \psi) = V_1(\psi(0)) + V_2(t, \psi)$, where $V_1(\psi(0)) = \psi^T(0)P\psi(0)$ and $V_2(t, \psi) = \sum_{i=1}^N \int_{-\delta_i(t)}^0 \psi^T(\theta)Q_i\psi(\theta)d\theta$. It follows from Equation (87) and Lemma 4 that the infinitesimal operator $\mathcal{L}V(t, S_t)$ associated with the stochastic process given by Equation (84) is given by

$$\mathcal{L}V(t, S_t) = \mathcal{L}V_1(S(t)) + \mathcal{L}V_2(t, S_t), \quad t \geq 0, \tag{99}$$

where

$$\mathcal{L}V_1(S(t)) = 2S^T(t)P(-LS(t) + Bf(\hat{S}(t))) + \text{tr}[\sigma(S(t))P\sigma(S(t))], \tag{100}$$

$$\mathcal{L}V_2(t, S_t) = \sum_{i=1}^N \left[S^T(t)Q_iS(t) - (1 - \dot{\delta}_i(t))S^T(t - \delta_i(t))Q_iS(t - \delta_i(t)) \right], \tag{101}$$

and $\hat{S}(t) \in \mathbb{R}^n$ is defined by Equation (85). Next, since $\mathbb{E}[V(t, S_t)] = V(0, S_0) + \mathbb{E}[\int_0^t \mathcal{L}V(u, S_u)du]$, it follows that $\mathbb{E}[dV(t, S_t)] = \mathbb{E}[\mathcal{L}V_1(S(t)) + \mathcal{L}V_2(t, S_t)]dt$.

To complete the proof, we show that $\mathcal{L}V_1(S(t)) + \mathcal{L}V_2(t, S_t) \leq 0, t \geq 0$, and $\mathcal{L}V_1(S(t)) + \mathcal{L}V_2(t, S_t) \equiv 0$ implies $MS(t) \equiv 0$. To see this, note that Equation (95) implies

$$\text{tr}[\sigma(S)P\sigma(S)] \leq k_1\text{tr}[\sigma(S)\sigma(S)] \leq k_1rS^T M^T MS, \quad S \in \mathbb{R}^n. \tag{102}$$

Furthermore, note that for every diagonal matrix $\Lambda \in \mathbb{R}^{n \times n}$ such that $\Lambda \geq 0$, it follows that for $f_i(\cdot), i = 1, \dots, n$, given by Equation (11) or Equation (12),

$$f^T(\hat{S}(t))\Lambda f(\hat{S}(t)) \leq \hat{S}^T(t)\Lambda\hat{S}(t), \quad t \geq 0. \tag{103}$$

Now, using Equations (96), (102), and (103), it follows from Equation (100) that

$$\begin{aligned} \mathcal{L}V_1(S(t)) &\leq 2S^T(t)P(-LS(t) + Bf(\hat{S}(t))) + k_1rS^T(t)M^T MS(t) \\ &\leq -2S^T(t)PLS(t) + S^T(t)RS(t) + f^T(\hat{S}(t))\Lambda f(S(t)) + k_1rS^T(t)M^T MS(t) \\ &\leq S^T(t)(-2PL + k_1rM^T M + R)S(t) + \hat{S}^T(t)\Lambda\hat{S}(t), \quad t \geq 0. \end{aligned} \tag{104}$$

Hence, since $\dot{\delta}_i(t) \leq h_1 < 1, t > 0$, it follows from Equations (101) and (104) that

$$\begin{aligned} \mathcal{L}V_1(S(t)) + \mathcal{L}V_2(t, S_t) &\leq S^T(t) \left[-2PL + k_1rM^T M + R + \sum_{i=1}^N Q_i \right] S(t) \\ &\quad + \sum_{i=1}^N \sum_{j=1}^N S^T(t - \delta_i(t))A_i^T \Lambda A_j S(t - \delta_j(t)) \\ &\quad - \sum_{i=1}^N (1 - h_1)S^T(t - \delta_i(t))Q_i S(t - \delta_i(t)) \\ &= \eta^T(t)\Omega_7\eta(t) + S^T(t)\Omega_8S(t) \\ &\leq S^T(t)\Omega_8S(t), \quad t \geq 0, \end{aligned}$$

where $\eta(t) \triangleq [S^T(t - \delta_1(t)), \dots, S^T(t - \delta_N(t))]^T$.

and, hence, is omitted. □

The following corollary to Theorem 9 is immediate.

Corollary 1 Consider the biological neural network given by Equation (84) with $f_i(\cdot), i = 1, 2, \dots, n$, given by either Equation (11) or Equation (12), and assume that $\dot{\delta}_i(t) \leq h_1 < 1$, and $h_2 \geq \delta_i(t) \geq 0, t \geq 0, i = 1, 2, \dots, N$, hold. If there exist a positive-definite matrix $P \in \mathbb{R}^{n \times n}$, nonnegative definite matrices $Q_i \in \mathbb{R}^{n \times n}, i = 1, 2, \dots, N$, and $R \in \mathbb{R}^{n \times n}$, and a nonnegative-definite diagonal matrix $\Lambda \in \mathbb{R}^{n \times n}$ such that Equation (96) holds, and $\Omega_7 < 0$ and $\Omega_8 < 0$, where Ω_7 and Ω_8 are given by Equations (97) and (98) with $k_1 \triangleq \lambda_{\max}(P)$, r is such that Equation (95) holds, M is given by Equation (60), $A_i, i = 1, \dots, N$, is defined in Equation (85), then Equation (84) is globally exponentially mean-square synchronized.

Proof. The result is a direct consequence of Theorem 9 by noting that if $\Omega_7 < 0$ and $\Omega_8 < 0$ hold, then there exists a sufficiently small $\varepsilon > 0$ such that $\Omega_9 \leq 0$ and $\Omega_{10} < 0$ hold, where Ω_9 and Ω_{10} are given by Equations (105) and (106). □

Remark 2 Note that Theorem 8 does not require that the time delays be bounded, whereas Theorem 9 and Corollary 1 hold for the case where the time delays are bounded.

Remark 3 It is important to note that if $f_i(\cdot), i = 1, 2, \dots, n$, in Equation (84) is replaced by Equation (10), then the results of Theorems 8 and 9 as well as Corollary 1 still hold.

Example 6 Consider the stochastic network system characterized by

$$dS_1(t) = (-S_1(t) + 0.2f_1(0.3S_2(t - \delta_1(t)) - 0.5S_3(t - \delta_2(t))))dt + 0.1(S_1(t) - S_2(t))dw_1(t), \quad S_1(\theta) = 2 + \sin \theta, \quad (107)$$

$$dS_2(t) = (-1.1S_2(t) + 0.3f_2(0.4S_1(t - \delta_3(t)) - 0.3S_3(t - \delta_4(t))))dt + 0.1(S_2(t) - S_3(t))dw_2(t), \quad S_2(\theta) = -3 + \cos \theta, \quad (108)$$

$$dS_3(t) = (-1.4S_3(t) + 0.5f_3(0.4S_1(t - \delta_5(t)) + 0.3S_2(t - \delta_6(t))))dt + 0.1(S_3(t) - S_1(t))dw_3(t), \quad S_3(\theta) = 1 - \theta, \quad (109)$$

where $\theta \in [-1, 0], \delta_1(t) = 1 + 0.1 \sin t, \delta_2(t) = 1 + 0.1t, \delta_3(t) = 0.5, \delta_4(t) = 0.1t, \delta_5(t) = 0.3, \delta_6(t) = 0.4, t \geq 0, f_i(\cdot), i = 1, 2, 3$, are defined by either Equation (11) or Equation (12), and $dw_i, i = 1, 2, 3$, are standard Gaussian white noise processes.

Using the MATLAB LMI Toolbox[®], it can be shown that

$$P = \begin{bmatrix} 205 & -2.69 & 0.13 \\ -2.69 & 169 & -1.96 \\ 0.13 & -1.96 & 122 \end{bmatrix}, \quad R = \begin{bmatrix} 101 & -1.85 & 0.12 \\ -1.85 & 107 & -2.00 \\ 0.12 & -2.00 & 118 \end{bmatrix}, \quad \Lambda = \begin{bmatrix} 81.2 & 0 & 0 \\ 0 & 114 & 0 \\ 0 & 0 & 144 \end{bmatrix},$$

$$Q_1 = \begin{bmatrix} 34.3 & 0.37 & 0.02 \\ 0.37 & 49.7 & 0.49 \\ 0.02 & 0.49 & 25.8 \end{bmatrix}, \quad Q_2 = \begin{bmatrix} 34.3 & 0.33 & 0.02 \\ 0.33 & 30.7 & 0.52 \\ 0.02 & 0.52 & 54.0 \end{bmatrix}, \quad Q_3 = \begin{bmatrix} 66.3 & 0.37 & 0.03 \\ 0.37 & 30.7 & 0.46 \\ 0.03 & 0.46 & 25.8 \end{bmatrix},$$

$$Q_4 = \begin{bmatrix} 34.3 & 0.33 & 0.02 \\ 0.3 & 30.7 & 0.52 \\ 0.02 & 0.52 & 45.9 \end{bmatrix}, \quad Q_5 = \begin{bmatrix} 76.9 & 0.38 & 0.03 \\ 0.38 & 30.7 & 0.46 \\ 0.03 & 0.46 & 25.8 \end{bmatrix}, \quad Q_6 = \begin{bmatrix} 34.3 & 0.39 & 0.02 \\ 0.39 & 60.8 & 0.52 \\ 0.02 & 0.52 & 25.8 \end{bmatrix},$$

satisfy Equations (96)–(98), with $r = 0.03$ and $A_i, i \in 1, 2, \dots, 6$, defined as in Equation (85), and hence, the conditions of Theorem 8 are satisfied. Next, define the synchronization error $e(t) \triangleq [(S_1(t) - S_2(t))^2 + (S_2(t) - S_3(t))^2]^{\frac{1}{2}}$. The trajectories of the state variables and the synchronization error with respect to time are shown in Figures 19 and 20, respectively. Note that even though some of the delays in this example are not bounded, that is, $\delta_i(t) \rightarrow \infty$ as $t \rightarrow \infty$ for $i \in \{2, 4\}$, the system is globally mean-square asymptotically synchronized.

Figure 19. State trajectories versus time for Example 6.

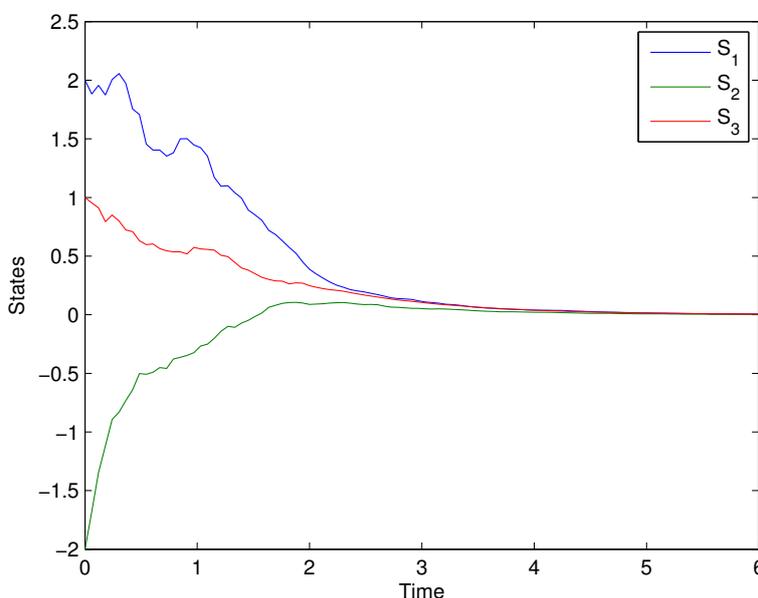
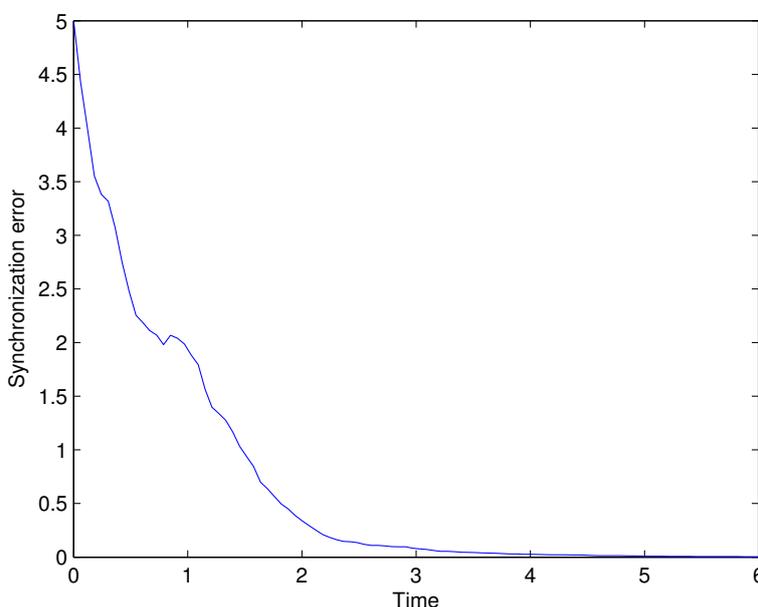


Figure 20. Synchronization error versus time for Example 6.



11. Thermodynamics, Neuroscience, Consciousness, and the Entropic Arrow of Time

In this and the next section, we present some qualitative insights from the fields of thermodynamics and electromagnetic field theory that can potentially be useful in developing mechanistic models [100] that can explain the underlying mechanisms of action for general anesthesia and consciousness. Specifically, by merging the two universalisms of thermodynamics and dynamical systems theory with neuroscience, one can provide key insights into the theoretical foundation for understanding the network properties of the brain by rigorously addressing large-scale interconnected biological neuronal network models that govern the neuroelectric behavior of biological excitatory and inhibitory neuronal networks. In addition, electrical signals in the brain can generate electromagnetic fields that can cause a shielding effect between the thalamus and frontal cortex, which in turn can lead to the emergence of unconsciousness. Both these paradigms are the subject of ongoing research.

In current clinical practice of general anesthesia, potent drugs are administered which profoundly influence levels of consciousness and vital respiratory (ventilation and oxygenation) and cardiovascular (heart rate, blood pressure, and cardiac output) functions. These variation patterns of the physiologic parameters (*i.e.*, ventilation, oxygenation, heart rate variability, blood pressure, and cardiac output) and their alteration with levels of consciousness can provide scale-invariant fractal temporal structures to characterize the degree of consciousness [101] in sedated patients. Here, we hypothesize that the degree of consciousness reflects the adaptability of the central nervous system and is proportional to the maximum work output under a fully conscious state divided by the work output of a given anesthetized state. A reduction in maximum work output (and cerebral oxygen consumption) or elevation in the anesthetized work output (or cerebral oxygen consumption) will thus reduce the degree of consciousness. Hence, the fractal nature (*i.e.*, complexity) of conscious variability is a self-organizing emergent property of the large-scale interconnected biological neuronal network since it enables the central nervous system to maximize entropy production and dissipate energy gradients. Within the context of aging and complexity in acute illnesses, variation of physiologic parameters and their relationship to system complexity and system thermodynamics have been explored in [102–107].

Complex dynamical systems involving self-organizing components forming spatio-temporal evolving structures that exhibit a hierarchy of emergent system properties form the underpinning of the central nervous system. These complex dynamical systems are ubiquitous in nature and are not limited to the central nervous system. Such systems include, for example, biological systems, immune systems, ecological systems, quantum particle systems, chemical reaction systems, economic systems, cellular systems, and galaxies, to cite but a few examples. The connection between the local subsystem interactions and the globally complex system behavior is often elusive. These systems are known as dissipative systems [108,109] and consume energy and matter while maintaining their stable structure by dissipating entropy to the environment.

In the central nervous system billions of neurons interact to form self-organizing dissipative nonequilibrium structures [108–110]. The fundamental common phenomenon among these systems are that they evolve in accordance to the laws of (nonequilibrium) thermodynamics, which are among the most firmly established laws of nature. System thermodynamics, in the sense of [110], involves open interconnected dynamical systems that exchange matter and energy with their environment in

accordance with the first law (conservation of energy) and the second law (nonconservation of entropy) of thermodynamics. Self-organization can spontaneously occur in such systems by invoking the two fundamental axioms of the science of heat. Namely, (i) if the energies in the connected subsystems of an interconnected system are equal, then energy exchange between these subsystems is not possible, and (ii) energy flows from more energetic subsystems to less energetic subsystems. These axioms establish the existence of a system entropy function as well as equipartition of energy [110–112] in system thermodynamics and synchronization [73] in neuronal networks; an emergent behavior in thermodynamic systems as well as neuroscience. Hence, in complex interconnected dynamical systems, self-organization is not a property of the systems parts but rather emerges as a result of the nonlinear subsystem interactions.

In recent research the authors in [110–114] combined the two universalisms of thermodynamics and dynamical systems theory under a single umbrella to develop a dynamical system formalism for classical thermodynamics so as to harmonize it with classical mechanics. While it seems impossible to reduce thermodynamics to a mechanistic world picture due to microscopic reversibility and Poincaré recurrence [110,115], the system thermodynamic formulation in [110,112] provides a harmonization of classical thermodynamics with classical mechanics. In particular, our dynamical system formalism captures all of the key aspects of thermodynamics, including its fundamental laws, while providing a mathematically rigorous formulation for thermodynamical systems out of equilibrium by unifying the theory of heat transfer with that of classical thermodynamics. In addition, the concept of entropy for a nonequilibrium state of a dynamical process is defined, and its global existence and uniqueness is established. This state space formalism of thermodynamics shows that the behavior of heat, as described by the conservation equations of thermal transport and as described by classical thermodynamics, can be derived from the same basic principles and is part of the same scientific discipline. Connections between irreversibility, the second law of thermodynamics, and the entropic arrow of time are also established in [110,112,113].

Building on the results of this paper, we propose to merge system thermodynamics with neuroscience to develop a theoretical foundation for the mechanisms of action of general anesthesia using the network properties of the brain by rigorously addressing the large-scale interconnected biological neuronal network model given by Equations (7) and (8). Even though simplified mean field models of the form given by Equations (24) and (25) have been extensively used in mathematical neuroscience literature to describe large neural populations, the complex large-scale interconnected system given by Equations (7) and (8) is essential in indentifying the mechanisms of action for general anesthesia. We postulate that unconsciousness is associated with reduced physiologic parameter variability, reflecting the inability of the central nervous system to adapt. The degree of consciousness is a function of the numerous coupling in the network properties of the brain that form a complex large-scale, interconnected system. Complexity here refers to the quality of a system wherein interacting subsystems self-organize to form hierarchical evolving structures exhibiting emergent system properties, and hence, a complex dynamical system is a system that is greater than the sum of its subsystems or parts. This complex system—involving numerous nonlinear dynamical subsystem interactions making up the system—has inherent emergent properties that depend on the integrity of the entire dynamical system and not merely a mean field simplified reduced-order model.

As in thermodynamics, neuroscience is a theory of large-scale systems wherein graph theory [116] can be used in capturing the (possibly dynamic) connectivity properties of network interconnections, with neurons represented by nodes, synapses represented by edges or arcs, and synaptic efficacy captured by edge weighting giving rise to a weighted adjacency matrix governing the underlying directed dynamic graph network topology. However, unlike thermodynamics, wherein energy spontaneously flows from a state of higher temperature to a state of lower temperature, neuron membrane potential variations occur due to ion species exchanges which evolve from regions of higher chemical potentials to regions of lower chemical potentials (*i.e.*, Gibbs' chemical potential [114]). And this evolution does not occur spontaneously but rather requires a hierarchical (*i.e.*, hybrid) continuous-discrete architecture for the opening and closing of specific gates within specific ion channels.

Merging our proposed dynamical neuroscience framework developed in Sections 2 and 3 with system thermodynamics [110,112,114] by embedding thermodynamic state notions (*i.e.*, entropy, energy, free energy, chemical potential, *etc.*) within our dynamical system framework can allow us to directly address the otherwise mathematically complex and computationally prohibitive large-scale dynamical model given by Equations (7) and (8). In particular, a thermodynamically consistent neuroscience model would emulate the clinically observed self-organizing, spatio-temporal fractal structures that optimally dissipate energy and optimize entropy production in thalamocortical circuits of fully conscious patients. This thermodynamically consistent neuroscience framework can provide the necessary tools involving multistability, synaptic drive equipartitioning (*i.e.*, synchronization across time scales), energy dispersal, and entropy production for connecting biophysical findings to psychophysical phenomena for general anesthesia. In particular, we hypothesize that as the model dynamics transition to an anesthetized state, the system will involve a reduction in system complexity—defined as a reduction in the degree of irregularity across time scales—exhibiting partial synchronization of neural oscillators (*i.e.*, thermodynamic energy equipartitioning). This would result in a decrease in system energy consumption (myocardial depression, respiratory depression, hypoxia, ischemia, hypotension, venodilation), and hence, a decrease in the rate of entropy production. In other words, unconsciousness is characterized by system decomplexification, which is manifested in the failure to develop efficient mechanisms to dissipate energy thereby pathologically retaining higher internal (or local) entropy levels.

The human brain is *isothermal* and *isobaric*, that is, the temperatures of the subnetworks of the brain are equal and remain constant, and the pressure in each subnetwork also remains constant. The human brain network is also constantly supplied with a source of (Gibbs) free energy provided by chemical nourishment of the blood to ensure adequate cerebral blood flow and oxygenation, which involves a blood concentration of oxygen to ensure proper brain function. Information-gathering channels of the blood also serve as a constant source of free energy for the brain. If these sources of free energy are degraded, then internal (local) entropy is produced.

In the transition to an anesthetic state, complex physiologic work cycles (cardiac respiratory pressure-volume loops, mitochondrial ATP production) necessary for homeostasis follow regressive diffusion and chemical reaction paths that degrade energy production and decrease the *rate of entropy production*. Hence, in an *isolated* large-scale network (*i.e.*, a network with no energy exchange between the internal and external environment) all the energy, though always conserved, will eventually be degraded to the point where it cannot produce any useful work (oxygenation, ventilation, heart rate

stability, organ function). Hence, all motion would cease leading the brain network to a state of unconsciousness (semistability) wherein all or partial [117] subnetworks will possess identical energies (energy equipartition or synchronization) and, hence, internal entropy will approach a local maximum or, more precisely, a saddle surface in the state space of the process state variables. Thus, the transition to a state of anesthetic unconsciousness involves an evolution from an initial state of high (external) entropy production (consciousness) to a temporary saddle state characterized by a series of fluctuations corresponding to a state of significantly reduced (external) entropy production (unconsciousness).

In contrast, in a *healthy* conscious human entropy dissipation occurs spontaneously and, in accordance with Jaynes' maximum entropy production principle [86,118], energy dispersal is optimal leading to a maximum entropy production rate. Hence, low entropy in (healthy) human brain networks is synonymous with consciousness and the creation of order (negative entropy) reflects a rich fractal spatio-temporal variability which, since the brain controls key physiological processes such as ventilation and cardiovascular function, is critical for delivering oxygen and anabolic substances as well as clearing the products of catabolism to maintain healthy organ function. And in accordance with the second law of thermodynamics, the creation and maintenance of consciousness (internal order—negative entropy) is balanced by the external production of a greater degree of (positive) entropy. This is consistent with the maxim of the second law of thermodynamics as well as the writings of Kelvin [119], Gibbs [120,121], and Schrödinger [122] in that the creation of a certain degree of life and order in the universe is inevitably coupled by an even greater degree of death and disorder [110].

In a network thermodynamic model of the human brain, consciousness can be equated to the brain dynamic states corresponding to a low internal system entropy. In recent research [112], the author shows that the second law of thermodynamics provides a physical foundation for the *arrow of time*. In particular, the author shows that the existence of a global strictly increasing entropy function on every nontrivial network thermodynamic system trajectory establishes the existence of a completely ordered time set that has a topological structure involving a closed set homeomorphic to real line, which establishes the emergence of the direction of time flow. Thus, awareness of the passage of time is a direct consequence of regressive changes in the continuous rate of entropy production taking place in the brain and eventually leading (in finite time) to a state of no entropy production (*i.e.*, death). Since these universal regressive changes in the rate of entropy production are spontaneous, continuous, and decreasing in time, human experience perceives time flow as unidirectional. However, since the rate of time flow and the rate of system entropy regression (*i.e.*, free energy consumption) are bijective (*i.e.*, one-to-one and onto), the human experience of time flow is subjective.

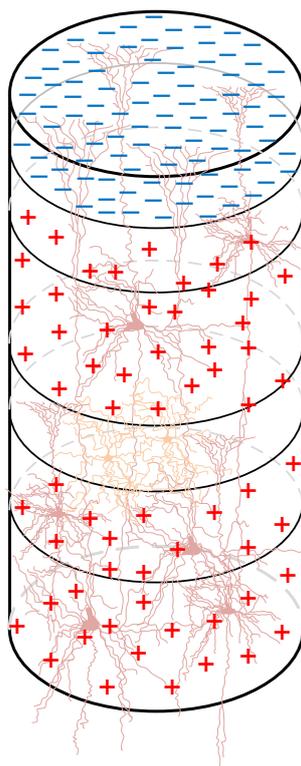
During the transition to an anesthetized state, the external and internal free energy sources are substantially reduced or completely severed in part of the brain leading the human brain network to a semistable state corresponding to a state of *local* saddle (stationary) entropy. Since all motion in the state variables (synaptic drives) ceases in this unconscious (synchronized) state, our index for the passage of time vanishes until the anesthetic wears off allowing for an immediate increase of the flow of free energy back into the brain and other parts of the body. This, in turn, gives rise to a state of consciousness wherein system entropy production is spontaneously resumed. Merging system thermodynamics with multistability theory and mathematical neuroscience with the goal of providing a mathematical framework for describing the anesthetic cascade mechanism is the subject

of current research. In addition, connections between thermodynamics, neuroscience, and the arrow of time [110,112,113] are also being explored to develop an understanding on how the arrow of time is built into the very fabric of our conscious brain.

12. An Electromagnetic Field Theory of Consciousness

The cerebral cortex can be modeled by a columnar topology, where thousands of 0.3–0.6 mm-wide *cortical macrocolumns* are compactly organized in a parallel configuration to create the approximately 4 mm-thick cerebral cortex [123]. Macrocolumns are bundles of approximately 100,000 neurons which are arranged in six distinguishable layers and create a six-layered structure of the cortex; see Figure 21. Layers I, II, III, V, and VI are mainly composed of excitatory pyramidal cells, whereas Layer IV is almost free of pyramidal cells and is mainly composed of stellate cells. The apical dendrites of pyramidal cells perpendicularly extend towards the cortex superficial layer (Layer I), where they receive inputs from other parts of the brain.

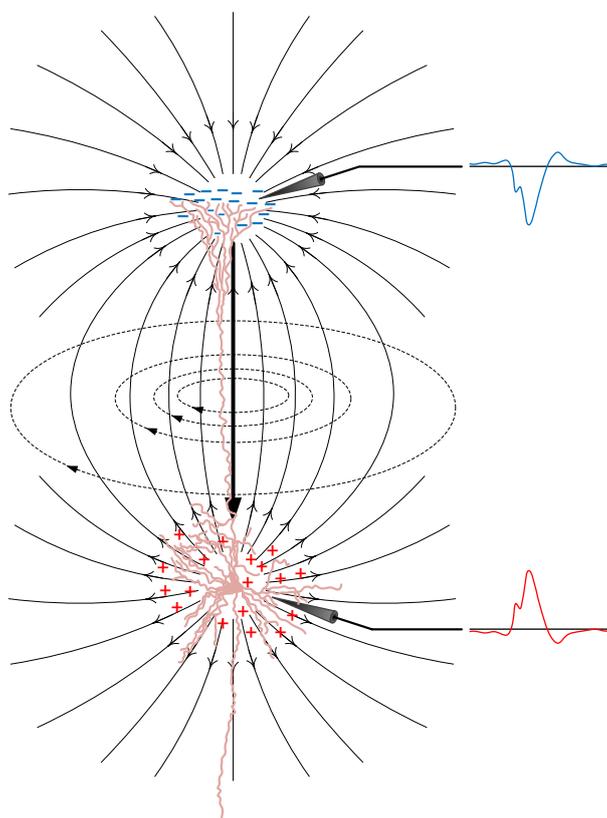
Figure 21. Schematic representing the connective topology within the six-layer cortical macrocolumn.



Information is communicated inside the brain by electrical signals, or more specifically, by the creation and transmission of ions. The electrical signal received by dendrites at synapses travels along the dendrites to the cell body, and the resulting cell-generated signal travels to another cell along the axon. Thus, the spatial topology of the cortex dictates the paths along which the ions can flow through and the locations where they can reside. Based on the specific columnar topology of the cerebral cortex described above, the major portion of the electric current inside the cortex, which is conducted by pyramidal cells, flows radially through the cortex.

The almost evenly distributed short-range current flow through stellate cells in Layer IV do not significantly contribute to the global orientation of the current flow through the cortex. When apical dendrites of pyramidal cells receive excitatory inputs through synapses in Layer I, positive ions flow through them from synaptic clefts toward cell bodies, creating a deficit of positive ions at the synaptic clefts. The positive ions flowing through apical dendrites generate a transient radial current and, due to the capacitive characteristics of the cell membrane, accumulate around the cell membrane. As a result, macroscopic electrical activities in cerebral cortex possess four global behaviors (see Figures 21 and 22). Namely, (i) at Layer I, the potential transition due to the activity of neurons is negative; (ii) at layers II, III, V, and VI, the potential transition due to the activity of neurons is positive; (iii) no significant potential transition is observed in Layer IV; and (iv) transient currents due to neuron activity flows radially inside the cortex.

Figure 22. Transient electromagnetic neuronal dipole. Solid lines represent the electric field and the dashed lines represent the magnetic field.



Sparse activities of neurons in a region of the cortex do not generate significant electromagnetic effect in the cortex. However, when a large portion of neurons in a column or adjacent columns are hyperactivated, that is, fire synchronously, the resulting electromagnetic field due to the structured movement and ion configuration described above is strong enough to create significant electromagnetic interference inside the cortex; see Figure 22.

The state of unconsciousness due to the induction of general anesthesia can potentially be due to the strong electromagnetic pattern formation in the cerebral cortex. Available physiological evidence strongly suggests that the excitatory input signals received from the *reticular activating system* [123]

are crucial in maintaining the conscious state of cerebral activities. Anesthetics are believed to prolong the post synaptic potential of inhibitory neurons, which mainly reside in Layer IV of the cortex [123]. Inhibitory neurons provide negative feedback signals for the population of neurons of cortical columns, and hence, can play an essential role in the stability of the highly interconnected network of neurons as well as the global behavior of these neurons. In particular, anesthetics can reduce the inflow of free energy to the brain leading to global or partial [117] synchronization of neuron firing.

Synchronization in the population of neurons present in Layers II, III, V, and VI of the cortex can create a strong electromagnetic pattern. The resulting electromagnetic field can affect the reception of signals from the reticular activating system by reducing the conduction velocity, partially or totally blocking, or even reflecting the incoming signals. Consequently, the cortex loses a large amount of its crucial background *wash* of input signals, which can result in loss of consciousness [15].

Even though electromagnetic field models have been proposed to explain consciousness (see [124] and the references therein), these models are largely based on empirical observation and lack precise mathematical formulations. A mathematical framework for fostering precision, completeness, and self-consistency in understanding the anesthetic cascade in the human brain using system thermodynamics and electromagnetic field theories are the subject of current research by the authors.

13. Conclusions

With advances in biochemistry, molecular biology, and neurochemistry there has been impressive progress in the understanding of the function of single neurons. Using the example of the mechanism of action of general anesthesia, the past decade has seen a remarkable explosion of our understanding of how anesthetic agents affect the properties of neurons. However, despite this advance, we still do not understand how molecular mechanisms translate into the induction of general anesthesia at the macroscopic level. In particular, there has been little focus on how the molecular properties of anesthetic agents lead to the observed macroscopic property that defines the anesthetic state, that is, lack of responsiveness to noxious stimuli. This clinical property leads to consideration of anesthesia as a nearly binary (on-or-off) variable, and the relationship between the concentration of an anesthetic agent in the central nervous system and the anesthetic state is described in terms of the probability of responsiveness as a function of anesthetic concentration [10]. In clinical studies, the typical observation is that at low concentrations of anesthetic agent the probability of responsiveness (to noxious stimuli) is high, possibly unity. Then as the anesthetic concentration increases there is a sharp transition to a probability of responsiveness that is low and possibly zero.

In this paper, we developed a synaptic drive firing rate model to model the central nervous system as a (possibly discontinuous) autonomous dynamical system and showed that the transition to the anesthetic state exhibits multistability; that is, the system exhibits multiple attracting equilibria under asymptotically slowly changing system parameters directly affected by anesthetic concentrations. In future research we plan to merge system thermodynamics, multistability theory, and dynamic network graph-theoretic notions to develop a framework for understanding central nervous system behavior characterized by abrupt transitions between mutually exclusive conscious and unconscious states.

Such phenomena are not limited to general anesthesia and can be seen in biochemical systems, ecosystems, gene regulation and cell replication, as well as numerous medical conditions (e.g., seizures, epilepsy, schizophrenia, hallucinations, *etc.*) and are obviously of great clinical importance but have been lacking rigorous theoretical frameworks. The primary impact of such frameworks will be to allow for the development of models that go beyond words to dynamic equations, leading to mathematical models with greater precision and self-consistency. Mathematical formulations enforce self-consistency and while “self-consistency is not necessarily truth, self-inconsistency is certainly falsehood.” And within the context of general anesthesia, a dynamical system formulation for neuroscience can foster the development of new frameworks that will allow us to interpret experimental and clinical results, connect biophysical findings to psychophysical phenomena, explore new hypothesis based on the cognitive neuroscience of consciousness and develop new assertions, and improve the reliability of general anesthesia. In addition, such a framework can establish a scientific basis for new metrics of anesthetic depth by making the assessment of consciousness a mechanistically grounded tool.

Acknowledgments

This research was supported in part by the Air Force Office of Scientific Research under Grant FA9550-12-1-0192 and QNRF under NPRP Grant 4-187-2-060.

Conflicts of Interest

The authors declare no conflict of interest.

References and Notes

1. Lapicque, L. Recherches quantitatives sur l' excitation électrique des nerfs traitée comme une polarisation. *J. Physiol. Gen.* **1907**, *9*, 620–635. (In French)
2. Hodgkin, A.L.; Huxley, A.F. A quantitative description of membrane current and application to conduction and excitation in nerve. *J. Physiol.* **1952**, *117*, 500–544.
3. Dayan, P.; Abbott, L.F. *Theoretical Neuroscience: Computational and Mathematical Modeling of Neural Systems*; MIT Press: Cambridge, MA, USA, 2005.
4. Ermentrout, B.; Terman, D.H. *Mathematical Foundations of Neuroscience*; Springer-Verlag: New York, NY, USA, 2010.
5. Mashour, G.A. Consciousness unbound: Toward a paradigm of general anesthesia. *Anesthesiology* **2004**, *100*, 428–433.
6. Zecharia, A.Y.; Franks, N.P. General anesthesia and ascending arousal pathways. *Anesthesiology* **2009**, *111*, 695–696.
7. Sonner, J.M.; Antognini, J.F.; Dutton, R.C.; Flood, P.; Gray, A.T.; Harris, R.A.; Homanics, G.E.; Kendig, J.; Orser, B.; Raines, D.E.; *et al.* Inhaled anesthetics and immobility: Mechanisms, mysteries, and minimum alveolar anesthetic concentration. *Anesth. Analg.* **2003**, *97*, 718–740.
8. Campagna, J.A.; Miller, K.W.; Forman, S.A. Mechanisms of actions of inhaled anesthetics. *N. Engl. J. Med.* **2003**, *348*, 2110–2124.

9. John, E.R.; Prichep, L.S. The anesthetic cascade: A theory of how anesthesia suppresses consciousness. *Anesthesiology* **2005**, *102*, 447–471.
10. Bailey, J.M.; Haddad, W.M. Drug dosing control in clinical pharmacology: Paradigms, benefits, and challenges. *Control Syst. Mag.* **2005**, *25*, 35–51.
11. Merkel, G.; Eger, E.I. A comparative study of halothane and halopropane anesthesia including method for determining equipotency. *Anesthesiology* **1963**, *24*, 346–357.
12. Vuyk, J.; Lim, T.; Engbers, F.H.M.; Burm, A.G.L.; Vietter, A.A.; Bovill, J.G. Pharmacodynamics of alfentanil as a supplement to propofol of nitrous oxide for lower abdominal surgery in female patients. *Anesthesiology* **1993**, *78*, 1936–1945.
13. Hameroff, S. The entwined mysteries of anesthesia and consciousness. Is there a common underlying mechanism? *Anesthesiology* **2006**, *105*, 400–412.
14. Hill, A.V. The possible effects of the aggregation of the molecules of hemoglobin on its dissociation curves. *J. Physiol.* **1910**, *40*, 4–7.
15. Steyn-Ross, M.L.; Steyn-Ross, D.A.; Sleight, J.W. Modelling general anesthesia as a first-order phase transition in the cortex. *Prog. Biophys. Mol. Biol.* **2004**, *85*, 369–385.
16. Angeli, D.; Ferrell, J.; Sontag, E.D. Detection of multistability, bifurcations, and hysteresis in a large class of biological positive-feedback systems. *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 1822–1827.
17. Angeli, D. New analysis technique for multistability detection. *IEE Proc. Syst. Biol.* **2006**, *153*, 61–69.
18. Angeli, D. Multistability in systems with counter-clockwise input-output dynamics. *IEEE Trans. Autom. Control* **2007**, *52*, 596–609.
19. Bhat, S.P.; Bernstein, D.S. Nontangency-based Lyapunov tests for convergence and stability in systems having a continuum of equilibria. *SIAM J. Control Optim.* **2003**, *42*, 1745–1775.
20. Haddad, W.M.; Chellaboina, V. *Nonlinear Dynamical Systems and Control: A Lyapunov-Based Approach*; Princeton University Press: Princeton, NJ, USA, 2008.
21. Carbb, R.; Mackey, M.C.; Rey, A.D. Propagating fronts, chaos and multistability in a cell replication model. *Chaos* **1996**, *6*, 477–492.
22. Rietkerk, M.; Dekker, S.; de Ruiter, P.; Koppel, J. Self-organized patchiness and catastrophic shifts in ecosystems. *Science* **2004**, *305*, 1926–1929.
23. Siegal-Gaskins, D.; Grotewold, E.; Smith, G.D. The capacity for multistability in small gene regulatory networks. *BMC Syst. Biol.* **2009**, *3*, 1–14.
24. Kiyamura, A.; Marszalec, W.; Yeh, J.Z.; Narahishi, T. Effects of halothane and propofol on excitatory and inhibitory synaptic transmission in rat cortical neurons. *J. Pharmacol.* **2002**, *304*, 162–171.
25. Hutt, A.; Longtin, A. Effects of the anesthetic agent propofol on neural populations. *Cogn. Neurodyn.* **2009**, *4*, 37–59.
26. Overton, E. *Studien über die Narkose Zugleich ein Beitrag zur allgemeinen Pharmakologie*; Verlag con Gustav Fischer: Jena, Germany, 1901. (In German)
27. Meyer, H. Welche Eigenschaft der Anästhetica bedingt ihre Narkotische Wirkung? *Naunyn-Schmiedebergs Arch. Exp. Path. Pharmacol.* **1889**, *42*, 109–118. (In German)

28. Ueda, I. Molecular mechanisms of anesthesia. *Anesth. Analg.* **2003**, *63*, 929–945.
29. Franks, N.P.; Lieb, W.R. Molecular and cellular mechanisms of general anesthesia. *Nature* **1994**, *367*, 607–614.
30. North, C.; Cafiso, D.S. Contrasting membrane localization and behavior of halogenated cyclobutanes that follow or violate the Meyer-Overton hypothesis of general anesthetic potency. *Biophys. J.* **1997**, *72*, 1754–1761.
31. Voss, L.J.; Sleight, J.W.; Barnard, J.P.M.; Kirsch, H.E. The howling cortex: Seizures and general anesthetic drugs. *Anesth. Analg.* **2008**, *107*, 1689–1703.
32. Gutkin, B.; Pinto, D.; Ermentrout, B. Mathematical neuroscience: From neurons to circuits to systems. *J. Physiol. Paris* **2003**, *97*, 209–219.
33. Bhatia, N.P.; Szegő, G.P. *Stability Theory of Dynamical Systems*; Springer-Verlag: Berlin, Germany, 1970.
34. Gerstner, W. Neural codes: Firing rates and beyond. *Proc. Natl. Acad. Sci. USA* **1997**, *94*, 12740–12741.
35. Haddad, W.M.; Chellaboina, V.; Hui, Q. *Nonnegative and Compartmental Dynamical Systems*; Princeton University Press: Princeton, NJ, USA, 2010.
36. Wilson, H.; Cowan, J. Excitatory and inhibitory interactions in localized populations of model neurons. *Biophys. J.* **1972**, *12*, 1–24.
37. Amari, S.I.; Yoshida, K.; Kanatani, K.I. A mathematical foundation for statistical neurodynamics. *SIAM J. Appl. Math.* **1977**, *33*, 95–126.
38. Sompolinsky, H.; Crisanti, A.; Sommers, H.J. Chaos in random neural networks. *Phys. Rev. Lett.* **1988**, *61*, 259–262.
39. Amit, D.J.; Gutfreund, H.; Sompolinsky, H. Spin-glass models of neural networks. *Phys. Rev. A* **1985**, *32*, 1007–1018.
40. Amit, D.J.; Brunel, N. Model of global spontaneous activity and local structured activity during delay periods in the cerebral cortex. *Cereb. Cortex* **1997**, *7*, 237–252.
41. Brunel, N.; Hakim, V. Fast global oscillations in networks of integrate-and-fire neurons with low firing rates. *Neural Comput.* **1999**, *11*, 1621–1671.
42. Gerstner, W.; Kistler, W. *Spiking Neuron Models: Single Neurons, Populations, Plasticity*; Cambridge University Press: Cambridge, UK, 2002.
43. Gerstner, W. Time structure of the activity in neural network models. *Phys. Rev. E* **1995**, *51*, 738–758.
44. Alternatively, the functional form of the postsynaptic potential can have the form $\alpha_i^{(E,I)}(t) = B_i^{(E,I)} te^{-\frac{t}{\lambda_i^{(E,I)}}}$.
45. Haddad, M.; Hui, Q.; Bailey, J. Multistability, bifurcations, and biological neural networks: A synaptic drive firing model for cerebral cortex transition in the induction of general anesthesia. In Proceedings of the 2011 50th IEEE Conference on Decision and Control and European Control Conference (CDC-ECC), Orlando, FL, USA, 12–15 December 2011; pp. 3901–3908.
46. Hui, Q.; Haddad, W.M.; Bailey, J.M.; Hayakawa, T. A stochastic mean field model for an excitatory and inhibitory synaptic drive cortical neuronal network. *IEEE Trans. Neural Netw.* **2014**, *25*, 751–763.

47. Filippov, A.F. *Differential Equations with Discontinuous Right-Hand Sides*; Kluwer: Dordrecht, The Netherlands, 1988.
48. Bacciotti, A.; Ceragioli, F. Stability and stabilization of discontinuous systems and nonsmooth Lyapunov functions. *ESAIM Control Optim. Calc. Var.* **1999**, *4*, 361–376.
49. Alternatively, we can consider Krasovskii solutions of Equation (26) wherein the possible misbehavior of the derivative of the state on null measure sets is not ignored; that is, $\mathcal{K}[f](x)$ is replaced with $\mathcal{K}[f](x) = \bigcap_{\delta>0} \overline{\text{co}}\{f(\mathcal{B}_\delta(x))\}$ and where f is assumed to be locally bounded.
50. Aubin, J.P.; Cellina, A. *Differential Inclusions*; Springer-Verlag: Berlin, Germany, 1984.
51. Paden, B.E.; Sastry, S.S. A calculus for computing Filippov's differential inclusion with application to the variable structure control of robot manipulators. *IEEE Trans. Circuits Syst.* **1987**, *34*, 73–82.
52. Haddad, W.M. Nonlinear differential equations with discontinuous right-hand sides: Filippov solutions, nonsmooth stability and dissipativity theory, and optimal discontinuous feedback control. *Commun. Appl. Anal.* **2014**, *18*, 455–522.
53. Leth, J.; Wisniewski, R. On formalism and stability of switched systems. *J. Control Theory Appl.* **2012**, *10*, 176–183.
54. Tangent cones are sometimes referred to as *contingent cones* in the literature.
55. Mashour, G.A. Consciousness and the 21st century operating room. *Anesthesiology* **2013**, *119*, 1003–1005.
56. Jordan, D.; Ilg, R.; Riedl, V.; Schorer, A.; Grimberg, S.; Neufang, S.; Omerovic, A.; Berger, S.; Untergehrer, G.; Preibisch, C.; *et al.* Simultaneous electroencephalographic and functional magnetic resonance imaging indicate impaired cortical top-down processing in association with anesthetic-induced unconsciousness. *Anesthesiology* **2013**, *119*, 1031–1042.
57. Lee, H.; Mashour, G.A.; Noh, G.J.; Kim, S.; Lee, U. Reconfiguration of network hub structure after propofol-induced unconsciousness. *Anesthesiology* **2013**, *119*, 1347–1359.
58. Teel, A.; Panteley, E.; Loria, A. Integral characterization of uniform asymptotic and exponential stability with applications. *Math. Control Signals Syst.* **2002**, *15*, 177–201.
59. Ryan, E.P. An integral invariance principle for differential inclusions with applications in adaptive control. *SIAM J. Control Optim.* **1998**, *36*, 960–980.
60. Clarke, F.H. *Optimization and Nonsmooth Analysis*; Wiley: New York, NY, USA, 1983.
61. Evans, L.C. *Partial Differential Equations*; American Mathematical Society: Providence, RI, USA, 2002.
62. Cortés, J.; Bullo, F. Coordination and geometric optimization via distributed dynamical systems. *SIAM J. Control Optim.* **2005**, *44*, 1543–1574.
63. Hui, Q.; Haddad, W.M.; Bhat, S.P. Semistability, finite-time stability, differential inclusions, and discontinuous dynamical systems having a continuum of equilibria. *IEEE Trans. Autom. Control* **2009**, *54*, 2465–2470.
64. Hartman, P. *Ordinary Differential Equations*; Wiley: New York, NY, USA, 1964.
65. Hale, J.K. *Ordinary Differential Equations*; Wiley: New York, NY, USA, 1980.
66. Aubin, J.P.; Frankowska, H. *Set-Valued Analysis*; Birkhäuser Boston: Boston, MA, USA, 1990.
67. Rockafellar, R.T.; Wets, R.J.B. *Variational Analysis*; Springer-Verlag: Berlin, Germany, 1998.

68. Guillemin, V.; Pollack, A. *Differential Topology*; Prentice-Hall: Englewood Cliffs, NJ, USA, 1974.
69. LaSalle, J.P. Some extensions of Lyapunov's second method. *IRE Trans. Circuit Theory* **1960**, *7*, 520–527.
70. Yoshizawa, T. *Stability Theory and the Existence of Periodic Solutions and Almost Periodic Solutions*; Springer-Verlag: New York, NY, USA, 1975.
71. Bhatia, N.P.; Hajek, O. *Local Semi-Dynamical Systems*; Springer-Verlag: Berlin, Germany, 1969.
72. Strogatz, S.H. From Kuramoto to Crawford: Exploring the onset of synchronization in populations of coupled oscillators. *Phys. D: Nonlinear Phenom.* **2000**, *143*, 1–20.
73. Brown, E.; Moehlis, J.; Holmes, P. On the phase reduction and response dynamics of neural oscillator populations. *Neural Comput.* **2004**, *16*, 673–715.
74. Gray, C.M. Synchronous oscillations in neuronal systems: Mechanisms and functions. *J. Comput. Neurosci.* **1994**, *1*, 11–38.
75. Bullock, T.H.; McClune, M.C.; Achimowicz, J.A.; Iragui-Madoz, V.J.; Duckrow, R.B.; Spencer, S.S. Temporal fluctuations in coherence of brain waves. *Proc. Nat. Acad. Sci. USA* **1995**, *92*, 11568–11572.
76. Buzsáki, G. *Rhythms of the Brain*; Oxford University Press: New York, NY, USA, 2006.
77. Tinker, J.H.; Sharbrough, F.W.; Michenfelder, J.D. Anterior shift of the dominant EEG rhythm during anesthesia in the Java monkey: Correlation with anesthetic potency. *Anesthesiology* **1977**, *46*, 252–259.
78. Rampil, I.J. A primer for EEG signal processing in anesthesia. *Anesthesiology* **1998**, *89*, 980–1002.
79. John, E.R.; Prichep, L.S. The anesthetic cascade: A theory of how anesthesia suppresses consciousness. *Anesthesiology* **2005**, *102*, 447–471.
80. Yu, W.; Cao, J.; Lu, J. Global synchronization of linearly hybrid coupled networks with time-varying delay. *SIAM J. Appl. Dyn. Syst.* **2008**, *7*, 108–133.
81. Wu, C.W.; Chua, L.O. Synchronization in an array of linearly coupled dynamical systems. *IEEE Trans. Circuits Syst.—I: Fundam. Theory Appl.* **1995**, *42*, 430–447.
82. Kuizenga, K.; Wierda, J.; Kalkman, C.J. Biphasic EEG changes in relation to loss of consciousness during induction with thiopental, propofol, etomidate, midazolam or sevoflurane. *Br. J. Anaesth.* **2001**, *86*, 354–360.
83. Deco, G.; Jirsa, V.K.; Robinson, P.A.; Breakspear, M.; Friston, K. The dynamic brain: From spiking neurons to neural masses and cortical fields. *PLoS Comput. Biol.* **2008**, *4*, e1000092:1–e1000092:35.
84. Buice, M.A.; Cowan, J.D. Field-theoretic approach to fluctuation effects in neural networks. *Phys. Rev. E* **2007**, *75*, 051919:1–051919:14.
85. Villarragut, V.M.; Novo, S.; Obaya, R. Neutral functional differential equations with applications to compartmental systems. *SIAM J. Math. Anal.* **2008**, *40*, 1003–1028.
86. Jaynes, E.T. Information theory and statistical mechanics. *Phys. Rev.* **1957**, *106*, 620–630.
87. Jaynes, E.T. Prior probabilities. *IEEE Trans. Syst. Sci. Cybern.* **1968**, *4*, 227–241.

88. Lyon, R.H. *Statistical Energy Analysis of Dynamical Systems: Theory and Applications*; MIT Press: Cambridge, MA, USA, 1975.
89. *E. T. Jaynes: Papers on Probability, Statistics and Statistical Physics*; Rosenkrantz, R.D., Ed.; Reidel: Boston, MA, USA, 1983.
90. Arnold, L. *Stochastic Differential Equations: Theory and Applications*; Wiley: New York, NY, USA, 1974.
91. Zhou, J.; Wang, Q. Stochastic semistability with application to agreement problems over random networks. In Proceedings of the American Control Conference, Baltimore, MD, USA, 30 June–2 July 2010; pp. 568–573.
92. Hui, Q.; Haddad, W.M.; Bhat, S.P. Finite-time semistability and consensus for nonlinear dynamical networks. *IEEE Trans. Autom. Control* **2008**, *53*, 1887–1900.
93. Bernstein, D.S. *Matrix Mathematics*, 2nd ed.; Princeton University Press: Princeton, NJ, USA, 2009.
94. Shen, J.; Hu, J.; Hui, Q. Semistability of switched linear systems with applications to distributed sensor networks: A generating function approach. In Proceedings of the 2011 50th IEEE Conference on Decision and Control and European Control Conference (CDC-ECC), Orlando, FL, USA, 12–15 December 2011; pp. 8044–8049.
95. Ash, R.B. *Real Analysis and Probability*; Academic: New York, NY, USA, 1972.
96. Hui, Q. Optimal semistable control for continuous-time linear systems. *Syst. Control Lett.* **2011**, *60*, 278–284.
97. Berman, A.; Plemmons, R.J. *Nonnegative Matrices in the Mathematical Sciences*; Academic Press, Inc.: New York, NY, USA, 1979.
98. Hui, Q.; Haddad, W.M.; Bailey, J.M. Multistability, bifurcations, and biological neural networks: A synaptic drive firing model for cerebral cortex transition in the induction of general anesthesia. *Nonlinear Anal.: Hybrid Syst.* **2011**, *5*, 554–572.
99. Mao, X. Exponential stability of stochastic delay interval systems with Markovian switching. *IEEE Trans. Autom. Control* **2002**, *47*, 1604–1612.
100. Aerts, J.M.; Haddad, W.M.; An, G.; Vodovt, Y. From Data Patterns to Mechanistic Models in Acute Critical Illness. *J. Crit. Care* **2014**, *29*, 604–610.
101. The term “degree of consciousness” reflects the intensity of a noxious stimulus. For example, we are often not aware (conscious) of ambient noise but would certainly be aware of an explosion. Thus, the term “degree” reflects awareness over a spectrum of stimuli. For any particular stimulus the transition from consciousness to unconsciousness is a very sharp transition which can be modeled using a very sharp sigmoidal function—practically a step function.
102. Macklem, P.T.; Seely, A.J.E. Towards a definition of life. *Prespect. Biol. Med.* **2010**, *53*, 330–340.
103. Seely, A.J.E.; Macklem, P. Fractal variability: An emergent property of complex dissipative systems. *Chaos* **2012**, *22*, 1–7.
104. Bircher, J. Towards a dynamic definition of health and disease. *Med. Health Care Philos.* **2005**, *8*, 335–341.

105. Goldberger, A.L.; Rigney, D.R.; West, B.J. Science in pictures: Chaos and fractals in human physiology. *Sci. Am.* **1990**, *262*, 42–49.
106. Goldberger, A.L.; Peng, C.K.; Lipsitz, L.A. What is physiologic complexity and how does it change with aging and disease? *Neurobiol. Aging* **2002**, *23*, 23–27.
107. Godin, P.J.; Buchman, T.G. Uncoupling of biological oscillators: A complementary hypothesis concerning the pathogenesis of multiple organ dysfunction syndrome. *Crit. Care Med.* **1996**, *24*, 1107–1116.
108. Kondepudi, D.; Prigogine, I. *Modern Thermodynamics: From Heat Engines to Dissipative Structures*; Wiley: New York, NY, USA, 1998.
109. Prigogine, I. *From Being to Becoming*; Freeman: New York, NY, USA, 1980.
110. Haddad, W.M.; Chellaboina, V.; Nersesov, S.G. *Thermodynamics: A Dynamical Systems Approach*; Princeton University Press: Princeton, NJ, USA, 2005.
111. Haddad, W.M.; Hui, Q. Complexity, robustness, self-organization, swarms, and system thermodynamics. *Nonlinear Anal.: Real World Appl.* **2009**, *10*, 531–543.
112. Haddad, W.M. Temporal asymmetry, entropic irreversibility, and finite-time thermodynamics: From Parmenides-Einstein time-reversal symmetry to the Heraclitan entropic arrow of time. *Entropy* **2012**, *14*, 407–455.
113. Haddad, W.M.; Chellaboina, V.; Nersesov, S.G. Time-reversal symmetry, Poincaré recurrence, irreversibility, and the entropic arrow of time: From mechanics to system thermodynamics. *Nonlinear Anal.: Real World Appl.* **2008**, *9*, 250–271.
114. Haddad, W.M. A Unification between dynamical system theory and thermodynamics involving an energy, mass, and entropy state space formalism. *Entropy* **2013**, *15*, 1821–1846.
115. Nersesov, S.G.; Haddad, W.M. Reversibility and Poincaré recurrence in linear dynamical systems. *IEEE Trans. Autom. Control* **2008**, *53*, 2160–2165.
116. Godsil, C.; Royle, G. *Algebraic Graph Theory*; Springer-Verlag: New York, NY, USA, 2001.
117. When patients lose consciousness other parts of the brain are still functional (heart rate control, ventilation, oxygenation, *etc.*), and hence, the development of biological neural network models that exhibit partial synchronization is critical. In particular, models that can handle synchronization of subsets of the brain with the non-synchronized parts firing at normal levels is essential in capturing biophysical behavior. The problem of partial system synchronization has not been addressed in the literature. This is the subject of current research by the authors.
118. Swenson, R. Emergent attractors and the law of maximum entropy production: Foundations to a theory of general evolution. *Syst. Res.* **1989**, *6*, 187–197.
119. Thomson, K.W. On a universal tendency in nature to the dissipation of mechanical energy. *Trans. R. Soc. Edina.* **1852**, *20*, 187–197.
120. Gibbs, J. On the equilibrium of heterogeneous substances. *Trans. Conn. Acad. Sci.* **1875**, *III*, 108–248.
121. Gibbs, J. On the equilibrium of heterogeneous substances. *Trans. Conn. Acad. Sci.* **1878**, *III*, 343–524.
122. Schrödinger, E. *What Is Life?*; Cambridge University Press: Cambridge, UK, 1944.

123. Kandel, E.R.; Schwartz, J.H.; Jessell, T.M.; Siegelbaum, S.A.; Hudspeth, A.J. *Principles of Neural Science*; McGraw-Hill: New York, NY, USA, 2013.
124. Pockett, S. The electromagnetic field theory of consciousness. *J. Conscious. Stud.* **2012**, *19*, 1191–1223.

© 2014 by the authors; licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/3.0/>).