Dynamics of coupled thalamocortical modules

**Jonathan D. Drover · Nicholas D. Schiff ·
Jonathan D. Victor**

Abstract We develop a model of thalamocortical dynamics using a shared population of thalamic neurons to couple distant cortical regions. Behavior of the model is determined as a function of the connection strengths with shared and unshared populations in the thalamus, either within a relay nucleus or the reticular nucleus. When the coupling is via the reticular nucleus, we locate solutions of the model where distant cortical regions maintain the same activity level, and regions where one region maintains an elevated activity level, suppressing activity in the other. We locate and investigate a region where both types of solutions exist and are stable, yielding a mechanism for spontaneous changes in global activity patterns. Power spectra and coherence are computed, and marked differences in the coherence are found between the two kinds of modes. When, on the other hand, the coupling is via a shared relay nuclei, the features seen with the reticular coupling are absent. These considerations suggest a role for the reticular nucleus in modulating long distance cortical communication.

Keywords Thalamocortical dynamics · Shared
inhibition · Reticular nucleus

J. D. Drover (✉) · N. D. Schiff · J. D. Victor
Weill Cornell Medical College of Cornell University,
New York, NY, USA
e-mail: jod2017@med.cornell.edu

1 Introduction

The brain is a complex, hierarchically organized system. The hierarchy spans a wide range of scales, each with its own class of phenomena. For example, understanding sensory processing (e.g. identifying a particular face) likely resides in the detailed activity of individual neurons, but overall state (e.g. wakefulness, drowsiness, and sleep) likely resides in the dynamics of large populations (10^7 or more) of neurons. The electroencephalogram (EEG), which consists of the scalp-recorded electrical signals generated by the brain, is well-suited to probing this level of organization. Because it is recorded on the scalp, it necessarily averages the activity of large populations of neurons— but does so with high temporal fidelity. Because it is non-invasive and simply obtained, it can be recorded in almost anyone, including subjects who are moving and patients whose medical condition requires close observation and/or support equipment. This makes it possible to obtain prolonged recordings spanning many behavioral states, even in patients who have major disturbances of consciousness due to severe neurologic or medical disease. As a result of such studies, there is extensive empirical knowledge of how various normal and pathological brain states are correlated with EEG observables, chiefly power spectra and coherence (e.g. Davey et al. 2000).

These correlations between EEG observables and behavioral state are altered when there is damage to either cortical or thalamic areas. Understanding how EEG recordings are related to underlying brain activity (specifically in the thalamus), and the ability to deduce how changes in EEG can be manifestations of disruption of this underlying activity (specifically the

interactions between populations of neurons) due to injury requires modeling. The use of EEG as an experimental background makes modeling at the population level an obvious choice. Previous studies using modeling at the population level (neural mass modeling or mean field modeling) were done by David and Friston (2003) as well as in the body of work (Rennie et al. 2002; Robinson et al. 1998, 2002). Each of these studies demonstrates the efficacy of modeling at the population level to obtain EEG characteristics. The studies also complement one another. The work in David and Friston (2003) models distant cortical regions, but without a thalamic component. On the other hand, the Robinson groups' model contains a thalamic component, but distant cortical populations are not considered. We discuss each of these models in more detail, as background for the present approach.

In David and Friston (2003) the authors develop a model capable of reproducing a range of spectral behaviors. In this paper, they utilize multiple cortical populations: a pair of excitatory populations (spiny stellate cells and pyramidal cells) and a population of inhibitory interneurons. They explore how the time constants within these populations, as well as the relative strength of output from each, influences the output frequency, analogous to an EEG output. They also determine the effect of different resonances in distant populations for different connection strategies (i.e. unidirectional and bidirectional), as well as the effect of propagation delays. However, this work does not take into account the role of the thalamus in the relationship between distant cortical populations; the model only considers the mentioned cortical populations. The behavior of the model is a stable periodic solution, they do not demonstrate the ability of the model to foster multiple attractors. Thus, they do not provide a mechanism by which the power spectral density can change without changing the parameters themselves.

The body of work represented by Rennie et al. (2002) and Robinson et al. (1998, 2002) uses a similar mathematical framework to model at the population level, but the authors consider the role of the thalamus—a relay nucleus and a population of reticular neurons—in addition to excitatory and inhibitory cortical populations. Also, rather than characterize behavior as a function of the various time constants and propagation delays as in David and Friston (2003), they demonstrate the capacity of this model to exhibit the spectral features seen in EEG for a number of behavioral states (awake with eyes open and closed and four sleep stages) as a function of the connectivity parameters. However, they do not consider distant cortical populations, and the thalamic populations have no

spatial structure. This body of work develops a structure we will refer to as a "thalamocortical module"— a cortex that is reciprocally connected to a thalamic relay nucleus, and a thalamic reticular nucleus that is reciprocally connected to the thalamic relay nucleus, and receives input from cortical populations (outlined in Fig. 1).

These studies, and the anatomy that they are based on, motivate a strategy for modeling: population-based models composed of several thalamocortical modules. The thalamocortical module is repeated multiple times in each hemisphere—for example, in the visual system (the relay nucleus is the lateral geniculate), in the somatosensory system (the relay nucleus is the ventralis posterolateralis/posteromedialis complex), and in the motor system (the relay nucleus is the ventralis anterior/ventralis posterior complex). Here we restrict our analysis to the simplest case, two modules that are coupled together.

While the fundamental importance of the thalamocortical module is well-recognized (Destexhe et al. 1998, 1999; Huguenard and McCormick 2007; Jones 2009), the distinctive role of the reticular component is less widely appreciated. In Destexhe et al. (1998, 1999), Huguenard and McCormick (2007), and Jones (2009), it is shown that the reticular nucleus is a key component in modulating thalamic synchrony within dorsal thalamic nuclei. Specifically, the reticular nucleus initiates a synchronizing mechanism of post-synaptic rebound in the thalamic relay nuclei (Destexhe et al. 1998). At the population level, we are primarily interested in the reticular nucleus because of its more widespread projections. The connections of the reticular nucleus
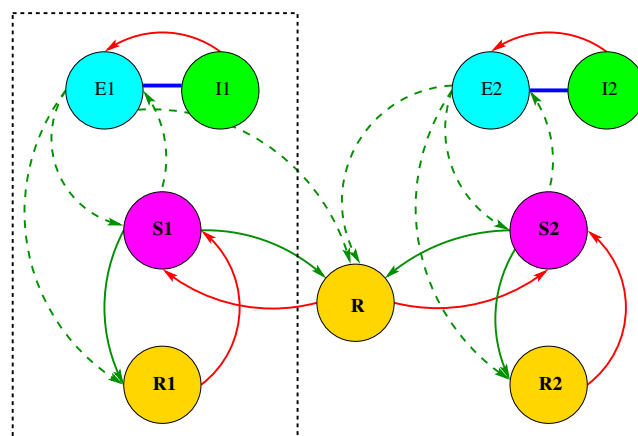

Fig. 1 The coupled network. The *red arrows* indicate inhibitory connections and the *green arrows* indicate excitatory connections. The *blue* between the cortical populations indicates the slaving of the inhibitory potential to that of the excitatory cortex. The *dotted rectangle* surrounds a single thalamocortical module.

(especially its projections into the anterior intralaminar regions of the thalamus (Velayos et al. 1989)) enable it to couple together thalamocortical modules and thereby induce changes in the coherence of neural activity in different cortical areas and possibly across the cerebral hemispheres (Raos and Bentivoglio 1993). The functional importance of this coupling is suggested by experimental and clinical observations (see Schiff and Purpura 2002 for a review). In the primate brain, inter-regional coherence is modulated during the successful performance of sensorimotor tasks. Moreover, inter-regional coherence occurs at the times during which transfer of information between regions is likely to occur (Rodriguez et al. 1999; Pesaran et al. 2002, 2008). Additionally, thalamic damage may selectively disrupt inter-regional coherence, leaving the local spectrum relatively intact (Davey et al. 2000).

These considerations motivate the present work, in which we analyze the dynamics of coupled thalamo-cortical modules. We use the Robinson model for the thalamocortical module as a starting point. In keeping with neuroanatomy (Crabtree and Isaac 2002; Rigas and Castro-Alamancos 2007; Truex and Carpenter 1969), we couple these modules by including a shared population of thalamic neurons (Fig. 1). Our goal is to understand the different behaviors that are possible when the modules are connected together via the thalamus. We focus on the case where the connection between modules is via a shared population of neurons in the reticular nucleus, and we find a number of interesting behaviors. For comparison, we also consider the case where a shared relay nucleus connects thalamocortical modules together. We consider each of these cases separately, so that the contribution of each is made clear.

In addition to thalamic connectivity, there are direct connections between the cortical populations. On smaller spatial scales, intra-cortical connections have been shown to synchronize thalamic oscillations (Destexhe et al. 1999). In this paper, these short range connections are included within the cortical populations as the excitatory (and inhibitory) connections (ν_{EE} and ν_{EI}). The connections between modules extend over a considerably larger distance. Since these long-range connections amount to a very small fraction of intra-cortical fibers (Barbas and Rempel-Clower 1997; Rigas and Castro-Alamancos 2007), they may be considered to be perturbations to the model with only thalamic connectivity between modules. That is, the behaviors seen with direct cortical connectivity superimposed on thalamic connectivity will be perturbations of those seen with thalamic connectivity alone. In other words, the qualitative behaviors possible with and without direct cortical connections will be the same, for realistic connection strengths. For this reason, we do not include direct cortical connections in our analysis.

The main goal of this work is to expand on the successes of the above modeling studies, focusing on what it predicts about interactions, primarily the possible role of the reticular nucleus as a coupling mechanism between distant cortical regions. We reduce the analysis of the coupling to a two variable bifurcation diagram where one axis reflects the strength of the contribution of the unshared thalamic population, and the other reflects the contribution of the shared (connecting) thalamic population.

A long term goal is to determine subcortical connectivity using EEG surface recordings, and so we also compute the power spectral density and coherence for each of the different relevant states that we encounter. We relegate the focus on PSD peak location to later studies, and consider here the qualitative differences that occur when the system switches from one region to another. Examples of these qualitative differences will be made clear in Section 3.2.

Our main result is that the simple addition of a shared population of neurons in the reticular nucleus leads to a rich dynamical repertoire within the physiologically meaningful parameter range. For uncoupled modules, the dynamics are characterized by a single stable fixed point, and their noise-excitation spectra are well-approximated by the spectra of the corresponding linearized system. In contrast, the coupled system has several modes of behavior, including a synchronous activity mode and a winner-take-all mode. For certain parameter values, these modes coexist, resulting in multistability. This suggests a mechanism for spontaneous state changes, and provides support for the concept that the reticular nucleus can control interactions between cortical regions. These behaviors were not seen using a shared relay nucleus.

The paper is organized as follows: In Section 2 we review the equations of the Robinson model and develop the model that we explore here. In this section we only consider the case where the modules are connected using shared reticular neurons. In Section 3, we perform the analysis of the model, focusing on the bifurcation diagram that characterizes the behavior of the model as a function of the coupling strengths with the different types of reticular populations, as well as the power spectra and coherence for the different model behaviors (both the linear approximations, as well as those determined by time series from simulations). In Section 3.3 we show that the various behaviors do not depend strictly on the chosen parameter set, as well as reveal a mechanism that

destroys multi-modality. In Section 3.4 we use a shared relay nucleus to connect modules together, and demonstrate that the behaviors seen with reticular coupling (shared inhibition), specifically the WTA behaviors and multi-stability, are not seen for this type of connection (shared excitation).

2 The model

In this section we present the dynamical system, a system of delay differential equations (DDEs), that we use to model the thalamocortical network. As in David and Friston (2003) and Robinson et al. (2002), each population is represented by two quantities—the average potential (mV) of the cells within each, and the average activity level within each (Hz). We describe the system using the thalamocortical module as a starting point.

A module has four populations, an excitatory cortex (E), an inhibitory cortex (I), a thalamic relay nucleus (S) and a thalamic reticular nucleus (R). The average potential for $a = E, S, R$ is modeled by the delay differential equation

$$DV_a = \sum_{i=E,I,S,R} \nu_{a,i} \phi_i(t - \tau_{a,i}). \tag{1}$$

As in Robinson et al. (2002), we slave the potential of the inhibitory cortical population to that of the excitatory population, setting $V_I = V_E$. The differential operator, D, is defined as

$$D = \frac{1}{\alpha\beta} \frac{d^2}{dt^2} + \left(\frac{1}{\alpha} + \frac{1}{\beta}\right) \frac{d}{dt} + 1. \tag{2}$$

In Eq. (1), ϕ_i is the activity variable, or average firing rate, of the population i (units Hz). The parameters $\nu_{a,b}$ are the strengths of the connections originating in population b and received in population a. These account for both the strength of an input from an individual neuron and the number of neurons in the population (units mVs). The delay term, $\tau_{a,b}$, is zero if both the sending and receiving populations are in the thalamus, or if both are in the cortex. For thalamocortical or corticothalamic pathways, the delay is nonzero. For Eq. (2), the parameters α and β determine the rise and decay times for the potential of the population when encountering a pulse input. For a detailed discussion of these parameters, see Debellis (2007).

In three of the populations, $a = I, S, R$, it is assumed that the spatial spread of activity is very fast. Consequently we model the activity variable, ϕ_a, simply as a saturating function of the potential, V_a. We use an increasing sigmoidal function, with saturation at $Q_{\max}$.

$$\phi_a = Q(V_a) = Q_{\max} \left[1 + \exp\left(-\frac{V_a - \theta}{\sigma}\right)\right]^{-1}. \tag{3}$$

where θ is the mean firing threshold for neurons in the population, $\sigma \frac{\pi}{3}$ is the standard deviation of this threshold and $Q_{\max}$ is the saturated firing rate. The sigmoidal shape of the graph of this function results from averaging over a population of individual neurons with step function thresholds.

Due to longer axonal length, the propagation within the excitatory cortical populations is slower than the others. Following Robinson, we model this using a wave equation and only consider spatially uniform solutions. This leads to

$$\left(\frac{1}{\gamma^2} \frac{d^2}{dt^2} + \frac{2}{\gamma} \frac{d}{dt} + 1\right) \phi_E = Q(V_E) \tag{4}$$

where γ is the ratio of propagation velocity to mean axonal length. Note that this is a difference between our approach and that in David and Friston (2003), where the translation from potential to activity level is taken to be immediate for all populations.

Input to the system from external sources (e.g. sensory input) is presented via the populations of relay neurons (S). This external input is written as

$$\nu_{S,N} \phi_N = N(\mu, \sigma_E^2), \tag{5}$$

where $N(\mu, \sigma_E^2)$ indicates normally distributed noise with mean μ and variance σ_E^2.[1] The mean is constant throughout this paper ($\mu = 2$), and σ_E^2 will be specified whenever the results of simulations are presented. We will also specify when we are treating this quantity as a constant ($\sigma_E^2 = 0$), such as when computing the location and stability of fixed points.

In this paper, we model a system that entails two cortical regions (including excitatory and inhibitory populations). Each cortical region belongs to a thalamocortical module, and thus has its own dedicated thalamic relay nucleus and population of reticular neurons. There are a number of potential strategies for connecting together modules. For the reasons mentioned in the introduction, here we initially focus on a shared population of reticular neurons. This population projects (GABA-ergic) inhibition onto the relay nuclei of each of the modules (Velayos et al. 1989), and receives input from cortical populations and relay nuclei of both modules (see Fig. 1).

---

[1] σ_E^2 is the effective variance. The actual variance used by the integrator is scaled by the square root of the time step, $\sqrt{dt}\sigma_E^2$.

Table 1 Coupling strengths ($\nu_{a,b}$ in mVs)

a \ b	E_1	I_1	S_1	R_1	E_2	I_2	S_2	R_2	R
E_1	1.7	−1.8	1.2						
S_1	1.0			$-\kappa_u$					$-\kappa_s$
R_1	$0.4\kappa_u$		$0.2\kappa_u$						
E_2					1.7	−1.8	1.2		
S_2					1.0			$-\kappa_u$	$-\kappa_s$
R_2					$0.4\kappa_u$		$0.2\kappa_u$		
R	$0.2\kappa_s$		$0.1\kappa_s$		$0.2\kappa_s$		$0.1\kappa_s$		

The connectivity parameters used are given in Table 1. The parameters κ_u, κ_s determine the balance between the contribution of the unshared reticular populations $R_{1,2}$ and the shared population R. For example, the when $\kappa_u = 1$ and $\kappa_s = 0$, the two modules are independent of one another, as there is no contribution from the shared reticular population. On the other hand, the case $\kappa_s = 1$ and $\kappa_u = 0$ corresponds to all reticular activity is in the shared population. The numerical values in Table 1 are taken from Debellis (2007), and are representative of parameters that yield spectra consistent with the eyes open state. Each of the relay populations receives independent external input, described by Eq. (5). Global parameters, those not related to connectivity, are given in Table 2.

For the next section, we use the first order formulation of the model, using well known changes of variables to reduce the second order operators (4) and (1). This first order system consists of 18 equations (7 first and 7 second derivatives for the potentials, and the first and second derivatives of the activity in the excitatory cortices), and we write this system as

$$\frac{d\mathbf{y}}{dt} = \mathbf{f}(\mathbf{y}(t)) + \mathbf{g}(\mathbf{y}(t - \tau)), \tag{6}$$

separating those terms that have delays and those that do not.

3 Results

Our immediate goal is to determine the possible behaviors of the model as a function of the coupling between the modules. We quantify this as the amount of cou-

Table 2 Global parameters

α	100.0 s^{-1}
β	400.0 s^{-1}
$Q_{\max}$	250.0 s^{-1}
τ	0.04 s
θ	15 mV
σ	3.3 mV
γ	100.0 s^{-1}

pling with an unshared reticular population compared with the strength of coupling with a shared population of reticular neurons.

3.1 Construction of the bifurcation diagram

To characterize the scope of possible behaviors of our model as a function of the reticular coupling strengths, we construct a bifurcation diagram. The bifurcation diagram will determine where, in a parameter space that we specify, the topological features such as number and stability of fixed points change. A fixed point, $\mathbf{y}_0$, is a point that satisfies

$$\mathbf{f}(\mathbf{y}_0) + \mathbf{g}(\mathbf{y}_0) = \mathbf{0} \tag{7}$$

for $\sigma_E^2 = 0$, with the $\mathbf{f}$ and $\mathbf{g}$ from Eq. (6). The diagram consists of curves that correspond to solutions, λ, of the equation

$$\det\left(\lambda I - J_{\mathbf{f}}(\mathbf{y}_0) - J_{\mathbf{g}}(\mathbf{y}_0)e^{-\lambda\tau}\right) = 0. \tag{8}$$

with zero real part. Here, I is the 18×18 identity matrix, and $J_{\mathbf{f}}$ and $J_{\mathbf{g}}$ are the Jacobian matrices for $\mathbf{f}$ and $\mathbf{g}$, respectively, evaluated at the fixed point $\mathbf{y}_0$. These curves are used to determine how, and for what parameter values, the number and stability of the fixed points changes. The diagram is shown in Fig. 2. Along the horizontal axis are values of the parameter κ_s from Table 1, and the vertical axis represents values for κ_u. We solve the transcendental equation for the rightmost roots as in Engelborghs et al. (2002), by solving the system

$$\left(\lambda I - J_{\mathbf{f}}(\mathbf{y}_0) - J_{\mathbf{g}}(\mathbf{y}_0)e^{-\lambda\tau}\right)\mathbf{v} = \mathbf{0} \tag{9}$$

for $\lambda \in \mathbb{C}$ and the eigenvector $\mathbf{v} \in \mathbb{C}^{18}$. Since the eigenvector can scale, this system is under determined. We add the additional constraint

$$\mathbf{c}^*\mathbf{v} = 1 \tag{10}$$

where $\mathbf{c} \in \mathbb{C}$ is chosen by the user. A Newton iteration is used to locate the roots of Eq. (9) subject to (10).

To find the various bifurcations, we alter the system slightly to reflect the form of the eigenvalue at the bifurcation. For a Hopf bifurcation the system we solve is given by

$$\left(\iota\omega I - J_{\mathbf{f}}(\mathbf{y}) - J_{\mathbf{g}}(\mathbf{y})e^{-\iota\omega\tau}\right)\mathbf{v} = \mathbf{0} \tag{11}$$

$$f(\mathbf{y}) + g(\mathbf{y}) = 0 \tag{12}$$

$$\mathbf{c}^*\mathbf{v} = 1 \tag{13}$$

where the variables are the eigenvector $\mathbf{v}$, $\omega \in \mathbb{R}$ and the free bifurcation parameter. To solve for a zero eigenvalue, we rewrite Eq. (11) as

$$\left(-J_{\mathbf{f}}(\mathbf{y}) - J_{\mathbf{g}}(\mathbf{y})\right)\mathbf{v} = \mathbf{0}. \tag{14}$$

Note that in this case, the eigenvector $\mathbf{v}$ is real and that delay terms (τ) do not influence these curves.

The important bifurcations that we consider are the Hopf bifurcation ($\lambda = \iota\omega$), the fold bifurcation($\lambda = 0$), and the pitchfork bifurcation (also $\lambda = 0$). For a detailed explanation of the difference between the pitchfork and fold bifurcation, as well as the conditions on the center manifold that categorize them unambiguously, see Wiggins (1990). The topological alterations by which these bifurcations manifest themselves are our major concern here.

We classify two types of fixed point solutions here. We refer to the first type as a symmetric solution. This type of solution is characterized by identical activity levels (and potentials) in each of the two modules. We refer to the second type of fixed point solution that we encounter, in which one modules has a higher activity level than the other, as a Winner-Take-All (WTA) solution. WTA solutions are characterized by a large difference in the activity levels between the two modules. Shared inhibition is known to result in this type of behavior, and for certain parameter ranges the shared inhibitory projections from the shared reticular population are sufficient. The bifurcation diagram will enable us to find out how the involvement from the different types of reticular populations (shared vs. unshared module) determines the type of fixed point solutions that exist and are stable.

Figure 2 shows, in the $\kappa_s - \kappa_u$ plane, where bifurcations occurs. In this paper, we are only interested in the behavior around stable fixed points. For this reason, we first eliminate some regions from consideration and analysis, and then return to a discussion of the regions of interest, which contain stable fixed points corresponding to symmetric and WTA activity. The first of these is the lower left corner of the figure. Here, there is not sufficient inhibition in the system, and the solutions grow to saturation. We refer to the dotted black line outlining this region as the "relevant boundary", since we are not interested in saturation. The other regions that we disregard are regions where the fixed point solution is destabilized in favor of periodic behavior via a Hopf bifurcation. One of our uses of the model involves capturing the spectral information present in the noisy perturbations near fixed points. One could do the same around a limit cycle, but the frequency of the periodic solution will dominate the power spectral
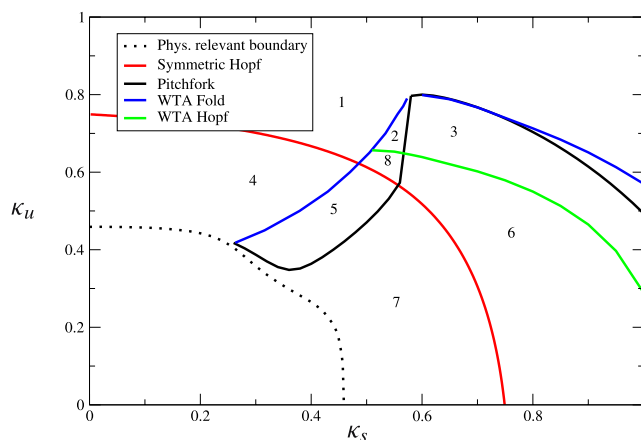

Fig. 2 The bifurcation diagram for the eyes open parameter values. On the *horizontal axis* is the shared coupling parameter κ_s. The *vertical axis* is the unshared reticular coupling, κ_u. The *red curve* is a curve of Hopf bifurcations. The *black curve* (Pitchfork) is a curve of subcritical pitchfork bifurcations that destabilize the symmetric solutions and generate unstable WTA solutions that act as separatices in the multistable region. The *blue curve* (WTAfold) indicates fold bifurcations. Stable WTA solutions are created at fold bifurcations, as well as unstable fixed points that eventually die on the pitchfork. The *green curve* is the curve where the WTA solutions undergo a Hopf bifurcation. Parameters below this curve, and to the right of the WTA-inducing fold curve, will result in an unstable WTA fixed point centering a stable periodic solution. Values of $\kappa_{s,u}$ that are below the curve marked "relevant boundary" result in a system with very little inhibition. Across this boundary, the symmetric fixed point is annihilated via a fold bifurcation, and the system approaches saturation

density. For this reason, we disregard the regions below the red curve of Hopf bifurcations. Similarly, we disregard parameter values that lie below the green curve, since this curve separates WTA fixed point behavior from WTA periodic behavior. Additionally, the Hopf bifurcations are not present for all delay lengths (τ). For delays that are less than approximately 20 ms there will not be a Hopf bifurcation curve.

This leaves the regions marked as 1, 2 and 3 for us to consider, as they contain the stable fixed points. The fixed points and their stability, across these three regions, are given as a function of κ_s (with κ_u fixed) in Fig. 3. The blue curve in Fig. 2 represents a curve of fold bifurcations. At these fold bifurcations, a new group of fixed points appears from left to right in the figure. In this case, four new fixed points are created: a stable and unstable winner and loser (WTA solutions), and the mirror (due to symmetry, either module can win) (see Fig. 3). The solid black curve is a curve of pitchfork bifurcations of the symmetric solution. These bifurcations do not impact either the existence or stability of the stable WTA fixed points. The unstable branches of the WTA (created at the fold bifurcation) coalesce
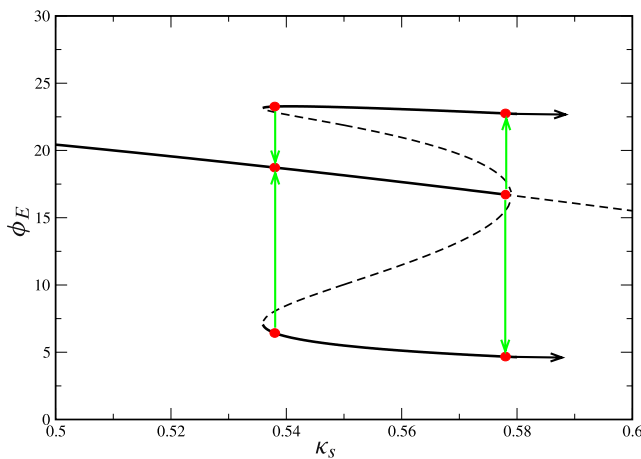
Fig. 3 The fold and pitchfork bifurcations. For this figure, $\kappa_u = 0.7$, and the bifurcation parameter is κ_s. The fold bifurcation, generating a pair of stable fixed points, is shown as well as the pitchfork bifurcation that destabilizes the symmetric solution. The *green arrows* show the most likely direction for spontaneous state changes at each end. Note that this figure also shows the relative activity levels for the WTA solutions over this range of κ_s

Fig. 4 Spontaneous switching in the multistable region. Parameter values are $\sigma_E^2 = 0.14$, $\kappa_u = 0.7$, and $\kappa_s = 0.558$

with the stable symmetric solution, destabilizing it (see Fig. 3). To the right of the black fold curve, the WTA fixed points are the only stable fixed points of interest.

The region marked region 2 is of interest because, for these values, there are multiple stable fixed points. That is, both the WTA solutions and the symmetric solution exist and are stable. This can be seen clearly in Fig. 3. In the range shown, the overall activity can switch modes of behavior spontaneously, without parametric alterations. There are two types of switching, either outward (from the symmetric solution to the WTA solution) or inward. Near the fold bifurcation, the inward switch is very likely, and the outward switch is highly unlikely. Near the pitchfork bifurcation, the outward switch is much more likely than the inward switch (arrows in Fig. 3). At intermediate values of κ_s, switching in either direction can occur (Fig. 4).

This suggests a mechanism for rapid, spontaneous changes in dynamics—which we refer to as "microstate changes". Specifically, the system state behaves like a particle in a multiple-well potential. There are three wells, one for each of the WTA fixed points and another for the symmetric state. These are separated by peaks, the unstable fixed point solutions born at the pitchfork bifurcation. A microstate change occurs when noisy perturbations are sufficient to bounce a ball out of one of the wells, over the separatix, and into another well. The larger the variance of the noisy input, the larger the likelihood of a switch, and so the variance of the input will have an effect on the dwell time in
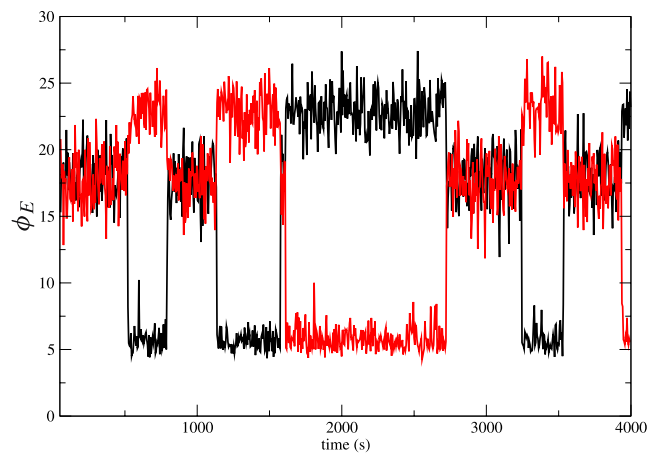
each of the microstates. In fact, for a given set of connection strengths, the average dwell times can be made arbitrarily long (short) by making the variance of the input signal small (large).

3.2 The power spectrum

It is well known that shared inhibition can result in winner-take-all behavior. With Fig. 2 we showed where these WTA solutions exist and how they arise (via a fold bifurcation). Of interest to us, as a link to EEG recordings in human subjects, are the spectral quantities predicted by the model while in both the symmetric and WTA regimes. Toward this end, we calculate the spectral information from the model. Of primary interest are power spectral density and coherence. We utilize two methods to calculate spectral quantities, though we will show that the linear analysis presented here is adequate for our purposes.

Because our goal is comparison to EEG recordings, we use the output from excitatory cortex. We designate, without loss of generality, one of the modules to be module A, and the other as module B. We linearize the system (6) around the fixed point of interest, take the Fourier transform and solve to obtain the transfer functions

$$L_{A,B}(\omega) = \frac{\hat{\phi}_{E,A}(\omega)}{\hat{\phi}_{N,B}(\omega)} \quad (15)$$

where $\hat{\phi}_{E,A}(\omega)$ is the Fourier transform of ϕ_E in module A, and $\hat{\phi}_{N,B}(\omega)$ is the Fourier transform of the noisy input (mean subtracted), to the relay population of module B. There are two inputs (the noise inputs via S_A and S_B) and two outputs ($\phi_{E,A}$ and $\phi_{E,B}$) and so we have a total of four transfer functions. We can use these

transfer functions to approximate the power spectral density (PSD)

$$|\gamma_A(\omega)|^2 = |L_{A,A}(\omega)|^2 + |L_{A,B}(\omega)|^2 \tag{16}$$

and if we assume that the noisy input is white noise, the cross spectrum is approximated by

$$CS(\omega) = L_{A,A}(\omega)L_{B,A}^*(\omega) + L_{A,B}(\omega)L_{B,B}^*(\omega) \tag{17}$$

where the superscript $*$ denotes the complex conjugate. The coherence, therefore, is given by

$$COH(\omega) = \frac{CS(\omega)}{\sqrt{|\gamma_A(\omega)|^2}\sqrt{|\gamma_B(\omega)|^2}} \tag{18}$$

Note that both the cross-spectrum and the coherence have non zero imaginary components. Therefore, the coherence has both an amplitude, given by the square root of the modulus, and a phase, given by the angle in the complex plane.

We can also compute approximations to the PSD and coherence for time series obtained by simulating the full model. We use the multitaper method for spectral analysis developed by Thompson ([1982]) with $\sigma_E^2 = 0.01$. For each type of solution we obtain both the linear spectral analysis and the approximate analysis using time series from the full model. These are shown in
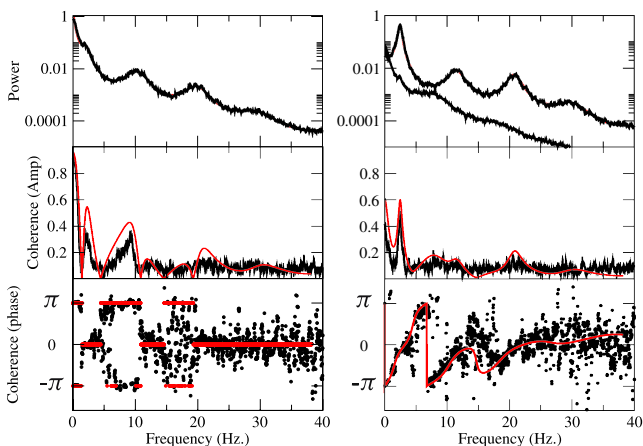

Fig. 5 The power spectral density and coherence. The *horizontal axis* for each of the plots is frequency, in Hz. The *left column* shows the power spectrum and coherence corresponding to the symmetric solution, and the *right column* shows the same information for the WTA solution. In each plot, the *black lines* represent the multitaper estimates and the *red lines* represent the linear approximations from Eqs. ([16])–([18]). The *top panel* in each column is the power spectral density, the *middle panel* is the amplitude of the coherence and plot at the *bottom* of each column is the phase of the coherence. The power spectral density plots demonstrate the difference in fluctuations between the winner and the loser. The coherence plots show that there is a drop off in the level of coherence when a transitions occurs from the WTA state to the symmetric state

Fig. 5. One can see from these figures that the spectral information drawn from the linear model is extremely close to the multitaper results for the small variance noise used.

When the system is near the symmetric solution, the spectral data are not surprising. The power spectral density is the same for each of the modules. The coherence is high at the largest spectral peak. The phase of the coherence is either zero or π for all frequencies, suggesting that neither of the modules leads or lags. The WTA solutions provide different results. Of particular interest, especially when comparing to human data, is the power difference between cortical modules when the system is in WTA mode. The winner has higher power level at all relevant frequencies (at the peaks). The slope of the coherence phase for WTA solutions suggests that the phase of the winner leads that of the loser. Perhaps the most relevant aspect of these results is the difference in coherence. We can see from Fig. 5 that there is a large drop in coherence in the 6–10 Hz range when the system switches from the symmetric state to the WTA state, as in Fig. 4.

3.3 Robustness to changes in intra-module connectivity

We have shown that there are a variety of different behaviors possible for the parameters given in Table 1, sliding only those connections to and from the reticular nucleus. Here we show that these behaviors are not specific to the values chosen for the other connections strengths in Table 1, which characterize the connectivity between the relay nuclei and the cortex. We use two complementary approaches: a theoretical argument based on continuity, and an empirical exploration of the parameter space.

Points in the bifurcation diagram bounded away from the bifurcation curves (i.e. Hopf curve, fold curve, pitchfork curve) correspond to hyperbolic fixed points—fixed points where the linearization has eigenvalues bounded away from the imaginary axis (zero real part). The right hand sides of the model ([6]) are continuously differentiable in each of the variables, and so there is an open ball around the parameter set given in Table 1 where the system will behave in a manner that is topologically equivalent to the original. This means that small enough perturbations to the parameters will generate roughly the same bifurcation diagram. This includes sufficiently small asymmetries between the modules. In the case of asymmetries in the model parameters for each module, there will no longer be a symmetric solution, but rather an asymmetric perturbation that separates smoothly as the parameters are varied.

Table 3 Connection parameters

	EO	EC	S2	S3
ν_{EE}	1.7	1.3	1.8	1.8
ν_{ES}	1.2	1.2	1.7	1.7
ν_{SE}	1.0	1.0	0.7	0.4
ν_{SR}	−1.0	−1.0	−0.8	−0.6
ν_{RS}	0.2	0.2	0.4	0.4
ν_{RE}	0.4	0.2	0.2	0.5
α	100.0	60.0	60.0	40.0

The topological equivalence near hyperbolic fixed points guarantees behavior for local subsets of parameter space. We want to know whether the behaviors exist for larger regions in parameter space. To determine the effect of larger changes in parameters, we simply repeat the bifurcation analysis for other parameter sets. We show the diagrams for two such parameter sets, given in Table 3, in Fig. 6. These parameter sets correspond to single-module parameters for the eyes-closed state (EC) and stage 2 sleep (S2), in terms of comparison to the spectral features of the EEG (Debellis 2007).

As shown, the diagrams in Fig. 6(b) and (c) are qualitatively identical to that shown in Fig. 2, and reproduced in Fig. 6(a). The key features are the generation of WTA solutions via the fold bifurcation (blue), the destabilization of the symmetric state at a pitchfork bifurcation, and the region of multi-stability that is bookended by these bifurcations. There are also curves of Hopf bifurcation points, both for the symmetric

and WTA solutions. Thus, the existence of the various features found for the EO parameters is not a local phenomenon, nor is it specific to the parameters in Table 1.

Qualitatively, the parameters in Table 3 differ from those in Table 1 in the thalamocortical loop, changing the excitatory feedback within cortex (lower for EC, higher for S2) as well as the connections to and from the relay nuclei. The balance between reticular connections to and from the cortex and those to and from the relay nuclei is also modified in each. The rise time parameter (α) is also changed, though this parameter does not impact the existence of the fixed point solutions.

We have shown that these features, specifically the multistable region, occur for a large range of parameters. It is important to note that they are not ubiquitous. There are intra-module parameter sets that do not admit bistability in the $\kappa_s - \kappa_u$ plane. An example is given in the S3 (stage 3 sleep) column of Table 3. The destruction of the bistable region is the result of the fold bifurcation colliding with the pitchfork, switching the pitchfork from subcritical to supercritical. This process is shown in Fig. 7. Crossing the curve of supercritical pitchfork points from left to right destabilizes the symmetric solution and brings into existence WTA states. The stable WTA states continuously exit symmetry from left to right, rather than quickly as they do with the subcritical pitchfork (see Fig. 8).
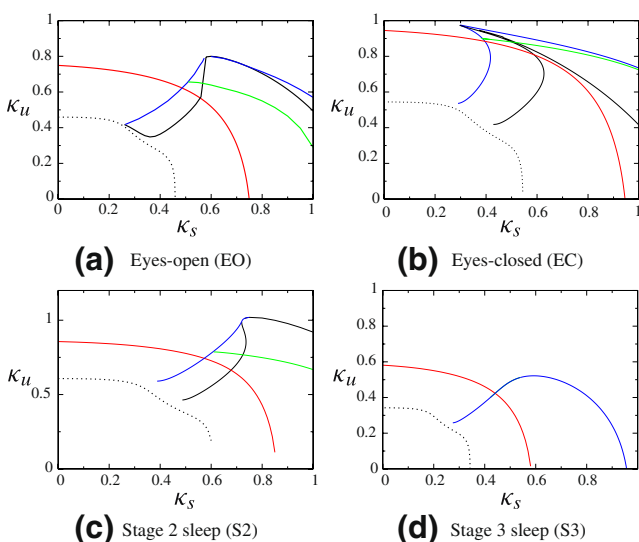

Fig. 6 Bifurcation diagrams for the parameter sets given in Table 3. Additional parameters common to all diagrams can be found in Tables 1 and 2. Figure 6(a) is identical to Fig. 2, and is reproduced here to facilitate comparison. The bifurcation curves are color coded as in Fig. 2, with the exception of Fig. 6(d), where the *blue line* indicates a *supercritical* pitchfork bifurcation
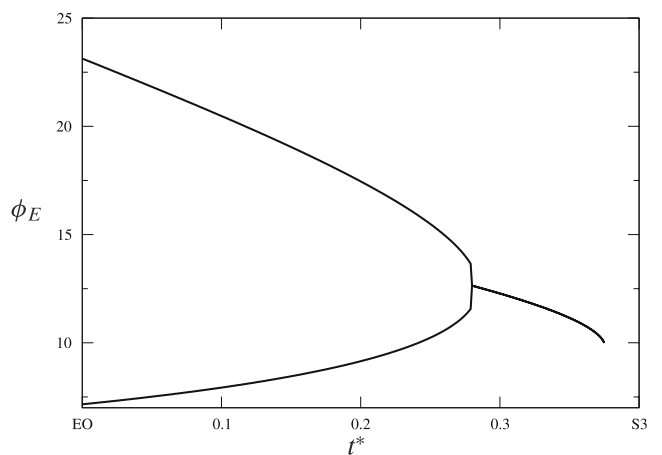

Fig. 7 Elimination of multi-stability. The *y*-axis shows the activity level at the elbow (the point where the stable and unstable WtA solutions collide, to the left of which there is no WTA solution—see Fig. 3) formed by the fold bifurcation. The *x*-axis slides the parameters from the values in Table 1 ($t^* = 0$) to the values in the S3 column of Table 3 ($t^* = 1$) linearly. The elbows of the fold bifurcation collapse onto the pitchfork bifurcation. In this figure, $\kappa_u = 0.7$, and so the pitchfork bifurcation eventually disappears since it does not exist for this value of κ_u in the S3 state (see Fig. 6(d))
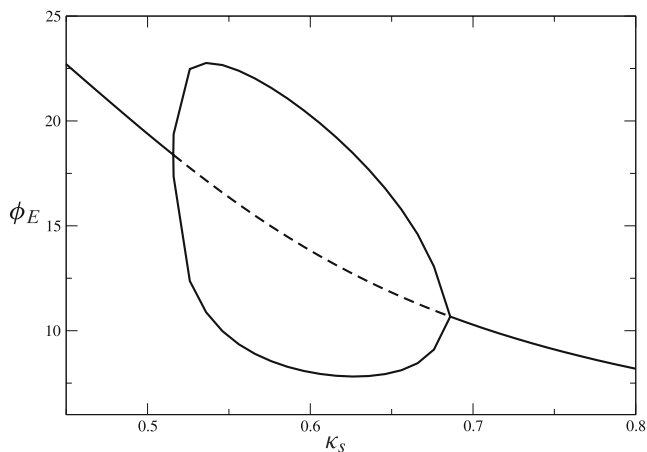
Fig. 8 The super-critical pitchfork bifurcations that result after the fold has been absorbed. This figure corresponds to the $\kappa_u = 0.5$ slice in Fig. 6(d). The *solid black curves* indicate stable fixed points, *dotted black lines* indicate unstable fixed points. The difference in amplitude of the WTA solutions begins at zero upon crossing the pitchfork bifurcation from left to right. Compare this to the sub-critical case (Fig. 3) where the difference between the winner and the loser is bounded below by some nonzero number

3.4 Other forms of coupling between the modules

The major claim of this paper is that the shared inhibition from the reticular nucleus is responsible for the existence of WTA solutions. As we have shown above, thalamocortical modules connected via shared inhibition from the reticular nucleus exhibit coexistence of symmetric and WTA solutions. To show that the nature of the coupling is critical we repeat the analysis using a shared relay nucleus instead of a shared reticular nucleus to couple the modules together.

To do this, we alter the model to incorporate a shared relay nucleus instead of the shared population of reticular neurons. We use the parameter set shown in Tables 1 and 3, with the κ's moved so that they scale the connections to and from the unshared and shared relay nuclei. The resulting bifurcation diagram is featureless (not shown). The symmetric solution does not destabilize for any value of the coupling. We were also unable to locate any fold bifurcations, the mechanism that creates the winner take all solutions when shared inhibition is present. We conclude that the WTA behaviors, and hence the multi-stability, are not possible when modules are connected via shared excitation.

Direct cortical connectivity is not included in this model. Nevertheless, we can take the effects of these connections into account because they are relatively weak (Barbas and Rempel-Clower 1997; Rigas and Castro-Alamancos 2007) compared to connections via the thalamus. One can consider the model as shown

above, with thalamic coupling, the limiting case for weak direct cortical connection strength. Nonzero corticocortical connection strength amounts to a perturbation of the results found for this limiting case. Thus, we expect that inclusion of these connections will not qualitatively change the results for either the shared reticular or the shared relay models. For coupling via the reticular nucleus, this implies that the WTA solutions will be maintained for sufficiently weak direct connection strengths. For coupling via a relay nucleus, one does not anticipate the generation of stable WTA solutions due to the presence of direct cortical connections.

4 Discussion

In this paper, we develop a population model for thalamocortical dynamics. Our motivation was to gain insight into the role of the thalamus in nonlocal corticocortical interactions, and to make contact with EEG recordings. For these reasons, the appropriate level of detail is at the population level, as opposed to individual neurons. We use the results of Robinson et al. to form the basic building block of our model, the thalamocortical module. We couple two of these modules together using a shared population of inhibitory thalamic neurons, for example located within the reticular nucleus. For comparison, we also consider the case where the coupling is purely excitatory, where the connecting population is a shared population of relay neurons.

Our goal in this paper was to expand upon the success of the models presented in David and Friston (2003), Rennie et al. (2002), and Robinson et al. (1998, 2002) with a focus on what is predicted about interactions between the populations. Our primary result is the characterization of possible behaviors in the model as a function of the coupling strengths. In particular, we classify two types of fixed point solution—a symmetric solution where both modules have the same activity level, and a winner-take-all solution that is characterized by a large disparity in the modules' activity level. The existence and stability of each of these, in the case of shared reticular populations, is altered by a combination of a fold bifurcation and a pitchfork bifurcation. For some parameter values, a WTA solution and a symmetric solution are both stable, allowing multi-stability. Since we are interested in EEG characteristics, we compute the power spectral densities for each of these types of solutions. We wanted to determine, qualitatively, how these different states manifest themselves in the power spectrum, as well as the coherence between distinct cortical populations.

We repeat the analysis with a shared relay nucleus, replacing the shared reticular nucleus. The functional difference between the two approaches is that the shared relay nuclei provide shared excitation, while the shared reticular population coupled using shared inhibition. We conclude that the shared inhibition, afferent (via relay nuclei) to cortical areas large distances apart, is necessary to have these winner take all solutions exist and be stable.

There are also direct cortical connections in the cortex, the myelinated white matter tracts. These connections are known to be weak relative to the thalamocortical connections. We view our model as the limiting case of weak direct cortical connections. Inserting these connections can be viewed as a perturbation to the model, where sufficiently small connection strengths will admit a topologically equivalent to those shown here.

The biological importance of the multi-modality arises in a comparison with experimental data (Freyer et al. 2009; Victor et al. 2009). In Victor et al. (2009), EEG recordings show that in some awake subjects microstates, or seemingly random jumps from one state to another, occur within a hemisphere. In Freyer et al. (2009) the authors show that there are fast switches between high and low amplitude modes in the alpha rhythm. For our model, multi-modal solutions are manifest in the coherence between distant electrodes, which is very different for symmetric and WTA states in the 6–10 Hz. range. These transitions occur on a much faster timescale than those corresponding to wakefulness and the various stages of sleep. How the presence of these transitions relates to the physiological status of the thalamocortical network is a subject of current investigation.

As with any neural model, our model is a simplified picture. One aspect of this simplification is that we model at the population level, and do not consider specific ionic conductances. An example is the post inhibitory rebound mechanism for thalamic synchrony of spindle oscillations (Destexhe et al. 1998) due to a T-current in the relay cells. These same ionic conductances play a role in cortical feedback that promotes thalamic synchrony (Destexhe et al. 1999). Nevertheless, the above modeling studies (Destexhe et al. 1998, 1999) support the use of a population-based model, because they show that the behavior of typical neurons within the population is represented by the average over the population.

Another aspect of our simplification is that not all anatomical connections are present in the model. The network modeled in Fig. 1 emphasizes a privileged component of the reticular nucleus that interacts with two relay nuclei that each have specialized reticular components that are not shared. This situation approximately models possible frontal cortical gating of reticular neurons (Crick 1984; Skinner and Yingling 1977). The network topology, however, does not include a specialized thalamic relay with disynaptic connection to another relay nucleus through the reticular nucleus (Crabtree and Isaac 2002), a known connection that may also influence coupling of modules. The cortical excitatory projections in this model do not distinguish between projections to both reticular and relay populations. Jones (2009) notes that such distinctions are present. While layer VI cells send collateral fibers to reticular neurons, layer V cells have a diffuse projection to relay and intralaminar nuclei that is thought to play a key role in large-scale synchronization. These connections and specificities are not modeled here and may change the dynamics when included.

The model suggests that the balance between shared and unshared inhibitory thalamic populations determines the modes of intracortical coherence patterns within an overall brain state, and enables spontaneous switching among these modes. These findings suggest a role for the reticular nucleus in modulating long distance cortical communication.

References

Barbas, H., & Rempel-Clower, N. (1997). Cortical structure predicts the pattern of corticocortical connections. *Cerebral Cortex, 7*, 635–646.

Crabtree, J. W., & Isaac, J. T. (2002). New intrathalamic pathways allowing modality-related and cross-modality switching in the dorsal thalamus. *Journal of Neuroscience, 22*(19), 8754–8761.

Crick, F. (1984). Function of the thalamic reticular complex: The searchlight hypothesis. *Proceedings of the National Academy of Sciences of the United States of America, 81*, 4586–4590.

Davey, M. P., Victor, J. D., & Schiff, N. D. (2000). Power spectra and coherence in the eeg of a vegetative patient with severe asymmetric brain damage. *Clinical Neurophysiology, 111*, 1949–1954.

David, O., & Friston, K. J. (2003). A neural mass model for meg/eeg coupling and neuronal dynamics. *NeuroImage, 20*, 1743–1755.

Debellis, R. (2007). *Dynamics of a mathematical model of thalamocortical interactions and implications on arousal and sleep states*. PhD thesis, Weill Medical College of Cornell University.

Destexhe, A., Contreras, D., & Steriade, M. (1998). Mechanisms underlying the synchronizing action of corticothalamic feedback through inhibition of thalamic relay cells. *Journal of Neurophysiology, 79*, 999–1016.

Destexhe, A., Contreras, D., & Steriade, M. (1999). Cortically-induced coherence of a thalamic-generated oscillation. *Neuroscience, 92*(2), 427–443.

Engelborghs, K., Luzyanina, T., & Roose, D. (2002). Numerical bifurcation analysis of delay differential equations using dde-biftool. *ACM Transactions on Mathematical Software, 28*(1), 361–385.

Freyer, F., Aquino, K., Robinson, P. A., Ritter, P., & Breakspear, M. (2009). Bistability and non-gaussian fluctuations in spontaneous cortical activity. *The Journal of Neuroscience, 29*(26), 8512–8524.

Huguenard, J. R., & McCormick, D. A. (2007). Thalamic synchrony and dynamic regulation of global forebrain oscillations. *Trends in Neurosciences, 30*(7), 350–356.

Jones, E. G. (2009). Synchrony in the interconnected circuitry of the thalamus and cerebral cortex. *Disorder of Consciousness: Annals of the New York Academy of Sciences, 1157*, 10–23.

Pesaran, B., Nelson, M. J., & Anderson, R. A. (2008). Free choice activates a decision circuit between frontal and parietal cortex. *Nature, 453*, 406–409.

Pesaran, B., Pezaris, J. S., Sahani, M., Mitra, P. P., & Anderson, R. A. (2002). Temporal structure in neuronal activity during working memory in macaque parietal cortex. *Nature Neuroscience, 5*, 805–811.

Raos, V., & Bentivoglio, M. (1993). Crosstalk between the two sides of the thalamus through the reticular nucleus: A retrograde and anterograde tracing study in the rat. *Journal of Comparative Neurology, 332*(2), 145–154.

Rennie, C. J., Robinson, P. A., & Wright, J. J. (2002). Unified neurophysical model of eeg spectra and evoked potentials. *Biological Cybernetics, 86*, 457–471.

Rigas, P., & Castro-Alamancos, M. A. (2007). Thalamocortical up states: Differential effects of intrinsic and extrinsic cortical inputs on persitent activity. *Journal of Neuroscience, 27*, 4261–4272.

Robinson, P. A., Rennie, C. J., & Rowe, D. L. (2002). Dynamics of large-scale brain activity in normal arousal states and eplieptic seizures. *Physical Review, E, 65*, 041924.

Robinson, P. A., Rennie, C. J., Wright, J. J., & Bourke, P. D. (1998). Steady states and global dynamics of electrical activity in the cerebral cortex. *Physical Review, E, 58*(3), 3557–3571.

Rodriguez, E., et al. (1999). Perception's shadow: Long-distance synchronization of human brain activity. *Nature, 397*, 430–433.

Schiff, N. D., & Purpura, K. P. (2002). Towards a neurophysiological foundation for cognitive neuromodulation through deep brain stimulation. *Thalamus and Related Systems, 2*, 55–69.

Skinner, J. E., & Yingling, C. D. (1977). Gating of thalamic input to the cerebral cortex by nucleus reticularis thalami. *Progress in Clinical Neurophysiology, 1*, 70–96.

Thompson, D. J. (1982). Spectrum estimation and harmonic analysis. *Proceedings of the IEEE, 70*, 1055–1096.

Truex, R. C., & Carpenter, M. B. (1969). *Human neuronanatomy*. The Williams & Wilkins Company.

Velayos, J. L., Jimenez-Castellanos, Jr., J., & Reinoso-Suarez, F. (1989). Topographical organization of the projections from the reticular thalamic nucleus to the intralaminar and medial thalamic nuclei in the cat. *Journal of Comparative Neurology, 279*(3), 457–469.

Victor, J. D., Williams, S. T., Conte, M. M., Drover, J. D., & Schiff, N. D. (2009). Fluctuating EEG coherence in chronic brain injury. (Abstract) Program No. 541.1.2009 Neuroscience Meeting Planner. Chicago, IL: Society for Neuroscience, 2009. Online.

Wiggins, S. (1990). *Introduction to applied nonlinear dynamical systems and chaos*. Heidelberg: Springer.