



Computational Neuroscience

Real-time automated EEG tracking of brain states using neural field theory

R.G. Abeysuriya ^{a,b,c,d,*}, P.A. Robinson ^{a,b,c,d}

^a School of Physics, University of Sydney, New South Wales 2006, Australia

^b Center for Integrative Brain Function, University of Sydney, New South Wales 2006, Australia

^c Neurosleep, 431 Glebe Point Rd, Glebe, New South Wales 2037, Australia

^d Brain Dynamics Center, Sydney Medical School – Western, University of Sydney, Westmead, New South Wales 2145, Australia

HIGHLIGHTS

- A real-time automated fitting system is developed to fit a neural field model to EEG.
- Inferred physiological parameters are objectively tracked over the sleep-wake cycle.
- Continuous trajectories supersede discrete Rechtschaffen–Kales sleep stages.

ARTICLE INFO

Article history:

Received 21 July 2015

Received in revised form

13 September 2015

Accepted 16 September 2015

Available online 31 October 2015

Keywords:

EEG

Neurophysiology

Modeling

Sleep

Neural field theory

ABSTRACT

A real-time fitting system is developed and used to fit the predictions of an established physiologically-based neural field model to electroencephalographic spectra, yielding a trajectory in a physiological parameter space that parametrizes intracortical, intrathalamic, and corticothalamic feedbacks as the arousal state evolves continuously over time. This avoids traditional sleep/wake staging (e.g., using Rechtschaffen–Kales stages), which is fundamentally limited because it forces classification of continuous dynamics into a few discrete categories that are neither physiologically informative nor individualized. The classification is also subject to substantial interobserver disagreement because traditional staging relies in part on subjective evaluations. The fitting routine objectively and robustly tracks arousal parameters over the course of a full night of sleep, and runs in real-time on a desktop computer. The system developed here supersedes discrete staging systems by representing arousal states in terms of physiology, and provides an objective measure of arousal state which solves the problem of interobserver disagreement. Discrete stages from traditional schemes can be expressed in terms of model parameters for backward compatibility with prior studies.

© 2015 Elsevier B.V. All rights reserved.

1. Introduction

Neural physiology and arousal state change significantly and continuously over the course of the sleep–wake cycle, but arousal state is typically analyzed using the Rechtschaffen and Kales (R&K) or American Academy of Sleep Medicine (AASM) classification schemes (Rechtschaffen and Kales, 1968; Iber et al., 2007). These schemes artificially force classification of continuous dynamics into a small selection of discrete population-averaged stages: wake (W); stage 1 sleep (called S1 in R&K, N1 in AASM), which corresponds to light sleep; stage 2 sleep (called S2 in R&K, N2 in AASM),

which is a deeper stage of sleep marked by K-complexes (transient waveforms typically marked by a large negative peak in the EEG, followed by a positive peak, similar to an evoked response) and sleep spindles (short bursts of activity at around 12–14 Hz); slow wave sleep (called S3 and S4 in R&K, N3 in AASM), which corresponds to deep sleep in which K-complexes and sleep spindles are sometimes present; and rapid eye movement (REM) sleep, which occurs during dreaming.

Although sleep stages can provide a useful qualitative summary, they have serious deficiencies when used to analyze brain states, dynamics, and physiology for several reasons (Abeysuriya et al., 2015). Real brain states vary continuously and cannot be accurately captured by a few discrete stages, and the small number of traditional sleep stages results in a wide range of different brain substates being grouped together into the same sleep stage. Traditional stages are also based on group averages of EEG and other

* Corresponding author at: School of Physics, University of Sydney, New South Wales 2006, Australia. Tel.: +61 2 9036 7274.

E-mail address: r.abeysuriya@physics.usyd.edu.au (R.G. Abeysuriya).

polysomnographic features, and do not account for the significant individual variation seen in experimental data. In some cases, subjective decisions contribute to determining the sleep stage, which adversely affects the objectivity and validity of the assigned sleep stage. This is reflected in the low interobserver agreement for the classified sleep stages; for AASM staging agreement is only 83% (Rosenberg and Van Hout, 2013), and Norman et al. (2000) reported interobserver agreement of just 73% for R&K staging. Finally, traditional stages are typically assigned to 30-s epochs based on the EEG features present within that epoch (e.g., sleep spindles), so the classification can be quite sensitive to the arbitrary timing of the epoch boundaries because this affects which epoch an EEG feature is assigned to.

Note that throughout this study we use the term 'state' to refer to the physiological state of the brain at an instant in time, and the term 'stage' to refer to R&K or AASM classifications. We relate each state to a single set of underlying physiological parameters in our model. Brain states evolve continuously (notwithstanding transitions between sleep and wake, which are rapid but still continuous) and are linked by trajectories in the parameter space, whereas assigned sleep stages change discontinuously and instantaneously.

The issues with sleep staging are illustrated in Fig. 1 (Abeysuriya et al., 2015). In Fig. 1(a), evolving brain states are represented schematically in terms of physiology, and continuous trajectories. In Fig. 1(b), traditional sleep stages are superimposed on the trajectories. In this frame, the stages are represented in terms of physiology because there are quantitative parameters associated with the trajectories, although the definitions of the stages from AASM or R&K are not based on physiology. The significant overlap between the stages reflects the fact that a single combination of parameters can be consistent with more than one assigned arousal stage. In Fig. 1(c), the arousal stages have been decoupled from the underlying physiology, and the degree of overlap between the stages can only be inferred via interobserver disagreement. Finally, Fig. 1(d) shows the current common usage of sleep staging, where each epoch of EEG is forced to be classified into a single sleep stage. Thus the true continuous trajectories in Fig. 1(a) have been replaced by discrete jumps between artificially defined stages, losing information about the physical processes underlying the change in brain state and resulting in inconsistency in assignment of stages (Abeysuriya et al., 2015).

Our central aim is to represent brain states using physiologically meaningful trajectories rather than sequences of arbitrary and unphysiologically discrete stages. In previous work, we showed that the physiologically meaningful parameters of an established neural field corticothalamic model (Rowe et al., 2004b; Robinson et al., 2001, 2002, 2004, 2005) are suitable quantities to reproduce Fig. 1(b) quantitatively (Abeysuriya et al., 2015). Moreover, the model reproduces a wide range of other phenomena such as the alpha rhythm (Robinson et al., 2003; O'Connor and Robinson, 2004), age-related changes to the physiology of the brain (van Albada et al., 2010), evoked response potentials (Rennie et al., 2002), sleep spindles (Abeysuriya et al., 2014a,b), and many other effects. Predictions from the model can be fitted to EEG spectra to estimate physiological parameters (van Albada et al., 2010, 2007; Rowe et al., 2004b; Robinson et al., 2003a, 2005), and these estimates are consistent with a range of EEG-related phenomena (Robinson et al., 2004; Rowe et al., 2004b). Overall, this represents a unified approach to brain dynamics, unlike traditional sleep staging which exists in isolation.

In a previous study (Abeysuriya et al., 2015) we investigated the relationship between sleep stages and the model's physiological parameter space for a population of subjects, corresponding to the colored regions in Fig. 1. Although it is important to be able to understand and interpret traditional sleep stages in terms of our new framework, ultimately it is the individual parameter

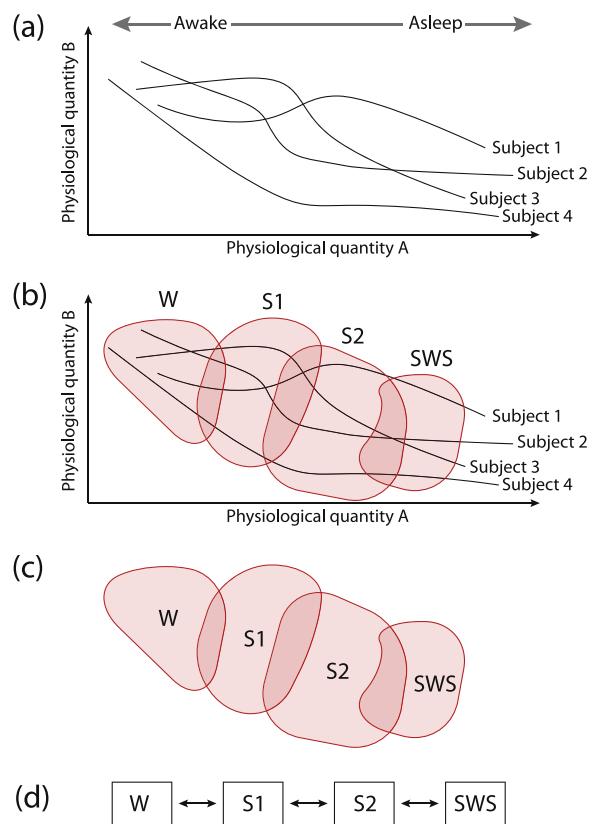


Fig. 1. Schematic illustration showing how physiological brain states are related to traditional sleep stages. (a) Brain states are differentiated by their physiology. Two quantities are shown here for clarity. Over time, brain states follow continuous trajectories. Both the states and the shape of the trajectories are individualized. (b) Traditional sleep stages are superimposed on the trajectories, showing their association with the underlying physiology. The overlap between stages can be quantified in terms of physiology. (c) Removing the physiological axes and trajectories shows only the sleep stages, without reference to the underlying physiology, but still acknowledging the overlap between stage assignments. (d) Common use of traditional sleep stages, with discrete classifications that permit no overlap. The arrows between the stages correspond to discrete jumps, that are the discrete analogs of the trajectories in (a). From Abeysuriya et al. (2015) with permission from Elsevier.

trajectories that are fundamental and take full advantage of our model-based approach, while discrete stages must be abandoned. In this study, we develop a real time, automated approach for fitting the model to an experimental EEG power spectrum and present a first analysis of full-night parameter trajectories. In other recent work, Dadok et al. (2014) examined fitting and tracking neural field parameters to sleep EEG data. Their work fitted two parameters (cortical excitatory feedback strength, and change in resting potential of cortical excitatory neurons) of a purely cortical model (Steyn-Ross et al., 2005) to features extracted from the EEG, and used a hidden Markov model to incorporate temporal continuity of brain states. Because their work did not provide a closed-form analytic expression for the EEG spectrum, their model is computationally expensive to simulate, and produces a stochastic output that has different EEG features each time it is run. In contrast, our approach fits an analytic power spectrum to the EEG spectrum directly, which removes the need to choose a set of extracted features for fitting, scales efficiently as the number of parameters is increased, and enables rapid comparison of different models. In Section 2 we present a brief overview of the model and its key elements. In Section 3 we develop our fitting approach, first for a single EEG power spectrum, and then for tracking the state over time. Finally, in Section 4 we analyze the fitted trajectories to

investigate the changes in the model parameters that take place over the sleep–wake cycle and their physiological interpretation.

2. Theory

In this section we present a brief overview of the relevant aspects of our neural field model (van Albada et al., 2010, 2007; Roberts and Robinson, 2012; Robinson et al., 2004, 2002, 1997; Abeysuriya et al., 2015).

2.1. Neural field model

Our model of the corticothalamic system consists of four neural populations with key connectivities shown schematically in Fig. 2 – the model has populations of excitatory (*e*) and inhibitory (*i*) cortical neurons, thalamic relay neurons (*s*), thalamic reticular neurons (*r*), and external sensory inputs (*n*).

Each population *a* has a mean firing rate Q_a which is related to the corresponding mean soma potential V_a , relative to resting, by a nonlinear sigmoid function that encapsulates the firing threshold, the threshold spread within a population of neurons, and saturation

$$Q_a = S(V_a) = \frac{Q_{\max}}{1 + \exp[-(V_a - \theta)/\sigma']}, \quad (1)$$

where θ is the mean threshold voltage, $\sigma'/\pi/\sqrt{3}$ is the standard deviation of the threshold distribution, and Q_{\max} is the maximum firing rate.

The potential V_a is dependent on contributions ϕ_b from presynaptic populations, the strength of the connections to populations *a* from *b*, and smoothing effects arising from synaptodendritic dynamics and soma capacitance, with

$$V_a(\mathbf{r}, t) = \sum_b V_{ab}(\mathbf{r}, t), \quad (2)$$

$$D_\alpha(t)V_{ab}(\mathbf{r}, t) = v_{ab}\phi_b(\mathbf{r}, t - \tau_{ab}), \quad (3)$$

$$D_\alpha(t) = \frac{1}{\alpha\beta} \frac{d^2}{dt^2} + \left(\frac{1}{\alpha} + \frac{1}{\beta}\right) \frac{d}{dt} + 1. \quad (4)$$

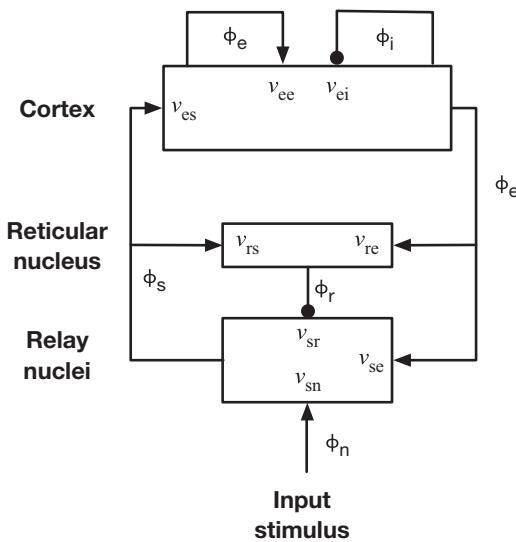


Fig. 2. Schematic diagram of our corticothalamic neural field model of the system. The neural populations shown are cortical excitatory *e* and inhibitory *i*, thalamic reticular *r* and thalamic relay nuclei *s*. The parameter v_{ab} quantifies the connection from population *b* to population *a*. Excitatory connections are shown with pointed arrowheads, and inhibitory connections are shown with round arrowheads. From Abeysuriya et al. (2015) with permission from Elsevier.

The quantity v_{ab} is defined as $v_{ab} = s_{ab}N_{ab}$ where N_{ab} is the mean number of synapses per neuron *a* from neurons of type *b*, and s_{ab} is the mean time-integrated strength of soma response per incoming spike (Rennie et al., 1999; Robinson et al., 2004; Roberts and Robinson, 2012). All connections in the model are excitatory with the exception of v_{ii} , v_{ei} , and v_{sr} , which thus have negative values.

The parameters $1/\beta$ and $1/\alpha$ are the characteristic rise time and decay time, respectively, of the response at the cell body. Although these can be different for each of the connections in the model, we take a weighted average over the corticothalamic system that provides a single effective value for α and β while preserving the key aspects of the system dynamics (Robinson et al., 2004). The time delays τ_{ab} arise due to discrete anatomical separations between populations. The only nonzero time delays in our model are $\tau_{es} = \tau_{is} = \tau_{se} = \tau_{re} = t_0/2$ which correspond to the corticothalamic and thalamocortical propagation times. The time t_0 is the total time delay for the corticothalamic feedback loop.

Each population generates outgoing action potentials at an average rate that propagates as a field $\phi_a(\mathbf{r}, t)$ when averaged over scales of order 0.1 mm. To a good approximation, ϕ_a obeys a damped wave equation with a source Q_a (Robinson et al., 1998, 2002, 1997), with

$$D_a(\mathbf{r}, t)\phi_a(\mathbf{r}, t) = Q_a, \quad (5)$$

$$D_a(\mathbf{r}, t) = \frac{1}{\gamma_a^2} \frac{\partial^2}{\partial t^2} + \frac{2}{\gamma_a} \frac{\partial}{\partial t} + 1 - r_a^2 \nabla^2, \quad (6)$$

where v_a is the propagation velocity, r_a is the mean range, and $\gamma_a = v_a/r_a$ is the temporal damping rate. In our model, only r_e is large enough to produce significant propagation effects. For the other populations, Eqs. (5) and (6) can be approximated by $\phi_a(\mathbf{r}, t) = S[V_a(\mathbf{r}, t)]$ (Robinson et al., 1997).

We make several simplifications to the model equations: First, we only include the connections shown in Fig. 2, so only 10 of the possible 16 connections between the 4 neural populations are nonzero. Second, we assume random intracortical connectivity so the number of connections between populations is proportional to the number of synapses (Robinson et al., 1998; Wright and Liley, 1994; Liley and Wright, 1994; Braatenberg and Schüz, 1998). This random connectivity assumption provides $N_{ib} = N_{eb}$ for all *b*, which implies that the connection strengths are also symmetric, with $v_{ee} = v_{ie}$, $v_{ei} = v_{ii}$ and $v_{es} = v_{is}$ (Robinson et al., 1998, 2004), giving the eight independent connection strengths shown in Fig. 2. The model equations (1)–(6) can be integrated numerically to obtain the EEG time series.

2.2. EEG power spectrum

Spatially uniform steady states of the model can be obtained by setting all of the time and space derivatives in Eqs. (1)–(6) to zero. The steady state value $\phi_e^{(0)}$ of ϕ_e is then given by solutions of

$$\begin{aligned} & S^{-1}(\phi_e^{(0)}) - (\nu_{ee} + \nu_{ei})\phi_e^{(0)} \\ &= \nu_{es}S \left\{ \nu_{se}\phi_e^{(0)} + \nu_{sr}S \left[\nu_{re}\phi_e^{(0)} + (\nu_{rs}/\nu_{es}) \right. \right. \\ & \quad \times \left. \left. \left\{ S^{-1}(\phi_e^{(0)}) - (\nu_{ee} + \nu_{ei})\phi_e^{(0)} \right\} \right] + \nu_{sn}\phi_n^{(0)} \right\}, \end{aligned} \quad (7)$$

where $\phi_n^{(0)}$ is the steady state component of the input stimulus (Robinson et al., 2004).

Previous work has revealed most normal brain states occur in the linear regime of the model (van Albada et al., 2010, 2007; Roberts and Robinson, 2012; Robinson et al., 2004, 2002, 1997, 2003, 2001, 2005; O'Connor and Robinson, 2004; Rowe et al.,

2004b). In this case, we only retain the first term in the Taylor expansion of Eq. (1), giving

$$Q_a(\mathbf{r}, t) - Q_a^{(0)} = S'(V_a^{(0)})[V_a(\mathbf{r}, t) - V_a^{(0)}]. \quad (8)$$

where S' is the derivative of the sigmoid function with respect to its argument. From this point on we will consider only perturbations from steady states, because in our model these are what give rise to time-dependent EEG signals in healthy resting brain states (Robinson et al., 1997, 1998, 2002). Hence, we simplify our notation by treating each dynamic quantity (ϕ_a, Q_a, V_a) as a perturbation from its resting state value. Eqs. (2), (3), and (5) are linear and are therefore unchanged aside from this reinterpretation. The linear approximation to (1) is

$$Q_a(\mathbf{r}, t) = \rho_a V_a(\mathbf{r}, t), \quad (9)$$

with $\rho_a = S'(V_a^{(0)})$. By taking the Fourier transform of Eqs. (2)–(6) and (9), we can express the firing rate ϕ_e in terms of the external signal ϕ_n . In particular, we find (Robinson et al., 2002)

$$\frac{\phi_e(\mathbf{k}, \omega)}{\phi_n(\mathbf{k}, \omega)} = \frac{G_{es}G_{sn}L^2 \exp(i\omega t_0/2)}{(1 - G_{srs}L^2)(1 - G_{ei}L)(k^2 r_e^2 + q^2 r_e^2)}, \quad (10)$$

$$q^2 r_e^2 = \left(1 - \frac{i\omega}{\gamma_e}\right)^2 - \frac{1}{1 - G_{ei}L} \left\{ L G_{ee} + \frac{[L^2 G_{ese} + L^3 G_{esre}] e^{i\omega t_0}}{1 - L^2 G_{srs}} \right\}, \quad (11)$$

$$L(\omega) = (1 - i\omega/\alpha)^{-1}(1 - i\omega/\beta)^{-1}, \quad (12)$$

where \mathbf{k} and ω are the wave vector (with magnitude $k = 2\pi/\lambda$ where λ is the wavelength) and angular frequency ($\omega = 2\pi f$ where f is the cyclic frequency in Hz), respectively. The gain $G_{ab} = \rho_a v_{ab} = \rho_a N_{ab} s_{ab}$ is the response in neurons a due to unit input from neurons b . The quantities $G_{ese} = G_{es}G_{se}$, $G_{esre} = G_{es}G_{sr}G_{re}$ and $G_{srs} = G_{sr}G_{rs}$ correspond to the overall gains for the excitatory corticothalamic, inhibitory corticothalamic, and intrathalamic loops, respectively. The input stimulus ϕ_n is taken to be white noise (Robinson et al., 2004, 2001, 1997, 2005; Roberts and Robinson, 2012; O'Connor et al., 2002), with uniform spectral power density $|\phi_n(\mathbf{k}, \omega)|^2 = \text{constant}$ for nonzero \mathbf{k} and ω . In the linear approximation, the value of $\phi_n(\mathbf{k}, \omega)$ only affects the normalization of the power spectrum. Since the steady state firing rates only depend on the product $v_{sn}\phi_n^{(0)}$, we set $\phi_n^{(0)} = 1 \text{ s}^{-1}$ without loss of generality when we calculate the steady state firing rates, allowing v_{sn} to vary.

The power spectrum $P(\omega)$ is calculated by integration of $\phi_e(\mathbf{k}, \omega)$ over \mathbf{k} . Because of the finite size of the brain, we approximate the cortex as a rectangular sheet. We impose periodic boundary conditions here to ease numerical investigation of the modal structure of the power spectrum, but this is not critical to the dynamics of the system (Robinson et al., 2001). With periodic boundary conditions, the power spectrum is given by (Robinson et al., 2001):

$$P(\omega) = \sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} |\phi_e(k_x, k_y, \omega)|^2 F(k) \Delta k_x \Delta k_y, \quad (13)$$

$$k_x = \frac{2\pi m}{L_x}, \quad k_y = \frac{2\pi n}{L_y}, \quad k = \sqrt{k_x^2 + k_y^2}. \quad (14)$$

where the size of the two-dimensional rectangular cortex is $L_x \times L_y$ with $L_x = L_y = 0.5 \text{ m}$ (Rowe et al., 2004b; van Albada et al., 2007, 2010). The filter function $F(k)$ approximates the low-pass spatial filtering that arises due to volume conduction by the cerebrospinal fluid, skull, and scalp, whose k dependence can be approximated by

(Robinson et al., 2001; Rowe et al., 2004b; van Albada et al., 2007, 2010)

$$F(k) = e^{-k^2/k_0^2}, \quad (15)$$

with a low-pass cutoff at $k_0 \approx 10 \text{ m}^{-1}$ based on the spherical harmonic head transfer function developed by Srinivasan et al. (1998).

The behavior of the power spectrum at low frequencies depends on the leading terms in the expansion of Eq. (11) in powers of ω (Robinson et al., 2001). If we write (11) as

$$q^2 r_e^2 = A_0 + A_1(-i\omega) + A_2(-i\omega)^2 + \dots, \quad (16)$$

the coefficients A_n are real (Robinson et al., 2001). When $A_0 = 0$ with $A_1 \neq 0$, $P(\omega) \propto \omega^{-1}$ which corresponds to typical waking states. When $A_0 = A_1 = 0$ with $A_2 \neq 0$, $P(\omega) \propto \omega^{-3}$ which gives the steep low frequency slope seen in deep sleep (Robinson et al., 2001; Roberts and Robinson, 2012). Thus the low frequency slope in the spectrum can be approximately inferred from the proximity to the surface defined by $A_0 = 0$ and to its intersection with the $A_1 = 0$ surface.

It should be noted that the model power spectrum depends only on the five gains $G_{ee}, G_{ei}, G_{ese}, G_{esre}$, and G_{srs} . Therefore, the spectrum alone is insufficient to determine the eight connection strengths v_{ab} in the absence of additional measures such as $\phi_a^{(0)}$ (Robinson et al., 2004).

2.3. Stability

Linear waves in the model obey the dispersion relation

$$(1 - G_{ei}L)(1 - G_{srs}L^2)(q^2 r_e^2 + k^2 r_e^2) = 0. \quad (17)$$

and the steady state is stable if all modes satisfy $\text{Im}(\omega) < 0$ (Robinson et al., 2002, 2004). Unstable steady states rapidly diverge to pathological firing rates and therefore do not correspond to normal healthy brain states (Robinson et al., 1997, 1998, 2002). We have previously associated these states with pathological conditions including seizures (Breakspear et al., 2006; Robinson et al., 2002).

Only a few types of global instabilities occur in the model, at low frequencies, which enables a three-dimensional representation of the model with axes (Robinson et al., 2002, 2005)

$$X = \frac{G_{ee}}{1 - G_{ei}}, \quad (18)$$

$$Y = \frac{G_{ese} + G_{esre}}{(1 - G_{srs})(1 - G_{ei})}, \quad (19)$$

$$Z = -G_{srs} \frac{\alpha\beta}{(\alpha + \beta)}. \quad (20)$$

The quantities X , Y , and Z parametrize cortical, corticothalamic, and intrathalamic loop gains, respectively (Robinson et al., 2002); this parametrization captures the most central dynamics of the model.

We can also derive a reduced model based entirely on the XYZ quantities by using the approximation $L \approx 1$ except in the factor $(1 - G_{srs}L^2)$, where it is essential to retain ω (Robinson et al., 2002). Using these definitions, the dispersion relation and power spectrum $P(\omega)$ for the system can be written as

$$\left[\left(1 - \frac{i\omega}{\gamma}\right)^2 - X + k^2 r_e^2 \right] (1 + Z'L^2) - Y(1 + Z') e^{i\omega t_0} = 0, \quad (21)$$

$$q^2 r_e^2 = \left(1 - \frac{i\omega}{\gamma}\right)^2 - X - \frac{Y(1 + Z')}{1 + Z'L^2} e^{i\omega t_0}, \quad (22)$$

$$P(\omega) = \frac{P_0}{|1 + Z'L^2|^2} \left(\frac{1}{|q^2|} \right)^2, \quad (23)$$

Table 1

Model parameter values and fit ranges. The values for unfitted parameters are based on previous work (Robinson et al., 2004, 2002; Roberts and Robinson, 2012; Rennie et al., 2002; van Albada et al., 2010; Rowe et al., 2004a). Fitted parameters are only allowed to vary between the specified bounds, which are hard constraints on the parameter values. Parameter limits for the gains are chosen to include all of the traditional sleep stages found in previous work (Abeysuriya et al., 2015).

Parameter	Description	Value	Min	Max	Initial step	Unit
γ_e	Cortical damping rate	116				s^{-1}
r_e	Excitatory axon range	86				mm
Q_{\max}	Maximum firing rate	340				s^{-1}
θ	Firing threshold	12.9				mV
σ'	Threshold spread	3.8				mV
ϕ_n	Input stimulus amplitude	1×10^{-5}				s^{-1}
G_{ee}	Excitatory cortical gain	–	0	20	0.4	–
$-G_{ei}$	Inhibitory cortical gain	–	0	40	0.4	–
G_{ese}	Positive corticothalamic loop gain	–	0	40	1	–
$-G_{esr}$	Negative corticothalamic loop gain	–	0	40	1	–
$-G_{srs}$	Intrathalamic loop gain	–	0	14	0.2	–
α	Decay rate of cell-body potential	–	10	100	5	s^{-1}
β	Rise rate of cell-body potential	–	100	800	40	s^{-1}
t_0	Corticothalamic loop delay	–	75	140	5	ms
A_{EMG}	Normalization of EMG power	–	0	1×10^{-12}	0.05×10^{-12}	s^{-1}
f_{EMG}	Characteristic EMG frequency	–	10	50	0.2	Hz

$$Z' = Z \frac{(\alpha + \beta)^2}{\alpha \beta}, \quad (24)$$

where P_0 is a normalization constant (Robinson et al., 2002).

2.4. EMG correction

Activity in pericranial muscles results in an electromyogram (EMG) signal that can cause an increase in EEG power at high frequencies (>25 Hz). This can be accounted for by including an EMG component P_{EMG} in the model power spectrum (Rowe et al., 2004a):

$$P_{\text{total}}(\omega) = P(\omega) + P_{EMG}(\omega), \quad (25)$$

$$P_{EMG}(\omega) = A_{EMG} \frac{(\omega/2\pi f_{EMG})^2}{[1 + (\omega/2\pi f_{EMG})^2]^2}. \quad (26)$$

The units of $P(\omega)$ in (13) are s^{-1} , because $\int P(\omega) d\omega$ has the same units as $\phi_e(t)^2$, and $\phi_e(t)$ has units s^{-1} . Thus the units of $P_{\text{total}}(\omega)$ and A_{EMG} are also s^{-1} . Because the experimental power spectrum has units $V^2 s$, $P_{\text{total}}(\omega)$ is related to the experimental power spectrum by multiplying it by a dimensional constant (Robinson et al., 2004).

In order for $P_{EMG}(\omega)$ to have a realistic effect on the power spectrum, it must be of similar magnitude to $P(\omega)$. However, Eqs. (10) and (13) show that $P(\omega)$ is proportionate to $|\phi_n(\mathbf{k}, \omega)|^2$. The magnitude of ϕ_n affects the size of nonlinear effects in the system, and we take $|\phi_n(\mathbf{k}, \omega)| = 1 \times 10^{-5} s^{-1}$ based on previous work investigating nonlinear effects in the EEG power spectrum (Abeysuriya et al., 2014a). This choice of ϕ_n results in A_{EMG} taking values on the order of $10^{-12} s^{-1}$, as shown in Table 1.

We note that the EMG signal measured by EEG electrodes reflects activity in different muscles from standard EMG measurements that typically measure muscle activity using an electrode on the chin. In particular, the EEG is likely to be more affected by extraocular muscle activity, which does not show significant atonia during REM sleep, unlike most other skeletal muscles (Caples and Somers, 2009). We therefore expect the EMG component estimated by the model to be relatively higher in REM, as opposed to the more general reduction in muscle activity in this stage.

3. Methods

In this section we develop a new algorithm to fit our model to experimental EEG data. In Section 3.1, we discuss the principal problems addressed by our method. Section 3.2 applies the

algorithm to a single power spectrum, and Section 3.3 extends it to time-varying spectra.

3.1. Fitting

Previous approaches to fitting the model power spectrum to an experimental spectrum (Rowe et al., 2004a; van Albada et al., 2010, 2007) used the Levenberg–Marquardt method (Press et al., 1992) to progressively refine an initial estimate of the parameter values. To obtain a better approximation to the lowest among what may be multiple local minimums, the fit was initialized from a number of different initial values until a predetermined number of successful fits were obtained. The mean and standard deviation of the parameters were then calculated from this set of successful fits. Although this algorithm is capable of obtaining very good fits to the experimental data, it has a few key limitations:

- (i) The optimized parameters depend somewhat on the choice of initial parameters, which in turn means that the overall parameter estimates depend on the distributions from which the initial parameters are sampled. The choice for these distributions biases the final fit, but there is no objective basis upon which to choose the distributions.
- (ii) The uncertainty in a single set of fitted parameters can be estimated by the Levenberg–Marquardt routine based on the curvature of the cost function in the vicinity of the fitted parameters. However, the overall parameter estimates are obtained by aggregating several fits to adequately cover the parameter space, which complicates propagation of uncertainties and makes interpretation of the output difficult.
- (iii) The Levenberg–Marquardt method does not directly incorporate parameter constraints. Typically ‘soft’ constraints are incorporated by artificially increasing the error in the fit to penalize physiologically unrealistic parameter values. However, the properties of the convergence of the algorithm depend on how the error is increased, particularly when the fit lies near a stability boundary in the model. Unstable parameter combinations must not be selected at all, but steep changes in the cost function to enforce this can affect the numerical convergence of the method.

In this study, we address these issues by developing a new fitting method based on the Metropolis–Hastings algorithm (Metropolis et al., 1953). We first apply this approach to a single EEG power spectrum, then extend the method to track parameters over time.

3.2. Single fit

We approach the fitting of model parameters as a statistical inference problem to determine the most likely parameter values \mathbf{x} that give rise to the experimental power spectrum P^{exp} , where \mathbf{x} is a vector of parameters. We write the *posterior distribution* of the parameter values as $p(\mathbf{x}|P^{\text{exp}})$, which is generally a distribution that is unknown in advance, but is maximal at the most likely value of \mathbf{x} . Our aim is to sample from this distribution to estimate $p(\mathbf{x}|P^{\text{exp}})$ and to infer the most likely values of \mathbf{x} , using the model to relate \mathbf{x} to the EEG power spectrum.

The Metropolis–Hastings (MH) algorithm is a Markov Chain Monte Carlo (MCMC) method that provides samples from general probability distributions. The critical feature of the MH algorithm is that it only requires the ability to evaluate the probability $p(\mathbf{x}|P^{\text{exp}})$ at a single point in \mathbf{x} -space, and it does not require previous knowledge of the shape of the distribution. This makes it especially useful when the dimension of \mathbf{x} is large, and when evaluation of $p(\mathbf{x}|P^{\text{exp}})$ is computationally intensive.

To apply the MH algorithm to curve fitting, we require that the posterior probability be maximal when the model generates an optimal fit to the data. We characterize the quality of the fit by taking a weighted squared fractional difference between the model prediction $P(\mathbf{x})$ and the experimental data P^{exp} :

$$\chi^2(\mathbf{x}) = \sum_j W_j \left| \frac{P_j(\mathbf{x}) - P_j^{\text{exp}}}{P_j^{\text{exp}}} \right|^2 \quad (27)$$

where j indexes frequency components from the fast Fourier transform of the experimental data. The weighting W_j is primarily incorporated to compensate for the large number of points at high frequencies compared to low frequencies, because EEG spectra typically exhibit significant features in each decade of frequency. The weights we use are:

$$W_j \propto f_j^{-1} \quad (28)$$

which provides equal weighting to each frequency decade (Rowe et al., 2004a). This ensures that sufficient weighting is given to low frequencies relative to high frequencies. We set $W_j = 0$ for $f < 1$ and $f > 45$ because the principal features of interest in the EEG power spectrum occur in this frequency range, and some data exhibit a roll-off in power spectral density outside this range.

Division by P_j^{exp} results in the difference between the experimental spectrum and the model spectrum being calculated as a fraction of the power in the experimental data. If this step is omitted, peaks in the spectrum are still fitted well; however, in the absence of peaks, the low frequency part of the spectrum in light sleep states is poorly fitted because the absolute difference in power at low frequencies is too small compared to the differences at high frequencies. The result is that the spectrum would be fitted marginally better at high frequencies, at the expense of large discrepancies at low frequencies. On the other hand, taking the logarithm of the powers before calculation of χ^2 results in the low frequency slope of the spectrum being fitted well, which is ideal for sleep states, but gives insufficient weight to peaks in wake states. The primary effect would be that the alpha peak is poorly fitted in wake, in favor of marginally better fitting of low frequencies in sleep. The normalization introduced in Eq. (27) provides a balance between these two extremes, and enables high quality fitting of both wake and sleep states. Finally, the model spectrum is rescaled by a multiplicative constant to the same total power as the experimental spectrum before χ^2 is calculated. Holding the total power constant reduces the number of fitted parameters, and results in a considerable improvement in convergence of the algorithm and quality of the fitted spectral features.

To perform model fitting, we first write our posterior distribution as

$$p(\mathbf{x}|P^{\text{exp}}) \propto L(\mathbf{x}). \quad (29)$$

We use this notation in preparation for the time-varying fits we examine in Section 3.3. We define the likelihood function for our problem as (Denison et al., 1998; Punskaya and Andrieu, 2002)

$$L(\mathbf{x}) = \exp \left[\frac{-\chi^2(\mathbf{x})}{2} \right]. \quad (30)$$

This function is maximized when χ^2 is minimized, thus satisfying the requirement that $p(\mathbf{x}|P^{\text{exp}})$ be maximal at the most likely parameter values.

We now use the MH algorithm to generate samples from $p(\mathbf{x}|P^{\text{exp}})$. The samples are generated using a random walk, in which the model parameters are repeatedly changed in small increments. Each step in the random walk x_{i+1} is generated by sampling from a ‘proposal’ distribution that depends on the current parameters x_i . Typically, a normal distribution centered at x_i is used for the proposal distribution. The proposed step is accepted with probability $p(x_{i+1})/p(x_i)$. Constraints on parameter values are implemented by setting the posterior density to 0 if the parameters are unphysiological. This hard constraint is particularly useful for dealing with stability boundaries in the model. The parameters are considered unphysiological if any of them lie outside the bounds listed in Table 1, or if the system is unstable because there are solutions to Eq. (17) with $\text{Im}(\omega) > 0$ and $\text{Re}(\omega) < 0$, as discussed in Section 2. In addition to the upper and lower bounds on the parameter values, we also incorporate additional constraints on pairs of parameter values. We require $G_{ee} + G_{ei} < 1$ to exclude unphysiologically high firing rates (Abeysuriya et al., 2015), $|G_{ee}/G_{ei}| > 0.5$ due to constraints on underlying synaptic strengths (Abeysuriya et al., 2015), and $\beta/\alpha < 20$ based on known synaptic dynamics and previous work on fitting (Robinson et al., 2004).

The sequence of parameters traversed by the random walk is referred to as the *chain*; each step in the random walk is added to the end of the chain. As the chain becomes longer, the distribution of the parameters in the chain converges to the posterior distribution (29). The rate of convergence depends on the proposal distribution. We use an adaptive proposal distribution (Roberts and Rosenthal, 2009; Rosenthal, 2011) with adaptive global scaling (Andrieu and Thoms, 2008) to improve convergence due to the high dimensionality (9 parameters) and interdependencies between the gains (these interdependencies are reflected in the fact that just the three quantities X , Y , and Z capture the main dynamics of the model).

To commence the random walk, the proposal distribution is initially chosen to be a normal distribution in each parameter, with the standard deviations specified in Table 1. The initial step size is chosen to be small relative to the parameter values so that the initial acceptance rate is high, thus more quickly initializing the covariance matrix used for the adaptive proposal distribution. The choice of initial step primarily affects the speed of the fitting, rather than the final result. After 100 accepted moves, the chain contains enough points to switch to the adaptive proposal distribution, which depends on the covariance of the chain (Haario et al., 2001). During this initialization, we follow the ‘greedy start’ procedure of only retaining accepted moves. This ensures that the initial covariance matrix computed based on the chain produces a sufficiently wide proposal distribution to ensure the random walk adequately covers the parameter space.

Convergence of the chain is accelerated by starting the random walk in a region of high probability. This is achieved by performing a preliminary fit to the set of existing parameters used in previous work (Abeysuriya et al., 2015). We compute χ^2 for a representative subset of the brain states associated with each of the

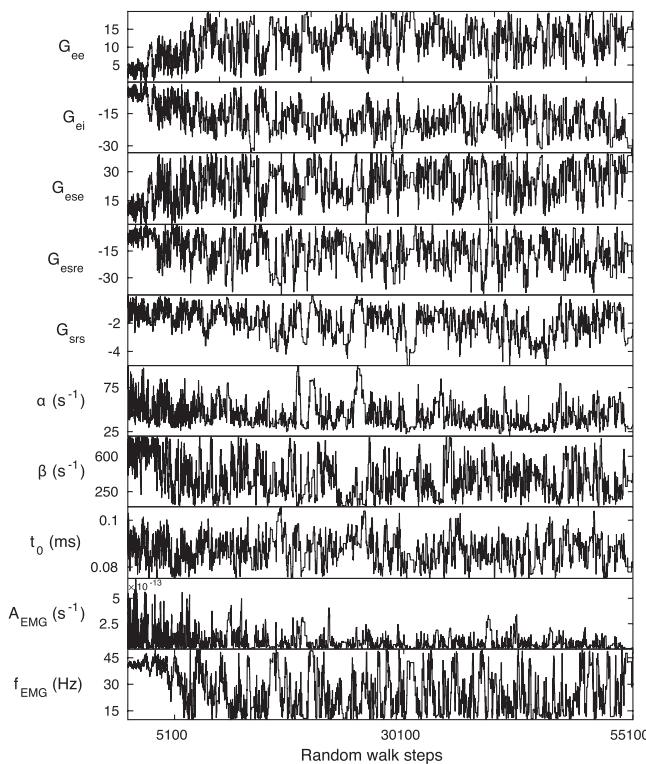


Fig. 3. Graphical plot of the chain for the fit shown in Fig. 6. Steps 0–5100 are discarded as burn-in. The sections from steps 5100–30,100 and 30,100–55,100 are used to generate the two distributions shown in Fig. 5. Note that steps 0–100 correspond to the initial steps using a fixed proposal distribution, after which an adaptive proposal distribution is used, as described in the text.

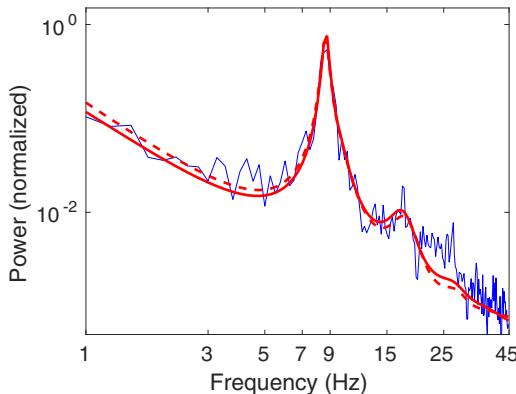


Fig. 4. (Color online) Experimental power spectrum (blue) and model power spectra for chains with 5×10^4 (solid red line) and 1×10^4 (red dashed line) samples. The chain with 5×10^4 samples was also used to produce Figs. 3 and 5. The fit obtained from the chain with 1×10^4 samples is discussed in Section 4.2. The experimental spectrum has been normalized to unit power.

population-averaged traditional arousal stages, and select the initial parameters on this basis.

The statistics of the chain are initially affected by the choice of the initial parameter values. We therefore discard the first 10% of the chain due to these transient variations in its statistical properties. This process is conventionally referred to as ‘burn-in’. Fig. 3 shows a trace plot for 5×10^4 samples obtained by fitting the model to the experimental power spectrum shown in Fig. 4. This trace plot verifies that the algorithm for tuning the proposal distribution results in adequate mixing of the Markov chain.

We consider the chain to be stationary when the statistics of the first half of the chain are the approximately the same as the statistics of the second half. The stationarity of the chain is illustrated in Fig. 5, which verifies that similar marginal distributions for each parameter are obtained using different subsets of the chain. These distributions peak at parameter values associated with low values of χ^2 . The width of the distribution provides an indication of the dependence of the fit on the parameter value. A narrow distribution shows that a good fit is only achievable for a small range of parameter values, whereas a wide distribution indicates that the data does not tightly constrain that parameter.

The final parameter estimate is obtained by selecting the point from the chain with the largest value of $p(\mathbf{x}|P_{\text{exp}})$. The fitted parameter values are marked as vertical bars in Fig. 5, and produce the model power spectrum shown in Fig. 4.

We can also display the fit in terms of X, Y, and Z, which are derived from the fitted parameters. Fig. 6 shows the marginal probability distributions of X, Y, and Z. These distributions are also illustrated in 3D as clouds, with high density regions corresponding to high probability.

We note that the exact shape of each marginal distribution is related to our choice of likelihood function (30), and marginal distributions would be different if an alternate measure of the quality of the fit were used. We use the form (30) in the present study based on previous work (Denison et al., 1998; Punskaya and Andrieu, 2002), but other forms could be used in our fitting routine.

3.3. Real-time state tracking

We now turn to fitting a sequence of spectra at successive times, to track temporal changes in parameter values. This offers considerable scope to improve the quality of the fit and parameter estimates, on the assumption that parameter values do not change excessively over a short period of time. We expect the parameters in our model to exhibit changes consistent with typical sleep timescales, so over a period of seconds we expect only small changes to the parameters. Therefore, we can use information from the spectrum at time t_i to improve the fit at time t_{i+1} . This approach is particularly powerful when there are multiple regions of parameter space that provide a good fit, because the requirement of temporal continuity rules out large parts of the parameter space.

3.3.1. Incremental priors

The key step in incorporating temporal constraints into the model fit is to use Bayes’s theorem, which relates the prior distribution (hereafter referred to as the prior) of parameters to the posterior distribution. In the previous section, we had $p(\mathbf{x}) \propto L(\mathbf{x})$, which implicitly assumes that all parameter values are equally probable, in the absence of any other information about the data being fitted. However, the marginal posteriors for each parameter estimated by the fitting routine are a suitable choice for the prior in subsequent fits, providing that the next fit is sufficiently close in time that the parameters are not too different. The effect of the prior is quantified using Bayes’s theorem, which yields

$$p_{t+1}(\mathbf{x}) \propto L(\mathbf{x})p_t(\mathbf{x}), \quad (31)$$

and the fitting routine thus optimizes for both spectral fit and probability of the parameters. Including the priors provides several advantages. First, using the information about the spectrum at previous times reduces sensitivity to noise, because rapid changes in the spectrum are more likely to be due to noise than parameter changes. Second, complex marginal distributions (e.g., bimodal ones) are correctly accounted for. For example, a bimodal distribution at time t_i will result in the fitting routine exploring parameters near both of the probable regions at time t_{i+1} . Third, the priors

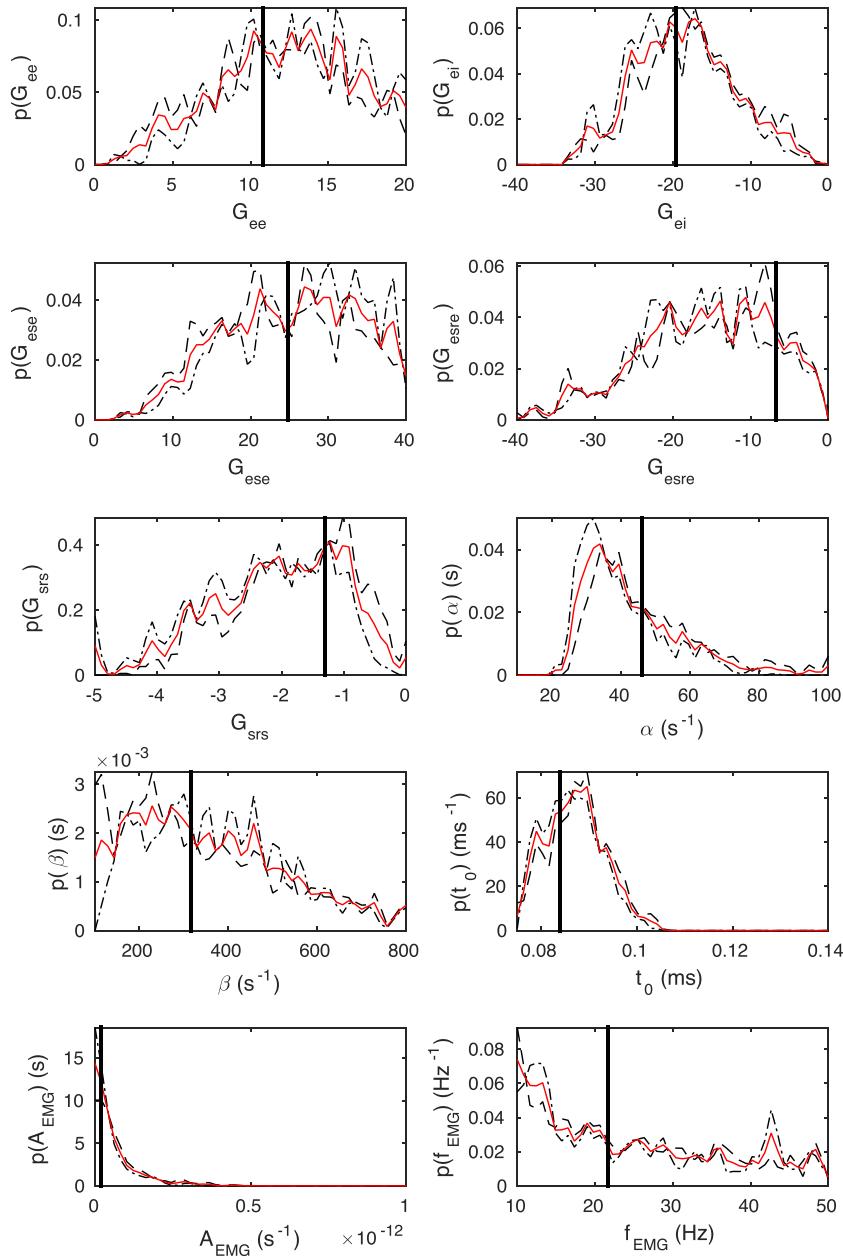


Fig. 5. (Color online) Marginal distributions of fitted parameters from the random walk shown in Fig. 3, as labeled under each panel. The distributions are calculated for different parts of the chain: first half (dashed lines), second half (solid lines), and the entire chain (red). The distributions are calculated for different parts of the chain: first half (dashed lines), second half (solid lines), and the entire chain (red). The vertical bars correspond to the fitted parameter values. The limits on the horizontal axes correspond to the allowed parameter ranges in Table 1.

increase the efficiency of the MCMC approach by reducing the parameter space being searched. This also causes the chain to converge to its stationary distribution more rapidly than when using uniform priors [i.e., using (29)]. Fig. 7 shows the marginal posterior distributions for the first and second halves of the chain for a fit at t_{i+10} (i.e., after 10 consecutive fits with updating of the priors at each successive fit). The distributions in Fig. 7 match much more closely compared to Fig. 5, for the same chain length. Finally, incorporating information from previous fits reduces the width of the marginal distributions for each parameter by enforcing temporal continuity in their estimates. The effect of the priors on the marginal distributions is shown in Fig. 8. Without updating the priors, the posterior distributions are almost unchanged at time t_i and time t_{i+10} . Updating the priors at each time step results in considerably narrower marginal distributions.

3.3.2. Tracking brain states

Model parameters can be tracked robustly over time by repeatedly applying the above procedure to incrementally update the priors and parameter estimates. The primary issue when fitting to real-world recordings is that of identifying and handling artifacts (motion-related or otherwise). Many healthy transient EEG events (such as sleep spindles) are short in duration, and their effect on the spectrum can be limited by calculating the spectrum over time periods that are long relative to the transient event (Abeysuriya et al., 2014b). Extremely large events, such as muscle artifacts, still affect the power spectrum even when long windows are used. Attempting to fit the spectrum when an artifact is present can result not only in an invalid estimate of the parameters, but in temporary corruption of the priors, which serves to degrade the quality of the next few fits as well. However, it

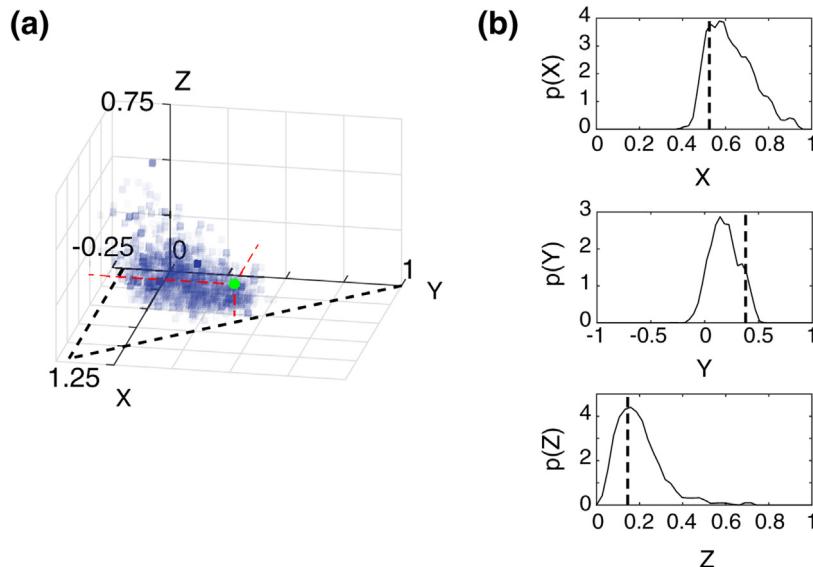


Fig. 6. (Color online) Probability distributions for XYZ derived from the fitted parameters. The distributions are shown in: (a) XYZ-space, in which the marginal posteriors are displayed by binning the chain in 3D, resulting in a ‘cloud’ with darker colors indicating a higher density of points and therefore higher parameter probabilities. The green dot shows the fitted parameters, with red dashed lines showing its position. (b) Individual marginal probability distributions. The vertical dashed line shows the fitted value.

is necessary to minimize the quantity of data rejected, firstly to prevent elimination of excessive numbers of valid data, and second because incrementally updating the priors is most effective when there is only a short time difference between consecutive spectra.

In this study, we use the following procedure to calculate a sequence of consecutive power spectra that are then fitted using the model. First, an initial set of spectra are calculated by using a 4 s window moved in 1 s increments. The 4 s blocks are marked as contaminated if more than 9 samples are within 3 μ V of the maximum recorded voltage, which may indicate that the range of the recording was insufficient. We also mark the block as contaminated if the power at frequencies below 4.5 Hz is more than 3 standard deviations higher than the mean over the recording, which is based on previous work by D'Rozario et al. (2014) that detected artifacts in slightly longer, non-overlapping sleep spectra. Some spectra are affected by high frequency artifacts, likely related to muscle activity. These are identified analogously, by excluding blocks where the power integrated from 30 to 45 Hz is more than 3 standard deviations higher than the mean. Finally, some of the recordings contain electrical artifacts which cause the recorded voltage to remain constant for short periods of time. We identify these artifacts by checking for periods of constant voltage lasting longer than 0.5 s. Blocks containing these artifacts are then marked as contaminated.

Next, a series of spectra are generated by averaging together the blocks contained within a 30 s sliding window. The 30 s window corresponds to 27 spectra, which provides sufficient reduction of noise to reliably fit the spectrum. The 30 s window also enables direct comparison with the manually-scored traditional sleep stage whenever the window lines up with the scored epochs. The temporal resolution of the method is affected by the choice of window, as stationarity is assumed over the time corresponding to any power spectrum. For each 30 s window, only clean 4 s blocks are averaged, so there may be fewer than 27 spectra contributing to the window. If less than 20 spectra in the window are clean, the decrease in signal to noise ratio can affect the fitting. In addition, a large number of contaminated blocks suggests an increased chance of additional blocks containing undetected artifacts, further degrading the fit. For these reasons, we reject spectra where there are less than 20

clean 4 s blocks when fitting, and instead use the last good estimates for the parameter values and their priors for the next clean 30 s spectrum. Sleep recordings typically begin with an extended period of wake either before the lights are turned off, or due to difficulty falling asleep in the lab. Because we are primarily interested in the dynamics of the sleep–wake cycle, we truncate the recording to start at most 5 min before sleep onset and then begin fitting at the first clean spectrum.

In total, around 2–5% of the final spectra are not fitted due to artifacts (with only 2 subjects out of 28 having more than 8% of their spectra rejected). We emphasize that the role of artifact rejection here is to prevent fitting our model to artifacts, and the fitting method developed in this study is not reliant upon the specific artifact detection methods presented here. This also means that the artifact detection routines can be adapted to different data sets if they contain types of artifacts not present here.

Not all of the model parameters can be reliably estimated in all brain states. In particular, the feature most prominently related to t_0 is the frequency of the alpha oscillation (Robinson et al., 2002). Although tightly constrained in wake, t_0 has a significantly weaker effect in sleep. Fitting t_0 in sleep states results in its marginal distribution widening to the entire range listed in Table 1, thereby discarding the information about t_0 obtained from wake states. Therefore, we only update the prior distribution for t_0 when an alpha peak is present [$R_\alpha > 5$ as defined by Abeysuriya et al. (2015)]. Fitting t_0 with a static prior allows us to incorporate its effect when fitting sleep states, while retaining the distribution determined from wake states in which t_0 can be accurately estimated.

4. Results

In this section, we present findings from initial application of the fitting algorithm to wake and sleep EEG data. This is primarily a demonstration and validation of the method, and thus only a preliminary analysis of the data set. In Section 4.1 we examine the dynamics of the model parameters over the sleep cycle. In Section 4.2 we examine the performance of the method and its application to real-time fitting. Finally, in Section 4.3 we test our

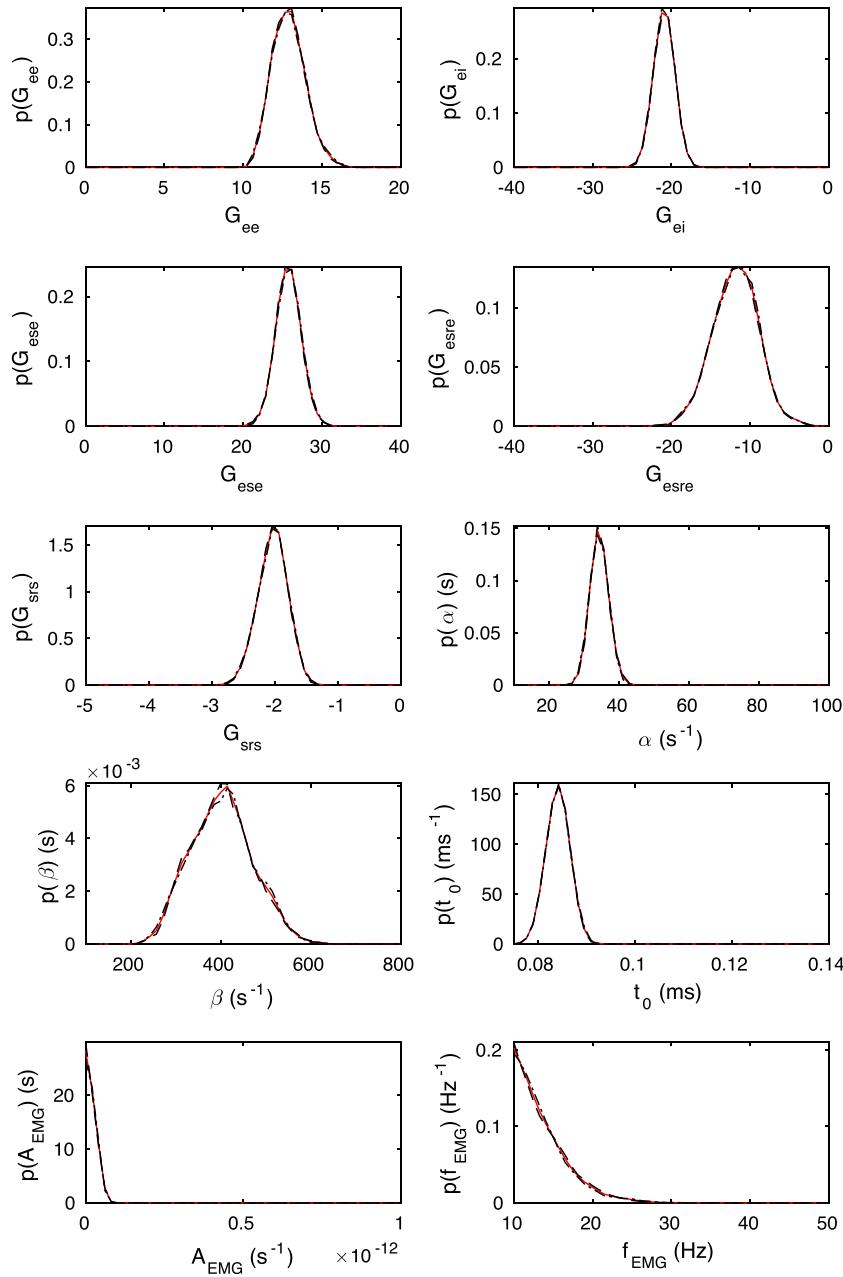


Fig. 7. (Color online) Marginal distributions for each fitted parameter at t_{t+10} with incremental updating of the priors, as labeled under each panel. The distributions are calculated for different parts of the chain: first half (dashed lines), second half (alternating dots and dashes), and the entire chain (red solid line). The limits on the horizontal axes correspond to the allowed parameter ranges in Table 1.

approach with a reduced model that has fewer parameters to assess the significance of parameter interdependencies in the model.

4.1. Parameter evolution and tracking

We have fitted the model to EEG data from 28 manually scored polysomnograms from healthy controls previously recorded as part of two clinical studies (D'Rozario et al., 2013; Wang et al., 2005). To minimize the effect of muscle artifacts, we use data from the Cz electrode where possible, and from C4 when Cz data is unavailable (van Albada et al., 2010; Saunders, 1979). Electrodes were referenced to the mastoid channels following the recording system standards. These data were scored using Rechtschaffen and Kales sleep stages. For simplicity, we group the S3 and S4 stages into a single slow wave sleep stage labeled 'SWS' in our analysis.

Fig. 9 shows time courses of each of the model parameters over a full night of sleep. The clearest variation over time occurs for the cortical excitatory gain G_{ee} , which is found to be larger in deep sleep stages. This is in agreement with our previous work associating strong cortical feedback with slow wave activity (Robinson et al., 2002; Abeysuriya et al., 2015). Our results are related to experimental evidence linking an increase in cortical synaptic strength to slow wave activity (Massimini et al., 2009; Huber et al., 2004, 2006; Liu et al., 2010; Vyazovskiy et al., 2008). These studies conclude that the times of highest synaptic strength coincide with the times of highest SWA during sleep (i.e., early in the night). Our prediction is of a different effect, in that we are describing the correlation between the gain and sleep stage rather than the synaptic strength over time. We also note that the gain G_{ee} depends on both the synaptic strength v_{ee} and the steady state firing rate $\phi_e^{(0)}$, so if

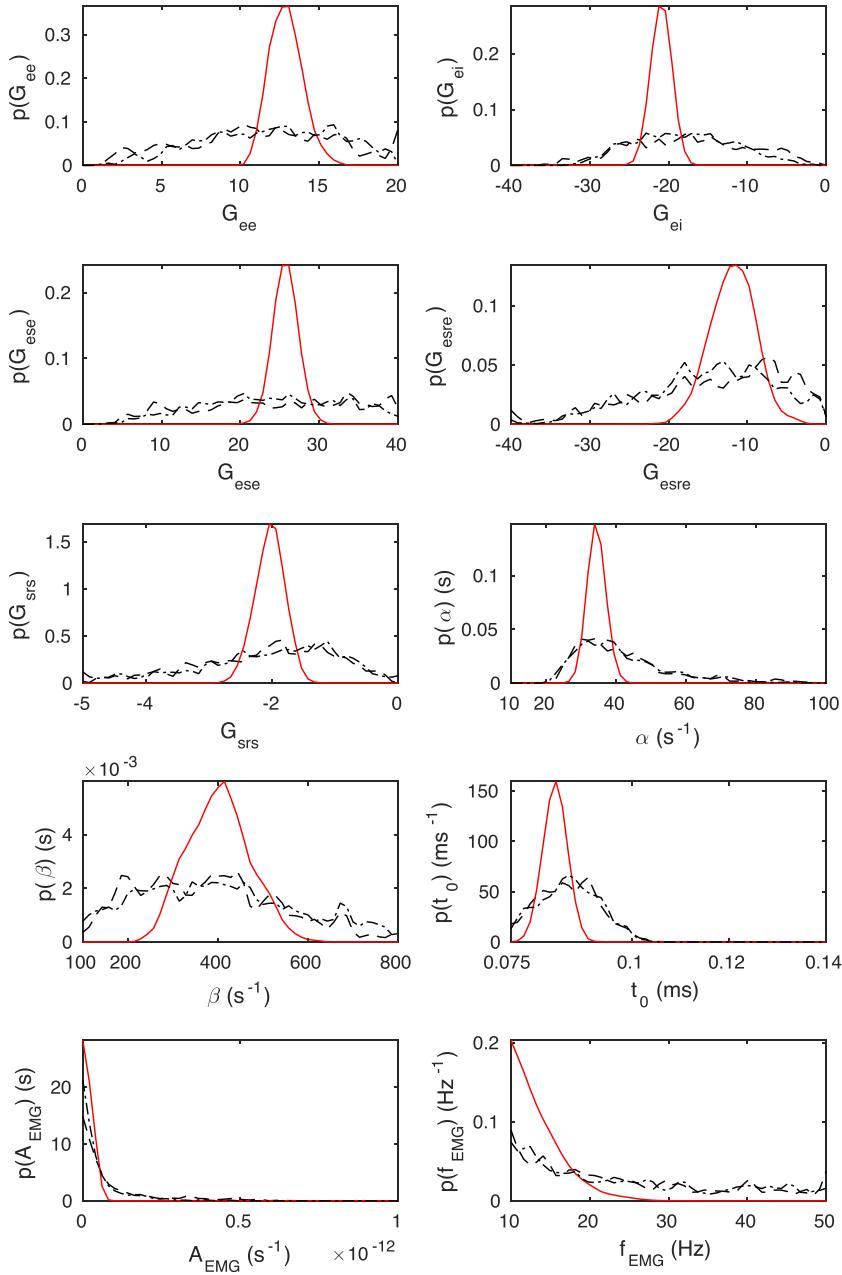


Fig. 8. (Color online) Comparison showing the effect of updating the priors when fitting a series of spectra. At time t_i , the initial fit with uniform priors (dashed line) has a wide distribution in most of the parameters. At time t_{i+10} the marginal distributions are calculated with (red solid line) and without (alternating dots and dashes) updating of the priors. The limits on the horizontal axes correspond to the allowed parameter ranges in Table 1.

$\phi_e^{(0)}$ increases, then G_{ee} could increase even if there is a decrease in synaptic strength. However, our results are not inconsistent with the experimental work, and lead to a similar observation that there is a link between strong intracortical connectivity and slow waves.

The increase in cortical excitation in deep sleep is accompanied by anticorrelated changes in G_{ei} reflecting a decrease in cortical inhibition that compensates for the decrease in corticothalamic feedback strength in sleep.

There are a number of correlated changes in model parameters over the sleep cycle. Fig. 10 shows the correlation matrix averaged over the 28 subjects in our study. There are three main blocks of correlated parameters – the cortical quantities G_{ee} and G_{ei} , the thalamic quantities G_{ese} , G_{esre} , and G_{srs} (and the EMG amplitude), and the time constants α and β .

The principal dynamics of the model are dependent on the combinations of gains present in X , Y , and Z . In particular, the anticorrelations between G_{ee} and G_{ei} , and between G_{ese} and G_{esre} reflect the structure of Eqs. (18) and (19), in which the these pairs of parameters have opposite effects on X and Y , respectively. In addition, there is enhanced correlation between G_{ese} , G_{esre} , and G_{srs} . These parameters represent corticothalamic and intrathalamic connectivity, and thus reflect the contribution of the thalamus to the dynamics of the system. The two blocks of enhanced correlation in the gains can thus also be interpreted as reflecting cortical and thalamic components in the model.

The intrathalamic gain G_{srs} is weakly correlated with G_{ee} . This reflects the fact that deep sleep states tend to have large X and small Z (Abeysuriya et al., 2015). A similar pattern can be observed

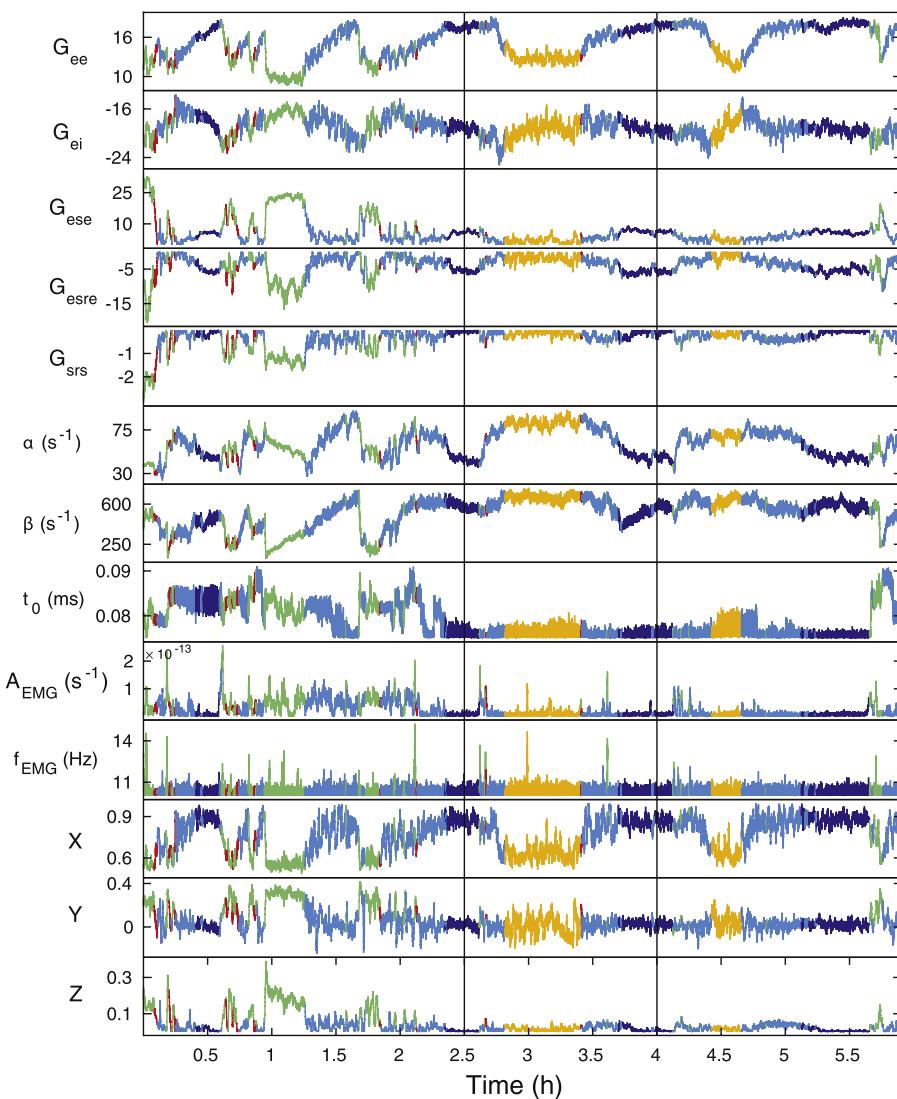


Fig. 9. (Color online) Parameter time courses over a full night of sleep for an individual subject obtained by fitting the model, as labelled on each panel. The color corresponds to the traditional sleep stage: wake (green), REM (yellow), S1 (red), S2 (light blue), and SWS (dark blue). The vertical bars indicate the 90 minute period examined in Fig. 12.

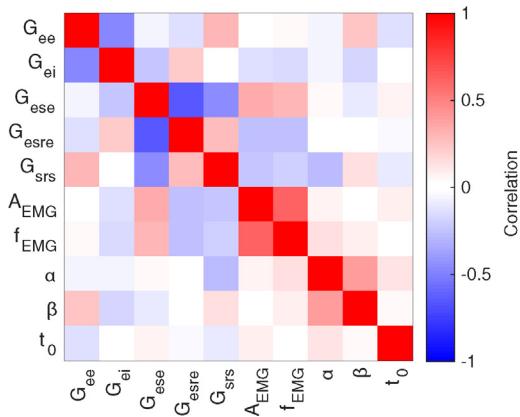


Fig. 10. Correlations between fitted model parameters for full night recordings, averaged over 28 subjects.

with G_{srs} and G_{ese} , which are strongly anticorrelated because the smallest values of Z occur at small Y . Interestingly, G_{srs} does not appear correlated with G_{ei} , which is unexpected given the strong anticorrelation between G_{ee} and G_{ei} . In contrast, G_{srs} is

correlated with G_{esre} , which reflects the anticorrelation between G_{ese} and G_{esre} .

Fig. 10 also shows that the EMG amplitude and frequency are most strongly correlated with the thalamic gains. As shown in Fig. 9, the largest change in EMG amplitude occurs between wake and sleep. The change in EMG amplitude between different sleep stages (e.g., REM and S2) is somewhat smaller. Sleep onset in the model is primarily associated with a decrease in Y (i.e., a decrease in $G_{ese} + G_{esre}$) with little change in X . Therefore, these are the quantities with which EMG amplitude is most strongly correlated. The ability to estimate EMG amplitude as part of the fitting routine also opens the possibility of replacing separate actigraphy measurements with muscle signals detected via EEG (Ancoli-Israel et al., 2003). The fitted EMG amplitude in wake states is considerably larger than in sleep states, and qualitatively resembles the time courses of typical actigraphy measurements (Martin and Hakim, 2011).

The rates α and β are those of the fall and rise of the synaptic response, respectively, and together account for low-pass filtering in the model due to synaptodendritic effects. Because the processes corresponding to α and β are physiologically related and these parameters have similar effects on the power spectrum, they are

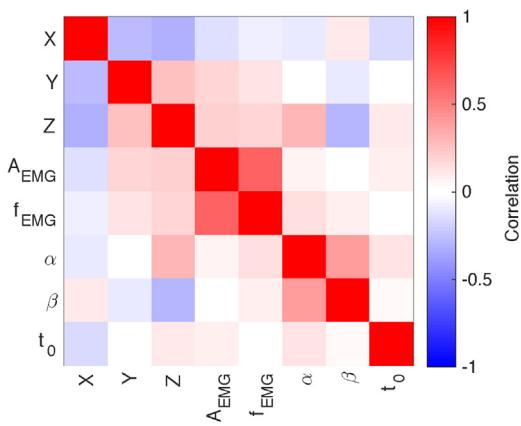


Fig. 11. Correlations between derived X, Y, and Z values and other fitted parameters, for full night recordings, averaged over 28 subjects.

also strongly correlated. This correlation is also expected based on physiology, which we have examined in previous work (Robinson et al., 2004). It is particularly interesting that despite the strong correlation between α and β , their individual correlations with the other model parameters are different. In particular, α is strongly anticorrelated with G_{srs} , whereas β is weakly correlated with G_{srs} . Similarly, β is moderately correlated with the cortical gains G_{ee} and G_{ei} , whereas α is much more weakly correlated with these parameters. These observations are likely a consequence of underlying neuromodulator differences between the brain states (Robinson et al., 2004).

Finally, t_0 shows little correlation with any of the other model parameters. This is expected, because the corticothalamic time delay is fixed for an individual and should not change over the sleep cycle. However, care must be taken in interpreting the correlations involving t_0 because we do not update the prior distribution for t_0 when the alpha peak is not present, which imposes a limit on the degree to which it can change during sleep.

The interdependencies in the gains motivate us to examine the correlations in XYZ-space, as shown in Fig. 11. This figure reveals a strong anticorrelation between X and Y, which arises because wake states with large values of Y require small values of X for the state to be linearly stable. This relationship between X and Y is also related to the observation in both theory and experiment that the brain operates near marginal stability, and thus the wake states lie close to the stability boundary $X+Y=1$ (Roberts and Robinson, 2012; Breakspear et al., 2006; Robinson et al., 2002; Stam, 2005; Stam et al., 1999). We also find that X is anticorrelated with Z, which arises because wake states occur at small X and relatively large Z, and sleep states occur at large X and small Z. In addition, Y is moderately correlated with Z, because wake states occur with large Y and relatively large Z, and sleep states require small Y and small Z. Together, X, Y, and Z form a closely correlated block, reflecting the overall structure of the model's parameter space. However, of these three quantities, Z is most strongly correlated with the other model parameters, particularly α and β . The α and β parameters show similar correlations with Z as with G_{srs} in Fig. 10, which is expected due to the relationship between G_{srs} and Z. Finally, the EMG amplitude and frequency are weakly correlated with Z. This likely reflects the fact that sleep stages that are found here to have larger EMG components, such as REM and S1, lie at larger Z than deep sleep stages (Abeysuriya et al., 2015). This correlation also matches the correlation between the EMG parameters and G_{srs} in Fig. 10.

The parameter trajectory can also be displayed directly in XYZ-space. Fig. 12 shows a 90 min trajectory during sleep, in which the subject started in SWS, then spent a period of time in REM, before

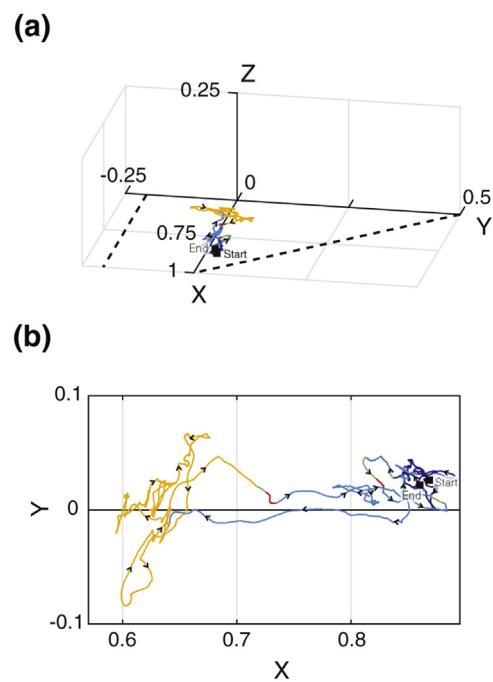


Fig. 12. (Color online) Smoothed parameter trajectory over 90 minutes, corresponding to the time period marked in Fig. 9. The color of the trajectory indicates the traditionally assigned sleep stage: wake (green), S1 (red), REM (yellow), S2 (light blue) and SWS (dark blue). The dashed lines correspond to $A_0 = 0$ and $A_1 = 0$ from Eq. (16). (a) Isometric view (b) Zoomed projection onto the XY-plane.

returning to SWS. Unlike sleep stages, the model parameters are highly descriptive and are able to characterize small changes in the power spectrum. Thus the model parameters show considerable variability on short timescales relative to sleep, which reflects a combination of fast fluctuations in physiology, and noise. Therefore, the trajectory in Fig. 12 has been smoothed using a five minute moving average to reduce the effect of fast fluctuations in the power spectrum and provide a clearer overall trajectory over longer periods. The features in the power spectrum are well-represented by the XYZ quantities and by the proximity to the locus of $A_0 = 0$ and $A_1 = 0$. This enables Fig. 12 to be related to features in the EEG power spectrum.

The present approach also enables localization of brain states in terms of model parameters for a single subject as shown in Fig. 13. Each set of fitted parameters is plotted in XYZ-space with color corresponding to the traditional sleep stage. This is an extension of previous work that examined the regions corresponding to traditional sleep stages based on EEG features averaged over multiple subjects (Abeysuriya et al., 2015). The location of the brain states in Fig. 13 is broadly consistent with the population-averaged parameter regimes identified in our previous work. However, unlike our previous work, the regions in Fig. 13 are individualized and reflect the spectral features of the individual's EEG, not just the typical features of each sleep stage. In addition, because each point represents a model fit and the fits are performed at regularly spaced time intervals, the density of the points in Fig. 13 is representative of the time spent near those parameter values. From Fig. 13, we can see that the majority of the regions corresponding to wake, REM, and NREM sleep (S2 and SWS) are well separated in terms of X, Y, and Z. The outer parts of these regions overlap, and points corresponding to S1 sleep tend to be located in these overlapping areas (most clearly visible in Fig. 13(b)). Wake states occurring at $Y \gtrsim 0.25$ are well separated from the sleep states in terms of both Y and Z. REM is primarily distinguished from S2 and SWS by X, and SWS is consistently located at smaller Z than S2.

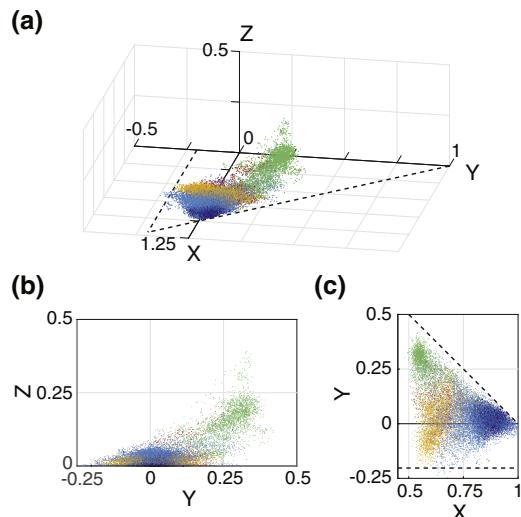


Fig. 13. (Color online) Fitted parameter values for a single subject over a complete night of sleep, shown in XYZ-space. The XYZ coordinates for each fit are plotted with color corresponding to the traditional sleep stage: wake (green), REM (yellow), S1 (red), S2 (light blue), and SWS (dark blue).

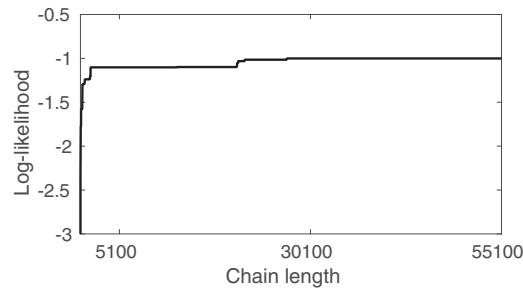


Fig. 14. Quality of the fit measured by the likelihood of the parameter estimate, for the chain shown in Fig. 3. As the chain becomes longer, points may be found that have a likelihood higher than the any of the previously visited parameter values. The likelihood of the parameter estimate therefore increases as the chain becomes longer.

4.2. Real-time performance

Our fitting method can be run in real time providing that the fit can be completed before the next power spectrum is acquired. The time taken to perform the fit is linear in the length of the chain. The time between successive power spectra can be arbitrarily chosen, except that as the time between spectra increases, so too does the size of possible changes in the underlying parameter values. The quality of the fit and efficiency of the algorithm depends on the choices for these two quantities. However, fitting the model in real-time requires a tradeoff between them because the time between successive power spectra imposes an upper limit on how much time is available to generate the chain.

In Section 3.2 we showed that a chain with 5×10^4 steps provides a reliable estimate of the posterior distributions. However, significantly fewer steps are sufficient if one only wishes to estimate the most likely parameter values. We find that the parameter estimate generated by our algorithm improves rapidly at the start of the random walk, as shown in Fig. 14. This means that high quality fits can be obtained even for very short chains. A representative example is shown in Fig. 4, in which a 80% reduction in the length of the chain still produces a good fit. Using a single core

of a 2.30 GHz Intel Xeon E5-2698 16 core processor, we are able to generate chains with 1×10^4 points in around 10 s.

Aside from shortening the chain to decrease the fitting time, we can also increase the time between successive fits to allow the algorithm to run longer on each fit. The trade-off is a broadening of the marginal posterior distributions as shown in Fig. 15, because parameter changes accumulate over time and are therefore larger when the time between fits is longer. Wider distributions reflect increased uncertainty in the parameter values. In addition, the convergence of the fitting algorithm is slowed because the routine explores a larger portion of the model's parameter space. However, Fig. 15 reveals that the peaks of the distributions still occur at similar values even though they are not as sharp; this demonstrates that increasing the gap between spectra to five or even 10 seconds still provides informative posterior distributions even though there is larger uncertainty in the parameter estimates.

Finally, we note that the chains can be sampled independently, which enables their generation to be easily parallelized. The order of the points in the chain does not affect our computation of the posterior distributions or the parameter estimate. Therefore, several chains can be assembled into a single data set, which is then processed in the same way as the serial case. This provides a performance increase that is linear in the number of processors used. However, the need to discard the transient burn-in period at the start of the chain provides a lower bound on the total length of the chain on each processor, because each processor must compute at least as many steps as required for the chain to be stationary, as shown in Fig. 3. This limits the theoretical maximum speedup available with parallelization; the minimum total time for the fit is the time taken for a single processor to compute the burn-in.

4.3. Model comparison

We have seen that the loop parameters X , Y , and Z offer a useful representation of model dynamics (Robinson et al., 2002). Hence, in this section, we examine the effect of reducing the number of model parameters fitted. In general, we always require α , β , and t_0 to be fitted in order to correctly model the synaptic and population delays. However, we are free to choose whether to fit some or all of the G_{abcd} , or alternatively to fit X , Y , and Z . Finally, we may fit the spectrum with or without EMG.

We approach the problem of model selection by examining the Bayesian information criterion (BIC) (Schwarz, 1978), the Akaike information criterion (AIC) (Akaike, 1974), and the AIC corrected for finite sample size (AICc) (Hurvich and Tsai, 1989). Each of these quantities measures the relative quality of a model based on the quality of the fit, but also including a penalty on the number of model parameters. Lower values indicate a better model in terms of likelihood and parsimony.

The models we compare are the *full model* given by Eq. (13) and the *reduced model* given by Eq. (23), an approximation to Eq. (13) obtained by setting $L_{sr} = 1$ and $L_{rs} = L^2$ which enables G_{se} and G_{esr} to be replaced by Y , thereby reducing the dimensionality of the model (Robinson et al., 2002).

We fit each of the models to a set of 2141 spectra from the Cz electrode in wake with eyes closed, obtained from the Brain Resource International Database, an archive of electrophysiological and psychophysiological measures, demographics, and psychometric tests (Gordon et al., 2005). Example fits are shown in Fig. 16. For each fit, we can then compute the BIC, AIC, and AICc.

Table 2 shows the mean values of the BIC, AIC, and AICc for each of the models, averaged over subjects. We find that all of these quantities are minimized for the reduced model, indicating that of the models tested, the reduced model offers the best balance

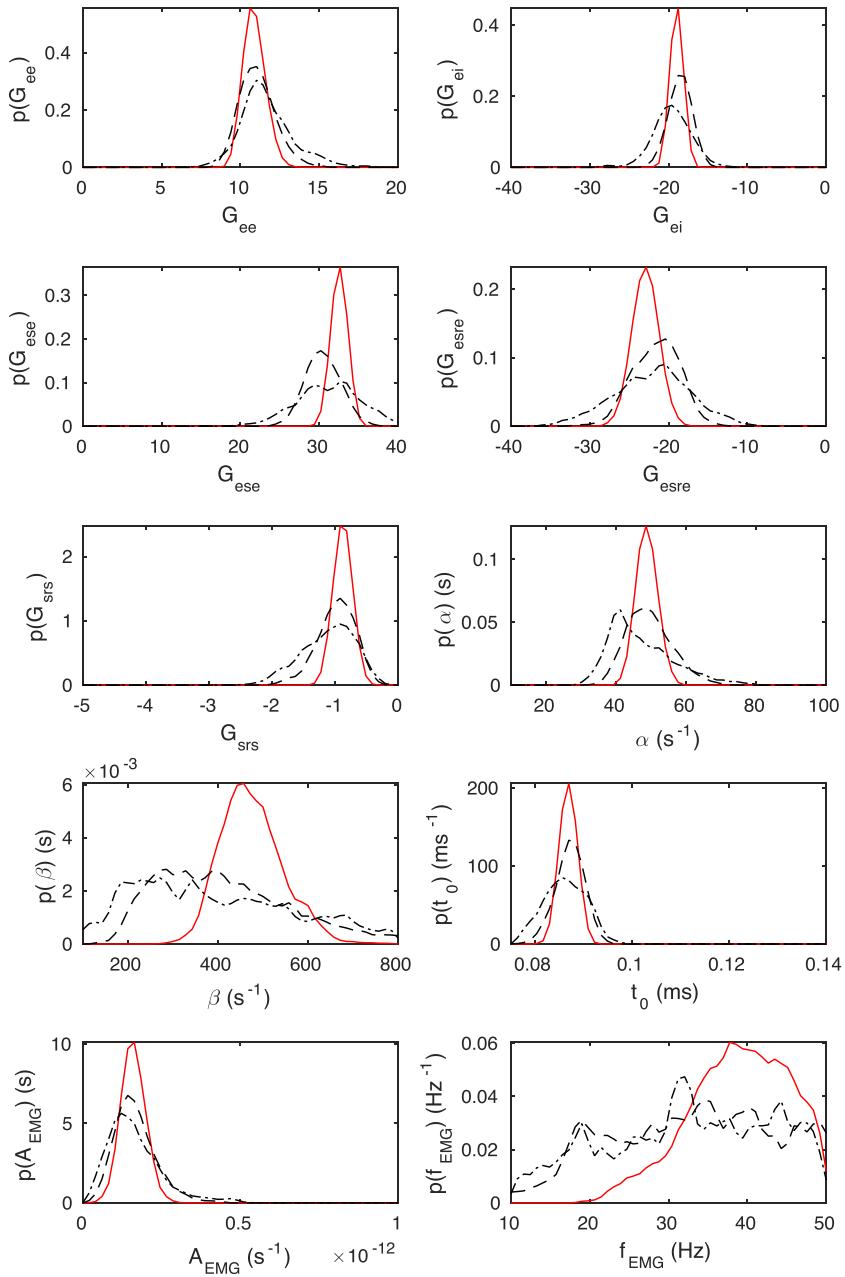


Fig. 15. (Color online) Comparison of marginal posterior distributions for each parameter after 20 s of data, with fitting and incremental updates of the prior every second (red solid line), every 5 s (dashed line), and every 10 s (alternating dots and dashes). As the time between fits increases, the posterior distributions widen. The limits on the horizontal axes correspond to the allowed parameter ranges in Table 1.

between the number of fitted parameters and the quality of the fit.

We can quantify the quality of the fitted spectrum in terms of χ^2 . Qualitatively, fits with $\chi^2 < 4$ correspond to acceptable fits in terms of reproducing key features of the power spectrum, such as the location and size of peaks, and overall slope. As shown in Table 2, the reduced model provides a good fit to the experimental data for around 80% of the subjects. For the remaining 20%, the main complication is that fitting the size and shape of the alpha peak is incompatible with fitting the low frequency part of the spectrum. This issue arises because the simplifications that reduce the number of fitted parameters in the reduced model result in stronger correlations between the low frequency power and the alpha peak than in the full model. In such cases, the weights in Eq. (28) can be adjusted to prioritize fitting particular features in the spectrum

(Rowe et al., 2004b; van Albada et al., 2010). For example, to obtain an accurate estimate of t_0 , it is more important to fit the alpha peak than to fit the low frequency slope. This can be achieved by increasing the weights for the frequencies corresponding to the alpha peak, which will produce a close fit to the alpha peak at the expense of fit accuracy at frequencies outside the alpha band. The adjustment to the weights could potentially be performed automatically based on features in the power spectrum, which would enable such modifications to be consistently applied across subjects and across wake and sleep, but this generalization is beyond the scope of the present study.

The full model is able to obtain high quality fits for around 99% of the subjects without adjusting the frequency weights, as shown in Table 2. However, this robustness comes at the expense of higher BIC and AIC, due to the increased number of parameters and model

Table 2

Comparison of mean BIC, AIC, and AICc values for a selection of models. Also shown are the number of fitted parameters, the mean value of χ^2 , and the percentage of fits with $\chi^2 < 4$.

Model	BIC	AIC	AICc	n	Mean χ^2	$\chi^2 < 4$ (%)
Full	53.4	21.7	23.1	10	1.7	99.2
Full (fixed f_{EMG})	48.3	19.8	20.9	9	1.8	99.0
Full (no EMG)	44.3	18.9	19.8	8	2.9	86.8
Reduced	44.6	19.2	20.1	8	3.2	77.1
Reduced (fixed f_{EMG})	39.3	17.1	17.8	7	3.1	78.7
Reduced (no EMG)	34.1	15.1	15.5	6	3.1	80.3

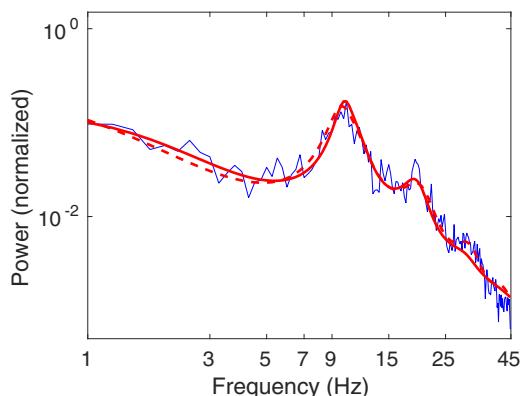


Fig. 16. (Color online) Comparison of fits using the full model (red, solid line) and the reduced model with no EMG (red, dashed line) to an experimental power spectrum in wake with eyes closed (blue). The experimental spectrum has been normalized to unit power.

complexity. The choice of model depends on the intended usage, and particularly the relative importance of increased robustness versus a smaller, more tractable parameter space.

5. Summary and discussion

We have developed a real-time system to fit the predictions of a physiologically-based neural field model of brain activity to experimental EEG. Real brain states change continuously, unlike the discrete transitions between traditional sleep stages. While sleep stages have historically served as a useful summary measure of sleep, they hinder investigation of the underlying dynamics of brain states. In this study, we have replaced sleep stages with physiologically meaningful quantities that vary continuously, reflecting the intrinsic continuous nature of brain states. Our main results are:

- (i) We have developed an automated EEG fitting routine based on Markov chain Monte-Carlo sampling. The key feature of our method is that temporal continuity of parameter values is enforced via Bayes's theorem. This significantly decreases the uncertainty in the fitted parameters because information about the system is accumulated over time.
- (ii) Our algorithm automatically estimates the size and shape of the parameter distributions, and does not require any advance knowledge of the EEG data. It is therefore able to adapt to different subjects and does not require customization to individuals. The adaptations reflect the uncertainty and rate of change of the model parameters for the individual, which means that they are physiologically informative. They also improve the computational efficiency of the method.
- (iii) The fitting routine robustly tracks arousal parameters over the sleep–wake cycle. Probability distributions for each parameter can be determined, providing detailed information about the likelihood of different parameter values in addition to the

fitted parameters. The probability distributions provide detailed information about the uncertainty in each fitted parameter, including parameter interdependencies and multimodal distributions.

- (iv) The fitting routine can be run in real-time, which opens new clinical applications. In this study, we have examined several options for adjusting the performance and accuracy of our algorithm depending on the application. We are able to obtain high-quality fits on a typical 2015 desktop computer in under 10 seconds, and high-quality posterior distributions in under 60 seconds. Our algorithm is readily parallelized, which can provide a further increase in speed if required.
- (v) Our initial analysis of sleep in healthy controls shows that a change in corticothalamic feedback strength is a key feature in the transition between wake states and sleep states. This adds to the body of work demonstrating that both cortical and thalamic populations need to be modeled to predict realistic EEG signals (Robinson et al., 2001, 2002; van Albada et al., 2010; Abeysuriya et al., 2014a).
- (vi) We have tested our algorithm on two different models, and by fitting different subsets of the model parameters. This verifies that our method can be generalized to other models of neural activity, and is not restricted to the specific ones studied here.
- (vii) We have demonstrated a framework for quantitative model comparison based on several information criterion measures of the goodness of fit. We find that a simplified version of our model is able to produce high quality fits to 80% of our data set. Fits in the remaining cases can be improved by manual adjustment of the frequency weights, and this could potentially be automated in future work. The full model is more robust, and produces high quality fits for up to 99% of our data set without manual adjustment. This comes at the expense of additional parameters, which also makes analysis of the model more difficult. The model comparison demonstrated here provides an example procedure to quantitatively select the most appropriate model depending on the research context.

Our approach provides a quantitative, physiologically based, and fully automated alternative to traditional sleep staging. Moreover, the ability to perform real-time tracking of brain states also has other potential clinical applications. A range of neurodegenerative diseases including Alzheimer's disease and Parkinson's disease are accompanied by changes in EEG activity (Benz et al., 2014; Babiloni et al., 2013). Our fitting approach enables the observed changes in EEG to be related to changes in physiology, which provides a new way to relate the mechanisms underlying these diseases to changes in neural activity. Similarly, we have previously associated with EEG changes that accompany states of reduced consciousness, such as vegetative states, comas, or anesthesia, with changes in model parameters (Robinson et al., 2002). These changes can now also be quantified and related to physiology on an individual basis using our new algorithm.

Real-time parameter tracking can also be applied to seizure prediction. The transition into a seizure corresponds to the model parameters crossing a stability boundary as they evolve over time (Robinson et al., 2002; Breakspear et al., 2006). Because crossing the stability boundaries of the model corresponds to seizure onset, the proximity to a stability boundary can be calculated from the fitted parameters and interpreted as a risk factor for a seizure. Further, the parameter trajectory provides information about the evolution of the system toward or away from the stability boundary. As a consequence of a specific neurological condition, seizures may always begin from the same region of parameter space. Using real-time tracking, susceptible individuals could thus be warned if their fitted parameters approached these 'danger zones'.

The fact that the fitting algorithm developed in this study can be applied to a range of different models opens a wide range of possibilities for future work. Not all brain states can necessarily be fully distinguished based on the EEG spectrum alone, but they may be more clearly separated when auxiliary measures such as electrooculography, electromyography, or actigraphy are considered. These measurements can be readily incorporated into the method developed in this study to further distinguish brain states. Notably, the fitted EMG component of the EEG power spectrum can be used to estimate muscle activity, and our preliminary analysis shows that fitting EMG could potentially provide an alternative to separate actigraphy measurements. Finally, future work will be able to use this method to fit alternate models, such as those with additional subthalamic populations. In particular, we have previously incorporated spatial variations in the model parameters to predict spatially localized EEG activity (O'Connor and Robinson, 2004, 2005). By introducing parameters for the spatial variations, the algorithm developed here could be used to fit local EEG phenomena, an aspect we are currently developing.

Acknowledgements

This work was supported by the Australian Research Council Center of Excellence for Integrative Brain Function (ARC Grant CE140100007), ARC Laureate Fellowship Grant FL1401000225, the National Health and Medical Research Council (through the Center of Research Excellence, NeuroSleep, and through the Center for Integrated Research and Understanding of Sleep), and the Westmead Millennium Institute. We acknowledge the support of the Brain Resource International Database (under the auspices of Brain Resource; www.brainresource.com) for data acquisition and processing. All scientific decisions are made independently of any BR commercial decisions via the independently operated scientific division, BRAINnet (www.brainnet.net), which is overseen by the Brain Dynamics Center at Westmead Hospital and scientist members.

References

- Abeysuriya RG, Rennie CJ, Robinson PA. Prediction and verification of nonlinear sleep spindle harmonic oscillations. *J Theor Biol* 2014a;344:70–7.
- Abeysuriya RG, Rennie CJ, Robinson PA. Physiologically based arousal state estimation and dynamics. *J Neurosci Methods* 2015;253:55–69.
- Abeysuriya RG, Rennie CJ, Robinson PA, Kim JW. Experimental observation of a theoretically predicted nonlinear sleep spindle harmonic in human EEG. *J Clin Neurophysiol* 2014b;125:2016–23.
- Akaike H. A new look at the statistical model identification. *IEEE Trans Autom Control* 1974;19:716–23.
- Ancoli-Israel S, Cole R, Alessi C, Chambers M, Moorcroft W, Pollak CP. The role of actigraphy in the study of sleep and circadian rhythms. *Sleep* 2003;26:342–92.
- Andrieu C, Thomas J. A tutorial on adaptive MCMC. *Stat Comput* 2008;18:343–73.
- Babiloni C, Carducci F, Lizio R, Vecchio F, Baglieri A, Bernardini S, Cavedo E, Bozzao A, Buttinelli C, Esposito F, Giubilei F, Guizzaro A, Marino S, Montella P, Quattrochi CC, Redolfi A, Soricelli A, Tedeschi G, Ferri R, Rossi-Fedele G, Ursini F, Scarscia F, Vernieri F, Pedersen TJ, Hardemark HG, Rossini PM, Frisoni GB. Resting state cortical electroencephalographic rhythms are related to gray matter volume in subjects with mild cognitive impairment and Alzheimer's disease. *Hum Brain Mapp* 2013;34:1427–46.
- Benz N, Hatz F, Bousleiman H, Ehrenspurger MM, Gschwandtner U, Hardmeier M, Ruegg S, Schindler C, Zimmermann R, Monsch AU, Fuhr P. Slowing of EEG background activity in Parkinson's and Alzheimer's disease with early cognitive dysfunction. *Front Aging Neurosci* 2014;6:1–6.
- Braitenberg V, Schüz A. Cortex: statistics and geometry of neuronal connectivity. 2nd ed. Berlin: Springer; 1998.
- Breakspear M, Roberts JA, Terry JR, Rodrigues S, Mahant N, Robinson PA. A unifying explanation of primary generalized seizures through nonlinear brain modeling and bifurcation analysis. *Cereb Cortex* 2006;16:1296–313.
- Caples SM, Somers VK. Autonomic dysregulation during REM Sleep. In: Stickgold R, Walker MP, editors. *The Neuroscience of Sleep*. London: Academic Press; 2009.
- Dadok VM, Kirsch HE, Sleigh JW, Lopour Ba, Szeri AJ. A probabilistic framework for a physiological representation of dynamically evolving sleep state. *J Comput Neurosci* 2014;37:105–24.
- Denison DGT, Mallick BK, Smith AFM. Automatic Bayesian curve fitting. *J R Stat Soc Series B Stat Methodol* 1998;60:333–50.
- D'Rozario AL, Dungan GC, Banks S, Liu PY, Wong KKH, Killick R, Grunstein RR, Kim JW. An automated algorithm to identify and reject artefacts for quantitative EEG analysis during sleep in patients with sleep-disordered breathing. *Sleep Breath* 2014;19:607–15.
- D'Rozario AL, Kim JW, Wong KKH, Bartlett DJ, Marshall NS, Dijk DJ, Robinson PA, Grunstein RR. A new EEG biomarker of neurobehavioural impairment and sleepiness in sleep apnea patients and controls during extended wakefulness. *Clin Neurophysiol* 2013;124:1605–14.
- Gordon E, Cooper N, Rennie CJ, Hermens D, Williams LM. Integrative neuroscience: the role of a standardized database. *Clin EEG Neurosci* 2005;36:64–75.
- Haario H, Saksman E, Tamminen J. An adaptive Metropolis algorithm. *Bernoulli* 2001;7:223–42.
- Huber R, Ghilardi MF, Massimini M, Ferrarelli F, Riedner BA, Peterson MJ, Tononi G. Arm immobilization causes cortical plastic changes and locally decreases sleep slow wave activity. *Nat Neurosci* 2006;9:1169–76.
- Huber R, Ghilardi MF, Massimini M, Tononi G. Local sleep and learning. *Nature* 2004;430:78–81.
- Hurvich CM, Tsai CL. Regression and time series model selection in small samples. *Biometrika* 1989;76:297.
- Iber C, Ancoli-Israel S, Chesson A, Quan SF. The AASM Manual for the Scoring of Sleep and Associated Events: Rules, Terminology and Technical Specification. Winchester, Illinois: American Academy of Sleep Medicine; 2007.
- Liley DTJ, Wright JJ. Intracortical connectivity of pyramidal and stellate cells: estimates of synaptic densities and coupling symmetry. *Netw: Computat Neural Syst* 1994;5:175–89.
- Liu ZW, Faraguna U, Cirelli C, Tononi G, Gao XB. Direct evidence for wake-related increases and sleep-related decreases in synaptic strength in rodent cortex. *J Neurosci* 2010;30:8671–5.
- Martin JL, Hakim AD. Wrist actigraphy. *Chest* 2011;139:1514–27.
- Massimini M, Tononi G, Huber R. Slow waves, synaptic plasticity and information processing: insights from transcranial magnetic stimulation and high-density EEG experiments. *Eur J Neurosci* 2009;29:1761–70.
- Metropolis N, Rosenbluth AW, Rosenbluth MN, Teller AH, Teller E. Equation of state calculations by fast computing machines. *J Chem Phys* 1953;21:1087–92.
- Norman RG, Pal I, Stewart C, Walsleben JA, Rapoport DM. Interobserver agreement among sleep scorers from different centers in a large dataset. *Sleep* 2000;23:901–8.
- O'Connor SC, Robinson PA. Spatially uniform and nonuniform analyses of electroencephalographic dynamics, with application to the topography of the alpha rhythm. *Phys Rev E* 2004;70:11911.
- O'Connor SC, Robinson PA. Analysis of the electroencephalographic activity associated with thalamic tumors. *J Theor Biol* 2005;233:271–86.
- O'Connor SC, Robinson PA, Chiang AKL. Wave-number spectrum of electroencephalographic signals. *Phys Rev E* 2002;66:1–12.
- Press WH, Teukolsky SA, Vetterling WT, Flannery BP. Numerical recipes in C. 2nd ed. Cambridge: Cambridge University Press; 1992.
- Punkskaya E, Andrieu C. Bayesian curve fitting using MCMC with applications to signal segmentation. *IEEE Trans Signal Process* 2002;50:747–58.
- Rechtschaffen A, Kales A. A manual of standardized terminology, techniques and scoring system for sleep stages of human subjects. Washington DC: US Government Printing Office, US Public Health Service; 1968.
- Rennie CJ, Robinson PA, Wright JJ. Effects of local feedback on dispersion of electrical waves in the cerebral cortex. *Phys Rev E* 1999;59:3320–9.
- Rennie CJ, Robinson PA, Wright JJ. Unified neurophysiological model of EEG spectra and evoked potentials. *Biol Cybernet* 2002;86:457–71.
- Roberts GO, Rosenthal JS. Examples of adaptive MCMC. *J Computat Graph Stat* 2009;18:349–67.
- Roberts JA, Robinson PA. Corticothalamic dynamics: structure of parameter space, spectra, instabilities, and reduced model. *Phys Rev E* 2012;85:011910.
- Robinson PA, Loxley PN, O'Connor SC, Rennie CJ. Modal analysis of corticothalamic dynamics, electroencephalographic spectra, and evoked potentials. *Phys Rev E* 2001;63:41909.
- Robinson PA, Rennie CJ, Rowe DL. Dynamics of large-scale brain activity in normal arousal states and epileptic seizures. *Phys Rev E* 2002;65:41924.
- Robinson PA, Rennie CJ, Rowe DL, O'Connor SC. Estimation of multiscale neurophysiological parameters by electroencephalographic means. *Hum Brain Mapp* 2004;23:53–72.
- Robinson PA, Rennie CJ, Rowe DL, O'Connor SC, Gordon E. Multiscale brain modelling. *Philos Trans R Soc B* 2005;360:1043–50.
- Robinson PA, Rennie CJ, Rowe DL, O'Connor SC, Wright JJ, Gordon E, Whitehouse RW. Neurophysical modeling of brain dynamics. *Neuropsychopharmacology* 2003a;28:S74–9.
- Robinson PA, Rennie CJ, Wright JJ. Propagation and stability of waves of electrical activity in the cerebral cortex. *Phys Rev E* 1997;56:826–40.
- Robinson PA, Rennie CJ, Wright JJ, Bahramali H, Gordon E, Rowe DL. Prediction of electroencephalographic spectra from neurophysiology. *Phys Rev E* 2001;63:21903.
- Robinson PA, Rennie CJ, Wright JJ, Bourke PD. Steady states and global dynamics of electrical activity in the cerebral cortex. *Phys Rev E* 1998;58:3557–71.
- Robinson PA, Whitehouse RW, Rennie CJ. Nonuniform corticothalamic continuum model of electroencephalographic spectra with application to split-alpha peaks. *Phys Rev E* 2003;68:21922.
- Rosenberg RS, Van Hout S. The American Academy of Sleep Medicine Inter-scorer Reliability Program: sleep stage scoring. *J Clin Sleep Med* 2013;9:81–7.

- Rosenthal JS. Optimal proposal distributions and adaptive MCMC. In: Brooks S, Gelman A, Jones GL, Meng XL, editors. *Handbook of Markov Chain Monte Carlo*. Boca Raton, FL: Chapman & Hall/CRC; 2011 [Ch. 4].
- Rowe DL, Robinson PA, Harris AW, Felmingham KL, Lazzaro IL, Gordon E. Neurophysiologically-based mean-field modelling of tonic cortical activity in post-traumatic stress disorder (PTSD), chronic schizophrenia, first episode schizophrenia and attention deficit hyperactivity disorder (ADHD). *J Integr Neurosci* 2004a;3:453–87.
- Rowe DL, Robinson PA, Rennie CJ. Estimation of neurophysiological parameters from the waking EEG using a biophysical model of brain dynamics. *J Theor Biol* 2004b;231:413–33.
- Saunders MG. Artifacts: activity of noncerebral origin in the EEG. In: Klass DW, Daly DD, editors. *Current practice of clinical electroencephalography*. New York: Raven; 1979. p. 37–67.
- Schwarz G. Estimating the dimension of a model. *Annal Stat* 1978;6:461–4.
- Srinivasan R, Nunez PL, Silberstein RB. Spatial filtering and neocortical dynamics: estimates of EEG coherence. *IEEE Trans Biomed Eng* 1998;45:814–26.
- Stam CJ. Nonlinear dynamical analysis of EEG and MEG: review of an emerging field. *Clin Neurophysiol* 2005;116:2266–301.
- Stam CJ, Pijn JPM, Suffczynski P, Lopes da Silva FH. Dynamics of the human alpha rhythm: evidence for non-linearity? *Clin Neurophysiol* 1999;110:1801–13.
- Steyn-Ross DA, Steyn-Ross ML, Sleigh JW, Wilson MT, Gillies IP, Wright JJ. The sleep cycle modelled as a cortical phase transition. *J Biol Phys* 2005;31:547–69.
- van Albada SJ, Kerr CC, Chiang AKI, Rennie CJ, Robinson PA. Neurophysiological changes with age probed by inverse modeling of EEG spectra. *Clin Neurophysiol* 2010;121:21–38.
- van Albada SJ, Rennie CJ, Robinson PA. Variability of model-free and model-based quantitative measures of EEG. *J Integr Neurosci* 2007;6:279–307.
- Vyazovskiy VV, Cirelli C, Pfister-Genskow M, Faraguna U, Tononi G. Molecular and electrophysiological evidence for net synaptic potentiation in wake and depression in sleep. *Nat Neurosci* 2008;11:200–8.
- Wang D, Teichtahl H, Drummer O, Goodman C, Cherry G, Cunningham D, Kronborg I. Central sleep apnea in stable methadone maintenance treatment patients. *Chest* 2005;128:1348–56.
- Wright JJ, Liley DTJ. A millimetric-scale simulation of electrocortical wave dynamics based on anatomical estimates of cortical synaptic density. *Netw: Computat Neural Syst* 1994;5:191–202.