Modeling the impact of impulsive stimuli on sleep-wake dynamics

B. D. Fulcher,[1] A. J. K. Phillips,[1,2] and P. A. Robinson[1,2,3]
[1]*School of Physics, University of Sydney, New South Wales 2006, Australia*
[2]*Brain Dynamics Center, Westmead Millenium Institute, Westmead Hospital and Western Clinical School, University of Sydney, Westmead, New South Wales 2145, Australia*
[3]*Faculty of Medicine, University of Sydney, New South Wales 2006, Australia*
(Received 12 June 2008; published 20 November 2008)

A neuronal population model of the sleep-wake switch is extended to incorporate impulsive external stimuli. The model includes the mutual inhibition of the sleep-active neurons in the hypothalamic ventrolateral preoptic area (VLPO) and the wake-active monoaminergic brainstem populations (MA), as well as circadian and homeostatic drives. Arbitrary stimuli are described in terms of their relative effects on the VLPO and MA nuclei and represent perturbations on the normal sleep-wake dynamics. By separating the model's intrinsic time scales, an analytic characterization of the dynamics in a reduced model space is developed. Using this representation, the model's response to stimuli is studied, including the latency to return to wake or sleep, or to elicit a transition between the two states. Since sensory stimuli are known to excite the MA, we correspondingly investigate the model's response to auditory tones during sleep, as in clinical sleep fragmentation studies. The arousal threshold is found to vary approximately linearly with the model's total sleep drive, which includes circadian and homeostatic components. This relationship is used to reproduce the clinically observed variation of the arousal threshold across the night, which rises to a maximum near the middle of the night and decreases thereafter. In a further application of the model, time-of-night arousal threshold and body temperature variations in an experimental sleep fragmentation study are replicated. It is proposed that the shift of the extrema of these curves to a greater magnitude later in the night is due to the homeostatic impact of the frequent nocturnal disturbances. By modeling the underlying neuronal interactions, the methods presented here allow the prediction of arousal state responses to external stimuli. This methodology is fundamentally different to previous approaches that model the clinical data within a phenomenological framework. As a result, a broader understanding of how impulsive external stimuli modulate arousal is gained.

DOI: 10.1103/PhysRevE.78.051920 PACS number(s): 87.10.Ed, 87.18.Sn, 87.18.Yt, 87.18.Vf

I. INTRODUCTION

It is currently understood that three main drives regulate sleep-wake dynamics: a *homeostatic* "sleep pressure" that builds during wake and is relieved with sleep, a ~24 h periodic *circadian* oscillation, which is entrained to the day-night cycle by photic input, and an *ultradian* rhythm, which controls the architecture of sleep, including transitions between rapid eye movement (REM) and nonrapid eye movement (NREM) sleep. Sleep-wake modeling derives much from the two-process model of Borbély [1] which, motivated by the variation of electroencephalogram (EEG) slow-wave activity (SWA) [2], combined the homeostatic and circadian drives into a conceptual framework that was able to explain the timing of sleep and wake. Quantitative implementations of the two-process model have since been able to reproduce a wide variety of sleep-wake phenomena, including subjective fatigue during sleep deprivation, internal desynchronization, fragmented sleep during continuous bedrest, and the sleep durations of shift workers [3], as well as ultradian variations in SWA [4]. Quantitative phenomenological sleep models proposed since build on the precedent established by the two-process model [5,6]. However, while successful for the protocols they have been calibrated against [7], the predictive power of these models for general regimens is limited [8].

By incorporating knowledge of the neuronal interactions that give rise to sleep-wake dynamics, physiologically based sleep models can be developed that provide insights into the neurological processes that underpin their behavioral predictions. Sleep-wake dynamics are characterized by rapid transitions between two distinct states: sleep and wake, and result from the mutual inhibition between sleep-active neurons in the ventrolateral preoptic area (VLPO) of the hypothalamus and wake-active monoaminergic neuronal populations (MA) in the brainstem [9]. The activity of each group suppresses the activity of the other, which reduces the subsequent inhibition on itself, i.e., the firing of each group indirectly reinforces its own firing. This behavior results in extended periods of either sleep (activated VLPO and suppressed MA) or wake (activated MA and suppressed VLPO), with rapid transitions between the two states dictated by the slowly varying circadian and homeostatic sleep drives; much like a "flip-flop" circuit [10]. The circadian signal is the result of molecular-level oscillations in the suprachiasmatic nucleus (SCN) [11], while the homeostatic sleep drive is believed to result from the accumulation of sleep-promoting metabolic by-products [12]. Physiologically based sleep models have recently been constructed that incorporate these neurological drives and interactions. Their parameters are in principle physiologically measurable and dynamical restrictions on model parameters can be used to infer physiological information. A number of such models have been proposed [13–15], but they tend to incorporate large numbers of neuronal populations and drives, with many resulting free parameters. The Phillips-Robinson sleep model [16] used in this paper set a more modest target of modeling the well-

substantiated physiology of the sleep-wake switch. Its parameters are fitted from a small number of experiments yet it is able to predict the results of many. Thus far, the model has been able to predict optimal recovery protocols from sleep deprivation, sleep latency times, and has offered qualitative descriptions of pathologies including narcolepsy [16,17].

External stimuli that disturb the normal sleep-wake cycle are prevalent in reality and it is therefore important that they are thoroughly understood. Such stimuli may promote arousal in the form of stimulants such as caffeine and sensory reactions to bright light or loud noise, or promote sleep in the form of inactivity, dim light, or sedative drugs including antihistamines. Despite the success of quantitative phenomenological models in reproducing a wide variety of clinical observations, the effects of general external stimuli have not yet received a detailed treatment. Indeed, the incorporation of external stimuli into phenomenological models poses a particular challenge because a physiological foundation is lacking. In contrast, as we will demonstrate, knowledge of the mechanistic pathway of an external stimulus provides a natural avenue for its inclusion into a physiologically based model. The major benefit of this approach is its ability to encompass a wide variety of stimuli on a common scale, in terms of their impact on the relevant arousal state nuclei. Ultimately, the behavioral outcome of a given stimulus can be understood in terms of its underlying neurological mechanism.

Our aims in this paper are threefold: (i) to present a rigorous investigation of the general dynamics of the Phillips-Robinson model, (ii) to illustrate its correspondence with several specific clinical measures, and (iii) to lay the groundwork for future, more clinically oriented investigations. In Sec. II, we summarize the relevant physiology and introduce the Phillips-Robinson model. A theoretical characterization of the model is undertaken in Sec. III, including a separation of model time scales that allows an equilibrium and linear stability analysis in a reduced two-dimensional model space. In Sec. IV, we consider the impact of general impulsive stimuli characterized by drives to the sleep-active VLPO and wake-active MA populations, explaining the results in terms of the dynamical features of the model. In Sec. V, we consider arousing impulses during sleep, replicating the time-of-night arousal threshold curves observed clinically. A subsequent calibration from the intensity of an auditory stimulus to the mean increase in MA cell-body potential required to provoke an arousal allows us to reproduce the arousal threshold and body temperature variation in another sleep fragmentation study.

II. PHYSIOLOGY AND MODELING

We begin with a brief summary of the relevant sleep-wake physiology and an overview of the Phillips-Robinson model [16].

A. Physiology

The overall arousal state of the brain is controlled by a series of brainstem nuclei, collectively termed the ascending arousal system (AAS), which diffusely project neuromodulators to the cerebrum [11,18,19]. Based on common effects and temporal patterns of activity, brainstem AAS nuclei can be broadly classified as monoaminergic (MA) or acetylcholine related (ACh). The MA group includes nuclei that use monoaminergic neurotransmitters: the histaminergic tuberomammillary nucleus (TMN), norepinephrinergic locus coeruleus (LC), serotoninergic dorsal raphé (DR) nucleus, and dopaminergic ventral tegmental area (VTA) [19–21]. The ACh group includes nuclei that express acetylcholine, including the pedunculopontine (PPT) and laterodorsal (LDT) tegmental nuclei in the mesopontine tegmentum [22,23]. The orexin population of the lateral hypothalamus excites the MA and ACh neuronal populations during wake, thereby acting to stabilize the waking state [24,25].

Circadian and homeostatic drives are integrated in the ventrolateral preoptic area (VLPO) of the hypothalamus, affecting the AAS via inhibitory GABA-ergic projections from there [11,26,27]. The 24 h periodic circadian signal originates in the suprachiasmatic nucleus (SCN) and maintains the entrainment of the sleep-wake cycle to the light cycle through projections to the VLPO, primarily via the dorsomedial nucleus of the hypothalamus (DMH) [11,27,28]. The homeostatic drive increases with time spent awake, which is believed to be due to a net build up of the somnogenic metabolic by-product adenosine in the basal forebrain [19,29,30], disinhibiting the VLPO [11,27]. During sleep, metabolic rates are low and adenosine is cleared faster than it is produced [30–32]. Mutual inhibition between the MA group and the VLPO gives rise to the sleep-wake "flip-flop," with each group disinhibiting its own activity [9,10]. During wake, the MA group is active and the VLPO is suppressed, while the converse occurs during sleep.

B. Phillips-Robinson model

Phillips and Robinson incorporated the above physiology into a neuronal population model [16], considering the average properties of large populations of neurons and their interactions, as has been done previously for the corticothalamic system [33–35]. Each population $j=m,v$, where m represents the MA and v represents the VLPO, has a mean cell-body potential $V_j(t)$ relative to resting and a mean firing rate $Q_j(t)$. The relationship of Q_j to V_j is approximated by a sigmoid [33], with

$$Q_j = S(V_j) = \frac{Q_{\max}}{1+\exp[-(V_j-\theta)/\sigma']}, \tag{1}$$

where $Q_{\max}$ is the maximum possible firing rate, θ is the mean firing threshold relative to resting, and $\sigma'\pi/\sqrt{3}$ is its standard deviation [35]. Due to the small spatial extent of the relevant nuclei, we assume spatial homogeneity of each population and neglect propagation delays.

Neuronal dynamics are represented by

$$\tau_v \dot{V}_v + V_v = \nu_{vm} Q_m + D_v, \tag{2}$$

$$\tau_m \dot{V}_m + V_m = \nu_{mv} Q_v + D_m, \tag{3}$$

where the ν_{jk} weight the input from population k to j, τ_j is the decay time for the neuromodulator expressed by group j, and D_j represents the external drive to population j. In the presence of external stimuli, the drive terms can be written in the form

$$D_v = D_v^0 + \Delta D_v, \tag{4}$$

$$D_m = D_m^0 + \Delta D_m, \tag{5}$$

where normal sleep-wake dynamics are governed by the drives D_j^0, and the ΔD_j represent perturbations. It is convenient to represent the drives in vector form so that the sum of the unperturbed drive $\mathbf{D}^0=(D_v^0, D_m^0)$ and the perturbation $\Delta\mathbf{D}=(\Delta D_v, \Delta D_m)$ gives the total drive $\mathbf{D}=(D_v, D_m)=\mathbf{D}^0+\Delta\mathbf{D}$. All types of perturbative external stimuli, including sensory and pharmaceutical effects, are encapsulated in $\Delta\mathbf{D}$.

Nominal input to the VLPO,

$$D_v^0 = \nu_{vc} C + \nu_{vh} H, \tag{6}$$

includes circadian (C) and homeostatic (H) components. The 24 h periodic circadian drive is taken to be well-entrained to the daily fluctuation in light intensity and is approximated as a sinusoidal function of time

$$C(t) = \sin \omega t + c_0, \tag{7}$$

where $\omega=(2\pi/24)$ h^{-1} and c_0 is a constant offset. An oscillation amplitude of unity is used without loss of generality—the actual amplitude is absorbed into ν_{vc}.

The homeostatic sleep drive is represented by the somnogen level H, which represents the concentration of extracellular adenosine in the basal forebrain. The somnogen clearance rate is assumed to be proportional to its concentration, with production approximated as a linear function of Q_m because MA activity is well-correlated with arousal [36] and, presumably, metabolism. Assuming that $Q_m=0$ results in negligible net production yields

$$\chi \dot{H} + H = \mu Q_m, \tag{8}$$

where χ is the characteristic somnogen clearance time and μ is a constant. The value of H grows during wake to a maximum at sleep onset, and decreases during sleep to a minimum at the end of a sleep period.

The unperturbed drive D_m^0 to the MA represents input from groups including acetylcholine and orexin. The ACh group is influenced by the homeostatic variation of adenosine concentration [12,37], and its interaction with the MA group may produce the ~90 min ultradian oscillation [11]. Orexinergic activity also varies, with factors including emotion, energy homeostasis, and reward systems [25]. However, in the current form of the model, the time-varying input from these groups is not explicitly considered, instead being averaged to a constant, intermediate level of activity $D_m^0=A$. This approximation allows us to focus on the "flip-flop" dynamics between sleep and wake. A dynamic drive D_m^0 would modulate the arousal state within sleep and wake and will be explored in a future version of the model.

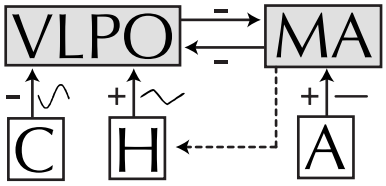

FIG. 1. Schematic of the Phillips-Robinson model of the sleep-wake switch. The mutually inhibitory VLPO and MA populations form the core of the model, which also includes circadian (C) and homeostatic (H) drives to the VLPO and the constant drive A to the MA. Excitatory (+) and inhibitory (−) interactions are shown with arrows and arousal state feedback from the MA to H is shown dashed. A schematic of the typical time evolution of each drive is also displayed [cf., Figs. 2(c) and 2(d)].

The above formulation is represented schematically in Fig. 1. Model parameters have been constrained elsewhere by both dynamics and physiology to produce realistic sleep-wake behavior [16,17], with current values shown in Table I. When external perturbations are excluded ($\Delta\mathbf{D}=0$) we obtain "normal" output, shown as a time series in Fig. 2. The arousal state alternates between periods of wake with high V_m, low V_v, and increasing H, and periods of sleep with high V_v, low V_m, and decreasing H. Plotting V_m or V_v against the oscillatory D_v, as shown in Fig. 3, reveals hysteresis—the mutual inhibition between the VLPO and MA groups gives rise to sleep and wake branches at high and low D_v, respectively, and an intermediate bistable region in which both states exist simultaneously. As D_v oscillates with a 24 h period, the arousal state evolves around the hysteresis loop, alternating between sleep and wake accordingly.

III. ANALYSIS OF THE MODEL

In this section, different representations of the model dynamics are explored. Following a description of the reduced manifold in V_v-V_m-D_v space, a linear stability analysis in V_v-V_m space is undertaken that exploits a separation of the model's intrinsic time scales. The analysis in this section provides a convenient framework with which to study the model's response to impulsive external stimuli, which is covered in Secs. IV and V.

A. Separation of time scales

In the absence of external stimuli ($\Delta\mathbf{D}=0$), the Phillips-Robinson model is formulated in terms of two distinct time scales: that of the sleep-wake transitions (~10 min) gov-

TABLE I. Nominal model parameter values obtained from[17].

Parameter	Value	Parameter	Value	Parameter	Value
$Q_{\max}$	100 s^{-1}	ν_{vm}	−2.1 mV s	μ	4.4 nM s
θ	10 mV	ν_{mv}	−1.8 mV s	χ	45 h
σ'	3 mV	ν_{vh}	1 mV nM^{-1}	τ_v, τ_m	10 s
A	1.3 mV	ν_{vc}	−2.9 mV	c_0	4.5

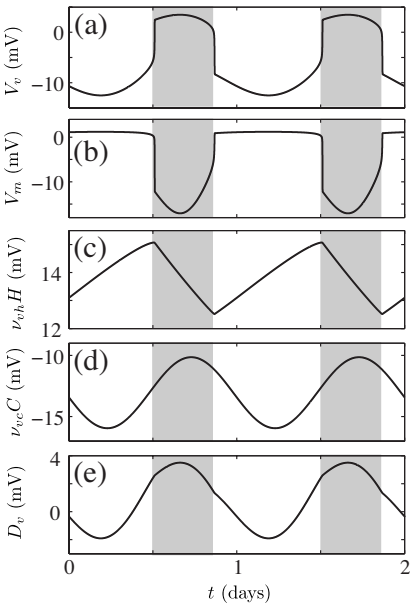

FIG. 2. Normal, unperturbed output from the Phillips-Robinson model. Time series are shown for (a) V_v [Eq. (2)], (b) V_m [Eq. (3)], (c) $\nu_{vh}H$ [cf., Eq. (8)], (d) $\nu_{vc}C$ [cf., Eq. (7)], and (e) $D_v=D_v^0=\nu_{vc}C+\nu_{vh}H$ [Eq. (6)]. Sleep periods are shaded.

erned by the time constants τ_m and τ_v, and that of the drive D_v ($\sim$1 day), governed by the time constants ω^{-1} and χ. Since τ_m, $\tau_v \ll \omega^{-1}$, χ, it follows that V_v [Eq. (2)] and V_m [Eq. (3)] evolve on a much faster time scale than the circadian [Eq. (7)] and homeostatic [Eq. (8)] drives. On short time scales we can therefore approximate $\dot{D}_v \approx 0$. In this regime, D_v can be treated as a control parameter and the system simplifies to the *layer problem* of Fenichel theory [38], which is characterized by rapid attraction onto the *reduced manifold* defined by $\dot{V}_v=\dot{V}_m=0$. On longer time scales we

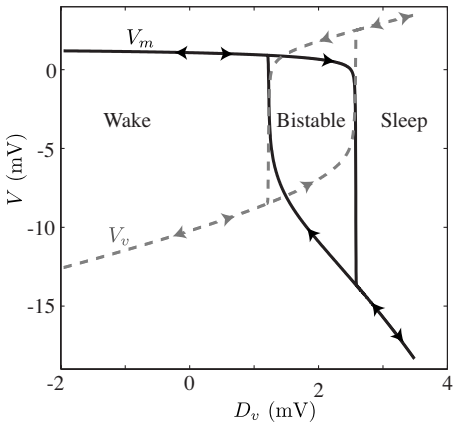

FIG. 3. Hysteresis in the Phillips-Robinson model. The mean cell-body potentials V_m (solid black line) and V_v (dashed gray line) are plotted against D_v, showing hysteresis. Wake (low D_v), sleep (high D_v), and bistable (intermediate) regions are labeled. The arousal state alternates between sleep and wake as the 24 h periodic sleep drive D_v oscillates with time. Note that this plot is equivalent to plotting the time series in Figs. 2(a) and 2(b), for V_v and V_m, respectively, against that shown in Fig. 2(e) for D_v.

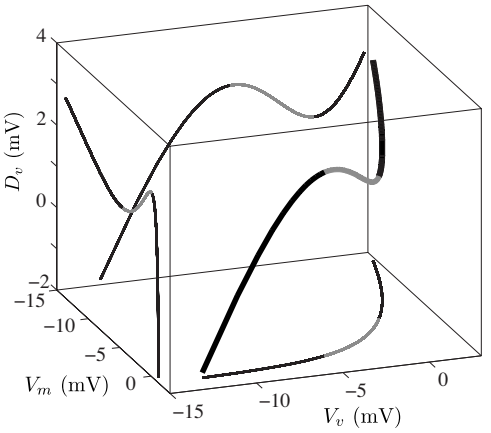

FIG. 4. The reduced manifold in V_v-V_m-D_v space. Stable sleep (high V_v, low V_m) and wake (low V_v, high V_m) branches (black lines) are connected by the intermediate unstable branch (gray line). Projections of the manifold onto the two-dimensional subspaces are also plotted.

have the *reduced problem*: attraction onto the reduced manifold is considered complete and evolution through V_v-V_m-D_v space is along the manifold. The reduced manifold consists of stable sleep and wake branches joined by an unstable branch, as shown in Fig. 4. Projections of the reduced manifold onto reduced two-dimensional subspaces yields bifurcation curves showing hysteresis for V_m and V_v as functions of D_v, and a monotonic curve in V_v-V_m space.

B. Reduction to V_v-V_m space

Treating D_v as a control parameter, we are motivated to explore slices of constant D_v in V_v-V_m space. Model dynamics in this space can be understood in terms of the nullclines, $\dot{V}_v=0$:

$$V_v = \nu_{vm}S(V_m) + D_v, \tag{9}$$

and $\dot{V}_m=0$:

$$V_m = \nu_{mv}S(V_v) + D_m. \tag{10}$$

The intersection of the nullclines defines an equilibrium point (V_v^*, V_m^*), which can be written implicitly in terms of V_v^* as

$$\tau_v \dot{V}_v = -V_v^* + \nu_{vm}S[\nu_{mv}S(V_v^*) + D_m] + D_v = 0, \tag{11}$$

or in terms of V_m^* as

$$\tau_m \dot{V}_m = -V_m^* + \nu_{mv}S[\nu_{vm}S(V_m^*) + D_v] + D_m = 0. \tag{12}$$

Movement through the V_v-V_m plane is determined by the velocity vector field

$$\dot{\mathbf{V}} = (\dot{V}_v, \dot{V}_m), \tag{13}$$

which has magnitude $\dot{V}=|\dot{\mathbf{V}}|=(\dot{V}_v^2+\dot{V}_m^2)^{1/2}$.

The drives are slowly varying and are treated as control parameters, hence $\dot{V}_v$ and $\dot{V}_m$ are functions of V_v and V_m

only. The Jacobian matrix $\mathbf{J}$ of the resulting two-dimensional system evaluated at an equilibrium (V_v^*, V_m^*) is

$$\mathbf{J}|_{(V_v^*,V_m^*)} = \left(\begin{array}{cc} \partial \dot{V}_v/\partial V_v & \partial \dot{V}_v/\partial V_m \\ \partial \dot{V}_m/\partial V_v & \partial \dot{V}_m/\partial V_m \end{array} \right) \Bigg|_{(V_v^*,V_m^*)} \tag{14}$$

$$= \left(\begin{array}{cc} -\tau_v^{-1} & \frac{\nu_{vm}}{\tau_v} S'(V_m^*) \\ \frac{\nu_{mv}}{\tau_m} S'(V_v^*) & -\tau_m^{-1} \end{array} \right), \tag{15}$$

where

$$S'(x) = \frac{d}{dx} S(x) = \frac{\exp\left(-\frac{x-\theta}{\sigma'}\right)}{Q_{\max}} S^2(x), \tag{16}$$

which is positive definite. The trace T and determinant Δ of the Jacobian are given by

$$T \equiv \mathrm{Tr}[\mathbf{J}] = -(\tau_m^{-1} + \tau_v^{-1}), \tag{17}$$

$$\Delta \equiv \det[\mathbf{J}] = \tau_m^{-1}\tau_v^{-1}[1 - \nu_{mv}\nu_{vm} S'(V_v^*) S'(V_m^*)], \tag{18}$$

and we have

$$T^2 - 4\Delta = (\tau_m^{-1} - \tau_v^{-1})^2 + 4\tau_m^{-1}\tau_v^{-1}\nu_{mv}\nu_{vm} S'(V_v^*) S'(V_m^*). \tag{19}$$

Since $T<0$ and $T^2-4\Delta>0$, the fixed points are either stable nodes ($\Delta>0$) or saddle points ($\Delta<0$) [39]. The eigenvalues $\lambda_\pm = (T \pm \sqrt{T^2-4\Delta})/2$ of the Jacobian are

$$\lambda_\pm = \frac{1}{2}[-(\tau_m^{-1} + \tau_v^{-1}) \pm \sqrt{(\tau_m^{-1} - \tau_v^{-1})^2 + 4\tau_m^{-1}\tau_v^{-1}\nu_{mv}\nu_{vm} S'(V_v^*) S'(V_m^*)}]. \tag{20}$$

For each equilibrium, there exist invariant linearized subspaces to which the corresponding invariant manifolds of the nonlinear system are tangent, according to the stable manifold theorem [40,41]. We proceed by solving the eigenvalue problem $(\mathbf{J}-\lambda\mathbf{I})\mathbf{v}=0$, where $\mathbf{I}$ is the 2×2 identity matrix and the eigenvector $\mathbf{v}$ is the column vector (v_1, v_2):

$$\left(\begin{array}{cc} -\tau_m^{-1} - \lambda_\pm & \frac{\nu_{mv}}{\tau_m} S'(V_v^*) \\ \frac{\nu_{vm}}{\tau_v} S'(V_m^*) & -\tau_v^{-1} - \lambda_\pm \end{array} \right) \mathbf{v} = \mathbf{0}. \tag{21}$$

Expanding the first row of this matrix equation yields

$$\frac{v_2}{v_1} = \frac{\tau_m(\tau_m^{-1} + \lambda_\pm)}{\nu_{mv} S'(V_v^*)}, \tag{22}$$

the sign of which is opposite to that of $\tau_m^{-1}+\lambda_\pm$ because $\nu_{mv}<0$ and $S'(V_v^*)$, $\tau_m>0$. From Eq. (20),

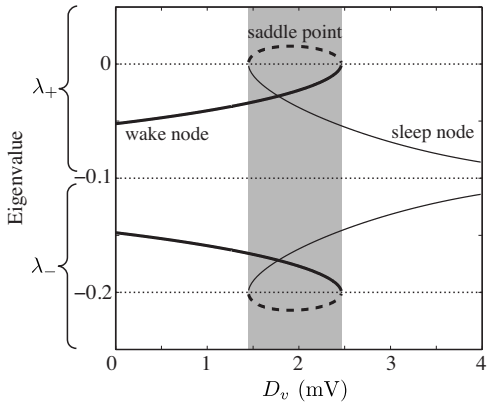

FIG. 5. Eigenvalues $\lambda_\pm$ of the Jacobian as a function of the control parameter D_v at a saddle point (dashed), wake node (thick), and sleep node (solid), as obtained from Eq. (24). The bistable region is shaded and the lines $\lambda=-\tau^{-1}=-0.1$, $\lambda=0$, and $\lambda=-0.2$ are dotted.

$$\tau_m^{-1} + \lambda_\pm = \frac{1}{2}[(\tau_m^{-1} - \tau_v^{-1}) \pm \sqrt{(\tau_m^{-1} - \tau_v^{-1})^2 + 4\tau_m^{-1}\tau_v^{-1}\nu_{mv}\nu_{vm} S'(V_v^*) S'(V_m^*)}]. \tag{23}$$

The magnitude of the square root term in Eq. (23) is greater than $|\tau_m^{-1}-\tau_v^{-1}|$, ensuring that $\tau_m^{-1}+\lambda_+>0$ and $\tau_m^{-1}+\lambda_-<0$. Therefore, from Eq. (22), the eigenvector $\mathbf{v}_+$ (corresponding to λ_+) has $v_2/v_1<0$, and $\mathbf{v}_-$ (corresponding to λ_-) has $v_2/v_1>0$. Also, because $T<0$ implies that $|\lambda_-|>|\lambda_+|$, trajectories evolve rapidly in the $\mathbf{v}_-$ eigendirection (with a positive gradient in the V_v-V_m plane) and approach fixed points along the weaker $\mathbf{v}_+$ eigendirection (with a negative gradient in the V_v-V_m plane). We refer to the fast and slow invariant manifolds of the nonlinear system as W^- and W^+, respectively.

In the current form of the model, τ_m is set equal to τ_v for simplicity (for $\tau_m \neq \tau_v$, the population with the smaller time constant slaves to the dynamics of the population with the larger time constant [16]). Writing $\tau \equiv \tau_v = \tau_m$, the eigenvalues in this case are given by

$$\lambda_\pm = \tau^{-1}[-1 \pm \sqrt{S'(V_v^*) S'(V_m^*)}], \tag{24}$$

and the ratio v_2/v_1 in the two eigendirections are negatives of each other:

$$\frac{v_2}{v_1} = \pm \sqrt{\frac{\nu_{vm} S'(V_v^*)}{\nu_{mv} S'(V_m^*)}}. \tag{25}$$

The eigenvalues $\lambda_\pm$ are plotted as a function of D_v in Fig. 5. Consistent with the hysteresis picture shown in Fig. 3, we have a pair of negative eigenvalues at low D_v corresponding to the stable "wake node," and at high D_v corresponding to the stable "sleep node," with a bistable region in which both stable nodes and a saddle point exist simultaneously. For normal parameters, there exist two saddle-node bifurcations that define the extent of the bistable region, associated with the annihilation of the sleep and wake nodes, respectively. Bifurcation points occur at an equilibrium: $\dot{V}_m(V_m)=0$ [Eq.

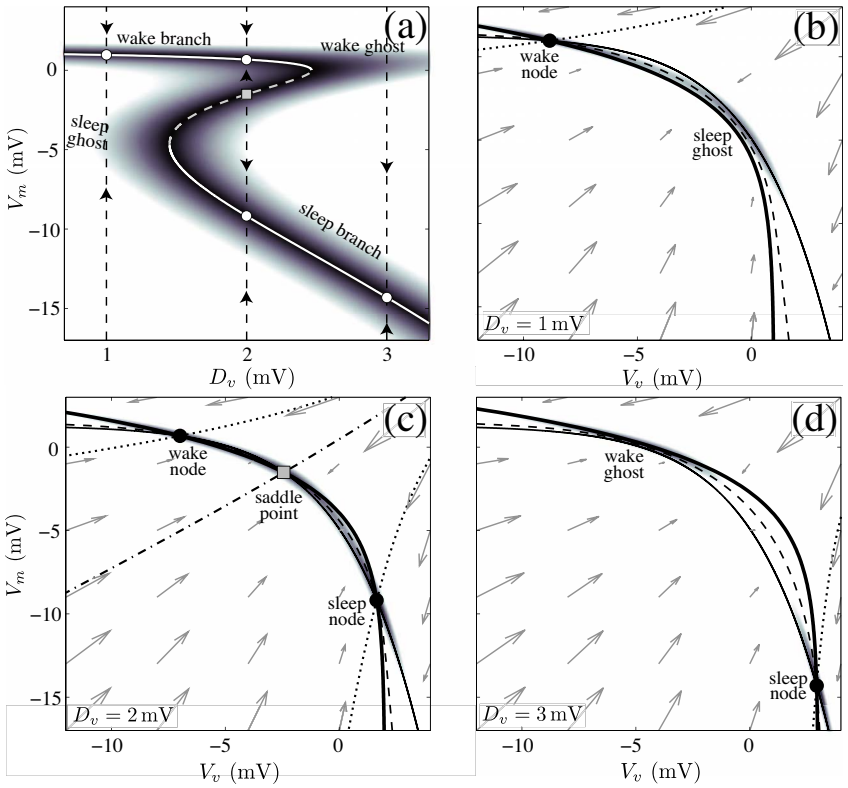

FIG. 6. (Color online) Correspondence between (a) the D_v-V_m hysteresis representation and (b)–(d) the V_v-V_m nullcline representation of the model. (a) Bifurcation diagram for V_m against the control parameter D_v, showing hysteresis. Arrows indicate the direction of $\dot{V}_m$ for $D_v = 1, 2, 3$ mV. Regions of low $\dot{V}$ have been shaded, from $\dot{V}=0$ (black) to $\dot{V}>0.05$ mV s^{-1} (white), where the minimum $\dot{V}$ (across variation in V_v) has been collapsed onto this picture. (b)–(d) V_v-V_m plots showing the V_v nullcline [Eq. (9)] (thick solid line), V_m nullcline [Eq. (10)] (thin solid line), stable nodes (circles), saddle points (squares), the W^+ manifold (dashed line), the W^- manifold of stable nodes (dotted line), the separatrix formed by the stable invariant manifold W^- of the saddle point (dot-dashed line), the velocity vector field $\dot{\mathbf{V}}$ [Eq. (13)] (arrows), and regions over which $\dot{V}<0.1$ mV s^{-1} (shading) for (b) $D_v=1$ mV, a stable wake node, (c) $D_v=2$ mV, in the bistable region, and (d) $D_v=3$ mV, a stable sleep node. Near-stable wake and sleep ghosts (cf., Sec. III D) are also labeled.

(12)] when $d/dV_m[\dot{V}_m(V_m)]=0$ [39]. Combining these conditions, we find that bifurcation points are solutions of

$$S'\left[S^{-1}\left(\frac{V_m - D_m}{\nu_{mv}}\right)\right]S'(V_m) - \nu_{mv}^{-1}\nu_{vm}^{-1} = 0, \qquad (26)$$

for V_m, where the inverse sigmoid $S^{-1}(x)$ is defined over $0<x<Q_{\max}$ by $S^{-1}(x)=-\sigma\ln(Q_{\max}/x-1)+\theta$. The V_v and D_v values corresponding to a bifurcation point can henceforth be determined from Eqs. (10) and (12), respectively. For the nominal model parameters used in this paper, the bistable region extends from $D_v=1.45$ mV (annihilation of the sleep node) to 2.46 mV (annihilation of the wake node). Note that the hysteresis loop attained by integrating the model equations (shown in Fig. 3) is wider than the theoretical extent [shown in Figs. 5 and 6(a)] because the nonzero time constants τ_v and τ_m cause the system to lag behind the equilibrium curve.

C. Nullcline dynamics

Dynamics in V_v-V_m space corresponding to the above analysis can be interpreted in terms of the time evolution of the nullclines. The drive D_v translates the $\dot{V}_v=0$ nullcline [Eq. (9)] horizontally and the drive D_m translates the $\dot{V}_m=0$ nullcline [Eq. (10)] vertically. Taking $\Delta\mathbf{D}=0$, the drive $D_m = D_m^0 = A$ to the MA is constant, and the drive $D_v = D_v^0$ to the VLPO—and hence the position of the V_v nullcline—is oscillatory. Normal dynamics therefore result from the 24 h periodic oscillation of the V_v nullcline relative to the stationary V_m nullcline. The correspondence between the hysteresis and nullcline representations of the model is illustrated in Fig. 6. These two-dimensional representations can be obtained from the reduced manifold in V_v-V_m-D_v space, shown in Fig. 4, by taking a projection onto the D_v-V_m plane [Fig. 6(a)], or slices through constant D_v [Figs. 6(b)–6(d)]. At low and high D_v, the nullclines intersect at a single point, representing the stable wake and sleep nodes, respectively. At intermediate D_v, corresponding to the bistable region of the hysteresis loop, the nullclines have three points of intersection: two stable nodes and one saddle point. The invariant manifolds plotted in Fig. 6 are obtained by following the flow $\dot{\mathbf{V}}$ after applying small perturbations in the directions given by Eq. (25). The W^+ manifolds from the two nodes and the saddle point connect with one another in the bistable region, with the stable invariant manifold W^- of the saddle point, shown as a dot-dashed line in Fig. 6(c), forming a separatrix be-

tween *sleep* and *wake basins* in V_v-V_m space. A trajectory with an initial condition in the sleep or wake basin will attract onto the sleep or wake node, respectively. These characteristic topological features of V_v-V_m space will facilitate an understanding of the model's response to impulsive external stimuli in Sec. IV.

D. Ghosts

Following the annihilation of the wake and sleep nodes due to saddle-node bifurcations, regions of near-stability are retained, in the form of low $\dot{V}$. This is a general property of the saddle-node bifurcation—the low velocity region is termed a saddle-node remnant, or *ghost* [39]. The *sleep* and *wake ghosts* of our model are illustrated in Figs. 6(b) and 6(d), respectively, and both ghosts are labeled in the hysteresis picture of Fig. 6(a). Because of the low $\dot{V}$ region, trajectories evolve slowly, through a bottleneck in the vicinity of the ghosts [39]. In our model, this behavior can be interpreted as brief awakenings during sleep (lingering in the wake ghost), or brief naps during wake (lingering in the sleep ghost). The time taken to pass through a bottleneck, and hence a measure of the duration of a brief awakening or nap, obeys a square-root scaling law: $|D_v - D_v^*|^{-1/2}$, where D_v^* is a bifurcation value [39]. Thus increasing proximity to a ghost increases the time taken to pass through it. This idea will be elaborated upon in Sec. IV B and the effect of the ghosts on the model dynamics in response to impulsive stimuli will be analyzed in Sec. V.

IV. IMPULSIVE EXTERNAL STIMULI

Having characterized the dynamics of the model in V_v-V_m space, we now explore its response to external stimuli, which are represented as time-dependent vectors $\Delta\mathbf{D}(t) = [\Delta D_v(t), \Delta D_m(t)]$. Pharmacological agents and sensory stimuli alike can be modeled by the relative effects of their mechanistic pathway on the VLPO and MA groups, with stimulants presumably characterized by $\Delta D_m \geqslant 0$ and $\Delta D_v \leqslant 0$, and sedatives by $\Delta D_m \leqslant 0$ and $\Delta D_v \geqslant 0$. The resulting perturbation on the sleep-wake dynamics can be deduced by solving the model equations. This method for treating external stimuli applies generally, but in this paper we focus on impulsive stimuli acting over short time scales. What is meant by an impulsive stimulus is explained in Sec. IV A, and the arousal state responses to such stimuli are explored in Sec. IV B. Then, in Sec. V, excitatory sensory stimuli applied during sleep are modeled.

A. Impact of stimulus duration

For a drive acting on a time scale $\ll$24 h, we can neglect variation in D_v, which is treated as a constant over the duration of the impulse. Such *impulsive* drives can be analyzed as trajectories through the V_v-V_m plane at a constant D_v. Further, for drives acting on a time scale $\ll \tau_m$, $\tau_v = 10$ s, trajectories through the V_v-V_m plane will be minimally affected by the velocity vector field $\dot{\mathbf{V}}$ during the impulse. In this high-amplitude, short-duration (δ-function) limit, the impact of a perturbative drive can be determined analytically. For a drive ΔD_m directed at the MA group, for example, we use Eq. (3):

$$\tau_m \dot{V}_m = -V_m + \nu_{mv} Q_v + D_m^0 + \Delta D_m. \tag{27}$$

Assuming that $|\Delta D_m| \gg |-V_m + \nu_{mv} Q_v + D_m^0|$, the approximate form $\tau_m \dot{V}_m \approx \Delta D_m$ is obtained. Integrating over the duration of the impulse, from t_1 to t_2, yields

$$\Delta V_m = V_m(t_2) - V_m(t_1) \approx \tau_m^{-1} \int_{t_1}^{t_2} \Delta D_m dt. \tag{28}$$

Similarly, a δ-function drive ΔD_v to the VLPO will produce a change in V_v of

$$\Delta V_v = V_v(t_2) - V_v(t_1) \approx \tau_v^{-1} \int_{t_1}^{t_2} \Delta D_v dt. \tag{29}$$

In combination as $\Delta\mathbf{D}$, the mean cell-body potentials V_m and V_v will change by the amounts given by Eqs. (28) and (29), respectively.

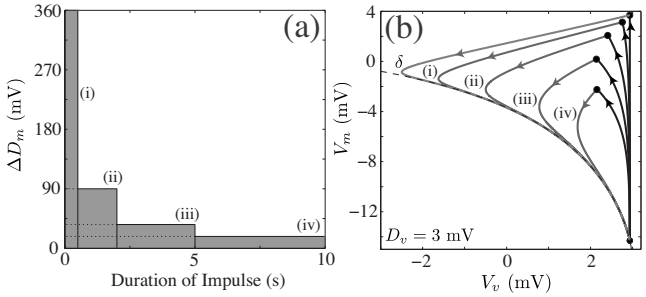

FIG. 7. The model's response discrete ΔD_m drives of different durations. (a) ΔD_m profiles for drives of duration (i) 0.5 s, (ii) 2 s, (iii) 5 s, and (iv) 10 s, with $\int \Delta D_m dt = 180$ mV s, a constant. (b) Corresponding trajectories through the V_v-V_m plane during an impulse (black) and during relaxation (gray), at $D_v = 3$ mV. As the impulse duration increases, the trajectories are increasingly distorted away from the δ-function limit (a vertical line).

In Fig. 7, ΔD_m impulses of varying duration are applied at $D_v = 3$ mV for a constant value of the integral $\int \Delta D_m dt = 180$ mV s. In the δ-function limit, a change $\Delta V_m = \tau_m^{-1} \int \Delta D_m dt = 18$ mV is produced. As illustrated in the figure, as the duration of the impulse increases, so does the trajectory's deviation from the δ-function limit due to the influence of $\dot{\mathbf{V}}$ [Eq. (13)], which includes the terms neglected in the limit. For the remainder of this paper we focus on δ-function impulses, which can be represented precisely in terms of the ΔV_v and ΔV_m they produce, using Eqs. (28) and (29). Realistic impulsive stimuli of nonzero duration are perturbed from this limit, but can be compared similarly, through a numerical mapping to the corresponding ΔV_v and ΔV_m they produce. We therefore assert that the trends obtained using δ-function stimuli are qualitatively comparable to those that would be obtained using impulsive stimuli of nonzero duration. Note that in the following, when $\Delta\mathbf{D} \neq 0$, we are careful to distinguish unperturbed drives D_v^0 and D_m^0 from the total drives D_v and D_m where appropriate; although this distinction is not essential when dealing with instantaneous δ-function stimuli.

B. Perturbative drives ΔD

In the presence of external stimuli, there will be variation in both ΔD_m, due to external sensory stimuli, for example (cf., Sec. V), and ΔD_v, due to a noisy circadian signal, for example. Whatever the cause of these fluctuations, we model a single impulsive stimulus using a δ-function drive $\Delta\mathbf{D}$, as described above. In V_v-V_m space, impulsive changes to an initial condition at a stable equilibrium (V_v^*, V_m^*) are given by Eqs. (28) and (29), perturbing it to $(V_v^*+\Delta V_v, V_m^*+\Delta V_m)$. We define the *sleep latency* $t_{\rm lat}$ as the time taken for the trajectory to return to a stable wake or sleep node. The threshold chosen to define a return to equilibrium is $\dot{V}<5\times10^{-3}$ mV s^{-1}. The problem of setting this threshold is akin to the clinical conventions for defining a sleep state [2], or an "arousal" [42] from what is essentially a continuous process. Measures of arousal are not always consistent across studies [43] and, as is done here, a reasonable value is chosen and applied consistently. Small changes in the $\dot{V}$ threshold would alter the quantitative $t_{\rm lat}$ values but not the qualitative trends of our analysis. In the following, we investigate the model's response to external stimuli at $D_v=1$ mV (wake), $D_v=2$ mV (bistable), and $D_v=3$ mV (sleep) as a representative sample of the model regimes, as shown in Fig. 8. We use the final steady state and the latency $t_{\rm lat}$ to characterize the impact of a stimulus and interpret the results in terms of the model dynamics.

In accordance with the linear stability analysis of Sec. III B, at the cessation of an impulse, trajectories through the V_v-V_m plane attract rapidly onto the W^+ manifold (shown dashed in Fig. 8), on which slower dynamics occur. Since $t_{\rm lat}$ is therefore dominated by movement along the W^+ manifold, patterns of almost equivalent $t_{\rm lat}$ can be seen along the fast eigendirections from the shading in Fig. 8. Impulses landing near the W^- of a stable node will rapidly attract onto that equilibrium, and are seen as black regions ($t_{\rm lat}\approx0$). On the other hand, a perturbation whose return trajectory passes through a low $\dot{V}$ region, like those surrounding a saddle point or a ghost, will take much longer to return to equilibrium and are seen as light regions ($t_{\rm lat}\gtrsim6$ min). Example trajectories through the V_v-V_m plane are shown in Fig. 8 to illustrate the dynamics.

The qualitative topological landscape of V_v-V_m space is similar at low and high D_v, where a single stable node equilibrium and a ghost exist, as shown in Figs. 8(a) and 8(b) for $D_v^0=1$ mV, and Figs. 8(g) and 8(h) for $D_v^0=3$ mV. Since only one stable equilibrium exists, no transient stimulus can produce a lasting change in state. The latency $t_{\rm lat}$ to return back to the initial equilibrium state is minimal for drives that perturb the initial condition to near the W^- manifold of the stable node, and maximal for trajectories passing through the near-stable ghost. In the presence of external perturbations $\Delta\mathbf{D}$, brief naps during wake bouts (by lingering in the sleep ghost) and brief awakenings during sleep bouts (by lingering in the wake ghost) can occur. Since the wake ghost "extends further" from the relevant saddle-node bifurcation than the sleep ghost, as shown in Fig. 6(a), brief awakenings during sleep are favored over brief naps during wake. This finding is qualitatively consistent with the occurrence of brief sleep-wake transitions throughout the sleep periods of mammalian species [44]. The presence of arousals during sleep are assumed to maintain the sleeper's association with their surrounding environment, reacting to possible indicators of danger [43]. The role of the wake ghost in stabilizing the waking state during sleep is discussed further throughout Sec. V, in which the slowing of trajectories near it, as brief awakenings, is demonstrated explicitly. Persistent stimuli $\Delta\mathbf{D}$ could act to

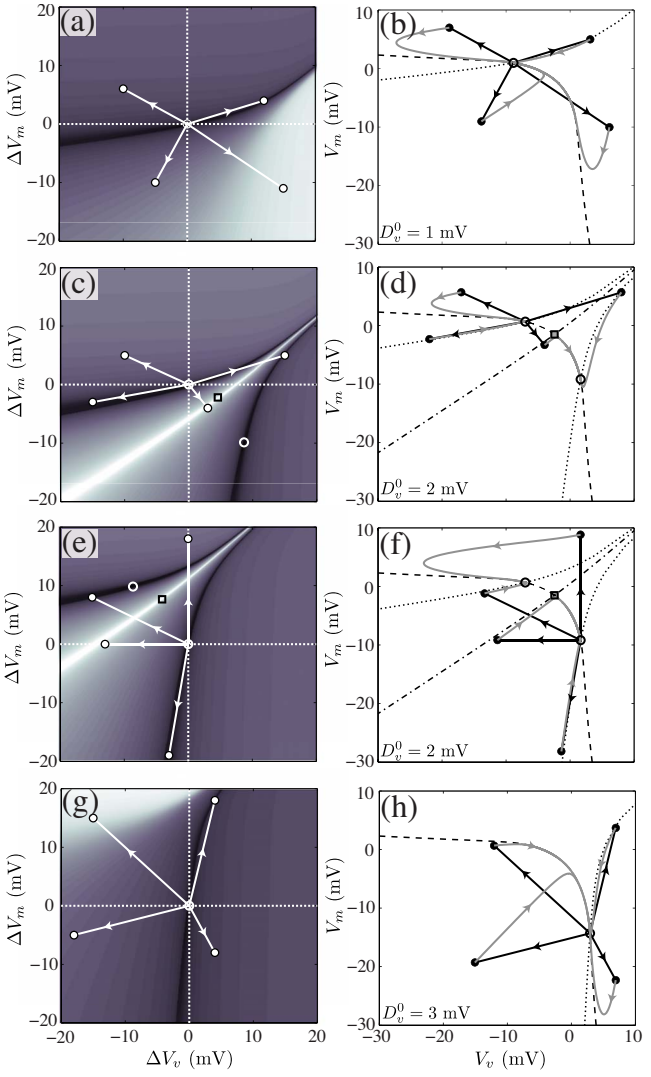

FIG. 8. (Color online) Model responses to δ-function $\Delta\mathbf{D}$ stimuli. ΔV_v-ΔV_m and V_v-V_m plots are shown for (a) and (b): $D_v^0=1$ mV (initial condition: wake node), (c) and (d): $D_v^0=2$ mV (initial condition: wake node), (e) and (f): $D_v^0=2$ mV (initial condition: sleep node), and (g) and (h): $D_v^0=3$ mV (initial condition: sleep node). In the ΔV_v-ΔV_m plots (defined relative to the initial equilibrium), the sleep latency $t_{\rm lat}$ is represented using shading, from 0 min (black) to >6 min (white). Sample δ-function impulses are shown as white arrows, with the corresponding trajectories through V_v-V_m space shown in the adjacent plots for the impulse (black) and during relaxation back to a stable node (gray). Equilibria (open white circles and black squares) and invariant manifolds (dotted, dashed, and dot-dashed lines) are labeled as per Fig. 6. Note that points in the ΔV_v-ΔV_m plane should be interpreted in terms of the δ-function drives $(\Delta D_v, \Delta D_m)$ that produce them (cf., Sec. IV A).

further stabilize the ghosts, producing longer episodes in them: as prolonged naps during normal waking hours, or extended periods of waking (sleep deprivation) during normal sleeping hours. For now, such effects are simply described qualitatively, as characteristic features of the model dynamics, and will be explored in more detail in future work.

At $D_v^0=2$ mV, in the bistable region, we investigate initial conditions at the wake node, shown in Figs. 8(c) and 8(d), and at the sleep node, shown in Figs. 8(e) and 8(f). A transient stimulus is able to elicit a change of state between sleep and wake if it perturbs the trajectory across the separatrix formed by the stable invariant manifold W^- of the saddle point (visible in the ΔV_m-ΔV_v and V_v-V_m plots of Fig. 8 as a white region and a dot-dashed line, respectively). If the return trajectory passes near the saddle point, it will take much longer to reach the relevant stable node because of the region of low $\dot{V}$ that surrounds it [cf., Fig. 6(c)]. On the other hand, a perturbation to near the fast manifold W^- of either stable node will result in rapid attraction onto it. Therefore to achieve a rapid change in state, an optimum set of drives $\Delta\mathbf{D}$ exist, corresponding to a perturbation onto the W^- of the alternate stable node. In summary, the behavior of the model can be explained in terms of regions of near-stability: ghosts and the low $\dot{V}$ region surrounding the saddle point, which lengthen t_{lat}, and the fast dynamics along the stable invariant manifolds W^- of the stable nodes, which shorten t_{lat}.

V. AROUSING STIMULI DURING SLEEP

Having characterized the dynamics of the model in response to general impulsive stimuli, we now consider the specific case of external sensory stimuli applied during sleep. This scenario corresponds to clinical sleep fragmentation protocols that use auditory stimuli to fragment sleep [45]. Auditory tones, like other external sensory stimuli, excite MA nuclei [36,46,47], perhaps via the orexin group [48,49] or from ACh [50,51]. Assuming no *direct* impact on the VLPO (i.e., only indirectly via increased inhibition from MA firing), we model a short, loud auditory tone by an excitatory δ-function drive ΔD_m, leaving $\Delta D_v=0$. In the bistable region ($1.45\ \text{mV}<D_v<2.46\ \text{mV}$), the arousing stimulus can cause a transition to the stable wake node, producing a permanent arousal. However, here we consider the range $D_v>2.46$ mV for which the stable sleep node and the wake ghost are the main dynamical features of the model; i.e., for the topology shown in Figs. 8(g) and 8(h). The impulse strength and time-of-night dependences of the model's response to such sensory stimuli are explored in Secs. V A and V B, respectively. This analysis is then used to simulate a sleep fragmentation study in Sec. V C.

A. Impulse strength

Figure 9(a) shows the monotonic increase of the time to return to sleep t_{lat} with stimulus strength ΔV_m for $D_v=3$ mV. The plot features two regions of steep gradient. The first, at low ΔV_m, is due to the region of low $\dot{V}$ surrounding the sleep node. As illustrated in Fig. 9(b), the rapid rise in t_{lat} at a higher stimulus strength corresponds to the return trajectory passing through the low $\dot{V}$ wake ghost. Increasing the intensity of the applied impulse past this point has relatively little effect on t_{lat} because increases in $\dot{V}$ beyond the wake ghost largely counteract the additional stimulus. Hence the time spent in the wake ghost dominates t_{lat}, which saturates at large ΔV_m.

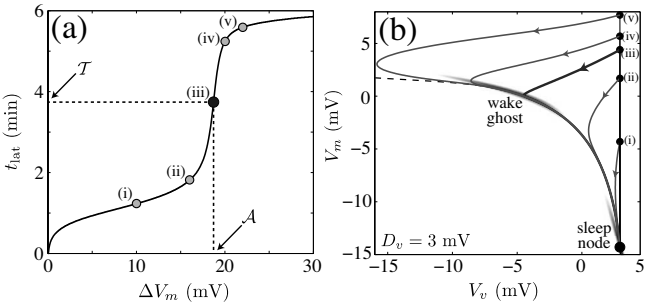

FIG. 9. The model's response to excitatory δ-function sensory stimuli ΔD_m of varying magnitude applied during sleep ($D_v=3$ mV). (a) The time to return to sleep t_{lat} following the impulse is plotted as a function of its strength ΔV_m. For the five points marked in (a): $\Delta V_m=$(i) 10 mV, (ii) 16 mV, (iii) 18.7 mV (the point of inflection), (iv) 20 mV, and (v) 22 mV, corresponding trajectories through the V_v-V_m plane are shown in (b), where regions of $\dot{V}<0.1\ \text{mV s}^{-1}$ have been shaded. The point of inflection is shown as a larger, closed circle in (a), with the arousal threshold $\mathcal{A}$ taken from the ΔV_m at this point and the critical sleep latency $\mathcal{T}$ taken from the t_{lat} at this point, as labeled.

In this model, the point of inflection of the ΔV_m-t_{lat} curve is taken to define the sleep-wake transition. The ΔV_m at this point defines the *arousal threshold* $\mathcal{A}$, which represents the mean increase in MA cell-body potential required to trigger an arousal from sleep. This is a suitable measure because stronger stimuli $\Delta V_m>\mathcal{A}$ cause a prolonged arousal and weaker stimuli $\Delta V_m<\mathcal{A}$ produce a rapid return to sleep, as illustrated in Fig. 9. Clinically, the arousal threshold is measured in terms of the magnitude of a given sensory stimulus, in units of auditory intensity or pressure deviation, for example. Our measure of the arousal threshold, $\mathcal{A}$, should therefore be comparable to clinical measures, as explored in Sec. V B below. We define the *critical sleep latency* $\mathcal{T}$ as the latency to return to sleep t_{lat} after receiving the critical impulse $\mathcal{A}$, as labeled in Fig. 9(a).

B. Time of night

Rather than explicitly investigating the impact of external stimuli as a function of the time since sleep onset, it is more convenient to study them as a function of the sleep drive D_v. Trends in D_v can then be mapped to trends in time using the variation in $D_v(t)$ across a normal night of sleep, shown in Fig. 2(e). Since $Q_m\approx 0\ \text{s}^{-1}$ during sleep, Eq. (8) becomes $\chi\dot{H}\approx -H$, and H approximates exponential decay. Therefore, during sleep, the drive $D_v(t)$ varies as the weighted sum of sinusoidal $C(t)$ and exponential decay $H(t)$ components, peaking near the middle of the night.

The model's response to $\Delta V_m=18$ mV stimuli as a function of D_v is shown in Fig. 10. The observed behavior can be

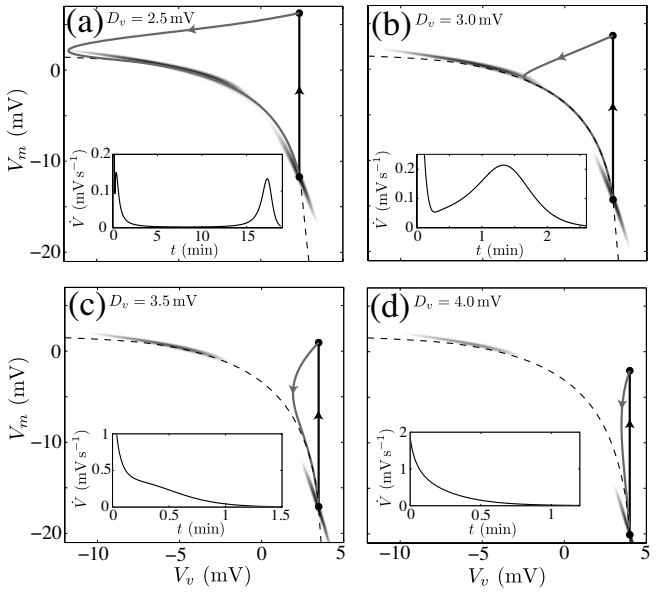

FIG. 10. The model's response to $\Delta V_m = 18$ mV δ-function impulses at different values of the sleep drive D_v. Trajectories through the V_v-V_m plane are plotted during excitation (black) and relaxation (gray) for D_v=(a) 2.5 mV, (b) 3.0 mV, (c) 3.5 mV, and (d) 4.0 mV. Regions of $\dot{V} < 0.1$ mV s^{-1} are shaded and plots of $\dot{V}$ as a function of time along each trajectory are shown inset. Note that the vertical scale of the inset figures is chosen to show the main features of the return to equilibrium, which does not always include the initial maximum in $\dot{V}$.

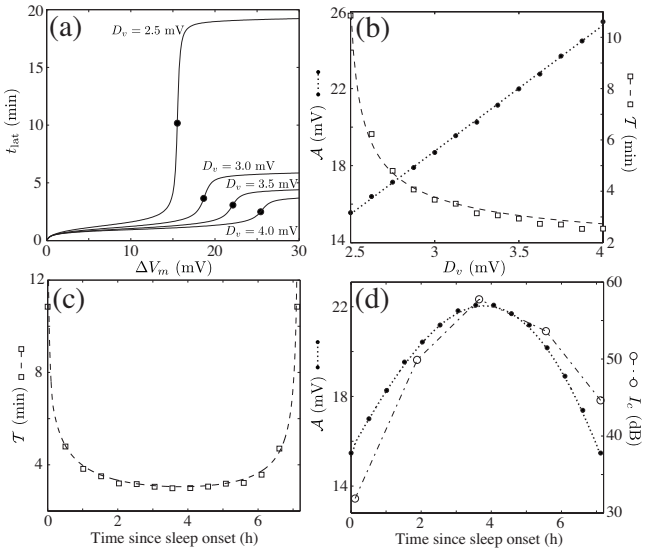

FIG. 11. Time-of-night variation in the arousal threshold $\mathcal{A}$ and critical sleep latency $\mathcal{T}$. (a) ΔV_m-$t_{\rm lat}$ curves are plotted for D_v=2.5, 3.0, 3.5, and 4.0 mV. Points of inflection are marked with circles. (b) $\mathcal{A}$ as a function of D_v (circles) with a fitted dotted line [Eq. (30)] and $\mathcal{T}$ as a function of D_v (squares) with a fitted dashed line [Eq. (31)]. (c) $\mathcal{T}$ and (d) $\mathcal{A}$ across the first ~7 h of sleep (the portion for which $D_v > 2.46$ mV), with the fits from (b) shown dashed and dotted, respectively. Clinical auditory arousal threshold I_c data from Bonnet *et al.* [53] is displayed using open circles and a dot-dashed line in (d).

explained in terms of both the minimum $\dot{V}$ in the wake ghost, which decreases with increasing D_v, and the position of the sleep node, which moves to a higher V_v and lower V_m with increasing D_v. At D_v=2.5 mV, near the saddle-node bifurcation of the wake node, the return trajectory lingers in the wake ghost for ~10 min before returning to sleep, as shown in Fig. 10(a). At D_v=3 mV, as shown in Fig. 10(b), the return trajectory passes through a lower $\dot{V}$ region of the wake ghost region, lingering there for only ~1 min. As D_v increases further, the ΔV_m=18 mV stimuli are no longer strong enough to perturb the system to near the wake ghost. Instead, the trajectories return ever more rapidly to sleep under the influence of the vector field $\dot{\mathbf{V}}$, as demonstrated in Figs. 10(c) and 10(d). Thus we find that at low D_v, an impulse has a much greater impact, producing a greater maximum V_m, and taking a longer time to return to sleep than the same impulse applied at high D_v. It follows that as D_v increases, stronger impulses are required to excite the system from the sleep node to the wake ghost—the arousal threshold increases with D_v.

The ΔV_m-$t_{\rm lat}$ curves discussed in Sec. V A also depend on D_v, as illustrated in Fig. 11(a) using the same D_v values as in Fig. 10. Consistent with the trends explained above, the ΔV_m-$t_{\rm lat}$ curves shift to greater $\mathcal{A}$ (higher arousal threshold) and lower $\mathcal{T}$ (the minimum $\dot{V}$ in the wake ghost is lower) as D_v increases. A plot showing the variation of both $\mathcal{A}$ and $\mathcal{T}$ with D_v is shown in Fig. 11(b). We find an approximately linear relationship between $\mathcal{A}$ and D_v [because the sleep branch and the minimum $\dot{V}$ in the wake ghost are approximately linear for $D_v > 2.46$ mV, as can be seen in Fig. 6(a)]:

$$\mathcal{A} \approx 6.5 D_v - 0.9 \text{ mV}. \tag{30}$$

Since the time spent in a ghost decreases as the inverse square root of the difference between the parameter and its bifurcation value [39], the critical sleep latency $\mathcal{T}$ is approximated by

$$\mathcal{T} \approx \left[111 \left(\frac{D_v}{1 \text{ mV}} - 2.46 \right)^{-1/2} + 74.0 \right] \text{ min} \tag{31}$$

for $D_v > 2.46$ mV, which diverges as D_v approaches the bifurcation value D_v=2.46 mV. Using the known variation of $D_v(t)$ across a normal sleep night, $\mathcal{A}(D_v)$ and $\mathcal{T}(D_v)$ are transformed to functions of the time since sleep onset, as shown in Figs. 11(c) and 11(d). The trend in $\mathcal{T}$ could be measured experimentally using sleep latency tests across the night, but to our knowledge this has not yet been performed. Because the $t_{\rm lat}$ and hence $\mathcal{T}$ values remain uncalibrated (since this depends on the choice of the threshold $\dot{V} < 5 \times 10^{-3}$ mV s^{-1} at which a trajectory is considered "close" to equilibrium), we seek to validate only the qualitative trends through experiment, which may help to calibrate this threshold.

The model predicts that the arousal threshold will peak near the middle of the night, as shown in Fig. 11(d). This is qualitatively consistent with some clinical studies [52–56] but is inconsistent with others that find a steady increase of arousal threshold across the night [57–59], as measured from the critical intensity I_c (dB) of the auditory stimulus required

to cause an arousal. However, a monotonically increasing arousal threshold is likely the result of progressive habituation to the auditory stimulus, which makes it less effective with increased exposure, and which we have not attempted to model here. A study reported by Bonnet *et al.* [53] aroused subjects only five to eight times each night, minimizing the effects of the frequent nocturnal disruptions. It is therefore a suitable study with which to compare the predictions of our model, which [by using $D_v(t)$ for normal sleep] assumes that sleep is minimally disturbed. The experimental I_c data adapted from the Bonnet *et al.* study is shown alongside our measure $\mathcal{A}$ in Fig. 11(d), showing good agreement. The systematic increase of the I_c data relative to $\mathcal{A}$ with time is likely the result of a small habituation effect. A linear fit between the two measures yields

$$I_c \approx \left[2.9 \frac{\mathcal{A}}{1\ \mathrm{mV}} - 7.1 \right] \mathrm{dB}, \tag{32}$$

giving a relationship between the intensity of an auditory stimulus I_c and the increase in MA cell-body potential $\mathcal{A}$ required to trigger an awakening from sleep. Combined with Eq. (30), $I_c(t)$ can then be inferred from $D_v(t)$ across *any* sleep period.

C. Modeling sleep fragmentation

The importance of the above calibration between I_c and $\mathcal{A}$ is demonstrated by simulating a sleep fragmentation protocol reported by Lammers *et al.* [55]. In this study, nine male subjects were required to make a microswitch closure in response to a series of 2 s long, 1000 Hz auditory tones applied 4 min after every entry into stage 2 sleep. Choosing a threshold $\dot{V} < 0.1$ mV s^{-1} to define a return to stage 2 sleep, we simulate this protocol by repeatedly applying discrete 2 s ΔD_m impulses 4 min after this threshold is crossed. The strength of the impulses is chosen to produce a ΔV_m equal to the arousal threshold $\mathcal{A}$, which is calculated from D_v using Eq. (30). However, the strength $\Delta V_m = \mathcal{A}$ given by Eq. (28) applies to δ-function impulses and needs to be increased to obtain the same ΔV_m for 2 s impulses [cf., Sec. IV A]; a magnitude $\Delta D_m = 1.1(\mathcal{A}/2)$ achieves this. The resulting fragmented time series for V_m and H are shown in Figs. 12(a) and 12(b), respectively. The V_m time series is characterized by recurring increases to waking levels, in contrast to unperturbed sleep in which V_m, and hence the firing rate Q_m, is low throughout the night, producing consolidated sleep. The increase in the average Q_m across the night feeds back onto the sleep cycle via the homeostatic source term μQ_m [cf., Eq. (8)]—adenosine production is increased and sleep is less restorative. The progressive increase in H over normal levels is shown in Fig. 12(b). Compared to normal sleep, an increased homeostatic component to the sleep drive $D_v = D_v^0 = \nu_{vc} C + \nu_{vh} H$ causes it to peak later in the night and at a higher value. Since $\mathcal{A}$ depends approximately linearly on D_v [Eq. (30)] and I_c depends approximately linearly on $\mathcal{A}$ [Eq. (32)], the arousal threshold I_c will also peak later in the night and at a higher value, as shown in Fig. 13. In this figure, the data from the Lammers *et al.* study [55] is shown alongside our model predictions, revealing a good agreement. However, in order to quantitatively reproduce the auditory arousal threshold I_c values, an additional offset of +9 dB relative to the previous calibration [Eq. (32)] was required. As well as inter-individual differences in experimental measures of the auditory arousal threshold, due to age and other effects, a large inter-study variation (e.g., from 36 to 91 dB in stage 4 sleep) is also evident, perhaps due to differing amounts of background noise in sleep laboratories or different clinical definitions of "arousal" [60]. In this light, the shift of only 9 dB used here seems reasonable.

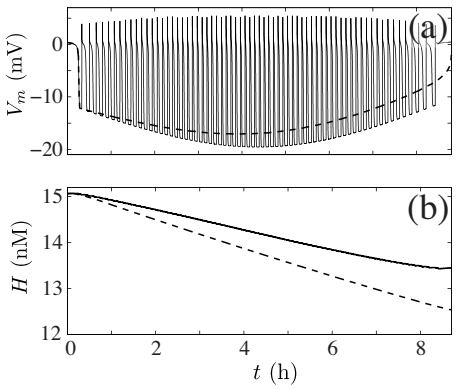

FIG. 12. The model's simulation of the sleep fragmentation protocol reported by Lammers *et al.* [55]. (a) V_m and (b) H as a function of time for fragmented (solid) and normal (dashed) sleep.

In addition to the arousal threshold, the Lammers *et al.* study simultaneously recorded body temperature, which is known to exhibit a circadian variation [61]. The authors proposed a negative relationship between body temperature and arousal threshold, implying that $-D_v$ in our model should reflect the body temperature variation. Indeed it does, allow-

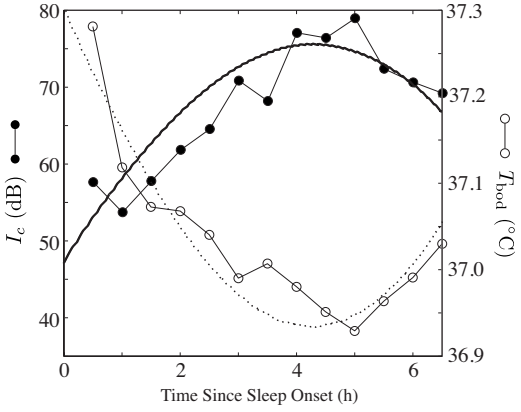

FIG. 13. The model's simulation of the sleep fragmentation protocol reported by Lammers *et al.* [55]. The auditory arousal threshold I_c (closed circles) and body temperature T_{bod} (open circles) data are plotted alongside the model's predictions for I_c [as inferred from $D_v(t)$ using Eqs. (30) and (32)] (solid line) and T_{bod} [as inferred from $D_v(t)$ using Eq. (33)] (dotted line). A 9 dB offset relative to the calibration to the Bonnet *et al.* data [53] is added to the model's I_c predictions. The model formulation emulates the clinical protocol: 2 s impulses of strength equal to the arousal threshold $\mathcal{A}$ are applied 4 min after every entry into Stage 2 sleep (defined as the $\dot{V} < 0.1$ mV s^{-1} region surrounding the sleep node).

ing us to deduce a linear fit between D_v and body temperature T_{bod}:

$$T_{\mathrm{bod}} \approx \left[-0.241 \frac{D_v}{1\ \mathrm{mV}} + 37.9 \right] ^\circ \mathrm{C}. \quad (33)$$

The predicted T_{bod} curve, as deduced from $D_v(t)$ using Eq. (33), is plotted as a dotted line in Fig. 13.

The model's ability to simultaneously reproduce the arousal threshold and body temperature curves during sleep fragmentation is a consequence of its physiological formulation, which allows us to exploit the known impact of sensory stimuli on MA nuclei to model sleep fragmentation. There are two free parameters: the $\dot{V}$ threshold to stage 2 sleep, which has a sensible value (near the sleep node), and the offset of 9 dB, which does not affect the qualitative trends and is modest compared to inter-study variability [60]. The success of this approach allows the arousal threshold and body temperature curves to be interpreted in terms of the physiological processes that produce them: the homeostatic and circadian components of the drive D_v to the VLPO. Frequent arousals reduce the clearance rate of H, diminishing the restorative value of sleep and causing D_v and hence I_c and T_{bod} to reach an extremum of greater magnitude later in the night.

VI. SUMMARY AND DISCUSSION

Within a quantitative, physiologically based model of the sleep-wake switch, we have used a separation of time scales and linear stability analysis to gain detailed insights into the impact of impulsive stimuli on arousal state. Stimuli acting on the system are represented as $\Delta\mathbf{D}(t)$ vectors, allowing a unified representation of diverse external influences. Through an analysis of the model dynamics, brief naps during wake and brief awakenings during sleep were shown to be a consequence of the model's sleep and wake ghosts. Brief awakenings during sleep are favored over brief naps during wake, a finding qualitatively consistent with experimental observations [44]. The set of drives that produce a change in state between sleep and wake perturb the system across the separatrix in the V_v-V_m plane, with fast and slow transitions characterized by trajectories following the invariant manifold W^- or the low $\dot{V}$ region surrounding the saddle point, respectively. In Sec. V, auditory stimuli applied during sleep were modeled according to their known excitation of MA brainstem populations: by drives $\Delta D_m > 0$. In the δ-function limit, these impulses produce impulsive changes ΔV_m to V_m. The sharp rise in the ΔV_m-t_{lat} curve, due to a lag associated with the wake ghost, allowed us to define the arousal threshold $\Delta V_m = \mathcal{A}$, which was found to depend approximately linearly on D_v. It therefore varies as a linear combination of sinusoidal C and approximately exponential H components, peaking near the middle of the night. This trend in $\mathcal{A}$ matches the clinical variation of the auditory arousal threshold I_c, against which it was subsequently calibrated. As a result, the time-of-night I_c variation for any arbitrary sleep-wake protocol can be inferred from D_v, as they are approximately linearly related. The significance of this finding was demonstrated by simulating a clinical sleep fragmentation protocol reported by Lammers *et al.* [55]. The observed time-of-night arousal threshold and body temperature curves were quantitatively reproduced, including the skew of the peak of the arousal threshold towards the latter half of the night and its increase in magnitude relative to normal sleep. We propose that an increase in the homeostatic sleep drive due to the frequent awakenings is the cause of this change.

A core idea that stems from the general formulation of stimuli as vectors $\Delta\mathbf{D}(t)$ is that diverse stimuli can be compared in the same space. For example, in addition to the more prevalent studies of auditory arousal thresholds considered in this work, arousal thresholds have also been measured for pain, pressure, temperature, light intensity, and olfactory stimuli [60]. In particular, the respiratory effort (measured in units of pressure) in obstructive sleep apnea-hypopnea syndrome (OSAHS) patients provides another well-studied measure of the arousal threshold, which presumably acts via ΔD_m [62,63]. Studies measuring the time-of-night arousal threshold of OSAHS patients, like that of Sforza *et al.* [64], show a qualitatively similar trend to our $\mathcal{A}$, but more detailed studies like that of Berry *et al.* [65] show an oscillation on a time scale of $\sim$128 min. This fine oscillatory variation is perhaps the result of ultradian dynamics which, if included in this model, would effectively constitute an oscillatory D_m^0 [11], modulating the arousal threshold accordingly. Within the Phillips-Robinson framework, a simple OSAHS model may consist of a monotonically increasing drive ΔD_m, triggered by a certain vicinity to the sleep node, and increasing until the arousal threshold $\mathcal{A}$, where the obstruction is resolved. Since the apneic drive ΔD_m acts over an extended time scale compared to the impulsive stimuli of the present work, the resulting dynamics will be different. Extending the current framework towards a predictive physiologically based model of OSAHS is a target for future work.

Although this paper focused on impulsive stimuli, which constitute merely transient disturbances to the sleep-wake cycle, drives acting over extended time scales can also be represented as $\Delta\mathbf{D}$. These persistent drives perturb the position of the equilibrium points themselves, altering the broader arousal state dynamics. Slowly varying $\Delta\mathbf{D}$ drives could be used to describe gradual changes in temperature, the impact of pharmaceuticals with extended biological half-lives, or neurological pathologies such as narcolepsy, for example. A persistent drive could also be used to model enforced waking protocols by maintaining the system in the wake ghost across a normal sleep period. The external drive required to keep the subject awake is small (owing to the small $\dot{V}$ in the wake ghost) and could be provided in the form of an arousing "effort," as a positive ΔD_m, for example, that may be cortical and/or orexinergic in origin [66–68]. The effort ΔD_m increases with D_v, as the wake ghost moves further from stability; consistent with the increasing difficulty in remaining awake at a high physiological sleep pressure. A simulation of subjective fatigue during sleep deprivation is one application of this approach and will be explored in detail in future work. The properties of the ghost states suggest

a qualitative reason why brief arousals during sleep are more prevalent than brief naps during wake (cf., Sec. IV B). A future task involves extending this approach towards a quantitative model that reproduces the observed scale-invariant trends [44], which might be achieved by adding noisy perturbations $\Delta\mathbf{D}$ to the system.

In summary, we have developed a framework for incorporating general external stimuli into a physiologically based model of the sleep-wake switch. Despite being a comparatively simple model, including only two neuronal populations and averaging over the ultradian rhythm, the orexin group, and cortical-level effects, we are able to gain insights into the basic neural mechanisms affecting arousal state responses to impulsive stimuli. This study represents an exploration of the impact of external stimuli on sleep-wake dynamics and takes an approach that is fundamentally different to previous attempts. Where past models framed sets of data into a phenomenological structure, our model is of the known physiology of the sleep-wake switch. Stimuli are modeled by their impact on the relevant neuronal populations, and behavioral outcomes are subsequently deduced. The model's predictions are directly testable through comparisons to both physiological and clinical data, as a way of justifying the methodology, or providing evidence against it. The agreement between the model's predictions and the available data, discussed in the latter sections of this paper, justifies our approach. Applying the above ideas to a broader range of stimuli, including pharmacological agents and other sensory stimuli is expected to guide developments in arousal state modulation in the future.

ACKNOWLEDGMENT

The Australian Research Council supported this work.

[1] A. A. Borbély, Hum. Neurobiol. **1**, 195 (1982).
[2] A. A. Borbély, F. Baumann, D. Brandeis, I. Strauch, and D. Lehmann, Electroencephalogr. Clin. Neurophysiol. **51**, 483 (1981).
[3] S. Daan, D. G. M. Beersma, and A. A. Borbély, Am. J. Physiol. Regulatory Integrative Comp. Physiol. **246**, 161 (1984).
[4] P. Achermann, D. Dijk, D. P. Brunner, and A. A. Borbély, Brain Res. Bull. **31**, 97 (1993).
[5] M. Nakao, A. Karashima, and N. Katayama, Cell. Mol. Life Sci. **64**, 1236 (2007).
[6] A. A. Borbély and P. Achermann, J. Biol. Rhythms **14**, 557 (1999).
[7] A. Gundel, K. Marsalek, and C. Thoren, Somnologie **11**, 148 (2007).
[8] H. P. A. Van Dongen, Aviat., Space Environ. Med. **75**, A15 (2004).
[9] T. Gallopin, P. Fort, E. Eggermann, B. Cauli, P. Luppi, J. Rossier, E. Audinat, M. Mühlethaler, and M. Serafin, Nature (London) **404**, 992 (2000).
[10] C. B. Saper, T. C. Chou, and T. E. Scammell, Trends Neurosci. **24**, 726 (2001).
[11] E. F. Pace-Schott and J. A. Hobson, Nat. Rev. Neurosci. **3**, 591 (2002).
[12] R. E. Strecker, S. Morairty, M. M. Thakkar, T. Porkka-Heiskanen, R. Basheer, L. J. Dauphin, D. G. Rainnie, C. M. Portas, R. W. Greene, and R. W. McCarley, Behav. Brain Res. **115**, 183 (2000).
[13] T. C. Chou, Ph.D. thesis, Harvard University, 2003.
[14] Y. Tamakawa, A. Karashima, Y. Koyama, N. Katayama, and M. Nakao, J. Neurophysiol. **95**, 2055 (2006).
[15] C. G. Diniz Behn, E. N. Brown, T. E. Scammell, and N. J. Kopell, J. Neurophysiol. **97**, 3828 (2007).
[16] A. J. K. Phillips and P. A. Robinson, J. Biol. Rhythms **22**, 167 (2007).
[17] A. J. K. Phillips and P. A. Robinson (unpublished).
[18] E. R. Kandel, J. H. Schwartz, and T. M. Jessell, *Principles of Neural Science*, 4th ed. (McGraw-Hill, New York, 1991).
[19] B. Jones, in *Principles and Practice of Sleep Medicine*, 3rd ed., edited by M. Kyger, T. Roth, and W. Dement (WB Saunders, Philadelphia, 2000), Chap. 10, pp. 134-154.
[20] G. Aston-Jones, C. Chiang, and T. Alexinsky, Prog. Brain Res. **88**, 501 (1991).
[21] R. Webster, *Neurotransmitters, Drugs and Brain Function* (Wiley, Chichester, 2002).
[22] A. I. Levey, A. E. Hallanger, and B. H. Wainer, Neurosci. Lett. **74**, 7 (1987).
[23] M. M. Thakkar, R. E. Strecker, and R. W. McCarley, J. Neurosci. **18**, 5490 (1998).
[24] T. Sakurai, Nat. Rev. Neurosci. **8**, 171 (2007).
[25] K. Ohno and T. Sakurai, Front Neuroendocrinol **29**, 70 (2008).
[26] J. E. Sherin, P. J. Shiromani, R. W. McCarley, and C. B. Saper, Science **271**, 216 (1996).
[27] C. B. Saper, T. E. Scammell, and J. Lu, Nature (London) **437**, 1257 (2005).
[28] J. J. Gooley, A. Schomer, and C. B. Saper, Nat. Neurosci. **9**, 398 (2006).
[29] T. Porkka-Heiskanen, R. E. Strecker, M. M. Thakkar, A. A. Bjørkum, R. W. Greene, and R. W. McCarley, Science **276**, 1265 (1997).
[30] J. T. McKenna, L. J. Dauphin, K. J. Mulkern, A. M. Stronge, R. W. McCarley, and R. E. Strecker, Sleep Res. Online **5**, 155 (2003).
[31] M. Dworak, P. Diel, S. Voss, W. Hollmann, and H. K. Strüder, Neuroscience **150**, 789 (2007).
[32] M. Díaz-Muńoz *et al.*, Sleep Res. Online **2**, 33 (1999).
[33] W. J. Freeman, *Mass Action in the Nervous System* (Academic Press, New York, 1975).
[34] P. L. Nunez, *Neocortical Dynamics and Human EEG Rhythms* (Oxford University Press, New York, 1995).
[35] P. A. Robinson, C. J. Rennie, and J. J. Wright, Phys. Rev. E **56**, 826 (1997).
[36] C. W. Berridge and B. D. Waterhouse, Brain Res. Rev. **42**, 33 (2003).
[37] D. G. Rainnie, H. C. Grunze, R. W. McCarley, and R. W. Greene, Science **263**, 689 (1994).

[38] C. K. R. T. Jones, in *Geometric Singular Perturbation Theory*, edited by R. Johnson, Lecture Notes in Mathematics Vol. 1609 (Springer, New York, 1995).
[39] S. H. Strogatz, *Nonlinear Dynamics and Chaos: With Applications to Physics, Biology, Chemistry, and Engineering* (Westview Press, Cambridge, MA, 1994).
[40] J. Guckenheimer and P. Holmes, *Nonlinear Oscillations, Dynamical Systems, and Bifurcations of Vector Fields* (Springer-Verlag, Berlin, 1983), Chap. 1.
[41] P. Glendinning, *Stability, Instability, and Chaos* (Cambridge University Press, Cambridge, England, 1994), Chap. 4.
[42] A. Report, Sleep **15**, 173 (1992).
[43] P. Halasz, M. Terzano, L. Parrino, and R. Bodizs, J. Sleep Res. **13**, 1 (2004).
[44] C. Lo, T. Chou, T. Penzel, T. E. Scammell, R. E. Strecker, H. E. Stanley, and P. C. Ivanov, Proc. Natl. Acad. Sci. U.S.A. **101**, 17545 (2004).
[45] M. H. Bonnet and D. L. Arand, Sleep Med. Rev. **7**, 297 (2003).
[46] J. Rajkowski, P. Kubiak, and G. Aston-Jones, Brain Res. Bull. **35**, 607 (1994).
[47] Y. Koyama, E. Joko, and Y. Kayama, Neuroscience **63**, 1021 (1994).
[48] B. Y. Mileykovskiy, L. I. Kiyashchenko, and J. M. Siegel, Neuron **46**, 787 (2005).
[49] K. Takahashi, J. S. Lin, and K. Sakai, Neuroscience **153**, 860 (2008).
[50] N. B. Reese, E. Garcia-Rill, and R. D. Skinner, Prog. Neurobiol. **47**, 105 (1995).
[51] B. E. Jones and H. H. Webster, Brain Res. **451**, 13 (1988).
[52] M. H. Bonnet, L. C. Johnson, and W. B. Webb, Psychophysiology **15**, 412 (1978).
[53] M. H. Bonnet, W. B. Webb, and G. Barnard, Sleep **1**, 271 (1979).
[54] M. H. Bonnet, Physiol. Behav. **45**, 1049 (1989).
[55] W. J. Lammers, P. Badia, R. Hughes, and J. Harsh, Psychophysiology **28**, 463 (1991).
[56] W. J. Lammers and P. Badia, Physiol. Behav. **50**, 867 (1991).
[57] W. B. Zimmerman, Psychophysiology **6**, 540 (1970).
[58] F. B. Keefe, L. C. Johnson, and E. J. Hunter, Psychophysiology **8**, 198 (1971).
[59] M. H. Bonnet, Sleep **8**, 11 (1985).
[60] M. Bonnet, *Biological Rhythms, Sleep, and Performance* (Wiley, New York, 1982), Chap. 8.
[61] R. A. Wever, *The Circadian System of Man* (Springer-Verlag, Berlin, 1979).
[62] R. L. Horner, Sleep **19**, 827 (1996).
[63] V. B. Fenik, R. O. Davies, and L. Kubin, Am. J. Respir. Crit. Care Med. **172**, 1322 (2005).
[64] E. Sforza, J. Krieger, and C. Petiau, Sleep **22**, 69 (1999).
[65] R. B. Berry, M. A. Asyali, M. I. McNellis, and M. C. K. Khoo, J. Appl. Physiol. **85**, 1434 (1998).
[66] T. E. Scammell and C. B. Saper, Nat. Neurosci. **8**, 1286 (2005).
[67] T. Sakurai *et al.*, Neuron **46**, 297 (2005).
[68] Y. Yoshida, N. Fujiki, T. Nakajima, B. Ripley, H. Matsumura, H. Yoneda, E. Mignot, and S. Nishino, Eur. J. Neurosci. **14**, 1075 (2001).