Journal of Computational Neuroscience 17, 13–29, 2004

The Influence of Spike Rate and Stimulus Duration on Noradrenergic Neurons

ERIC BROWN* AND JEFF MOEHLIS†
Program in Applied and Computational Mathematics, Princeton University, Princeton, NJ 08544, USA
ebrown@math.princeton.edu

PHILIP HOLMES
Program in Applied and Computational Mathematics, Department of Mechanical and Aerospace Engineering, Princeton University, Princeton, NJ 08544, USA

ED CLAYTON, JANUSZ RAJKOWSKI AND GARY ASTON-JONES
Department of Psychiatry and Laboratory of Neuromodulation and Behavior, University of Pennsylvania, Philadelphia, PA 19104, USA

Received June 11, 2003; Revised December 22, 2003; Accepted February 19, 2004

Action Editor: Peter Dayan

Abstract. We model spiking neurons in *locus coeruleus* (LC), a brain nucleus involved in modulating cognitive performance, and compare with recent experimental data. Extracellular recordings from LC of monkeys performing target detection and selective attention tasks show varying responses dependent on stimuli and performance accuracy. From membrane voltage and ion channel equations, we derive a phase oscillator model for LC neurons. Average spiking probabilities of a pool of cells over many trials are then computed via a probability density formulation. These show that: (1) Post-stimulus response is elevated in populations with lower spike rates; (2) Responses decay exponentially due to noise and variable pre-stimulus spike rates; and (3) Shorter stimuli preferentially cause depressed post-activation spiking. These results allow us to propose mechanisms for the different LC responses observed across behavioral and task conditions, and to make explicit the role of baseline firing rates and the duration of task-related inputs in determining LC response.

Keywords: *locus coeruleus*, response dynamics, phase density, phase oscillators, cognitive performance, neuromodulator, phasic and tonic states, conductance-based neuron models

1. Introduction and Background

The *locus coeruleus* (LC) is a brainstem nucleus containing approximately 15,000 neurons in monkeys (35,000 in humans), each of which can make 100,000 or more synapses with its widespread target regions, including the cerebral cortex (Moore and Bloom, 1979; Foote et al., 1983). LC neurons release norepinephrine, which is known to modulate brain processes including the sleep/wake cycle and arousal (Foote et al., 1983; Aston-Jones et al., 2001a). Recent data indicate that the LC regulates attention and behavioral flexibility (Aston-Jones et al., 1994; Usher et al., 1999; Aston-Jones et al., 2000). Specifically, increased LC

*To whom correspondence should be addressed.
†Present address: Department of Mechanical and Environmental Engineering, University of California, Santa Barbara, CA 93106, USA.

activity may increase the responsivity of decision networks following salient stimuli, hence improving accuracy. Conversely, lower baseline LC activity would reduce responsiveness to distractors (Usher et al., 1999; Servan-Schreiber et al., 1990).

Neurons in the alert monkey LC exhibit two distinct modes: *phasic* and *tonic* (Grant et al., 1988; Aston-Jones et al., 1994; Usher et al., 1999). In the latter, associated with labile behavior and poor performance on tasks requiring focused attention, cells fire at relatively high rates with little synchrony; in the former, associated with good performance, firing rates are lower but display greater synchrony. The phasic mode also produces greater response to stimuli, as detailed below. We emphasize that the phasic and tonic modes are defined based on good vs. poor task performance alone, but that this performance correlates strongly with baseline LC firing rates, as detailed below (also see Fig. 1A of Usher et al., 1999).

Two previous modeling studies have proposed mechanisms for the different firing properties of the phasic and tonic modes. The computational model of Usher et al. (1999) used a pool of coupled integrate-and-fire neurons and found, via numerical simulation, that increased electronic coupling promotes synchrony and enhanced responses to task stimuli. More recently, experimental and computational studies of paired *in vitro* LC cells have shown that decreased baseline activity can enhance the synchronizing effects of fixed-strength electrotonic coupling (Alvarez et al., 2002) (cf. Chow and Kopell, 2000); however, the different patterns of LC response to stimuli in the phasic vs. tonic modes were not addressed in these studies. Here, we show how decreased baseline spiking can, via different mechanisms, cause not only the partial synchronization but also the amplified response to exogenous inputs observed in the phasic mode. Hence we propose decreased baseline spiking rates in the phasic mode, resulting from altered exogenous input to the LC, as a new mechanism contributing to the phasic/tonic transition (as elaborated in the Discussion, this does not exclude other effects such as coupling). We suggest that these lower rates may result from decreased excitatory or enhanced inhibitory input from brain areas afferent to the LC (including the anterior cingulate cortex (ACC), a region previously implicated in cognitive control) or from neuromodulators such as corticotropin releasing factor (CRF). Neural evidence for these possibilities is elaborated upon in the Discussion.

Recent data indicates that LC responses differ not only between LC modes, as just discussed, but also among different psychological tasks. In Section 2, we present data demonstrating this difference for the target identification vs. Eriksen flanker tasks. This data also indicates different LC responses in trials in which correct vs. incorrect behavioral responses to task stimuli were obtained. In this paper, we show how differences in LC responses among task types and behavioral conditions can be accounted for by assuming different time-courses of inputs to the LC in the two tasks, as well as greater variation of input arrival times in incorrect vs. correct conditions, in accord with behavioral data. Indeed, these different LC inputs are a *prediction* of our model.

Below we develop a mathematical model for a pool of LC neurons, reduce it to differential equations for individual neuron phases, thereby retaining spike timing information, and analyze spiking probabilities in response to stimuli. This elucidates the dependence of spike histograms on model parameters and reveals how timescales in the neural substrate interact with those in the stimuli. Section 2 reviews relevant experimental data and provides an overview of its relationship with the modelling results that follow. We describe the neural model and probabilistic analysis in Sections 3 and 4, fit parameters and compare model results with data in Section 5, and discuss the work in Section 6.

2. Experimental Data and Modelling Overview

2.1. Experimental Results and Methods

Activities of individual neurons were obtained from behaving monkeys using extracellular recording techniques, as described previously (Aston-Jones et al., 1994). Animals were trained to continuously depress a pedal and visually fixate a centrally located spot on a video monitor. In the target detection task, after successful fixation, target (20% of trials) or non-target cues (80% of trials) were displayed singly in random order across trials, with random inter-trial intervals (1.65 sec on average). Release of the pedal within 650 msec after target cue onset was rewarded by juice. Four response categories are possible: correct detection (hit), correct rejection, incorrect detection (false alarm), and incorrect omission (miss).

The second task, the Eriksen flanker paradigm (Eriksen and Eriksen, 1974), requires greater

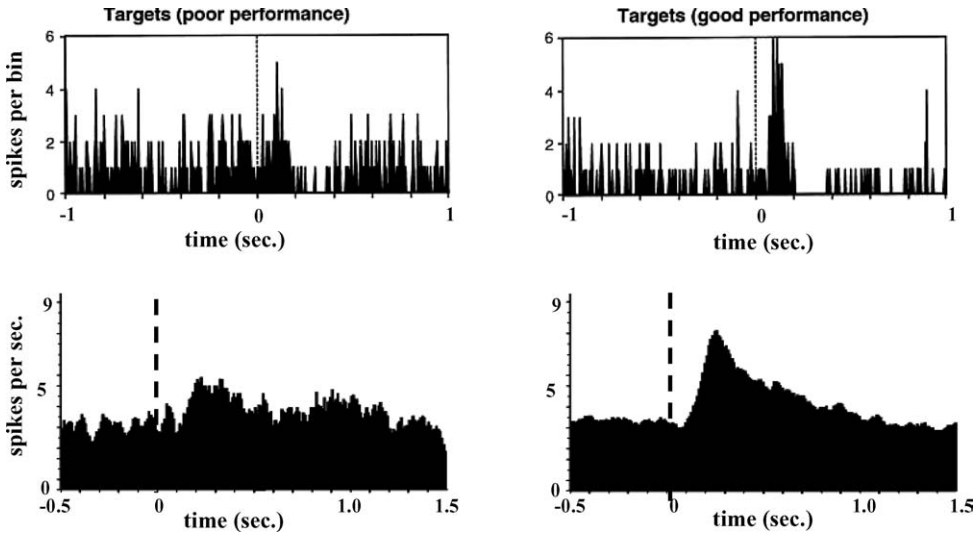

Figure 1. Peri-stimulus time histograms of LC activity for poor (left) and good (right) performance periods during the target identification task (top) and for incorrect (left) and correct (right) responses in the Eriksen task (bottom). Former are from single neurons, averaged over ∼100 trials; latter are cumulative PSTHs from multiple neurons averaged over ∼600 and 6000 trials respectively. Histograms are aligned at visual stimulus onset, marked by dashed line.

attentional focus. The display comprises five icons, with two 'flankers' on each side of the central cue, each selected at random with 50% probability from two possible cues ('left' or 'right'). The subject was trained to respond by releasing the left or right of two pedals according only to the central icon. The distracting flankers were either all identical to the central cue (congruent stimulus) or identical to the opposite, nondisplayed cue (incongruent stimulus).

Extracellular recordings from LC neurons were obtained from microwire electrodes positioned within the brain via a stereotaxically implanted guide cannula. LC neurons were identified during recording sessions by electrophysiological criteria, and continuous monitoring of eye position and pupil diameter was performed, as previously described (Aston-Jones et al., 1994). Baseline activity was calculated as an average spike rate during 500 msec epochs immediately preceding stimuli. Peri-stimulus time histograms (PSTHs) were produced and population PSTHs generated by aligning visual stimulus onsets and averaging across multiple sessions, or selected portions thereof. Histograms were smoothed via averages of spike counts in neighboring bins (using the program SigmaPlot) to facilitate superposition of cumulative PSTHs in a single figure while preserving response pattern and timing.

Figure 1 shows examples of the resulting PSTH data. These histograms reveal LC responses to stimuli for both tasks. As previously reported (Usher et al., 1999), in the target detection task, response relative to baseline is greater during good (phasic LC mode) compared to poor (tonic LC mode) performance, and a period of depressed spiking follows the response before activity returns to baseline in both cases. The phasic mode also displays greater synchrony (cf. Fig. 7 below). Reduced spiking following the LC response is *not* observed for the Eriksen task; instead, near-monotonic decay occurs following phasic activation, and the magnitude is much reduced for incorrect responses. Furthermore, in this latter task pre-stimulus spike rates are similar for both correct and incorrect responses (Clayton et al., 2004).

2.2. *Assumptions and Modelling Overview*

Figure 1 shows three ways in which LC response differs among task and behavioral conditions: between the phasic and tonic modes in the target detection task, between the target detection and the Eriksen flanker tasks, and between correct and incorrect trials in the Eriksen task. In this paper we develop a model of LC response to stimuli and use it to propose mechanisms for these differences. Our model is based on the following assumptions: (A1) Different levels of baseline

current input to LC neurons determine the different distributions of baseline firing rates reported in Usher et al. (1999) (p. 550, col. 1) for the phasic (slower, more tightly distributed rates: 2 ± 1.1 Hz, (mean $\pm$ st. dev.)) vs. tonic (faster, more broadly distributed: 3 ± 1.6 Hz) modes. (Note that these values are reported in Usher et al. (1999) as mean $\pm$ SEM.) (A2) In contrast to Usher et al. (1999) and to clearly separate the effects of baseline firing rate, coupling strengths are chosen to be identical in the phasic vs. tonic modes. (A3) The strength and duration of stimulus-related input to the LC may differ between the target detection and Eriksen tasks, as described in Section 3.2. (A4) Reflecting the greater variability in reaction times on incorrect vs. correct trials, onsets of stimulus-evoked inputs to the LC may also be more variable for incorrect trials. This latter assumption is justified under the hypothesis that pulsed inputs to the LC are driven by the (stochastic) decision process (Usher et al., 1999); more-variable timing of responses (reaction times) then implies increased variability of "triggering times" for LC inputs.

Analysis of our model yields three main mathematical results, presented in italics in the text. These are (1) Maximum LC response is elevated in populations with slower baseline firing rates, (2) response decays exponentially or faster with time due to noise and heterogeneous frequencies, and (3) in systems with narrow frequency distributions, short inputs necessarily lead to intervals of depressed firing following enhanced spiking and stimulus offset.

The first of these findings explains the influence of baseline spike rate on response of LC neurons to exogenous stimuli, and suggests that any factor leading to decreased baseline rate contributes to stronger responses in the phasic vs. tonic mode; see Fig. 4 (top right panel vs. bottom left) and Fig. 8 (top). The second finding tells us that, in order to produce the protracted LC responses seen in the Eriksen data, Eriksen task stimuli must elicit protracted inputs to the LC (since the impact of brief inputs decays quickly). The third finding implies that, in contrast to the Eriksen task, inputs to the LC elicited by target detection task stimuli must be punctate, because the interval of depressed firing observed in the data (Fig. 1, top) can occur only following the offset of (pulsed) inputs to the LC. Therefore the second and third findings address influence of stimulus duration on response of LC neurons. We also show via additional simulations that varying onset times of stimulus-evoked LC inputs in incorrect vs. correct Eriksen trials (in proportion to reaction time variability under these conditions) reproduces the trend in Fig. 1 (bottom).

3. A Mathematical Model for LC Neurons

3.1. A Conductance-Based Model and Phase Reduction

LC neurons possess calcium- and voltage-dependent potassium currents (e.g., 'A-currents'), which are largely responsible for their slow ($\lesssim$8 Hz) firing rate (Williams et al., 1984) and their resulting classification as 'Type I' cells (Ermentrout, 1996). We base our model of individual LC neurons on the original model by Connor et al. (1977) (cf. Rush and Rinzel (1995)) for a multi-ion-channel 'Type I' neuron including the A-current. This is a generic choice, intended to capture the essence but not necessarily the biophysical detail of LC dynamics. For ease of computation, we exploit a further simplification by Rose and Hindmarsh (1989), who used differing timescales and approximate relationships among state variables to reduce the Connor et al. system to two variables:

$$\begin{aligned}\dot{V}_i &= \big[I_i^b - g_{\mathrm{Na}} m_\infty(V_i)^3(-3(q_i - Bb_\infty(V_i)) \\ &\quad + 0.85)(V_i - V_{\mathrm{Na}}) \\ &\quad - g_K q_i(V_i - V_K) - g_L(V_i - V_L) + I_i^{\mathrm{ext}}\big]/C, \quad (1) \\ \dot{q}_i &= (q_\infty(V_i) - q_i)/\tau_q(V_i).\end{aligned}$$

Rose and Hindmarsh show that this reduction compares favorably with the original system. Here V_i is the voltage of neuron i ($i = 1, \ldots, N$ for an N-cell model), q_i is a collective gating variable, C is cell membrane conductance, g_{Na}, g_K, and g_L are maximum conductances for sodium, potassium, and leak currents, and I_i^b is the baseline inward current, which effectively sets spike frequency. I_i^{ext} denotes extracellular currents described below, and the other terms are channel gating variables, $m_\infty(V)$ and $b_\infty(V)$ denoting equilibrium levels for fast sodium and potassium channels. Functional forms and parameter values for this Rose-Hindmarsh model are given in the Appendix.

LC neurons are coupled by: (1) voltage differences between cells in electrical contact at electrotonic or gap junctions; (2) neurotransmitter release across synaptic clefts following presynaptic spikes. These effects, along with excitatory currents $I(t)$ representing inputs

due to external stimuli, enter I_i^{ext}:

$$I_i^{\text{ext}} = \frac{\beta_e}{N}\sum_{j=1}^{N}(V_j - V_i) + \frac{\beta_s}{N}\sum_{j=1}^{N}\sum_{k} A(t - t_{jk})(V_K - V_i) + I(t) + \sigma\eta_i(t). \quad (2)$$

Here uniform all-to-all coupling is assumed, β_e and β_s denote electrotonic and synaptic coupling strengths, and the 'alpha function' describes the postsynaptic excitation at neuron i after neuron j fires at time(s) $t_{jk} < t$: $A(t - t_{jk}) = [(t - t_{jk} - t_d)/\tau_A]\cdot \exp(-(t - t_{jk} - t_d)/\tau_A)$, where $\tau_A = 30$ msec is the synaptic time constant and $t_d = 25$ msec is an estimate of the alpha 2 NE receptor-mediated synaptic delay (Williams et al., 1991, Fig. 1). The white noise term $\sigma\eta_i(t)$ represents unmodeled 'fast' synaptic inputs.

Figure 2 shows the orbit in (V_i, q_i)-space of an isolated Rose-Hindmarsh neuron with I_i^b set to produce periodic spiking, and subject to a tonic stimulus $I(t)$ of greater strength than those employed below, superposed on the unperturbed trajectory ($I_i^{\text{ext}} \equiv 0$). Like most conductance-based neural models in repetitive firing modes, (1) possesses a strongly attracting, normally hyperbolic limit cycle (Guckenheimer and Holmes, 1983), implying that in the presence of moderate perturbations due to coupling and input currents, solutions remain confined to a small neighborhood of the original orbit. This permits reduction of (1) to phase variables, by defining nonlinear polar coordinates and projecting along isochrons onto the unperturbed limit cycle (Fig. 2) (Winfree, 2001):

$$d\theta_i = \left[\omega_i + z(\theta_i)\left(I(t) + \frac{\beta_e}{N}\sum_{j=1}^{N}(V(\theta_j) - V(\theta_i)) + \frac{\beta_s}{N}\sum_{j=1}^{N}\sum_{k} A(t - t_{jk})(V_K - V(\theta_i))\right) + \frac{\sigma^2}{2}z(\theta_i)z'(\theta_i)\right]dt + \sigma z(\theta_i)dW_i(t). \quad (3)$$

Here the $\sigma W_i(t)$ are independent Wiener processes with variance $\sigma^2 t$, and ω_i is the frequency of the ith LC neuron, which may vary slowly, e.g. via I_i^b, but is assumed constant over each experimental trial (see below). The phase θ is defined to increase at a constant rate ω_i in the absence of coupling and external inputs, with voltage peak (spike) at $\theta = 0$. The phase response curve (PRC) $z(\theta_i)$ (Winfree, 2001; Tass, 1999; Ermentrout, 1996), encoding the phase shift due to instantaneous perturbations, multiplies the stimulus and the external noise term; the $\mathcal{O}(\sigma^2)$ term is the 'Ito correction' resulting from changing variables from the stochastic differential equation (2) (Gardiner, 1985). The functions $V(\theta_i)$ in (3) are computed from the unperturbed voltage profile as $V(\theta_i) = V(\omega_i t)$.

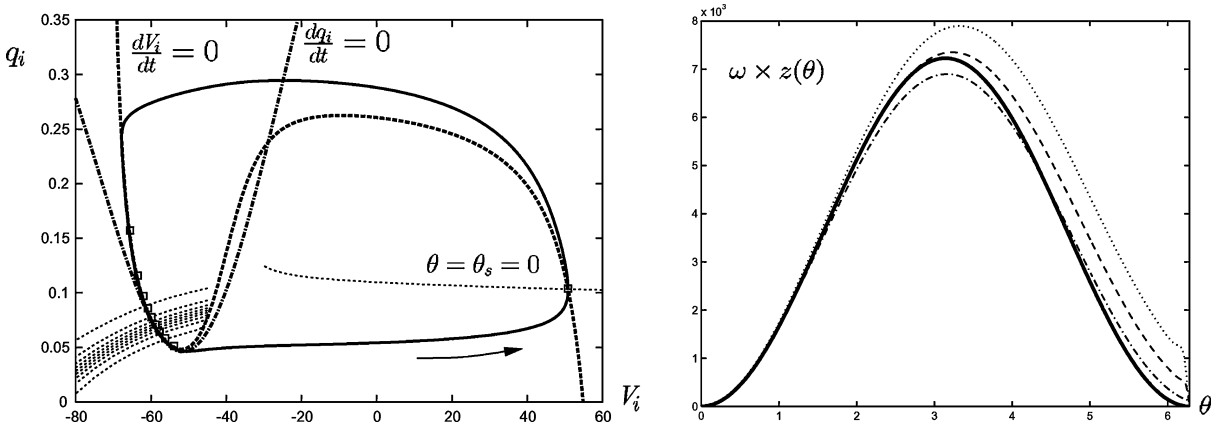

Figure 2. (Left) Phase space structure for a repetitively spiking Rose-Hindmarsh neuron (1), $I_i^b = 5\ \mu A/\text{cm}^2$. Attracting limit cycle for $I_i^{\text{ext}} = 0$ shown solid. Initial conditions on a given isochron (shown dashed) asymptotically approach the same point on the limit cycle as $t \to \infty$; isochrons are equally spaced in phase by $2\pi/10$, with $\theta = 0$ at action potential peak. The thick dashed and dash-dotted lines are nullclines for $\dot{V}_i = 0$ and $\dot{q}_i = 0$, respectively, and squares show points on perturbed limit cycle, equally spaced in time, under tonic stimulus of $I_i^{\text{ext}} = 1\ \mu\text{A/cm}^2$. (Right) PRCs for the Rose-Hindmarsh model (1) at frequencies $\omega/2\pi \approx 5$ Hz (dotted), $\omega/2\pi \approx 3.2$ Hz (dashed), $\omega/2\pi \approx 1.6$ Hz (dot-dashed). PRCs plotted as $\omega \times z(\theta)$ vs. θ to illustrate that $z(\theta) = \frac{c}{\omega}[1 - \cos\theta]$ with $c = 0.0036\ (\text{mV}\cdot\text{msec})^{-1}$ (solid) provides an acceptable fit, improving as ω decreases.

As I_i^b increases, Type I neurons undergo a transition from excitability, with a stable hyperpolarised rest point, to repetitive spiking via a saddlenode bifurcation on a limit cycle (Guckenheimer and Holmes, 1983). Normal form theory may be used to derive the PRC approximation $z(\theta) = C(\omega)[1-\cos(\theta)]$ near the bifurcation point (Ermentrout, 1996; Brown et al., 2004). This approximation is reasonable in the frequency range of interest (1–5 Hz); moreover, $C(\omega) = c/\omega$ where $c = 0.0036\,(\text{mV}\cdot\text{msec})^{-1}$: see Fig. 2. Other neuron models yield different PRC forms and ω-scaling (Ermentrout, 1996). In particular, the two-compartment LC neuron of Alvarez et al. (2002) yields PRCs similar to those of Fig. 2, although the lack of an explicit form for $z(\theta)$ precludes analysis of the type done below.

3.2. *Modeling LC Modes, Frequency Variability, and Stimuli*

To further develop our model, we review the data analysis leading to Fig. 1. Target detection PSTHs were obtained by averaging single-cell recordings over one session (≈100 trials), after separating epochs of good and poor behavioral performance according to error rates. These epochs correspond to phasic vs. tonic LC modes, respectively (Usher et al., 1999, Fig. 1A). Eriksen PSTHs derive from single- or multi-unit recordings, and multiple sessions. No clear tonic episodes were identified in the Eriksen data, although significant frequency variations were seen in individual cells over time, and among multiple cells at any given time (see Fig. 3). To reproduce the experimental data, frequencies ω will be drawn from appropriate narrow (for target detection) and broad (for Eriksen) distributions, noise variances fitted to match interspike interval distributions, and coupling strengths chosen to reproduce correlograms.

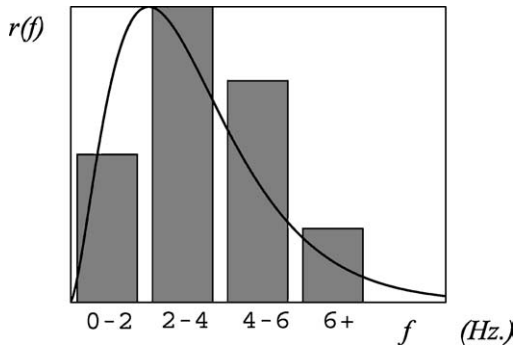

Figure 3. Bars: Estimated distribution of LC spiking frequencies at a fixed time. Solid line: Gamma distribution fit: $\Gamma(f) = \frac{1}{\Gamma(\beta)\phi^\beta} f^{\beta-1}\exp(-\frac{f}{\phi})$, with $\phi = 1.2, \beta = 2.75$ chosen to match mean and minimize least-squares difference within each quartile.

The correlation data of Usher et al. (1999, Fig. 4A) indicates only *partial* synchrony, even in the phasic mode. We find below that this can be captured by coupling terms that are weak compared with the stimulus and noise (cf. Fig. 7). Neglecting such weak coupling, any given cell is approximately governed by:

$$d\theta = \left[\omega + z(\theta)I(t) + \frac{\sigma^2}{2}z(\theta)z'(\theta)\right]dt + \sigma z(\theta)\,dW(t) \triangleq v(\theta,t)\,dt + \sigma z(\theta)\,dW(t). \quad (4)$$

Here and henceforth we drop the subscripts i and let θ and ω represent the phase and frequency of a typical neuron. Via comparison with simulations of the fully coupled biophysical model (1–2), we demonstrate below that this greatly reduced equation provides an adequate model (cf. Figs. 4 and 8).

Since decisions take longer in more complex tasks, we assume that LC inputs due to stimuli are briefer and more intense in target identification than in the Eriksen task. We take a simple square wave input of intensity $\bar{I}_{TD}$ and duration d_{TD} in the former case, and in the latter, a function that rises exponentially towards $\bar{I}_E$ for a period d_E and decays exponentially thereafter. Moreover, the Eriksen data does not indicate performance-dependent variations in baseline LC activity, and incorrect PSTHs keyed on response (rather than stimulus, as in Fig. 1) have peak activities similar to corrects, but it does display significantly broader reaction time distributions (Clayton et al., 2004). We therefore ascribe differences between the correct and incorrect PSTHs of Fig. 1 to variable input latencies in this more complex task. These and all other parameter choices are specified in Section 5.1, below.

4. Probabilistic Analysis

4.1. *Phase Density Equation*

We now describe how PSTHs such as those of Fig. 1 emerge from averages of single trials with appropriate initial conditions (as in, e.g. Herrmann and Gerstner, 2001; Fetz and Gustaffson, 1983). In this and the fol-

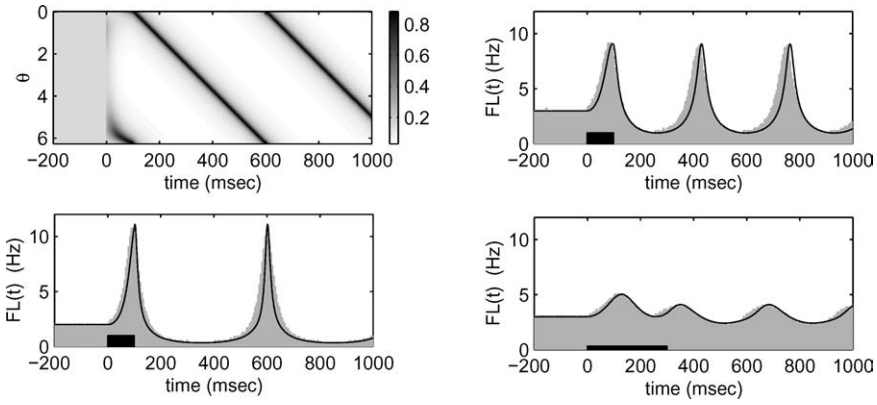

Figure 4. (Top and bottom left) Phase density $\rho(\theta, t)$ and flux $FL(t)$ computed from Eqs. (10) and (11) with $\omega/2\pi = 2$ Hz, $\bar{I} = 0.1\ \mu\text{A/cm}^2$, $d = 100$ msec. (Top and bottom right) Fluxes $FL(t)$ for $\omega/2\pi = 3$ Hz, $\bar{I} = 0.1\ \mu\text{A/cm}^2$, $d = 100$ msec (top) and $\bar{I} = 0.0333\ \mu\text{A/cm}^2$, $d = 300$ msec (bottom). Stimuli indicated by black bars. The 'charge' $\bar{I}d = 10\ \mu\text{A} \cdot \text{msec/cm}^2$ in both cases. Gray bars show spike rates computed directly from Rose-Hindmarsh equations (1).

lowing subsection we suppress explicit dependence on oscillator frequency ω, assuming that baseline frequency remains constant for each trial. Frequency variability is introduced later.

Let $\rho(\theta, t)$ denote the probability density of solutions of (4); thus $\rho(\theta, t)d\theta$ is the probability that a neuron's phase lies in the interval $[\theta, \theta + d\theta]$ at time t. The evolution of ρ is governed by the forward Kolmogorov or Fokker-Planck equation (Arnold, 1974):

$$\frac{\partial \rho(\theta, t)}{\partial t} = -\frac{\partial}{\partial \theta}[v(\theta, t)\rho(\theta, t)] + \frac{\sigma^2}{2}\frac{\partial^2}{\partial \theta^2}[z^2(\theta)\rho(\theta, t)]. \tag{5}$$

Here $v(\theta, t)$ denotes the deterministic vectorfield of (4) and σ is the r.m.s noise strength. Since (5) is linear, histograms representing many trials may be produced by averaging over suitable initial conditions. Specifically, pre- and post-stimulus, v lacks explicit t-dependence, so the probability that the phase lies in $[\theta, \theta + d\theta]$ is proportional to $T(\theta)$, the time spent in this interval during each cycle. Neglecting noise, this implies $\rho(\theta, 0) = \rho_0 \propto 1/v(\theta)$; for constant $v = \omega$, normalization gives the uniform distribution $\rho_0 \equiv 1/2\pi$. PSTHs may be extracted from (5) by noting that the probability for an arbitrary neuron to spike at time t is the rate at which its phase passes through $\theta = 0$, i.e., the flux $FL(t) \triangleq v(0, t)\rho(0, t)$. Similar phase density formulations appear in Tass (1999) and in Ritt (2003); Stein (1965), Herrmann and Gerstner (2001), Fetz and Gustaffson (1983), Omurtag et al. (2000), Nykamp and Tranchina (2000) and references therein, for example, develop voltage density descriptions. The formulation below is generalized to other neuron models in Brown et al. (2004).

4.2. *Spiking Probabilities in the Absence of Noise*

For noise-free ($\sigma = 0$) systems, (5) becomes a standard advection equation that may be solved explicitly for piecewise constant stimuli of duration $d = t_2 - t_1$: $I(t) = \bar{I}$ for $t_1 \le t \le t_2$ and $I(t) = 0$ otherwise. Since (5) then propagates the function $\rho(\theta, t)$ at speed $v(\theta, t)$, the method of characteristics ((Whitham, 1974), pp. 97–100 of Evans (1998)) yields:

$$\begin{aligned}\rho(\theta, t) &= \rho_0(\Theta_{\theta,t}(0)) \exp\left(-\int_0^t \frac{\partial}{\partial \theta} v(\Theta_{\theta,t}(t'), t')\, dt'\right) \\ &= \frac{1}{2\pi} \exp\left(-\bar{I}\int_{t_1}^{\tilde{t}_2} z'[\Theta_{\theta,t}(s)]\, ds\right),\end{aligned} \tag{6}$$

where $t \ge t_1, \tilde{t}_2 = \min(t, t_2)$ and we take the initial condition $\rho_0 = \rho(\theta, 0) = 1/2\pi$. Here, $\Theta_{\theta,t}(s)$ lies on the characteristic curve given by

$$\frac{d}{ds}\Theta_{\theta,t}(s) = v(\Theta_{\theta,t}(s), s), \tag{7}$$

with 'endpoint' condition $\Theta_{\theta,t}(t) = \theta$. To obtain the terms in (6), we integrate (7) backward in time from the

final condition at $s = t$ until $s = t_1$ or $s = \tilde{t}_2$, giving

$$\Theta_{\theta,t}(\tilde{t}_2) = \theta - \omega(t - \tilde{t}_2), \tag{8}$$

$$\Theta_{\theta,t}(t_1) = 2\arctan\left\{\sqrt{\frac{\omega}{b}}\tan\left[\arctan\left(\sqrt{\frac{b}{\omega}}\tan\left[\frac{\Theta_{\theta,t}(\tilde{t}_2)}{2}\right]\right) - \frac{1}{2}(\tilde{t}_2 - t_1)\sqrt{\omega b}\right]\right\}, \tag{9}$$

where $b = \omega + 2c\bar{I}/\omega$. For $t < t_1$, $\rho(0, t) \equiv 1/2\pi$.

Using the fact that $v(\Theta_{\theta,t}(s), s) = \omega + \bar{I}z(\Theta_{\theta,t}(s))$ for $t_1 \le s \le t_2$ and changing variables from s to $z(\Theta_{\theta,t}(s))$, the integral in (6) may be evaluated to give:

$$\rho(\theta, t) = \frac{1}{2\pi}\left[\frac{\omega + \bar{I}z(\Theta_{\theta,t}(t_1))}{\omega + \bar{I}z(\Theta_{\theta,t}(\tilde{t}_2))}\right]. \tag{10}$$

The top left panel of Fig. 4 shows $\rho(\theta, t)$ computed from (8–10); note the constant speed post-stimulus propagation.

Since $v(0, t) = \omega + z(0)I(t) \equiv \omega$ at the spike phase, we obtain the flux:

$$FL(t) = \omega\rho(0, t) = \frac{\omega}{2\pi}\left[\frac{\omega + \bar{I}z(\Theta_{0,t}(t_1))}{\omega + \bar{I}z(\Theta_{0,t}(\tilde{t}_2))}\right]. \tag{11}$$

The right hand panels of Fig. 4 show $FL(t)$ for two different inputs. The short, strong (target detection) inputs yield post-peak intervals of depressed firing and substantial 'ringing,' while the protracted input gives less ringing. We also show histograms computed via direct numerical simulations of the Rose-Hindmarsh equations (1), indicating that, apart from a slight time stretch due to the PRC approximation, the reduction to a phase equation is remarkably accurate.

Building on these examples, we now elucidate the parameter dependence hidden in (11). First we recall that, during the input, $\Theta_{0,t}(\tilde{t}_2) = \Theta_{0,t}(t) = 0$, so that the following special case of Eq. (11) is valid:

$$FL(t) = \frac{\omega}{2\pi}\left[\frac{\omega + \bar{I}z(\Theta_{0,t}(t_1))}{\omega + \bar{I}z(0)}\right] = \frac{1}{2\pi}[\omega + \bar{I}z(\Theta_{0,t}(t_1))] \ge \frac{\omega}{2\pi}; \tag{12}$$

where we consecutively use the facts $z(0) = 0$ and $z(\theta) \ge 0\ \forall\ \theta \in [0, 2\pi)$. The inequality implies that firing rates do not dip below baseline while the stimulus is on.

After the stimulus, i.e. absent $I(t)$, (5) supports traveling waves of unchanging shape; hence, for $t > t_2$, $FL(t) = \omega\rho(0, t) = \omega\rho(-\omega(t - t_2), t_2)$: the distribution $\rho(\theta, t_2)$ determines the entire $2\pi/\omega$-periodic post-stimulus response. We characterize $\rho(\theta, t_2)$ as a function of frequency ω, stimulus duration $d = t_2 - t_1$ and strength $\bar{I}$. From the remarks above, it suffices to describe the evolution of $\rho(\theta, t)$ from $\rho(\theta, t_1) \equiv 1/2\pi$ through the stimulus duration d, governed by (5) with $v(\theta, t) \equiv \omega + \bar{I}z(\theta)$. From (8)–(10), $\rho(\theta, t)$ is periodic during stimulus (cf. Tass, 1999), with 'response period'

$$P = \frac{2\pi}{\sqrt{\omega b}} = \frac{2\pi}{\sqrt{2c\bar{I} + \omega^2}}, \tag{13}$$

and, between returns to the pre-stimulus distribution $1/2\pi$, $\rho(\theta, t)$ develops a peak and a trough that assume extremal values at 'half integer' points $nP + P/2$, $n = 0, 1, 2, \ldots$. From (9)–(10), these extrema are:

$$\rho_{\max} = \rho(0, t_1 + nP + P/2) = \frac{(\omega + \bar{I}z(\pi))}{2\pi\omega};$$
$$\rho_{\min} = \rho(\pi, t_1 + nP + P/2) = \frac{\omega}{2\pi(\omega + \bar{I}z(\pi))}, \tag{14}$$

where we use the facts that, during inputs, $\Theta_{\theta,t}(\tilde{t}_2) = \Theta_{\theta,t}(t) = \theta$, and $z(0) = 0$ and $z(\pi) = 2c/\omega$ are respectively the minimum and maximum of the approximate cosine-fitted PRC of Fig. 2. Thus stimuli of durations nP leave no trace (cf. Tass, 1999) after stimulus offset, while those with $d = nP + P/2$ have maximal effect.

To compare responses for different values of ω, $\bar{I}$ and d, we define the peak and refractory indices $R_p(d)$ and $R_r(d)$ as

$$R_p(d) = \frac{FL_{\max} - FL_{\text{base}}}{FL_{\text{base}}};$$
$$R_r(d) = \frac{FL_{\text{base}} - FL_{\min}}{FL_{\text{base}}}, \tag{15}$$

where the baseline value is $FL_{\text{base}} = \omega/2\pi$. Figure 5 (right) illustrates the d-dependence of R_p and R_r; in particular, Eqs. (14) and (15) yield:

$$R_p^{\max} = R_p(nP + P/2) = \frac{2c\bar{I}}{\omega^2};$$
$$R_r^{\max} = R_r(nP + P/2) = \frac{2c\bar{I}}{2c\bar{I} + \omega^2}, \tag{16}$$

as shown in Fig. 5 (left and center). Note that $R_p^{\max}$ is proportional to stimulus strength over frequency squared, which quantifies our first main result: *Maximum LC response is elevated in populations with*

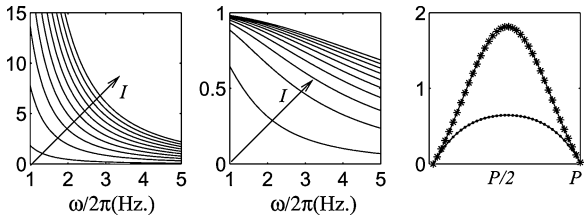

Figure 5. (Left, center) $R_p^{\max}$ and $R_r^{\max}$ as functions of ω for values of $\bar{I}$ evenly spaced between 0.01 and 0.30 μA/cm^2. (Right) $R_p(d)$ (stars) and $R_r(d)$ (dots) for $\omega/2\pi = 3.2$ Hz, $\bar{I} = 0.1$ μA/cm^2 and d ranging between 0 and P. The greatest post-stimulus ringing of the firing rate will occur for values of d around $P/2$ (where $R_{p,r}(d)$ are largest).

baseline slower firing rates. Density formulations derived from integrate and fire models (e.g. Fetz and Gustaffson, 1983; Herrmann and Gerstner, 2001), establish analogous inverse relationships between peak firing rates and baseline frequency. This effect, which in our model primarily derives from the c/ω factor in the PRC, is clear in the bottom left and top right panels of Fig. 4; note that (16) implies $FL_{\max} - FL_{\text{base}} \sim 1/\omega$.

4.3. *Effects of Noise and Distributed Frequencies*

We now restore two effects thus far neglected: noise and heterogeneous oscillator frequencies, and determine how these effects damp post-stimulus "ringing" of the LC response, thereby generalizing formulas in Tass (1999) to populations with distributed frequencies. Following stimulus offset and neglecting coupling, from (4) individual phases are described by

$$d\theta = \left[\omega + \frac{\sigma^2}{2} z(\theta) z'(\theta)\right] dt + \sigma z(\theta)\, dW(t). \quad (17)$$

In the limit of small r.m.s. noise strength σ, stochastic averaging (Zhu, 1988; Freidlin and Wentzell, 1998, Theorem 3.1) may be applied to (17). As in the averaging theory of deterministic dynamical systems (Guckenheimer and Holmes, 1983), this amounts to replacing the small deterministic term $\frac{\sigma^2}{2} z(\theta) z'(\theta)$ by its average $\frac{1}{2\pi}\int_0^{2\pi} \frac{\sigma^2}{2} z(\theta) z'(\theta) d\theta = 0$; additionally, the coefficient of the noise term, $\sigma z(\theta)$, is replaced by its r.m.s. value $\sigma \hat{z}_\omega = (\frac{1}{2\pi}\int_0^{2\pi} \sigma^2 z(\theta)^2 d\theta)^{1/2} = \sqrt{\frac{3}{2}}\frac{\sigma c}{\omega}$, yielding

$$d\theta = \omega dt + \sigma \hat{z}_\omega dW(t). \quad (18)$$

(Recall from Section 3.1 that the PRC has the explicit form $z(\theta) = \frac{c}{\omega}[1 - \cos(\theta)]$, and the subscript on $\hat{z}_\omega$ indicates this frequency dependence.) The corresponding Fokker-Planck equation

$$\frac{\partial \rho(\theta, t, \omega)}{\partial t} = -\frac{\partial}{\partial \theta}[\omega \rho(\theta, t, \omega)] + \frac{\sigma^2 \hat{z}_\omega^2}{2} \frac{\partial^2}{\partial \theta^2}[\rho(\theta, t, \omega)] \quad (19)$$

may then be Fourier transformed for each ω, as in Tass (1999), to yield

$$\rho(\theta, t, \omega) = \sum_{n=-\infty}^{\infty} a_n(t, \omega) \exp(in\theta), \quad \text{where}$$

$$\dot{a}_n = -i\omega n a_n - \frac{\sigma^2 \hat{z}_\omega^2}{2} n^2 a_n. \quad (20)$$

Solving the latter equations with 'initial' values $a_n(t_2, \omega)$ representing the state at stimulus end, we get:

$$FL(t, \omega) = \omega \sum_{n=-\infty}^{\infty} a_n(t_2, \omega) \times \exp\left[-\left(i\omega n + \frac{\sigma^2 \hat{z}_\omega^2 n^2}{2}\right)(t - t_2)\right]. \quad (21)$$

Additionally, as noted above (Fig. 3), there is significant baseline spike rate variation both among cells and in single cells over multiple trials. The PSTH data of Fig. 1 is thus effectively averaged over a frequency distribution $r(\omega)$. Carrying this out, we obtain:

$$\langle FL(t) \rangle = \int r(\omega) FL(t, \omega) d\omega = \int r(\omega)\omega \sum_{n=-\infty}^{\infty} a_n(t_2, \omega) \times \exp\left[-\left(i\omega n + \frac{\sigma^2 \hat{z}_\omega^2 n^2}{2}\right)(t - t_2)\right] d\omega. \quad (22)$$

so that

$$\langle FL(t) \rangle - \frac{\langle \omega \rangle}{2\pi} = \int_{-\infty}^{\infty} r(\omega)\omega \sum_{n=-\infty, n\neq 0}^{\infty} a_n(t_2, \omega) \times \exp\left[-\left(i\omega n + \frac{\sigma^2 \hat{z}_\omega^2 n^2}{2}\right)(t - t_2)\right] d\omega, \quad (23)$$

where $\langle\omega\rangle$ denotes the mean of $r(\omega)$. We will now estimate how the average spike rate $\langle FL(t)\rangle$ relaxes to its baseline value $\frac{\langle\omega\rangle}{2\pi}$.

Choosing a 'maximal' frequency ω_m beyond the essential support of the integrand, breaking the integral into pieces, and applying the triangle inequality, (23) yields

$$\begin{aligned}\left|FL(t)-\frac{\langle\omega\rangle}{2\pi}\right| \le & \left|\int_0^{\omega_m} r(\omega)\omega \sum_{n=-\infty,n\neq 0}^{\infty} a_n(t_2,\omega)\right. \\ & \left.\times \exp\left[-\left(i\omega n+\frac{\sigma^2\hat{z}_\omega^2 n^2}{2}\right)(t-t_2)\right]d\omega\right| \\ & +\left|\int_{\omega_m}^{\infty} r(\omega)\omega \sum_{n=-\infty,n\neq 0}^{\infty} a_n(t_2,\omega)\right. \\ & \left.\times \exp\left[-\left(i\omega n+\frac{\sigma^2\hat{z}_\omega^2 n^2}{2}\right)(t-t_2)\right]d\omega\right| \qquad (24)\end{aligned}$$

Noting that the exponential in the second term is bounded in modulus by 1, and that $\sum_{n=-\infty,n\neq 0}^{\infty} a_n(t_2,\omega) = FL(t_2,\omega) - \omega/2\pi$, we can bound the second term of (24) as follows:

$$\begin{aligned}&\left|\int_{\omega_m}^{\infty} r(\omega)\omega \sum_{n=-\infty,n\neq 0}^{\infty} a_n(t_2,\omega)\right. \\ &\qquad \left.\times \exp\left[-\left(i\omega n+\frac{\sigma^2\hat{z}_\omega^2 n^2}{2}\right)(t-t_2)\right]d\omega\right| \\ &\le \int_{\omega_m}^{\infty}\left|r(\omega)\omega \sum_{n=-\infty,n\neq 0}^{\infty} a_n(t_2,\omega)\right. \\ &\qquad \left.\times \exp[-(i\omega n)(t-t_2)]d\omega\right| \\ &= \int_{\omega_m}^{\infty} r(\omega)\omega\left|\tilde{\rho}(\theta=0,t-t_2,\omega)-\frac{1}{2\pi}\right|d\omega \\ &\le \int_{\omega_m}^{\infty} r(\omega)\omega \max_\theta\left|\tilde{\rho}(\theta,t_2,\omega)-\frac{1}{2\pi}\right|d\omega \triangleq E.\end{aligned}$$

Here, $\tilde{\rho}(\theta, t-t_2, \omega)$ is the density function that would result from evolution at frequency ω with $\sigma = 0$. Because $z_\omega(\theta) \sim 1/\omega$, the effective stimulus strength declines with ω, so that for sufficiently large ω the perturbation of $\rho(\theta, t_2, \omega)$ from equilibrium $\rho \equiv 1/2\pi$ is negligible. Because, additionally, $\omega r(\omega)$ decays for sufficiently large ω (since r is a PDF), ω_m may be chosen so that E is arbitrarily small. As for the first term in (24), noting that $\hat{z}_\omega$ decays with ω we can replace $\hat{z}_\omega$ in the exponential with its least value $\hat{z}_{\omega_m}$ and also replace n by 1 to obtain an upper bound on the decaying flux. This allows us to remove the exponential from the integral, giving

$$\begin{aligned}\left|FL(t)-\frac{\langle\omega\rangle}{2\pi}\right| \le & \exp\left[-\frac{\sigma^2\hat{z}_{\omega_m}^2(t-t_2)}{2}\right] \\ & \times\left|\int_0^{\omega_m} r(\omega)\omega \sum_{n=-\infty,n\neq 0}^{\infty} a_n(t_2,\omega)\right. \\ & \left.\times \exp[-(i\omega n)(t-t_2)]d\omega\right| + E. \qquad (25)\end{aligned}$$

Next, we note that

$$\begin{aligned}&\left|\int_0^{\omega_m} r(\omega)\omega \sum_{n=-\infty,n\neq 0}^{\infty} a_n(t_2,\omega)\right. \\ &\qquad \left.\times \exp[-(i\omega n)(t-t_2)]d\omega\right| \\ &= \left|\int_0^{\infty} r(\omega)\omega \sum_{n=-\infty,n\neq 0}^{\infty} a_n(t_2,\omega)\right. \\ &\qquad \times \exp[-(i\omega n)(t-t_2)]d\omega \\ &\qquad -\int_{\omega_m}^{\infty} r(\omega)\omega \sum_{n=-\infty,n\neq 0}^{\infty} a_n(t_2,\omega) \\ &\qquad \left.\times \exp[-(i\omega n)(t-t_2)]d\omega\right| \\ &\le \left|\int_0^{\infty} r(\omega)\omega \sum_{n=-\infty,n\neq 0}^{\infty} a_n(t_2,\omega)\right. \\ &\qquad \left.\times \exp[-(i\omega n)(t-t_2)]d\omega\right| + E\end{aligned}$$

which, together with (25), implies

$$\begin{aligned}\left|FL(t)-\frac{\langle\omega\rangle}{2\pi}\right| \le & \exp\left[-\frac{\sigma^2\hat{z}_{\omega_m}^2(t-t_2)}{2}\right] \\ & \times\left|\int_0^{\infty} r(\omega)\omega \sum_{n=-\infty,n\neq 0}^{\infty} a_n(t_2,\omega)\right. \\ & \left.\times \exp[-(i\omega n)(t-t_2)]d\omega\right|. \qquad (26)\end{aligned}$$

up to the arbitrarily small error $2E$. For each n, the integral in (26) is the Fourier transform of $r(\omega)\omega a_n(t_2,\omega)$

evaluated at $[n(t-t_2)]$. This integral decays for sufficiently large $t-t_2$ (by the Riemann-Lebesgue Lemma), giving an additional decay factor. Thus, (26) supplies our second main finding: *Response decays exponentially or faster with t due to noise and heterogeneous frequencies.*

In the case that $r(\omega)$ is Gaussian and varies rapidly compared with $\omega a_n(t_2, \omega)$, for each n contributing significantly to the sum, we have

$$r(\omega)\omega a_n(t_2,\omega) \approx \frac{1}{\sqrt{2\pi\gamma}} \exp\left(-\frac{(\omega-\langle\omega\rangle)^2}{2\gamma}\right)\langle\omega\rangle a_n(t_2,\langle\omega\rangle), \quad (27)$$

and the integral in (26) may be evaluated to give the following upper bound on decay rate:

$$\begin{aligned}\left|\langle FL(t)\rangle - \frac{\langle\omega\rangle}{2\pi}\right| &\le \exp\left[-\frac{\sigma^2\hat{z}^2_{\omega_m}(t-t_2)}{2} - \frac{\gamma(t-t_2)^2}{2}\right] \times \left|\sum_{n=-\infty, n\neq 0}^{\infty}\langle\omega\rangle a_n(t_2,\langle\omega\rangle)\right| \\ &= \exp\left[-\frac{\sigma^2\hat{z}^2_{\omega_m}(t-t_2)}{2} - \frac{\gamma(t-t_2)^2}{2}\right] \times \left|\langle FL(t_2,\langle\omega\rangle)\rangle - \frac{\langle\omega\rangle}{2\pi}\right|. \end{aligned} \quad (28)$$

Here $\langle FL(t_2, \langle\omega\rangle)\rangle$ is the value of FL at time t_2 under the condition that $r(\omega) = \delta(\omega - \langle\omega\rangle)$. When we fit parameters below, a (narrow) Gaussian distribution for which (27) holds is found to be appropriate, and Fig. 8, below, illustrates that (28) provides a good decay estimate.

4.4. Stimulus Duration and Depressed Firing

In Section 4.2, we showed that if inputs due to stimuli are sufficiently short compared with the (ω-dependent) response period:

$$d < P(\omega) = \frac{2\pi}{\sqrt{2c\bar{I}+\omega^2}}, \quad (29)$$

then $\rho(\theta, t_2, \omega)$ necessarily exhibits a peak and a trough, so that successive episodes of enhanced and depressed spiking ensue following stimulus offset. Longer inputs may or may not have this effect: they can end near 'integer points' $d \approx nP$, leaving $\rho(\theta, t_2, \omega) \approx 1/2\pi$, or at $d \approx nP + P/2$, leaving stronger post-stimulus effects (cf. Fig. 4). Furthermore, during the stimulus itself, firing rates do not dip below their baseline, as per the inequality in (12) and as demonstrated in Fig. 4d.

Noise and frequency heterogeneity cause additional decay of firing rates to their baseline level, cf. (28). For example, typical variations in $P(\omega)$ for the broad distribution of Fig. 3 range from 145 to 205 msec, leading to significantly differing $\rho(\theta, t_2, \omega)$'s, and differing propagation speeds. However, for tight distributions $r(\omega)$, $P(\omega)$ varies little and $\rho(\theta, t_2, \omega)$ travel at approximately the same speed, so the leading peak and depression can be expected to survive averaging over mild oscillator heterogeneity. This, together with the inequality in (12), leads to the finding: *In systems with narrow frequency distributions, short inputs necessarily lead to intervals of depressed firing following enhanced spiking and stimulus offset.*

This effect is further magnified if we normalise to maintain fixed 'synaptic charge' $\bar{I}d = S$. Now $\bar{I} \propto 1/d$ and (16) (Fig. 5) shows that brief inputs are yet further enhanced over longer, more diffuse ones. In this case, eliminating $\bar{I}$ from (29) yields an explicit input duration for maximal effect:

$$d \approx \frac{P}{2} = \frac{1}{\omega^2}\left(\sqrt{c^2S^2+\pi^2\omega^2} - cS\right). \quad (30)$$

5. Comparison with Experimental Data

5.1. Parameter Fitting

To compare model predictions with data, we first determine appropriate frequency distributions $r(\omega)$ for *single* neurons recorded over long durations, characterized by mean μ^ω and variance γ^ω, and r.m.s. noise strength σ, by seeking parameter values for which model realizations match both an empirical interspike interval (ISI) histogram and correlations between neighboring ISIs:

$$r_1 \triangleq \frac{\mathbb{E}\{(y_j - m)(y_{j+1} - m)\}}{\mathbb{E}\{(y_j - m)^2\}}. \quad (31)$$

Here subsequent ISIs are labeled y_j and $\mathbb{E}$ denotes expectation, $m = \mathbb{E}\{y_j\}$. The process $\{y_j\}$ is assumed stationary so r_1 and m are independent of j. Variability is assumed due to: (1) slow drift in baseline

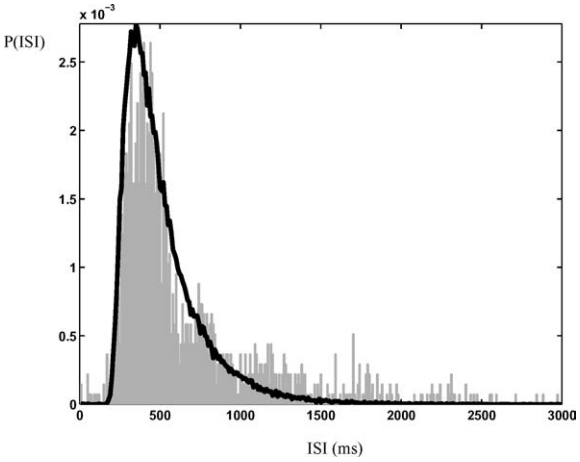

Figure 6. Empirical (gray bars) and model (solid line) histograms of baseline ISIs for a single neuron during the Eriksen task. For empirical data, windows of 1.5 sec. following each stimulus are removed to isolate baseline frequencies. Histogram bins are 10 ms wide.

frequency ω, and (2) rapid input current fluctuations modeled through the $\sigma z(\theta)dW(t)$ term in (3). Thus $y_j = y_j^\omega + \eta_j$, where y_j^ω are the noise-free drift values, and η_j causes additional variance $(\gamma^\sigma)^2$ due to rapid noise. If the drift is sufficiently slow then $\mathbb{E}\{(y_j^d - m)(y_{j+1}^d - m)\} \approx \mathbb{E}\{(y_j^d - m)^2\} \triangleq (\gamma^\omega)^2$, and (31) becomes

$$r_1 \approx \frac{(\gamma^\omega)^2}{(\gamma^\sigma)^2 + (\gamma^\omega)^2} = 0.1, \tag{32}$$

where we appeal to independence of the η_j and insert the numerical value derived from the data of Fig. 6, recorded from a single LC neuron in an Eriksen session.

Equation (32) constrains the ratio of slow $((\gamma^\omega)^2)$ to fast $((\gamma^\sigma)^2)$ ISI variances, the breadth of the ISI histogram constrains the magnitude of these variances, and the mean frequency μ^ω may be estimated directly from the ISI mean m. Guided by this and by analytical expressions for, e.g., barrier hitting times relating σ and γ^σ, Monte-Carlo simulations suggest a Gaussian distribution $r(\omega)$ with mean 1.69 Hz and standard deviation 0.47 Hz, and r.m.s. noise strength $\sigma = 0.45$. This yields the model ISI distribution of Fig. 6. To match the baseline data for the single neuron PSTHs of the target detection task, we rescale the center frequencies to 2 and 3 Hz respectively for the phasic and tonic modes, while keeping the ratio of mean to standard deviation constant. For the multi-neuron Eriksen data, we use the broader Gamma distribution of Fig. 3. (Recall that these different frequency distributions are realized by different distributions of baseline currents I_i^b, see Discussion.) We maintain $\sigma = 0.45$ throughout.

Synaptic and electrotonic coupling strengths were chosen to qualitatively capture the experimental cross-correlograms for phasic and tonic episodes given in Usher et al. (1999) (see Fig. 7). As described in Section 2.2 (Usher et al., 1999) reports that baseline firing rates (among the entire LC population) are not only slower than in the tonic mode but are also more tightly distributed. Thus, in the phasic (tonic) mode, we draw spike frequencies from a Gamma distribution with $\beta = 3, \phi = 0.667(\beta = 3, \phi = 1)$, giving mean 2 Hz (3 Hz) and standard deviation 1.16 Hz (1.73 Hz) (cf. Fig. 3). (However, as described in the previous paragraph, only some of these frequencies contribute to the single-neuron PSTHs of the target detection task.) We then require that a central subgroup of oscillators are largely synchronous (asynchronous). This yields $\beta_s = 0.01, \beta_e = 0.05$. Note that, unlike (Usher et al.,

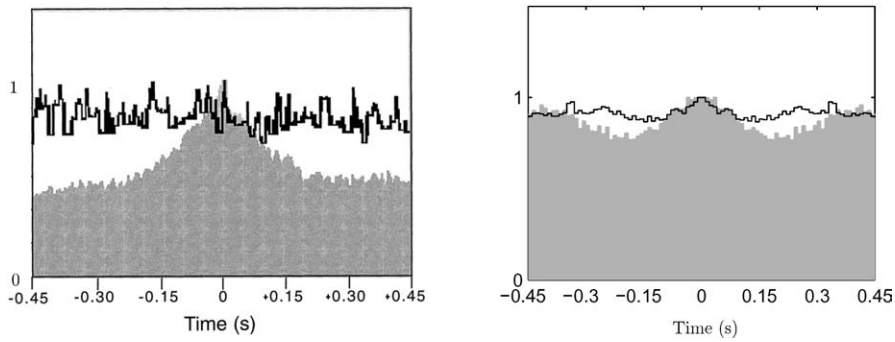

Figure 7. Normalized cross correlograms for phasic LC mode (filled histogram) and tonic mode (solid line). (Left) from Usher et al. (1999, Fig. 4) for two simultaneously recorded LC neurons. (Right) model results derived from mean ± 1 standard deviation of 100 oscillator population. In both cases, central peak indicates increased synchrony in phasic mode.

1999), we take the *same* coupling strengths for phasic and tonic modes, showing that increased synchrony can result solely from a tighter distribution of phasic frequencies.

Finally, we found via interactive simulations the following appropriate inputs $I(t)$. For the target detection task we found that a square wave of height $\bar{I} = 0.125\ \mu\text{A/cm}^2$ and duration $d = 110$ msec was satisfactory. Since the Eriksen data is averaged over all conditions (congruent and incongruent stimuli) and presumably involves more complex cognitive processing, a more diffuse input is appropriate. We adopted an exponentially rising and falling function, with rise duration $d = 180$ msec and rise and fall time constants 75 and 90 msec respectively and maximum height $\bar{I} = 0.22\ \mu\text{A/cm}^2$. Moreover, reaction time (RT) distributions have significantly greater standard deviation than for target detection: 114 and 241 msec for correct and incorrect respectively (Clayton et al., 2004), compared to ≈34 and 53 msec for phasic and tonic modes respectively (Usher et al., 1999). We therefore averaged over Gaussian distributions of onset times with standard deviations of 38 and 80 msec in the Eriksen task, assuming that variability in input arrival times at LC contributes about one third of total RT variability. Because our simulations indicate that the much smaller RT variability in the target detection task produces only minor effects (see below), we used fixed latencies in modeling this task. In all cases, since LC input lags visual stimulus, we include a time delay of 90 msec (Aston-Jones et al., 1994). Thus, as noted in Section 2.2, the results of Section 4 *predict* that the target detection and Eriksen inputs must differ qualitatively.

5.2. *Comparison of Model and Empirical PSTH Data*

Figure 8 shows model PSTH data for the target detection and Eriksen tasks, obtained in three ways: (1) by numerical solution of (5) in the presence of noise ($\sigma \neq 0$), followed by averaging over the frequency distributions derived above; (2) via direct simulations of a set of $N = 100$ globally-coupled Rose-Hindmarsh equations (1) representative of the same distributions, excited by independent Brownian noise currents of appropriate strength; and (3) directly from the noise-free expression (11) averaged over the same frequency distributions. The probabilistic effects considered above are clear: population averaging and noise combine to damp the periodic ringing of the noise-free single-frequency data of Fig. 4 (cf. the decay rate bound (28)).

These results confirm that reduction to a single phase equation (4) and the probabilistic theory developed above provide good descriptions of the coupled Rose-Hindmarsh system, and that the decay rate bounds are reasonable. The noise-free limit (11) is a useful qualitative estimator of PSTHs, although target detection phasic/tonic response ratios are significantly less than 9/4 predicted by (16), due to the high noise level that

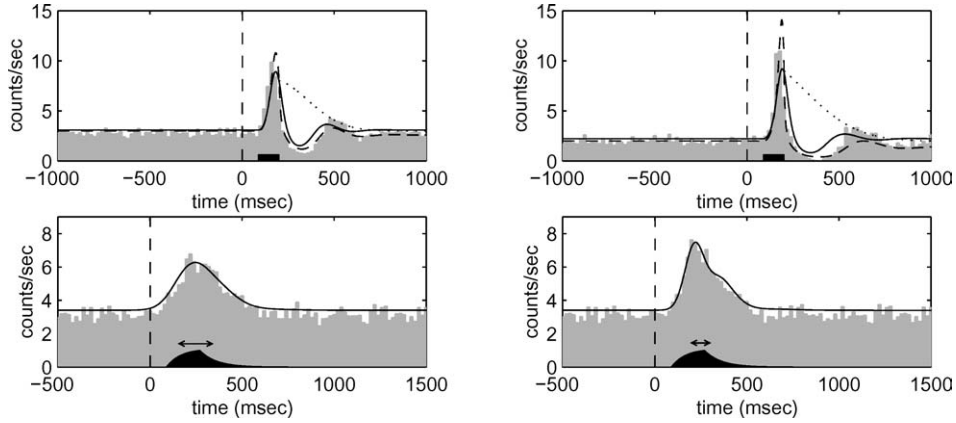

Figure 8. Model PSTHs computed from solution of Eq. (5) (solid) and from Eq. (11) (dashed), averaged over neuron frequency distributions and with stimuli (shown as filled black; arrows above stimuli for Eriksen task indicate variability in stimulus onset) and all other parameters as described in text. Gray bars show results of simulating 100 Rose-Hindmarsh neurons for multiple trials. Decay bound of (28) shown dotted. Top row: target detection task for poor performance/tonic mode (left) and good performance/phasic mode (right); bottom row: Eriksen task for incorrect (left) and correct (right) responses. Analytical results (11) do not apply to the varying magnitude Eriksen Stimuli; hence no dashed curves appear on bottom row.

selectively damps the sharply peaked phase densities arising at low frequencies.

The model results of Fig. 8 qualitatively reproduce the PSTHs of Fig. 1, with the major quantitative discrepancy that enhancement of response magnitude for phasic relative to tonic states in target detection captures only a part of that reported in Usher et al. (1999). In terms of the measure R_{mag} that characterizes enhanced spiking following stimulus (Aston-Jones et al., 1994), our model predicts a ratio R_{mag} (phasic)/R_{mag} (tonic) ≈ 1.3, compared to the value 3.4 of Usher et al. (1999). Hence, additional mechanisms, beyond the frequency effects studied here, must be operative in the phasic/tonic transition. For example, noise levels σ may be elevated in the tonic mode (in addition to mean current values I_b). Our simulations (not reported here) confirm that this increases the relative magnitude of phasic mode responses, as predicted in Herrmann and Gerstner (2001). Averaging over a slightly broader distribution of input onsets in the tonic mode than in the phasic mode, as suggested by reaction time distributions in the two tasks, further enhances phasic vs. tonic responses, although the small RT variance shows that this is a minor effect in target detection. Electrotonic coupling changes may also play a role as in Usher et al. (1999). We note that recent additional analyses of target detection data, in which recordings were grouped by baseline rate without reference to tonic and phasic behavioral modes, revealed small differences in R_{mag} similar to those reported here.

Our model reveals that the tonic/phasic frequency difference contributes to the variation between PSTHs for poor and good target identification performance, while in the Eriksen data, for which baseline frequencies are similar, PSTH differences can be accounted for by variations in stimulus arrival times originating in earlier processing. Moreover, diffuse stimuli in the latter case eliminate the depressed post-activation spiking seen in target identification.

6. Discussion

We have shown that a biophysical model of coupled LC neurons can be reduced to a stochastic differential equation for the phase of a given cell, and that a probabilistic formulation and averaging over suitable frequency distributions allows one to model and analyze peri-stimulus time histograms derived from single and multi-cell LC recordings. Our model supplements that of Usher et al. (1999), and our analysis reveals explicit parameter dependencies, including the effects of stimuli appropriate to two different cognitive tasks.

In Usher et al. (1999), electrotonic coupling variations were proposed as the cause for transitions between tonic and phasic LC modes, and hence for differences in PSTHs associated with poor and good target detection performance. In the model presented here, while coupling clearly affects synchrony, the key factor influencing PSTHs averaged over many trials is the LC spike rate, governed by the baseline currents I_i^b. In any case, our baseline rate explanation differs from the electrotonic coupling mechanism of Usher et al. (1999); also, in that paper the I_i^b were set in the excitable range, so that noise and other external inputs were *necessary* for spiking. From simulations of subthreshold networks of coupled Rose-Hindmarsh neurons with noise-driven firing at 2–3 Hz, we found that reproducing post-stimulus periods of depressed activity requires strong collateral coupling among LC neurons in both phasic and tonic LC modes. The same conclusion held for solutions of the corresponding (coupled) phase density equations derived from the full 'theta model' (Ermentrout, 1996) (not reported here). A study of response dynamics in this high noise, high coupling regime, in which different mechanisms for the phasic to tonic transition may dominate, is in progress.

Since we assume here that frequencies are distributed more tightly in the slower phasic mode, we obtain enhanced phasic mode synchrony *without* changing coupling strength (Fig. 7); this differs from the subtler mechanism of Alvarez et al. (2002). In sum, we see synchrony as a *correlate* of elevated LC response, rather than its primary cause. For the LC *in vivo*, the synchronizing effects identified here and in Usher et al. (1999) and Alvarez et al. (2002) may all be relevant. Additional effects of stronger coupling terms, noise, and subthreshold neurons may also be important and are under investigation.

Possible explanations for decreased I_i^b include reduced inputs from other neurons afferent to the LC. The anterior cingulate cortex (ACC), a prefrontal area involved in cognitive control, has recently been shown to have excitatory (presumably glutamatergic) projections to the LC (Rajkowski et al., 2000; Jodo and Aston-Jones, 1997; Jodo et al., 1998). Our findings suggest that the ACC may send decreased excitation to the LC in the phasic vs. tonic mode. Intriguingly, small inhibitory (GABA-ergic) neurons have been found among in a peri-LC region and are known to project to LC neurons and dendrites (Aston-Jones et al., 2001b).

Several areas, including the prefrontal cortex, innervate this region, suggesting a pathway by which input currents I^b, and hence baseline firing, may be regulated by increased inhibition in the phasic mode.

Decreased firing rates in the phasic vs. tonic mode could also result from neuromodulators. For example, in some cases direct application of the neuropeptide corticotropin releasing factor (CRF) increases LC baseline activity and simultaneously decreases responses to sensory stimuli (Valentino and Foote, 1987). It has also been found that that the alpha2 adrenoceptor agonist clonidine (or ST-91) can decrease baseline activity and increase response (Aston-Jones et al., 1991); the neuromodulator corresponding to this drug is norepinephrine, which presumably could be sent to the LC from other noradrenergic brain areas. Many other examples of such 'modulatory' effects of neurotransmitters or exogenous inputs exist for neurons in other brain areas (Aston-Jones et al., 2001a). Finally, we note that since synaptic coupling among LC neurons is inhibitory, it transiently reduces net input currents, thus effectively decreasing I_i^b if LC neurons are sufficiently decorrelated.

The present analysis therefore provides a simple explanation for how mechanisms which set baseline firing rate, such as background levels of exogenous input, neuromodulators, or pharmacological agents, will influence the response of a neural population to pulsed stimuli (cf. Herrmann and Gerstner, 2001; Fetz and Gustaffson, 1983): indeed, it shows that this dual effect on baseline rates and evoked response is intrinsic to the dynamics of neural groups. The LC phasic and tonic modes are an example, but the dual effect occurs in numerous other brain areas and neurons (Aston-Jones et al., 2001a). This intrinsic baseline rate mechanism joins a list of others: in addition to altered electrotonic coupling (Usher et al., 1999), other mechanisms for simultaneous effects on baseline and stimulus-evoked firing have been proposed, including simultaneous transmitter actions at multiple receptors (Aston-Jones et al., 1994), alterations in specific second messenger pathways and ion conductances (Moore and Bloom, 1979; Foote et al., 1983).

In addition to the predictions regarding *in vivo* baseline LC inputs just described, our analysis also provides a prediction about inputs evoked by task-related stimuli. That is, neurons that project to the LC and evoke responses should remain active longer following stimuli in complex tasks such as the Eriksen paradigm than in simpler ones like target detection. This is consistent with the notion introduced above (assumptions (A3) and (A4) of Section 2.2) that the LC is driven by accumulating activity in decision areas, as this activity may be expected to accumulate more gradually in complex decision tasks.

In recent related work (Usher and Davelaar, 2002; Gilzenrat et al., 2002) abstracted models of LC population activity that modify gains in connectionist networks have been shown to capture neuromodulatory effects on cognitive performance. The present LC model, derived from the neural substrate, offers simplification comparable to Usher and Davelaar (2002) and Gilzenrat et al. (2002) as well as suggesting, in the coupled multi-unit phase model of Eq. (3), a middle ground between those abstractions and the complexity of the integrate-and fire pool of Usher et al. (1999) or the full Rose-Hindmarsh system of (1)–(2).

In summary, we have shown how: (1) post-stimulus LC response is elevated in populations with slower firing rates; (2) response decays exponentially or faster due to noise and heterogeneous neuron frequencies; and (3) 'focused' stimuli tend to lead to intervals of depressed spiking. The analytical tools developed here apply to rather general systems of limit cycle oscillators occurring in neural and other applications (e.g. Tass, 1999; Winfree, 2001). In addition to the Type I Rose-Hindmarsh cells studied here, we have found that populations of weakly-coupled neurons modeled by the Hodgkin-Huxley and Fitzhugh-Nagumo equations show qualitatively similar behavior, suggesting that the present analysis may apply to other (non-LC) neurons (Brown et al., 2004).

Acknowledgments

This work was partially supported by DoE grant DE-FG02-95ER25238, and PHS grants MH58480 and MH62196 (Cognitive and Neural Mechanisms of Conflict and Control, Silvio M. Conte Center). EB was supported under a National Science Foundation Graduate Fellowship and a Burroughs-Wellcome Training Grant in Biological Dynamics: 1001782 and JM by a National Science Foundation Postdoctoral Research Fellowship. We thank Jonathan Cohen and Jason Ritt for useful discussions. The Rose-Hindmarsh code was run on a Beowulf parallel cluster provided by the PICASso program at Princeton, funded under NSF IGERT grant DGE-9972930. In particular, we used the parallel wrapper mw.cpp by Steven H. Kleinstein, and relied on his very effective help in adapting this software to our

application. We thank the reviewers for their careful comments and suggestions, which led to substantial improvement of the paper.

Appendix: The Rose-Hindmarsh Equations

Parameter values and function definitions of the Rose-Hindmarsh model (1) are as follows:

$$
\begin{aligned}
V_{\mathrm{Na}} &= 55\ \mathrm{mV},\quad V_K = -72\ \mathrm{mV},\quad v_L = -17\ \mathrm{mV},\\
g_{\mathrm{Na}} &= 120\ \mathrm{mS/cm^2},\quad g_K = 20\ \mathrm{mS/cm^2},\\
g_L &= 0.3\ \mathrm{mS/cm^2},\quad g_A = 47.7\ \mathrm{mS/cm^2},\\
C &= 1\ \mu\mathrm{F/cm^2},\quad I_i^b = 5\ \mu\mathrm{A/cm^2},\\
\gamma_b &= 0.069\ \mathrm{mV}^{-1},\quad T_b = 1\ \mathrm{msec},\\
T_n &= 0.52\ \mathrm{msec},\quad B = 0.21\, g_A/g_K.
\end{aligned}
$$

$$
\begin{aligned}
q_\infty(v) &= n_\infty(v)^4 + Bb_\infty(v),\\
b_\infty(v) &= (1/(1+\exp(\gamma_b(v+53.3))))^4,\\
m_\infty(v) &= \alpha_m(v)/(\alpha_m(v)+\beta_m(v)),\\
n_\infty(v) &= \alpha_n(v)/(\alpha_n(v)+\beta_n(v)),\\
\tau_q(v) &= (\tau_b(v)+\tau_n(v))/2,\\
\tau_n(v) &= T_n/(\alpha_n(v)+\beta_n(v)),\\
\tau_b(v) &= T_b(1.24 + 2.678/\\
&\qquad (1+\exp((v+50)/16.027))),\\
\alpha_n(v) &= 0.01(v+45.7)/(1-\exp(-(v+45.7)/10)),\\
\alpha_m(v) &= 0.1(v+29.7)/(1-\exp(-(v+29.7)/10)),\\
\beta_n(v) &= 0.125\exp(-(v+55.7)/80),\\
\beta_m(v) &= 4\exp(-(v+54.7)/18).
\end{aligned}
$$

References

Alvarez V, Chow C, van Bockstaele E, Williams J (2002) Frequencydependent synchrony in locus coeruleus: Role of electronic coupling. Proc. Nat. Acad. Sci. USA 99: 4032–4036.

Arnold L (1974) Stochastic Differential Equations. John Wiley, New York.

Aston-Jones G, Akaoka H, Charlety P, Chouvet G (1991) Serotonin selectively attenuates glutamate-evoked activation of locus coeruleus neurons in vivo. J. Neurosci. 11: 760–769.

Aston-Jones G, Chen S, Zhu Y, Oshinsky ML (2001a) A neural circuit for circadian regulation of arousal. Nature Neurosci. 4: 732–738.

Aston-Jones G, Rajkowski J, Cohen J (2000) Locus coeruleus and regulation of behavioral flexibility and attention. Prog. Brain Res. 126: 165–182.

Aston-Jones G, Rajkowski J, Kubiak P, Alexinsky T (1994) Locus coeruleus neurons in the monkey are selectively activated by attended stimuli in a vigilance task. J. Neurosci. 14: 4467–4480.

Aston-Jones G, Zhu Y, Card P (2001b) Gabaergic afferents to locus coeruleus (LC) from the peri-lc region: Possible LC interneurons. Soc. Neurosci. Abst. 27: 373.8.

Brown E, Moehlis J, Holmes P (2004) On the phase reduction and response dynamics of neural oscillator populations. Neural Comp. 16(4): 673–715.

Chow C, Kopell N (2000) Dynamics of spiking neurons with electrotonic coupling. Neural Comp. 12: 1643–1678.

Clayton E, Rajkowski J, Cohen JD, Aston-Jones G (2004) Decision-related activation of monkey locus coeruleus neurons in a forced choice task. In preparation.

Connor J, Walter D, McKown R (1977) Neural repetitive firing: Modifications of the Hodgkin-Huxley axon suggested by experimental results from crustacean axons. Biophys. J. 18: 81–102.

Eriksen BA, Eriksen CW (1974) Effects of noise letters upon the identification of target letters in a non-search task. Perception and Psychophysics 16: 143–149.

Ermentrout B (1996) Type I membranes, phase resetting curves, and synchrony. Neural Comp. 8: 979–1001.

Evans L (1998) Partial Differential Equations. American Mathematical Society, Providence.

Fetz E, Gustaffson B (1983) Relation between shapes of post-synaptic potentials and changes in firing probability of cat motoneurones. J. Physiol. 341: 387–410.

Foote SL, Bloom FE, Aston-Jones G (1983) Nucleus locus coeruleus: New evidence of anatomical and physiological specificity. Physiol. Rev. 63(3): 844–914.

Freidlin M, Wentzell A (1998) Random Perturbations of Dynamical Systems. Springer, New York.

Gardiner C (1985) Handbook of Stochastic Methods. Springer, New York.

Gilzenrat MG, Holmes BD, Rajkowski J, Aston-Jones G, Cohen JD (2002) Simplified dynamics in a model of noradrenergic modulation of cognitive performance. Neural Networks 15: 647–663.

Grant SJ, Aston-Jones G, Redmond DE (1988) Responses of primate locus coeruleus neurons to simple and complex sensory stimuli. Brain Res. Bull. 21(3): 401–410.

Guckenheimer J, Holmes PJ (1983) Nonlinear Oscillations, Dynamical Systems and Bifurcations of Vector Fields. Springer-Verlag, New York.

Herrmann A, Gerstner W (2001) Noise and the PSTH response to current transients: I. General theory and application to the integrate-and-fire neuron. J. Comp. Neurosci. 11: 135–151.

Jodo E, Aston-Jones, G (1997) Activation of locus coeruleus by prefrontal cortex is mediated by excitatory amino acid inputs. Brain Res. 768: 327–332.

Jodo E, Chiang C, Aston-Jones G (1998) Potent excitatory influence of prefrontal cortex activity on noradrenergic locus coeruleus neurons. Neuroscience 83: 63–80.

Moore RY, Bloom FE (1979) Central catecholamine neuron systems: Anatomy and physiology of the norepinephrine and epinephrine systems. Ann. Rev. Neurosci. 2: 113–168.

Nykamp D, Tranchina D (2000) A population density approach that facilitates large-scale modeling of neural networks: Analysis and application to orientation tuning. J. Comp. Neurosci. 8: 19–50.

Omurtag A, Knight BW, Sirovich L (2000) On the simulation of large populations of neurons. J. Comp. Neurosci. 8: 51–63.

Rajkowski J, Lu W, Zhu Y, Cohen J, Aston-Jones G (2000) Prominent projections from the anterior cingulate cortex to the locus coeruleus (LC) in rhesus monkey. Soc. Neurosci. Abst. 26: 838.15.

Ritt J (2003) A Probabilistic Analysis of Forced Oscillators, with Application to Neuronal Response Reliability. PhD thesis, Boston University.

Rose R, Hindmarsh J (1989) The assembly of ionic currents in a thalamic neuron I. The three-dimensional model. Proc. R. Soc. Lond. B 237: 267–288.

Rush M, Rinzel J (1995) The potassium A-current, low firing rates and rebound excitation in Hodgkin-Huxley models. Bull. Math. Biol. 57: 899–929.

Servan-Schreiber D, Printz H, Cohen JD (1990) A network model of catecholamine effects: Gain, signal-to-noise ratio, and behavior. Science 249: 892–895.

Stein R (1965) A theoretical analysis of neuronal variability. Biophys. J. 5: 173–194.

Tass P (1999) Phase Resetting in Medicine and Biology. Springer, New York.

Usher M, Cohen JD, Servan-Schreiber D, Rajkowsky J, Aston-Jones G (1999) The role of locus coeruleus in the regulation of cognitive performance. Science 283: 549–554.

Usher M, Davelaar EJ (2002) Neuromodulation of decision and response selection. Neural Networks 15: 635–645.

Valentino RJ, Foote SL (1987) Corticotropin-releasing factor disrupts sensory responses of brain noradrenergic neurons. Neuroendocrinology 45(1): 28–36.

Whitham GB (1974) Linear and Nonlinear Waves. Wiley, New York.

Williams J, North R, Shefner A, Nishi S, Egan T (1984) Membrane properties of rat locus coeruleus neurons. Neuroscience 13: 137–156.

Williams JT, Bobker DH, Harris GC (1991) Synaptic potentials in locus coeruleus neurons in brain slices. Prog. Brain Res. 88: 167–172.

Winfree AT (2001) The Geometry of Biological Time. 2nd edition. Springer, New York.

Zhu WQ (1988) Stochastic averaging methods in random vibration. Appl. Mech. Rev. 41: 189–199.