The influence of spike rate and stimulus duration on noradrenergic neurons

E. Brown\textsuperscript{1*}, J. Moehlis\textsuperscript{1}, P. Holmes\textsuperscript{1,2}, E. Clayton\textsuperscript{3}, J. Rajkowski\textsuperscript{3}

and G. Aston-Jones\textsuperscript{3}

\textsuperscript{1} Program in Applied and Computational Mathematics,
\textsuperscript{2} Department of Mechanical and Aerospace Engineering,
Princeton University, Princeton, NJ 08544, U.S.A.
\textsuperscript{3} Department of Psychiatry
and Laboratory of Neuromodulation and Behavior,
University of Pennsylvania, Philadelphia, PA 19104, U.S.A.

June 11, 2003

Abstract

We model spiking neurons in \textit{locus coeruleus} (LC), a brain nucleus involved in modulating cognitive performance, and compare with experimental data. Extracellular recordings from LC of monkeys performing target detection and selective attention tasks show varying responses dependent on stimuli and task performance level. 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 findings provide an explanation for modulatory effects of neurotransmitters, and may also apply to non-LC neurons.

\textbf{Key words:} \textit{locus coeruleus}, response dynamics, phase density, cognitive performance, neuromodulator

\textsuperscript{*To whom correspondence should be addressed:} ebrown@math.princeton.edu, PH: (609) 258-2185, FAX: (609) 258-1735
Introduction

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., 2001]. 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 activity may increase the responsibility 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 two modes also produce differing responses to stimuli, as detailed below.

Here we develop a biophysical model for LC neurons that explores the role of baseline frequency and stimulus duration in these and other observations. We reduce the model 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. We begin by reviewing relevant experimental data. Next, we describe the model and probabilistic analysis, fit parameters, and compare model results with data.

A previous LC model [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. Recently, experimental and computational studies of paired in vitro LC cells showed that decreased baseline activity can enhance the synchronizing effects of fixed-strength electrotonic coupling [Alvarez et al., 2002], cf. [Chow and Kopell, 2000]. The present paper demonstrates how decreased baseline spiking can, via different mechanisms, cause both partial synchronization and amplified response to exogenous inputs.
Experimental results

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 attentional focus. The display comprises five icons, with two 'flankers' on each side of the central cue. The subject was trained to respond on one 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, non-displayed 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 to facilitate superposition of cumulative PSTHs in a single figure while preserving response pattern and timing.

Fig. 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. 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. Furthermore, in this latter task pre-stimulus spike rates are similar for both correct and incorrect responses [Rajkowski et al.].
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.

Mathematical model, analysis, and results

Neuron model and phase oscillators

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]. Connor et al. [Connor et al., 1977], cf. [Rush and Rinzel, 1995] introduced a multi-ion-channel model for a ‘Type I’ neuron including the A-current, and Rose and Hindmarsh [Rose and Hindmarsh, 1989] exploited differing timescales and approximate relationships among state variables to reduce this system to two variables:

\[
\begin{align*}
\dot{v}_i &= \left[I_{i}^b - g_{Na}m_{\infty}(v_i)^3(-3(q_i - Bb_{\infty}(v_i)) + 0.85)(v_i - v_{Na}) \right], \\
&-g_K q_i (v_i - v_K) - g_L (v_i - v_L) + I_{ext}^i)/C \\
\dot{q}_i &= (q_{\infty}(v_i) - q_i)/\tau_q(v_i).
\end{align*}
\]
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 $I^b_i$ is the baseline inward current, which effectively sets spike frequency. $I^e_i$ 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 (RH) 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 depolarization. These effects, along with excitatory currents $I(t)$ representing inputs due to external stimuli, enter $I^e_i$:

$$I^e_i = \frac{\alpha_e}{N} \sum_{j=1}^{N} (v_j - v_i) + \frac{\alpha_s}{N} \sum_{j=1}^{N} \sum_{k} A(t - t_{jk})(v_K - v_i) + I(t) + \sigma \eta_i(t). \quad (2)$$

Here uniform all-to-all coupling is assumed, $\alpha_e$ and $\alpha_s$ 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 $t_d = 25$ msec is the synaptic delay and $\tau_A = 30$ msec the synaptic time constant. The white noise term $\sigma \eta_i(t)$ represents unmodeled ‘fast’ synaptic inputs.

Fig. 2 shows the orbit in $(v_i, q_i)$-space of an isolated RH neuron with $I^b_i$ 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 ($I^e_i \equiv 0$). 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).
Figure 2: (Left) Phase space structure for a repetitively spiking RH neuron \((1)\), \(I_I^b = 5 \mu A/cm^2\). Attracting limit cycle for \(I_I^{ext} = 0\) 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 \(\dot{\theta}_i = 0\) and \(\dot{q}_i = 0\), respectively, and squares show points on perturbed limit cycle, equally spaced in time, under tonic stimulus of \(I_I^{ext} = 1 \mu A/cm^2\). (Right) PRCs for the RH 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. \(\theta\) to illustrate that \(z(\theta) = \frac{\omega}{\omega}[1 - \cos \theta]\) with \(c = 0.0036\) (mV \cdot msec)\(^{-1}\) (solid) provides an acceptable fit, improving as \(\omega\) decreases.

[Winfree, 2001]:

\[
\frac{d\theta_i}{dt} = \left(\omega_i + z(\theta_i) \left( I(t) + \frac{\alpha_e}{N} \sum_{j=1}^{N} (v(\theta_j) - v(\theta_i)) + \frac{\alpha_s}{N} \sum_{j=1}^{N} \sum_{k} A(t - t_{jk})(v_K - v(\theta_i)) \right) + \frac{\sigma^2}{2} z(\theta_i) z'(\theta_i) \right) dt + \sigma z(\theta_i)dW_i(t).
\]

Here the \(\sigma W_i(t)\) are independent Wiener processes with variance \(\sigma^2 t\), and \(\omega_i\) is the frequency of the \(i\)th LC neuron, which may vary slowly, e.g. via \(I_I^b\), but is assumed constant over each experimental trial (see below). Phase \(\theta\) is defined to increase at a constant rate \(\omega_i\) 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 \(O(\sigma^2)\) 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 \(v(\theta_i) = v(\omega_i t)\).

As \(I_h^b\) increases, Type I neurons undergo a transition from excitability, with a stable hyperpolarized rest point, to repetitive spiking via a saddle-node 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]. This approximation is reasonable in the frequency range of interest (1-5 Hz); moreover, \(C(\omega) \sim 1/\omega\): see Fig. 2. Other neuron models yield different PRC forms and \(\omega\)-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.

**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 \((\approx 100\) 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 \(\omega\) 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.

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
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.

governed by:

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

$$\triangle = v(\theta, t)dt + \sigma z(\theta)dW(t). \tag{4}$$

Here and henceforth we drop the subscripts $i$ and let $\theta$ and $\omega$ 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. Fig. 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 $I_{TD}$ and duration $d_{TD}$ in the former case, and in the latter, a function that rises exponentially towards $I_E$ 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 have peak activities similar to corrects, but it does display significantly broader reaction time distributions. 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 the section on parameter fitting.
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 subsection we suppress explicit dependence on oscillator frequency $\omega$, 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 $\rho$ is governed by the forward Kolmogorov or Fokker-Planck equation [Arnold, 1974]:

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

Here $v(\theta, t)$ denotes the deterministic vectorfield of (4) and $\sigma$ 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 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 $FL(t) \equiv v(0, t) \rho(0, t)$. Similar phase density formulations appear in [Tass, 1999] and in work by Jason Ritt; 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.

Spiking probabilities in the absence of noise

For noise-free ($\sigma = 0$) systems, (5) may be solved exactly for piecewise constant stimuli of duration $d = t_2 - t_1$: $I(t) = \bar{I}$ for $t_1 \leq t \leq t_2$ and $I(t) = 0$ otherwise. Noting that (5) simply propagates the function $\rho(\theta, t)$ at speed $v(\theta, t)$, the method of characteristics yields:

$$\rho(\theta, t) = \rho_0(\Theta_{\theta, t}(0)) \exp \left( -\int_{0}^{t} \frac{\partial}{\partial \theta} v(\Theta_{\theta, t}(\theta), t')d\theta' \right)$$

$$= \frac{1}{2\pi} \exp \left( -\bar{I} \int_{t_1}^{t_2} \frac{\partial}{\partial s} \Theta_{\theta, t}(s) ds \right),$$ (6)
where \( t \geq t_1, \tilde{t}_2 = \min(t, t_2) \) 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

\[
\frac{d}{ds} \Theta_{\theta,t}(s) = v(\Theta_{\theta,t}(s), s),
\]

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

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

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 \( t_1 \leq s \leq t_2 \) and changing variables from \( s \) to \( z(\Theta_{\theta,t}(s)) \), the integral in (6) may be evaluated to give:

\[
\rho(\theta, t) = \frac{1}{2\pi} \left[ \frac{\omega + \bar{I} z(\Theta_{\theta,t}(t_1))}{\omega + \bar{I} z(\Theta_{\theta,t}(t_2))} \right].
\]

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:

\[
FL(t) = \omega \rho(0, t) = \frac{\omega}{2\pi} \left[ \frac{\omega + \bar{I} z(\Theta_{0,t}(t_1))}{\omega + \bar{I} z(\Theta_{0,t}(t_2))} \right].
\]

The right hand panels of Fig. 4 show \( FL(t) \) for two different inputs. The short, strong (target detection) inputs yields 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 RH Eqns. (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). 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 \( \omega \), 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
Figure 4: (Top and bottom left) Phase density $\rho(\theta,t)$ and flux $FL(t)$ computed from (10) and (11) with $\omega/2\pi = 2\text{ Hz}$, $\bar{I} = 0.1 \mu A/cm^2$, $d = 100\text{ msec}$. (Top and bottom right) Fluxes $FL(t)$ for $\omega/2\pi = 3\text{ Hz}$, $\bar{I} = 0.1 \mu A/cm^2$, $d = 100\text{ msec}$ (top) and $\bar{I} = 0.0333 \mu A/cm^2$, $d = 300\text{ msec}$ (bottom). Stimuli indicated by black bars. The ‘charge’ $\bar{I}d = 10 \mu A \cdot \text{msec}/cm^2$ in both cases. Gray bars show spike rates computed directly from RH Eqns. (1).

$d$, 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’

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

(12)

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:

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

$$\rho_{\text{min}} = \rho(\pi,t_1 + nP + P/2) = \frac{(\omega + \bar{I}z(\pi))}{2\pi(\omega + \bar{I}z(\pi))},$$

(13)

where we use the facts that, during stimulus presentation, $\Theta_{\theta,t}(\bar{I}_2) = \Theta_{\theta,t}(t) = \theta$, 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]), while those with $d = nP + P/2$ have maximal effect.
Figure 5: (Left, center) $R_p^{\text{max}}$ and $R_r^{\text{max}}$ as functions of $\omega$ for values of $\bar{I}$ evenly spaced between 0.01 and 0.30 $\mu$A/cm\(^2\). (Right) $R_p(d)$ (stars) and $R_r(d)$ (dots) for $\omega/2\pi = 3.2$ Hz, $\bar{I} = 0.1$ $\mu$A/cm\(^2\) and $d$ ranging between 0 and $P$.

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

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

(14)

where the baseline value is $FL_{\text{base}} = \omega/2\pi$. Fig. 5 (right) illustrates the $d$-dependence of $R_p$ and $R_r$; in particular, Eqns. (13-14) yield:

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

(15)

as shown in Fig. 5 (left and center). Note that $R_p^{\text{max}}$ is proportional to stimulus strength over frequency squared, which quantifies our first main result: **Maximum LC response is elevated in populations with slower firing rates.** Density formulations derived from integrate and fire models, e.g. [Fetz and Gustaffson, 1983, Hermann and Gerstner, 2001], establish analogous inverse relationships between peak firing rates and baseline frequency. This effect, which in our model primarily derives from the $c/\omega$ factor in the PRC, is clear in the bottom left and top right panels of Fig. 4; note that (15) implies $FL_{\text{max}} - FL_{\text{base}} \sim 1/\omega$.

**Effects of noise and distributed frequencies**

We now restore two effects thus far neglected: noise and heterogeneous oscillator frequencies. In the limit of small r.m.s. noise strength $\sigma$, stochastic
averaging [Freidlin and Wentzell, 1998] may be applied to (4), replacing
the $\theta$-dependent PRC by
\[
\bar{\varepsilon} = \left( \frac{1}{2\pi} \int_0^{2\pi} z(\theta)^2 d\theta \right)^{1/2};
\]
the Itô correction term vanishes under this average. Eqn. (5) may then be Fourier transformed for
each $\omega$ (cf. [Tass, 1999]), and solved for Fourier coefficients with ‘initial’
values $a_n(t_2, \omega)$ representing the state at stimulus end, to yield:
\[
FL(t) = \omega \sum_{n=-\infty}^{\infty} a_n(t_2) \exp \left[ - \left( i\omega n + \frac{\sigma^2 \bar{\varepsilon}^2 n^2}{2} \right) (t - t_2) \right].
\]  
(16)
Additionally, as noted in 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:
\[
\langle FL(t) \rangle - \frac{\langle \omega \rangle}{2\pi} = \int_{-\infty}^{\infty} r(\omega) \omega \times \sum_{n=-\infty, n\neq 0}^{\infty} a_n(t_2, \omega) \exp \left[ - \left( i\omega n + \frac{\sigma^2 \bar{\varepsilon}^2 n^2}{2} \right) (t - t_2) \right] \, d\omega,
\]  
(17)
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 $\omega_m$ beyond the essential support of the integrand, breaking the integral into pieces, and repeatedly applying the triangle inequality, (17) yields
\[
\left| \langle FL(t) \rangle - \frac{\langle \omega \rangle}{2\pi} \right| \leq \exp \left[ -\frac{\sigma^2 \bar{\varepsilon}^2 \omega_m^2 (t - t_2)}{2} \right] \times \int_{-\infty}^{\infty} r(\omega) \omega \sum_{n=-\infty, n\neq 0}^{\infty} a_n(t_2, \omega) \exp \left[ -i\omega n(t - t_2) \right] \, d\omega
\]  
(18)
up to an arbitrarily small error depending on $\omega_m$. For each $n$, the integral in (18) 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, (18) 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,
\[
r(\omega) \omega a_n(t_2, \omega) \approx \frac{1}{\sqrt{2\pi} \gamma} \exp \left( -\frac{(\omega - \langle \omega \rangle)^2}{2\gamma} \right) \langle \omega \rangle a_n(t_2, \langle \omega \rangle),
\]  
(19)
and the integral in (18) may be evaluated to give an upper bound on decay rate:

$$\left|\frac{\langle FL(t) \rangle - \langle \omega \rangle}{2\pi}\right| \leq \exp\left[ -\frac{\sigma^2 \omega_m^2 (t - t_2)}{2} - \frac{\gamma (t - t_2)^2}{2} \right] \left|\frac{\langle FL(t_2, \langle \omega \rangle) \rangle - \langle \omega \rangle}{2\pi}\right|.$$  

(20)

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 (19) holds is found to be appropriate, and Fig. 8, below, illustrates that (20) provides a good decay estimate.

**Stimulus duration and depressed firing**

In the above section on spike probabilities in the absence of noise, we showed that if inputs due to stimuli are sufficiently short compared with the ($\omega$-dependent) response period:

$$d < P(\omega) = \frac{2\pi}{\sqrt{2cI + \omega^2}},$$  

(21)

then $\rho(\theta, t_2, \omega)$ necessarily exhibits a peak and a trough, so that successive episodes of enhanced and depressed spiking ensue. 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 effects; cf. Fig. 4.

Noise and frequency heterogeneity cause additional overall decay, cf. (20). 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 leads to our third finding: In systems with narrow frequency distributions, short inputs necessarily lead to intervals of depressed firing following enhanced spiking.

This effect is further magnified if we normalise to maintain fixed ‘synaptic charge’ $\bar{I} d = S$. Now $\bar{I} \propto 1/d$ and (15) (Fig. 5) shows that brief inputs are yet further enhanced over longer, more diffuse ones. In this case, eliminating $\bar{I}$ from (21) yields an explicit input duration for maximal effect:

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

(22)
Parameter fitting

To compare model predictions with data, we first determine appropriate frequency distributions \( r(\omega) \), characterized by mean \( \mu^\omega \) and variance \( \gamma^\omega \), and r.m.s. noise strength \( \sigma \), by seeking parameter values for which model realizations match both an empirical interspike interval (ISI) histogram and correlations between neighboring ISIs:

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

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 frequency \( \omega \), and 2) rapid input current fluctuations modeled through the \( \sigma z(\theta) dW(t) \) term in (3). Thus \( y_j = y^d_j + \eta_j \), where \( y^d_j \) are the noise-free drift values, and \( \eta_j \) causes additional variance \( (\gamma^\sigma)^2 \) due to rapid noise. If the drift is sufficiently slow then \( \mathbb{E}\{ (y^d_j - m)(y^d_{j+1} - m) \} \approx \mathbb{E}\{ (y^d_j - m)^2 \} \triangleq (\gamma^\omega)^2 \), and (23) becomes

\[
    r_1 \approx \frac{(\gamma^\omega)^2}{(\gamma^\sigma)^2 + (\gamma^\omega)^2} = 0.1, \tag{24}
\]

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

Eqn. (24) 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 \( \mu^\omega \) may be estimated directly from the ISI mean \( m \). Guided by this and by analytical expressions for, e.g., barrier hitting times relating \( \sigma \) and \( \gamma^\sigma \), 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 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. 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, Fig. 4]; see Fig. 7. Specifically, we suppose that in the phasic (tonic) mode, spike frequencies are drawn 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. We then require that a central subgroup of oscillators are largely synchronous (asynchronous). This yields $\alpha_s = 0.01, \alpha_c = 0.05$. Note that, unlike [Usher et al., 1999], we take the same coupling strengths for phasic and tonic modes, so that increased synchrony is due to the tighter distribution of phasic frequencies.

Finally, we specify appropriate inputs $I(t)$. For the target detection task we found that a square wave of height $I = 0.125 \mu A/cm^2$ 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 $I = 0.22 \mu A/cm^2$. Moreover, reaction time (RT) distributions have significantly greater standard deviation than for target detection: 114 and 241 msec for correct and incorrect respectively [Rajkowski et al.], compared to $\approx 34$ 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
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.

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].

Comparison of model and empirical PSTH data

Fig. 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 (σ ≠ 0), followed by averaging over the frequency distributions derived above; 2) via direct simulations of a set of \( N = 100 \) globally-coupled RH 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 (20)).

These results confirm that reduction to a single phase equation (4) and the probabilistic theory developed above provide good descriptions of the
Figure 8: Model PSTHs computed from solution of Eqn (5) (solid) and from Eqn. (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 RH neurons for multiple trials. Decay bound of (20) shown dotted. Top row: target detection task for poor performance/tonic mode (left) and good performance/phonetic mode (right); bottom row: Eriksen task for incorrect (left) and correct (right) responses.

coupled RH 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 (15), due to the high noise level that 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 $R_{mag}(\text{phasic})/R_{mag}(\text{tonic}) \approx 1.3$, compared to the value 3.4 of [Usher et al., 1999].) Recent additional analyses of target detection data revealed a similar small difference in $R_{mag}$ when measured as a function of baseline rate, as opposed to as a function of cognitive performance (good vs. poor) as in [Usher et al., 1999]. Hence, other mechanisms, beyond the frequency effects studied here, must also be
operative. For example, noise levels \( \sigma \) may be elevated in the tonic mode (in addition to mean current values \( I_0 \)). This would increase the relative magnitude of phasic mode responses as in [Hermann and Gerstner, 2001]. Additionally, 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 mode responses relative to those in the tonic mode. However, the small RT variance shows that this is a minor effect in target detection. Finally, coupling effects may play a role as in [Usher et al., 1999].

Our model reveals that the tonic/phasic frequency difference contributes to the variation between PSTDs for poor and good target identification performance, while in the Eriksen data, for which baseline frequencies are similar, PSTD 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.

**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 PSTDs associated with poor and good target detection performance. In the model presented here, while coupling clearly affects synchrony, the key factor influencing PSTDs averaged over many trials is the LC spike rate, governed by the baseline currents \( I_0 \). Possible explanations for decreased \( I_0 \) include reduced inputs from other neurons afferent to the LC as well as neuromodulatory effects. Additionally, since synaptic coupling is inhibitory, it transiently reduces net input currents, thus effectively decreasing \( I_0 \) if LC neurons are sufficiently decorrelated. Our baseline rate explanation differs from the electrotonic mechanism of [Usher et al., 1999]; also, in that paper the \( I_0 \) were set in the excitable range, so that noise and
other external inputs were necessary for spiking. From simulations of sub-
threshold networks of coupled RH neurons with noise-driven firing at 2 – 3
Hz., as well as solutions of the corresponding (coupled) phase density equa-
tions derived from the full ‘theta model’ [Ermentrout, 1996] (not reported
here), we found that, in the absence of explicitly imposed cellular refractory
periods as in [Usher et al., 1999], modeling empirical post-stimulus periods
of depressed activity requires strong collateral coupling among LC neurons.
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. In the in vivo LC the synchron-
izing effects identified here and in [Usher et al., 1999, Alvarez et al., 2002]
may all be relevant. Additional effects of stronger coupling terms, noise,
and subthreshold neurons may be important and are under investigation.

In recent related work [Usher and Davilaar, 2002, Gilzenrat et al., 2002]
abstracted models of LC population activity that modify gains in connec-
tionist 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 Davilaar, 2002,
Gilzenrat et al., 2002] as well as suggesting, in the coupled multi-unit phase
model of Eqn. (3), a middle ground between those abstractions and the com-
plexity of the integrate-and fire pool of [Usher et al., 1999] or the full RH
system of (1-2).

In addition to the three findings highlighted above, our model makes
two predictions about in vivo LC activity. First, neurons that project to
the LC and evoke responses should remain active longer following stimulus
in complex tasks such as the Eriksen paradigm than in simpler ones like
target detection. Secondly, any mechanism that decreases the LC spike rate
should increase response to exogenous inputs. Indeed, other studies have
reported modulators that produce similar effects on LC activity. For ex-
ample, in some cases direct application of the neuropeptide corticotropin
releasing factor (CRF) increases LC baseline activity and simultaneously
decreases responses to sensory stimuli [Moore and Bloom, 1979], while the
alpha2 adrenoceptor agonist clonidine (or ST-91) can decrease baseline ac-
tivity and increase response [Foote et al., 1983]. Many other examples of
such ‘modulatory’ effects of neurotransmitters or exogenous inputs exist for
neurons in other brain areas [Aston-Jones et al., 2001].

The present analysis therefore provides a simple explanation for how neurotransmitters can affect not only the activity of a target neuron but also its response to other inputs, via the same mechanism revealed by e.g. [Herrmann and Gerstner, 2001, Fetz and Gustaffson, 1983] in different models. The LC phasic and tonic modes exemplify this dual effect [Usher et al., 1999], which occurs in numerous other brain areas and neurons [Aston-Jones et al., 2001]. Several mechanisms for this neuromodulatory effect 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], or altered electrotonic coupling [Usher et al., 1999]. The results presented here add variation in baseline spike rate to these mechanisms. For example, the effects of gain modulation on signal/noise discrimination of [Servan-Schreiber et al., 1990] could be produced by changes in baseline rate, \( \omega \), of the relevant cells. (Finding 1 above shows that transient increases in response to inputs are inversely proportional to \( \omega \) and proportional to input strength \( I \); thus, gain increase in [Servan-Schreiber et al., 1990, Fig. 1] is analogous to decrease of \( \omega \).) Moreover, baseline rates are a characteristic of any regularly firing neuron. In addition to the Type I RH 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.

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]).

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.
Appendix: the Rose-Hindmarsh equations

Parameter values and function definitions of the RH model (1) are as follows:

\[ V_{Na} = 55 \text{ mV}, \; V_K = -72 \text{ mV}, \; v_L = -17 \text{ mV}, \; g_{Na} = 120 \text{ mS/cm}^2, \]
\[ g_K = 20 \text{ mS/cm}^2, \; g_L = 0.3 \text{ mS/cm}^2, \; g_A = 47.7 \text{ mS/cm}^2, \]
\[ C = 1 \mu\text{F/cm}^2, \; I_b^0 = 5 \mu\text{A/cm}^2, \; \gamma_b = 0.069 \text{ mV}^{-1}, \]
\[ T_b = 1 \text{ msec}, \; T_n = 0.52 \text{ msec}, \; B = 0.21 g_A/g_K. \]

\[ q_\infty(v) = n_\infty(v)^4 + B b_\infty(v), \; b_\infty(v) = (1/(1 + \exp(\gamma_b(v + 53.3))))^4, \]
\[ m_\infty(v) = \alpha_m(v)/(\alpha_m(v) + \beta_m(v)), \; n_\infty(v) = \alpha_n(v)/(\alpha_n(v) + \beta_n(v)), \]
\[ \tau_q(v) = (\tau_0(v) + \tau_n(v))/2, \; \tau_n(v) = T_n/(\alpha_n(v) + \beta_n(v)), \]
\[ \gamma_b(v) = T_b(1.24 + 2.678/(1 + \exp((v + 50)/16.027))), \]
\[ \alpha_n(v) = 0.01(v + 45.7)/(1 - \exp(-(v + 45.7)/10)), \]
\[ \alpha_m(v) = 0.1(v + 29.7)/(1 - \exp(-(v + 29.7)/10)), \]
\[ \beta_n(v) = 0.125\exp(-(v + 55.7)/80), \; \beta_m(v) = 4\exp(-(v + 54.7)/18). \]

References


