



*Center for the Study of  
Brain, Mind & Behavior*



# **The locus coeruleus, adaptive gain, and the optimization of simple decision tasks**

Eric T. Brown   Mark S. Gilzenrat   Jonathan D. Cohen

Center for the Study of Brain, Mind and Behavior  
Princeton University  
*and*  
Program in Applied and Computational Mathematics  
Department of Psychology

Princeton University  
Princeton, NJ 08544

Technical Report #04-02

## Abstract

We show that adaptive gain changes, by hypothesis mediated by the locus coeruleus (LC), can help optimize performance on simulated sensory discrimination tasks, even when no knowledge of stimulus timing is assumed. The metric of performance used here is the rate of correct responses (or ‘reward rate’) achieved by the simulated decision network. The primary model that we study has two layers: the first integrates sensory input directly and the second accumulates this filtered input (as well as noise from other brain areas) and translates it into motor responses via threshold crossing events. Gain transients occur after a physiologically-motivated delay following threshold crossings in the first layer – in this sense, gain schedules adapt according to accumulated sensory information.

We adopt a linearization and reduction of the two layer model that allows a clear understanding of parameter effects and which is used to obtain a simpler set of optimization problems without sacrificing the generality of their solutions. By comparing optimal model reward rates in the presence of simulated LC-mediated gain changes with the (separately) optimized reward rates in the absence of such gain changes, the extent to which these gain changes contribute to enhanced task performance is determined, all for a ‘standard parameter set’ derived from fits to experiments. The results indicate that a significant improvement in reward (12 – 24%) is attributable to the LC-mediated gain mechanism. Additionally, the statistical variations in the optimal model gain transients from trial-to-trial agree with trends reported in recent experimental studies involving direct recordings from the LC. This provides converging evidence for the hypothesis that the LC plays a part in optimizing the dynamics of simple decision tasks.

# 1 Introduction and background

Neurons of the noradrenergic nucleus locus coeruleus (LC) selectively fire following the presentation of salient stimuli in decision tasks [1, 2, 3]. The phasic release of norepinephrine associated with this LC response has diverse effects on widely distributed efferent brain areas. Among these effects is thought to be potentiation of cells to inputs [4]. Thus, the LC can provide transient facilitation in processing, time-locked to the presence of behaviorally salient information in motor or decision areas. This facilitation has been modelled as a change in the gain (i.e., sensitivity to inputs) of the neural populations that determine the timing and accuracy of simple decisions [5, 1, 6, 7].

Previous work shows that, under certain modelling assumptions, advance knowledge of the timing and rate with which stimuli will appear on each trial may be exploited to design LC-mediated gain signals that implement the statistically optimal strategy in simple alternative decision tasks [8]. However, the complementary situation, in which stimuli appear with no cue (spatial or temporal), requires an entirely new analysis. Here, we assess the extent to which the LC, via a simplified model of its effects on processing units, can optimize decision tasks in the absence of explicit knowledge about stimulus schedules. That is, we require that all LC-mediated gain transients be adaptively determined by the incoming signal on a trial-to-trial basis. In particular, we assume that the LC is driven by the accumulation of firing rates past critical levels in the neural integrators simulating the decision process. Activating the LC is then in turn simulated as a transient increase in gain in all units of the model. We then explore via computational simulation whether a system equipped with such an intra-trial “feedback” control of its own gain parameter exhibits enhanced task performance in comparison to a system that maintains a static, but separately optimized, level of gain.

Three previous modelling studies [1, 6, 7] have addressed the effects of adaptive LC-mediated dynamic gain on decision tasks, offering intriguing results and suggesting several important questions. Foremost among these is the question of optimality. The work of [1, 6, 7] showed that, by allowing greater transients in gain schedules, error rates in a simulated target detection task could be reduced while reaction times held relatively constant or also decreased. This characterizes improved cognitive performance rather than a simple shift in the speed-accuracy tradeoff. However, the improved performance observed in these models may be at least partially due to a confound of specific values of decision thresholds (held fixed as gain schedules varied in [1, 6, 7]) or the specific gain schedules chosen for comparison. Thus, to establish that in fact there exist general performance gains due to the flexibility afforded by dynamic vs. constant gain schedules, it is necessary to compare compare the best-possible levels of performance across the whole class of dynamic vs. constant gain schedules and over the whole range of free parameters. This approach is taken below.

Although an optimized decision model equipped with LC-mediated dynamic gain trajectories clearly will perform no worse than an equivalent optimized model for which gain values are restricted to be constant, it is not obvious that it will perform

significantly better. One reason for this is that realistic gain schedules implemented in this adaptive fashion are subject to latencies: experiments by Waterhouse and collaborators have shown that increases in the sensitivity of cortical neurons to sensory inputs follow LC activation by more than 100 ms and reach a plateau at maximum effect 200-300 ms following this activation [4]. Here we model this via a delay of  $\approx 150$  ms; see below for details. Coupled with the latency from stimulus onset to triggering of the LC, gain changes therefore take effect significantly after stimuli are presented, which may minimize the benefits of these changes. Additionally, we consider only piecewise-constant gain trajectories here, which require only relatively simple threshold-based mechanisms to implement. Further, as with any adaptive gain scheduler but with particular relevance to the class of ‘jumped’ gain and stimulus trajectories considered here, gain transients triggered at the wrong moments can lead to erroneous responses, possibly obviating benefits derived on correct trials. This points to an overarching ‘circularity’ problem that is revisited in the Discussion of Section 7: due to the undesirable effects of prematurely changing gain, will so much information about stimulus presence be needed in order to beneficially increase gain that the stimulus could be reliably identified (as alternative 1 or alternative 2) before gain changes typically take effect?

Another unresolved question surrounding transient gain schedules and simple cognitive tasks surrounds the design of these tasks. In [1, 6, 7], a target detection task was studied, in which only one of two possible stimuli (the target) demands a behavioral response. For models in which transient gain changes are driven by partial evidence of the behaviorally relevant cues, this target detection task (compared with the two-alternative choice task studied here, in which both stimuli require responses) requires fewer thresholds for the triggering of gain transients and hence presents fewer opportunities for system gain to be turned up too early and evoke an erroneous response. The results below address whether performance benefits are still seen in two-alternative choice tasks.

Additionally, the role of the neural architecture with which decision tasks are solved remains to be investigated. Schall ([9] and references therein), and Reddi [10] have presented neural evidence that at least two stages of neural processing contribute significantly to variability in behavioral response times. Motivated by these studies, by the structure of the decision models of [1, 6, 7], and by the intuitive conclusion that a flexible system mapping sensory inputs to varied motor responses (cf. [9]) must involve separate centers of sensory integration and motor processing, the present paper primarily considers two-layered decision networks. However, we also consider the effects of adaptive gain schedules on a one layer system for comparison.

A final open question is suggested by the experiments and analysis of [3]. Figure 1 reproduces some of this data, which illustrates that transients in LC firing rate for the two-alternative choice task are more tightly correlated with behavioral responses than stimulus presentations. At optimal performance in our task model, will transients in gain trajectories (and hence transients in LC firing rates) be more tightly locked to stimulus presentations or (simulated) behavioral responses? To the extent to which the models capture the dynamics of neural decision processes, an answer in favor of locking to responses would support the hypothesis that the LC

serves to optimize decision tasks.

The modelling reported in this paper finds the following answers to the questions just introduced. For two layer decision models, LC mediated gain transients yield moderate performance enhancements relative to their counterparts without the capacity to dynamically adjust gain. Furthermore, optimized gain transients are more strongly correlated with responses, rather than stimulus onsets. A two layered architecture is required for this result: in single layer systems, adaptive adjustments gain cannot enhance reward rate beyond the best values achievable with optimized constant gain values.

The balance of the paper develops these results as follows. The following Section 2.1 defines the linearized decision model used here. Section 2.2 then introduces reward rate, the figure of merit for the different decision models we study, and discusses the effect of ‘premature’ responses on this quantity. Thus equipped, in Section 2.3 we describe the method, central to the present analysis, of comparing optimal task performance of a model with LC-mediated adaptive gain against optimal task performance of a model restricted to have constant gain. Next, in Section 2.4, we simplify the model by eliminating redundant parameters. Having thus refined our problem, in the following Section 3 we present the results of numerical reward rate optimization. Next, Section 5 considers two modifications to the model defined in Section 2.1 which are motivated by the underlying neurobiology, and Section 4 investigates how the number of layers in a decision model affects the performance that it can achieve when optimized. We discuss the results in Section 7.

## 2 The RR optimization problem

### 2.1 The two layer decision model

Our decision model, as shown in Figure 2, consists of two layers, each containing two mutually inhibitory neural subpopulations whose spike rates evolve under a ‘firing rate model’ of the general form derived in [11], cf. [12, 13, 14]. The first ‘decision’ layer receives sensory inputs and noise:

$$\tau dy_1 = [-y_1 + f_{\tilde{g}_y(t)}(-\beta y_2 + ka_1(t))] dt + \tilde{g}_y(t) \frac{kc\sqrt{\tau}}{\sqrt{2}} dW_t^{1'}, \quad (1)$$

$$\tau dy_2 = [-y_2 + f_{\tilde{g}_y(t)}(-\beta y_1 + ka_2(t))] dt + \tilde{g}_y(t) \frac{kc\sqrt{\tau}}{\sqrt{2}} dW_t^{2'}. \quad (2)$$

Here, the  $y_j$  are the firing rates of population  $j$  and other terms are as above, and the time scale  $\tau$  determines the rate at which these rates approach their equilibrium values or respond to transient inputs. The parameter  $\beta$  sets the strength of mutual inhibition between the populations, where  $f_{g(t)}(\cdot)$  is the sigmoidal ‘activation’ (or ‘frequency-current’ or neural ‘input-output’) function to be described shortly. Sensory inputs are represented by  $ka_j(t)$  and noise by  $\frac{kc\sqrt{\tau}}{\sqrt{2}} dW_t^j$ ; here  $k$  is a parameter scaling the strengths of these quantities, and the  $dW_t^j$  are increments of standard Wiener processes. We assume that the strength of firing rate fluctuations in response

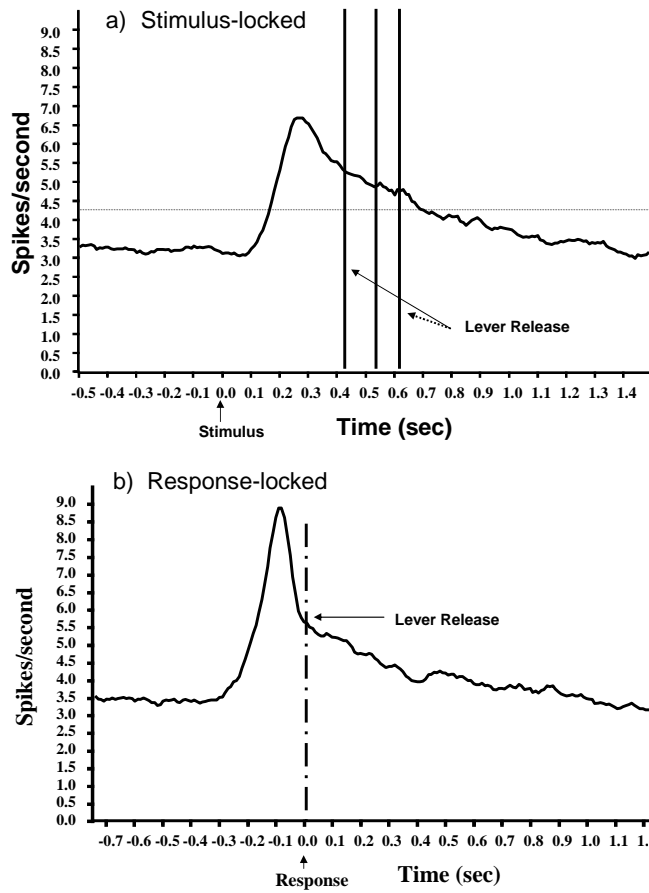


Figure 1: Peri-event time histograms (PETHs) of LC firing rates compiled across many experimental trials of the Eriksen flanker sensory discrimination task. Note that the observed post-stimulus transient increase in LC discharge is more tightly correlated with behavioral responses (b) than stimulus presentations (a). Data kindly provided by the authors of [3].

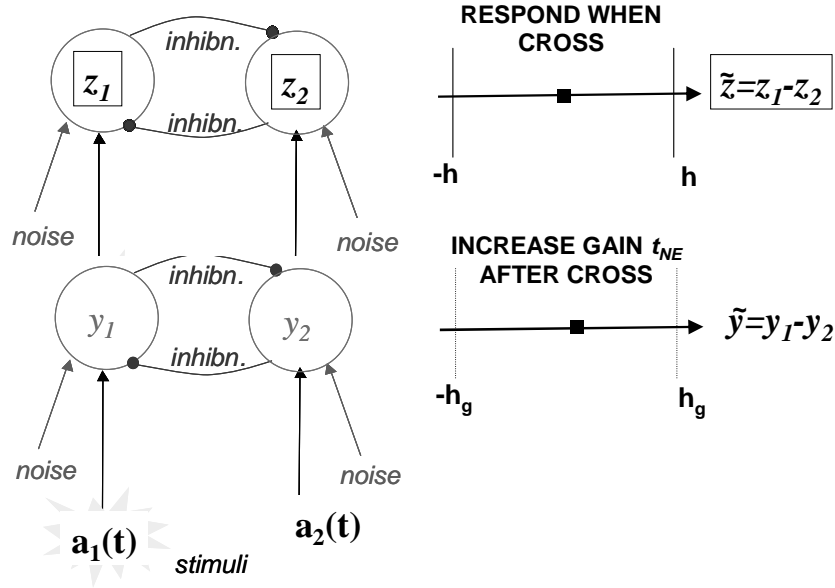


Figure 2: The architecture of the decision model, which consists of two layers of mutually inhibitory neural populations. The schematic to the right indicates the flow of stimulus information through the network, and the fact that we collapse each layer to a single variable ( $y$  or  $z$ ) characterizing the difference between firing rates in the layer. When the difference  $|y|$  exceeds threshold value  $h_g$ , gain levels are adjusted (following the delay  $\tau_{NE}$ ); when  $|z|$  exceeds  $h$ , task responses are made.

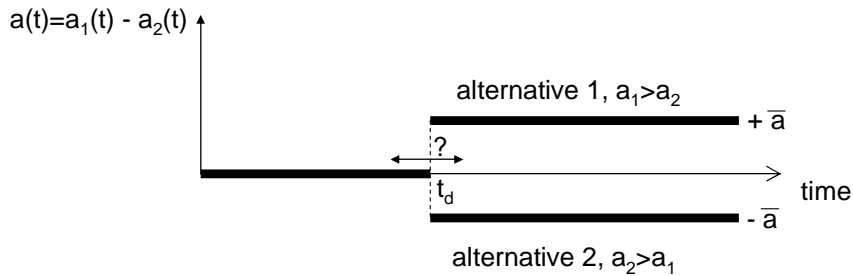


Figure 3: The type of stimuli considered here. Saliency (i.e., the difference between strengths of inputs favoring one of the two alternatives) develops at a randomly distributed time  $t_d$ . The task of the decision maker is to determine which of the two alternative stimuli has occurred, balancing speed vs. accuracy in this discrimination in order to maximize reward rate (see text).

to noise in inputs scales with  $g(t)$  (i.e., with the maximal sensitivity of firing rates to the deterministic component of the input). We take initial conditions  $y_j(0) = 0$ .

The subscript in  $f_{g(t)}(\cdot)$  indicates dependence on the time-varying gain, or sensitivity,  $g(t)$  of the neural populations: gain sets the slope of the activation function. For example, the logistic function

$$f_{g(t)}(x) = \frac{1}{1 + \exp(-4g(t)(x - b))} = \frac{1}{2} [1 + \tanh(2g(t)(x - b))] \quad (3)$$

has maximal slope  $g(t)$ . While this specific form is not required for the results derived below, we do assume that  $f_{g(t)}$  takes its time-dependent maximal slope  $g(t)$  at some time-independent point, as for (3).

The second ‘response’ layer receives inputs from the first (weighted by  $w$ ) as well as independent noisy inputs (presumably from afferents not related to the decision task being modeled):

$$\tau dz_1 = [-z_1 + f_{\tilde{g}_y(t)}(-\beta z_2 + w y_1(t))] dt + \tilde{g}_z(t) \frac{kc\sqrt{\tau}}{\sqrt{2}} dW_t^{1''}, \quad (4)$$

$$\tau dz_2 = [-z_2 + f_{\tilde{g}_y(t)}(-\beta z_1 + w y_2(t))] dt + \tilde{g}_z(t) \frac{kc\sqrt{\tau}}{\sqrt{2}} dW_t^{2''} \quad (5)$$

(the final terms in (1)-(5) are independent Wiener increments, and likewise below).

We next define the firing rate differences  $\tilde{z} = z_1 - z_2$ ,  $\tilde{y} = y_1 - y_2$  as well as  $a = a_1 - a_2$  and study the (linearized) model involving only their differences:

$$\tau d\tilde{z} = [-\tilde{z} + \tilde{g}_z \tilde{\beta} \tilde{z} + \tilde{w} \tilde{g}_z \tilde{y}] dt + \tilde{g}_z kc\sqrt{\tau} dW_t^1 \quad (6)$$

$$\tau d\tilde{y} = [-\tilde{y} + \tilde{g}_y \tilde{\beta} \tilde{y} + \tilde{g}_y ka] dt + \tilde{g}_y kc\sqrt{\tau} dW_t^2. \quad (7)$$

The difference variables  $\tilde{y}$  and  $\tilde{z}$  alone are sufficient to fully characterize the dynamics of the original four-dimensional model if the probability density of solutions to (1-2) and (4-5) has mostly collapsed along a one-dimensional ‘decision manifold [15, 8].’ While the papers just cited demonstrate that this assumption holds for fairly broad sets of parameters (see also [16, 17]), below we will simply assume that this collapse has occurred. Thus, from this point forward we consider the model (6-7) in which decisions are determined entirely by the relative (i.e., subtracted) evidence accumulated between two possible alternatives, without explicit regard for the absolute (i.e., summed) evidence.

We consider time-dependent stimuli  $a(t)$  of the type illustrated in Figure 3:  $a(t) = 0$ ,  $t < t_d$  and  $a(t) = \pm \bar{a}$ ,  $t \geq t_d$ , where  $t_d$  is the randomly distributed time of stimulus onset (see below).

Furthermore, we take initial conditions  $y(0) = z(0) = 0$  at the beginning of the preparatory phase, at which time we assume firing rates are reset (e.g. by inhibitory afferents originating in the prefrontal cortex). Additionally,  $\tilde{w}$  is the weight associated with inputs to the response layer from the decision layer.  $ka$  is the rate of incoming currents associated with sensory ‘evidence’ of task stimuli, and  $kc$  sets the strength of noise in afferent currents to the two populations;  $k > 0$  is an

arbitrary scale factor setting the magnitude of inputs to the neural populations.  $\tau$  sets the timescale of neural integration, and  $\beta$  sets the strength of inhibition between the pair of competing neural subpopulations at each layer. The reason for the tilde notation will become clear below.

We assume that the gain values  $\tilde{g}_z, \tilde{g}_y$  undergo transient changes in both layers at a (physiologically determined) delay  $t_{NE}$  following the first passage time  $T_y = \inf\{t : |y(t)| > \tilde{h}_g\}$  of  $\tilde{y}$  from the interval  $[-\tilde{h}_g, \tilde{h}_g]$ . Specifically, we assume that the gain values jump from values  $\tilde{g}_z^{pre}, \tilde{g}_y^{pre}$  to  $\tilde{g}_z^{post}, \tilde{g}_y^{post}$  at this time, and we refer to  $\tilde{h}_g$  as the ‘gain threshold.’ This mechanism for adjusting gain generalizes the approach of [1, 6, 7] from the target identification task (in which gain transients are driven only by firing rates in the population responsive to target stimuli) to the two-alternative forced choice task (in which gain transients may be driven by the population responsive to either of the two possible task stimuli). We assume that NE levels have decayed to their baseline levels  $\tilde{g}_z^{pre}, \tilde{g}_y^{pre}$  by the time each subsequent trial begins.

A response corresponding to alternative 2 (i.e., that “ $a < 0$ ”) is made at the passage time  $T_z = \inf\{t : |\tilde{z}(t)| > \tilde{h}\}$  if  $\tilde{z}$  first exits the interval  $[-\tilde{h}, \tilde{h}]$  by crossing the barrier  $-\tilde{h}$  (i.e.,  $\tilde{z}(T_z) = -\tilde{h}$ ), and vice-versa for alternative 1 ( $a > 0$ ) and the barrier at  $\tilde{h}$ ; we call  $\tilde{h}$  the response threshold. To determine optimal performance of the models under different architectures and assumptions about the time dependence of gain, we shall allow gains  $\tilde{g}_y$  and  $\tilde{g}_z$  and thresholds  $\tilde{h}$  and  $\tilde{h}_g$  to vary freely, as specified in greater detail below.

## 2.2 Task setup and reward rate

We model a task in which the objective is to correctly identify which of two alternative stimuli have been presented on each trial so as to maximize the reward rate ( $RR$ ) [18, 15], or the rate at which correct responses are made. This is a natural metric of task performance, as it also determines the expected number of correct (i.e., rewarded) responses made in any fixed interval. Trials occur in a long sequence, with randomized delays between each behavioral response and the presentation of the subsequent stimulus. We assume that these stimulus presentations are accompanied by a step in signal coherence  $a$  entering the first layer (cf. Figure 3) but with variable onset times  $t_d$ . We assume that no other changes in signal coherence  $a$  occur. To make contact with typical behavioral experiments, we take  $t_d$  to be uniformly distributed between 1 and 3 seconds for all of the results reported below.

In defining  $RR$ , we must specify how ‘premature’ responses made in the period  $[0, t_d]$  (that is, between the previous response and the presentation of the current sensory cue) are scored as correct vs. incorrect. Furthermore, we must determine what effect, if any, these responses will have on the delay before presentation of the next sensory cue. We adopt the following protocol, based on task designs in common use (including in [3]): all premature responses are counted as errors, and the next trials are started immediately. In this case, we relate the reward rate  $RR$  to random

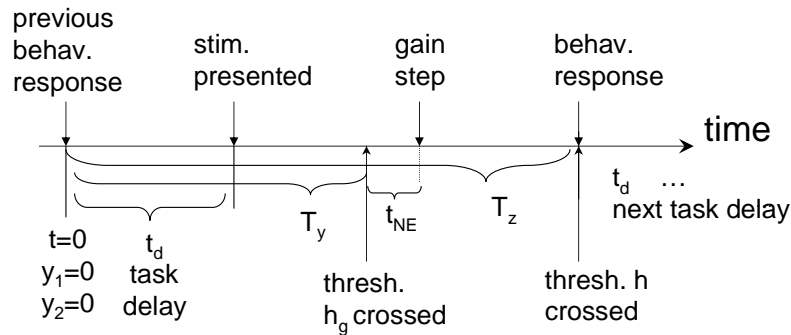


Figure 4: Schematic of the timing of the model decision task. Following a behavioral response on the ‘previous’ trial, model time  $t$  is reset to  $t = 0$ , initial conditions  $y_1(0) = y_2(0) = 0$  applied, and firing rates begin evolving under Eqns. (6)-(7). The sensory stimulus is presented following a (randomly distributed) delay  $t_d$ . At a delay  $t_{NE}$  following the first passage time  $T_y$  of the firing rate difference  $y$  across thresholds  $\pm h_g$ , model gains undergo a step change. Finally, the behavioral response on the ‘current’ trial is made at time  $T_z$ . Although the ordering of  $t_d$ ,  $T_y$ , and  $T_z$  displayed here is typical, it is not enforced: for example, on some trials gain threshold crossings at  $T_y$  could occur before stimulus presentations at  $t_d$ .

variables characterizing the decision process as follows:

$$RR = \frac{\mathbb{E} C(\tilde{z}(T_z))}{\mathbb{E} T_z}, \quad (8)$$

where the function  $C(\tilde{z}(T_z))$  measures the correctness of the firing rate difference  $\tilde{z}(T_z)$  at the response time  $T_z$  by taking value 1 (resp. 0) if, on a trial,  $a > 0$ ,  $\tilde{z}(T_z) = \tilde{h}$ , and  $T_z \geq t_d$  (resp  $a < 0$ ,  $\tilde{z}(T_z) = -\tilde{h}$ , and  $T_z \geq t_d$ );  $\mathbb{E}$  stands for expected value. See Fig. 4 for a schematic of the various times and delays that contribute to the  $RR$ . Thus defined,  $RR$  is the expected fraction of correct responses divided by the expected time elapsed for each response.  $RR$  is therefore the rate at which correct responses are made and has the units 1/time. We note that  $RR$  cannot be written as an expectation of a single function of firing rates at the hitting time  $T_z$  (unlike other objective functions typically used in sequential analysis, such as Bayes risk [19]), but rather takes quotient form. The  $RR$  measure adopted here is, however, directly related to the cognitive tasks we consider, in which participants are typically instructed to make as many correct responses as possible in a fixed interval.

Finally, we note that simulations were also performed with ranges of  $t_d$  other than uniform distribution from 1 to 3 seconds described above (including fixing  $t_d$  to take only a single value, nonetheless assumed unknown to the decision maker) and other protocols for treating premature responses (such as simply ignoring any threshold crossing prior to stimulus presentation time  $t_d$  at which responses are immediately

recorded if the firing rate difference  $|z|$  remains above threshold). In our separate modelling of these cases, the qualitative relationship between optimal performance of the constant vs. adaptive gain models that we next describe is preserved.

### 2.3 Optimizing decision models under adaptive vs. constant gain schedules

We now state the optimization problem of tuning model parameters to achieve maximal reward rate using LC-mediated adaptive gain:

$$RR_d = \max RR(\tilde{g}_y^{pre}, \tilde{g}_z^{pre}, \tilde{g}_y^{post}, \tilde{g}_z^{post}, \tilde{h}, \tilde{h}_g) \text{ under Eqns. (6)-(7)} . \quad (9)$$

The subscript  $d$  indicates that gain is ‘dynamic:’ allowed to adaptively change as described above.

Next, we determine the extent to which LC-mediated adaptive gain contributes to reward rates found by solving (9). As mentioned in the Introduction, we accomplish this by simply asking what the best performance would be in the absence of such an adaptive gain mechanism (but allowing freedom in choosing all other parameters). That is, we determine the optimal reward rate  $RR_c$  for constant gain schedules by restricting  $\Delta g = 0$ :

$$RR_c = \max RR(g_y^{pre}, g_z^{pre}, h, h_g) \text{ under Eqns. (6)-(7) with } \Delta g = 0 . \quad (10)$$

In Sections 3-4 below, we directly compare (refined versions of)  $RR_d$  and  $RR_c$  under various assumptions on task protocols and model architecture.

### 2.4 Eliminating redundant parameters

We now eliminate redundant parameters in the model to clarify the assumptions implicit in the model and hence the generality of the results that derive from optimizing it. First we show that the scale factor  $k$  (which sets the absolute magnitude of stimulus and noise inputs but not their ratio) may be eliminated by rescaling thresholds  $\tilde{h}$  and  $\tilde{h}_g$ . Dividing (6)-(7) through by  $k$  and redefining  $\tilde{\tilde{z}} = \tilde{z}/k$  and  $\tilde{\tilde{y}} = \tilde{y}/k$  we obtain

$$\tau d\tilde{\tilde{z}} = \left[ -\tilde{\tilde{z}} + \tilde{g}_z \tilde{\beta} \tilde{\tilde{z}} + \tilde{w} \tilde{g}_z \tilde{\tilde{y}} \right] dt + \tilde{g}_z c \sqrt{\tau} dW_t^1 \quad (11)$$

$$\tau d\tilde{\tilde{y}} = \left[ -\tilde{\tilde{y}} + \tilde{g}_y \tilde{\beta} \tilde{\tilde{y}} + \tilde{g}_y a \right] dt + \tilde{g}_y c \sqrt{\tau} dW_t^2 . \quad (12)$$

After scaling thresholds  $\tilde{h}$  and  $\tilde{h}_g$  by the (positive) factor  $1/k$  (matching the scaling of variables  $\tilde{z}$ ,  $\tilde{y}$ ), we obtain exactly the same statistics for (11)-(12) as for the original system (6)-(7); furthermore, since thresholds are free parameters in the optimization problem (9), we have lost no generality in eliminating the parameter  $k$ , and see that only the signal to noise ratio  $a/c$  affects optimal reward rates.

The rate  $a$  has units of  $1/t$ , and  $c$  has units of  $1/\sqrt{t}$ . We choose units of time to be seconds, and fix the ratio  $(\bar{a}/c)^2 = 8 \text{ sec.}^{-1}$ , a value derived from maximum

likelihood fits of reaction time distributions from two-alternative choice task experiments in [15]. Furthermore, we fix  $\tau = 1$ , noting that the results we derive from the model with  $\tau = 1$  correspond to results from  $\tau \neq 1$  and a suitably rescaled signal to noise ratio (following division of (11)-(12) by  $\tau$  and another rescaling of the firing rate variables).

Additionally making the definitions  $g_y = \tilde{\beta}\tilde{g}_y$ ,  $g_z = \tilde{\beta}\tilde{g}_z$ ,  $y = \tilde{\beta}\tilde{y}$ ,  $z = \tilde{\beta}\tilde{z}/w$ , and  $w = \tilde{w}/\tilde{\beta}$ , (11)-(12) become

$$dz = [-z + g_z z + g_z y] dt + g_z \frac{c}{w} dW_t^1 \quad (13)$$

$$dy = [-y + g_y y + g_y a] dt + g_y c dW_t^2 . \quad (14)$$

Recall that the variables  $\tilde{g}_y$  and  $\tilde{g}_z$  can take arbitrary (positive) values in the optimization scheme, so that rescaling them by the constant  $\tilde{\beta}$  to form  $g_z, g_y$  has no effect on the optimization problem. Furthermore, since the thresholds  $\tilde{h}$ ,  $\tilde{h}_g$  are also free variables, the rescaling of variables  $y$ ,  $z$  is again without consequence for the optimization problem: the rescaled system (13)-(14) with free parameters  $g_z$ ,  $g_y$  as well as thresholds  $h_g$  and  $h$  produces the same optimal error rates and reaction times (and hence reward rates) as the original system (6)-(7) with free parameters  $\tilde{g}_z$ ,  $\tilde{g}_y$ ,  $\tilde{h}_g$  and  $\tilde{h}$ .

The only parameter in (13)-(14) that remains to be defined is the weight  $w$ , which we set as follows. Recall that the variable  $z$  represents firing rates in the ‘response’ cortical populations which drive the motor neurons enacting task decisions. In any flexible neural architecture these cortical populations receive inputs not only from the decision layer populations responsive in the specific two-alternative task modelled here, but also from a diverse set of other neurons and brain areas. It is fluctuations in these inputs that are modelled by the input noise  $\frac{c}{w}dW_t^1$ . The magnitude of (mean values of)  $y$ , which is the ‘signal’ component of the input to the response layer, are determined by  $a$ : thus, the signal to noise ratio for this layer is  $\approx aw/c$ . For simplicity, we set  $w = 1$  so that signal to noise ratios in both layers are the same. Therefore, the final system of equations is

$$dz = [-z + g_z z + g_z y] dt + g_z c dW_t^1 \quad (15)$$

$$dy = [-y + g_y y + g_y a] dt + g_y c dW_t^2 , \quad (16)$$

and we study the optimization problems equivalent to (9), (10):

$$RR_d = \max RR(g_y^{pre}, g_z^{pre}, g_y^{post}, g_z^{post}, h, h_g) \text{ under Eqns. (15)-(16) } , \quad (17)$$

$$RR_c = \max RR(g_y^{pre}, g_z^{pre}, h, h_g) \text{ under Eqns. (15)-(16) with } \Delta g = 0 . \quad (18)$$

The extent to which norepinephrine adjusts the gain of a particular cortical population depends on the density and type of norepinephrine receptors in that population (as well as, e.g., connectivity within and afferent inputs to that population) [4]. However, for simplicity we assume here and below that the phasic LC impulse results in gain changes of the same magnitude in both the first and second

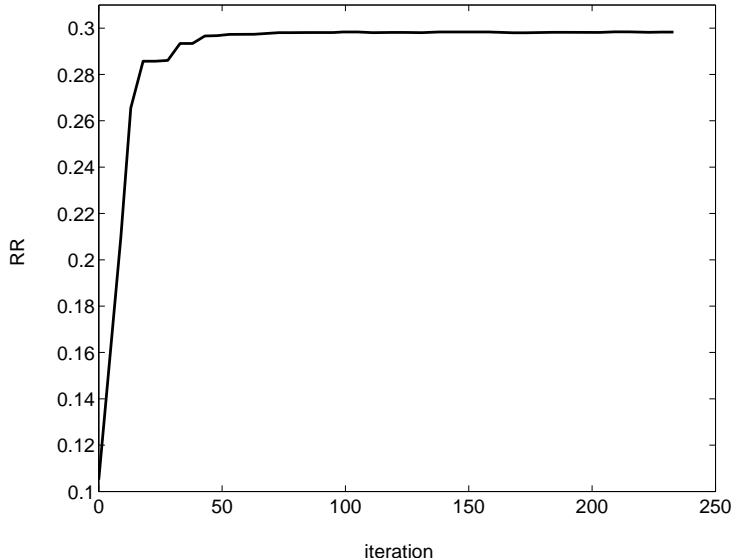


Figure 5: Typical convergence of the SUBPLEX algorithm on parameters yielding an (approximately) optimal RR.

layers. That is, a delay  $t_{NE}$  following the time  $T_y$  at which thresholds are crossed in the first layer,  $g_y^{pre} \rightarrow g_y^{pre} + \Delta g$  and  $g_z^{pre} \rightarrow g_z^{pre} + \Delta g$ . This reduces by one the number of free parameters in the optimization problem (but does not otherwise affect the generality of the results with regard to redundant parameters), giving the modified dynamic gain optimization problem:

$$RR_d = \max RR(g_y^{pre}, g_z^{pre}, \Delta g, h, h_g) \text{ under Eqns. (15)-(16)} . \quad (19)$$

Here,  $RR_d$  denotes the best possible reward rate that can be achieved with this adaptive gain scheme. Simulations indicate that the problems (19) and the slightly more general (17) give the same optimal RR's within  $\sim 2\%$ .

### 3 Numerical optimization

#### 3.1 Algorithm

We use the SUBPLEX optimization algorithm [20], a generalization of the Nelder-Mead Simplex method well suited for noisy objective functions, to solve the problem of optimizing RR under various conditions. The evaluation of RR (8) for at each set of parameters sampled by the algorithm is done by Monte-Carlo simulation of 200,000 ‘trials’ of the SDEs (15)-(16), each with simulated inter-trial interval  $t_d$  (i.e. ‘prepcycle’ duration) drawn from a uniform distribution between 1 and 3 seconds. The standard deviation in  $RR$  for a typical evaluation (at fixed parameter values) with this number of trials is  $0.0003 \text{ sec}^{-1}$ . Convergence of the SUBPLEX algorithm for a randomly chosen set of initial conditions is shown in Figure 5.

### 3.2 The standard parameter set

We now collect the values of model parameters that we hold fixed in the evaluation of  $RR$  via (15)-(16). These values, motivated in the text above, are:  $\bar{a} = 2 \text{ sec}^{-1}$ ,  $c = 1/\sqrt{2} \text{ sec}^{-1/2}$ , and delays  $t_d$  between behavioral responses and presentation of the next sensory cues are uniformly distributed between 1 and 3  $\text{sec}$ . All of these quantities may be derived from task design and behavioral data under certain assumptions (see above).

### 3.3 Range of reward rate values

It will be instructive to study how the reward rates achieved by the (linearized) two-layer model (15)-(16), with or without dynamic gain, compare with the reward rate achievable by an optimal decision maker. We begin by recalling that the optimal decision maker *equipped with trial-by-trial "advance" knowledge of stimulus onset time  $t_d$*  performs the sequential probability ratio test, which assumes knowledge of the time course of signal-to-noise ratios [21, 22, 18, 15]. From the results in [8], a one-layer linearized network can perform the SPRT if it is provided with perfect apriori knowledge of stimulus onset times and the capacity to instantaneously adjust gain to exploit this knowledge. Again drawing on [8], the optimal gain for the piecewise-constant stimuli considered here is:  $g(t) = 0, t < t_d, g(t) = 1, t \geq t_d$ . In this case error rate and reaction time depend on the model parameters as follows ([23, 15]):

$$ER = \frac{1}{1 + \exp(2h\bar{a}/c^2)} \quad (20)$$

$$RT = \frac{h}{\bar{a}} \tanh\left(\frac{h\bar{a}}{c^2}\right) . \quad (21)$$

As shown in [15], the unique optimal value of  $h$  optimizing

$$RR = \frac{1 - ER}{RT + \mathbb{E}(t_d)} \quad (22)$$

is given by the solution to

$$\exp(2h\bar{a}/c^2) - 1 = 2\beta(\mathbb{E}(t_d) - h/\bar{a}) . \quad (23)$$

Solving (23) using Newton's method (for the standard parameter set, that is, for  $\mathbb{E}(t_d) = 2, \bar{a} = 2, c = 1/\sqrt{2}$ ), we obtain  $h = 0.212$ . Using this to find  $ER$  and  $RT$  and then inserting these into (22) gives the value  $RR_{SPRT} = 0.440 \text{ sec}^{-1}$  achieved by the SPRT with optimal thresholds.

For the case in which the stimulus onset times are unknown, as assumed for the decision models introduced in Section 2, we are not aware of a single statistical test that is known to be optimal for a broad class of decision problems. However, the CUSUM algorithm (originally proposed by Page and analyzed extensively by Lorden) is known to have certain optimality properties in asymptotic limits of low error rates (i.e., large thresholds), and, in more general parameter regimes, compares favorably with other statistical tests in numerical studies (see Chapter 5 of

[24] and references therein). Therefore, we adopt the “two-sided” version of this algorithm (i.e., with separate thresholds for the detection of each of the two possible alternatives, Eqn. 2.2.24 of [24]) as the primary benchmark against which to evaluate our decision models. With stimulus detection thresholds numerically optimized (using SUBPLEX) to optimize reward rate for the standard parameter set, we find  $RR_{CUSUM} = 0.344 \text{ sec}^{-1}$ .

We note that the CUSUM algorithm does not have a simple implementation via gain transients and mutual inhibition in, for example, the reduced firing rate model, lending impetus to the question of how closely our neurally motivated models, with and without dynamic gain adjustments, can approach its performance. To accomplish this, we define a suitable range of  $RR$  to facilitate comparison with the range of possible decision makers. Motivated by the discussion above, we take  $RR_{CUSUM}$  to be a ‘ceiling’ value  $RR_{ceil}$  of this range. We next compute a complementary floor value for  $RR$ .

We let the floor value of  $RR_{floor}$  be achieved by chance guessing, assuming that a decision maker employing this strategy chooses hypothesis 0 or 1 without paying any attention to task stimuli, at times separated from the previous response by a uniform distribution between 0 and 3 seconds (i.e., up to the maximum possible value of  $t_d$ ). Since  $t_d > 1 \text{ sec.}$ , a third of these responses will occur before the stimulus has been presented, and hence are guaranteed to result in errors. The rest of the responses will be likewise anticipatory half of the time; when they follow the stimulus, they are correct at the chance rate of 50%. Thus,  $1 - ER = \frac{1}{3} * 0 + \frac{2}{3} * \frac{1}{2} * \frac{1}{2} = \frac{1}{6}$ . The mean elapsed time between responses using this strategy is  $\frac{3}{2} \text{ sec.}$  Thus the floor value produced is  $RR_{floor} = \frac{1}{6} * \frac{2}{3} \approx 0.111 \text{ sec}^{-1}$ .

In summary, the range of  $RR$  is  $[RR_{floor}, RR_{ceil}] = [.111, .344] \text{ sec}^{-1}$  for the standard parameter set; the ‘dynamic range’ is defined to be  $.344 - .111 = .233 \text{ sec}^{-1}$ .

### 3.4 Optimal reward rates for the ‘standard’ parameter set

The results of the optimization problems (19)- (10) for the standard parameter set are shown in Figure 6. The best reward for the dynamic gain problem (19) is  $RR_d=0.299 \text{ sec.}^{-1}$  while  $RR_c=0.267 \text{ sec.}^{-1}$  is achieved for the constant gain version (10). Thus, transient modulation of the decision process enabled by the LC allows the decision maker to enhance rewards by 12 percent (on average) during the interval over which the task is performed.

We now view these results in terms of the range of  $RR$  values derived above in Section 3.3. Letting  $RR = RR_{floor}$  be 0% through the range and  $RR = RR_{ceil}$  be 100% through the range, the optimized two-layer constant gain model achieves a  $RR$  67.0% of the way through the range; this improves to 80.7% with the addition of dynamic gain.

As Figure 6 indicates, the SUBPLEX algorithm converged to slightly different  $RR$  values each time it was run (with randomly sampled initial values for all free parameters). Figure 7 shows the values of two of the free parameters following optimization for the dynamic gain optimization problem (19). There is a family of

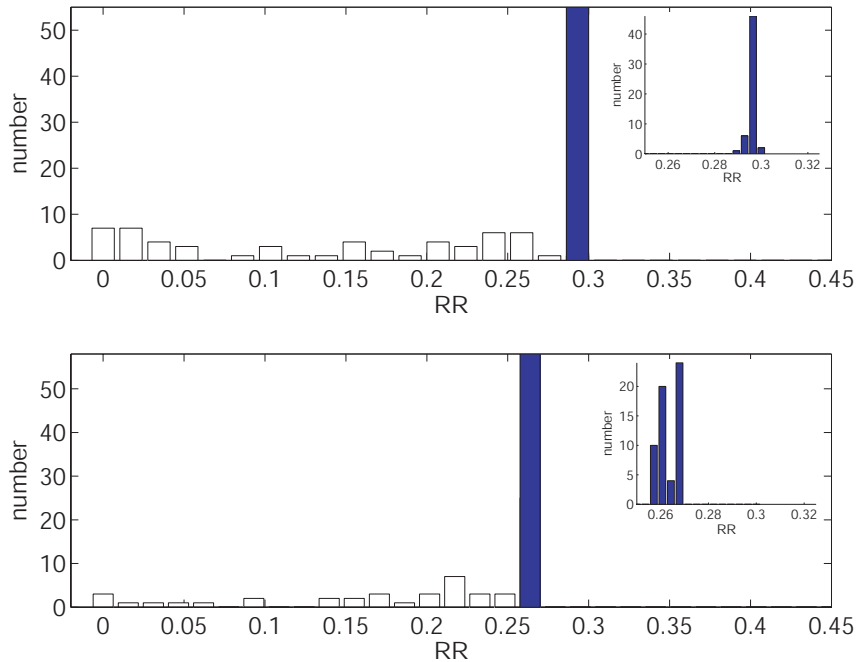


Figure 6: Filled bars–histogram of reward rates found using the SUBPLEX optimization algorithm; the vertical axis measures the number of runs of the SUBPLEX algorithm (each with a different, randomly chosen set of initial parameter values) that converged to particular  $RR$  values. (top) for the dynamic, adaptive gain optimization problem (17); (bottom) for the constant gain optimization problem (18). Insets give zoomed view with smaller histogram bin size. As reported in Section 3.4, the maximal reward rate obtained via SUBPLEX optimization for the adaptive gain case was  $RR_d=0.299 \text{ sec.}^{-1}$  (rightmost filled bar, top insert) and the maximal reward rate for the constant gain case was  $RR_c=0.267 \text{ sec.}^{-1}$  (rightmost filled bar, bottom insert). Outlined bars–for comparison, histogram of reward rates for randomly chosen, non-optimized parameter values. As expected, reward rates are much lower for these parameters.

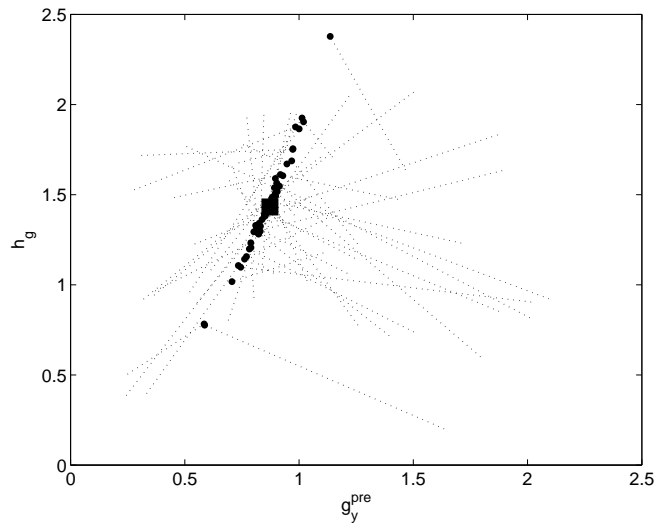


Figure 7: Values of two of the parameters to which the optimization algorithm converges when solving the dynamic gain optimization problem (19). Initial values of the parameters for each randomly initialized run of the algorithm lie at one end of the dotted lines, with final values indicated by the dots at the other end. The black square indicates the parameter values for the best-obtained  $RR$ .

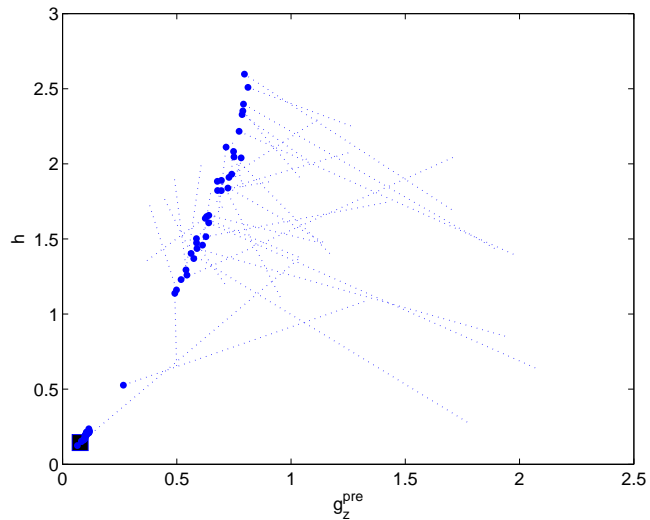


Figure 8: Values of two of the parameters to which the optimization algorithm converges when solving the constant gain optimization problem (10). Initial values of the parameters for each randomly initialized run of the algorithm lie at one end of the dotted lines, with final values indicated by the dots at the other end. The black square indicates the parameter values for the best-obtained  $RR$ .

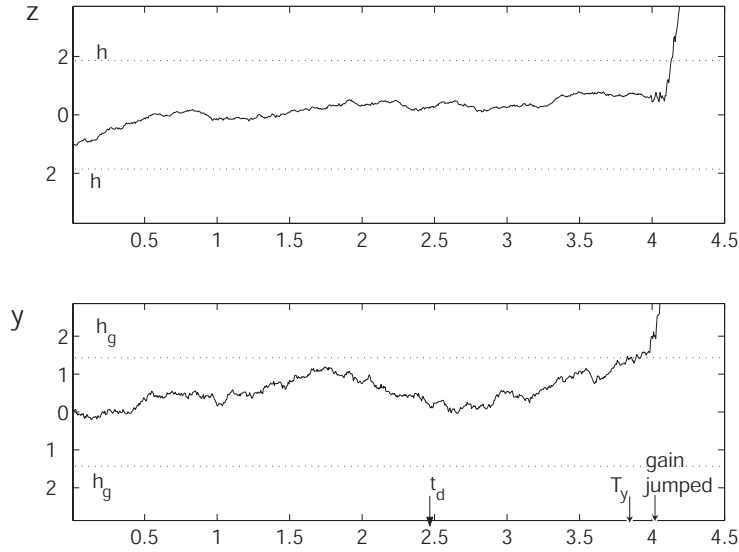


Figure 9: Typical trajectories for the optimized two-layer decision model (15)-(16) with dynamic gain.

different  $g_y^{pre}$  and  $h_g$  pairs (for example) that give similar  $RR$  values (the  $RR$ 's reported in this paper are the highest produced by any parameter values belonging to this family). Values of  $h_g$  depend monotonically on  $g_y^{pre}$  in this ‘nearly optimal’ regime, preserving the extent of evidence in favor of one or the other hypothesis necessary to trigger an increase in gain. Figure 8 demonstrates the analogous relationship between  $h_g$  and  $g_z^{pre}$  for the constant gain problem (10).

### 3.5 Predicted behavioral data for the standard parameter set

We now describe the behavioral statistics that characterize optimized performance for two layer models with dynamic gain. These statistics were produced for the parameter set giving the best  $RR$  among all parameter sets to which the SUBPLEX algorithm converged (corresponding to the black square in Figure 7):  $g_y^{pre} = 0.873$ ,  $g_z^{pre} = 0.474$ ,  $\Delta g = 3.33$ ,  $h_g = 1.43$ , and  $h = 1.86$ .

Typical trajectories for the firing rate equations (15) -(16) with these optimal parameters are shown in Fig. 9. Note that, due to the relatively large value of  $\Delta g$ , (15)-(16) become strongly unstable Ornstein-Uhlenbeck processes following the jump in gain (i.e., a delay  $t_{NE}$  after the thresholds  $\pm h_g$  are crossed). Figure 10 displays the reaction time distribution (stimulus-locked times  $T_z - t_d$ ) resulting from an ensemble of such trials, 16.8% of which resulted in premature responses and 2.0% of which were errors made following stimulus presentation. Recall that the decision model assumes no explicit or implicit cue indicating stimulus onset, leading to the high level of premature responses at optimal performance.

Finally, Figure 11 demonstrates that the times  $T_y$  at which the thresholds  $\pm h_g$

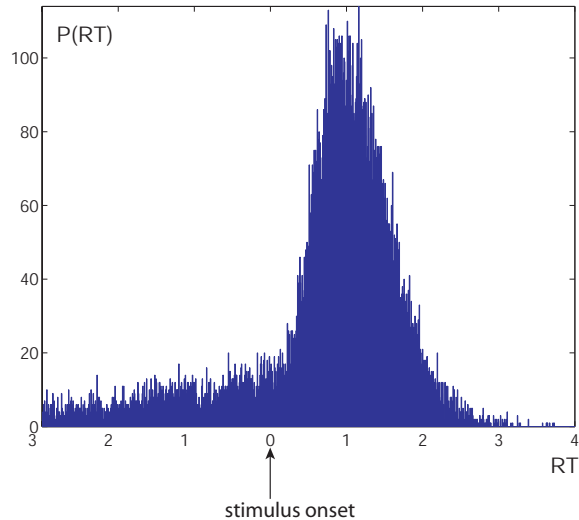


Figure 10: Distribution of reaction times for the optimized two-layer decision model with dynamic gain, relative to stimulus presentation; negative values correspond to premature responses.

were crossed are more tightly correlated with behavioral response than with stimulus onset times. According to the model of LC-mediated adaptive gain adopted here, the LC commences a period of increased firing at these threshold crossing times (which are followed  $t_{NE}$  later by the gain increase  $\Delta g$ ). Thus, the results of Figure 11 may be interpreted as histograms of predicted onset times for bursts of LC activity. These histograms, when locked to simulated behavioral response times  $T_z$  vs. stimulus onset times  $t_d$ , display the same qualitative narrow vs. broad trend as the experimental histograms of trial-averaged LC firing rates displayed in Fig. 1.

## 4 One layer decision models

In this section we investigate the role that the two-layer architecture of the model has in determining the processing benefits of LC-mediated transient gain over the ‘control’ hypothesis of constant gain. We collapse the two-layer model of Fig. 2 to an effectively single-layer decision model by assuming that both gain effects and behavioral responses are driven by threshold crossing in the first layer. In other words, gain changes continue to be triggered exactly as above, but responses are made at the passage time  $T = \inf\{t : |y(t)| > h\}$ . See Fig. 12.

The optimization problem for the one layer model is

$$RR_d^{1\text{ layer}} = \max RR(g_y^{pre}, \Delta g, h_g, h) \text{ under Eqn. (16) .} \quad (24)$$

for dynamic gain and

$$RR_c^{1\text{ layer}} = \max RR(g_y^{pre}, h) \text{ under Eqn. (16) .} \quad (25)$$

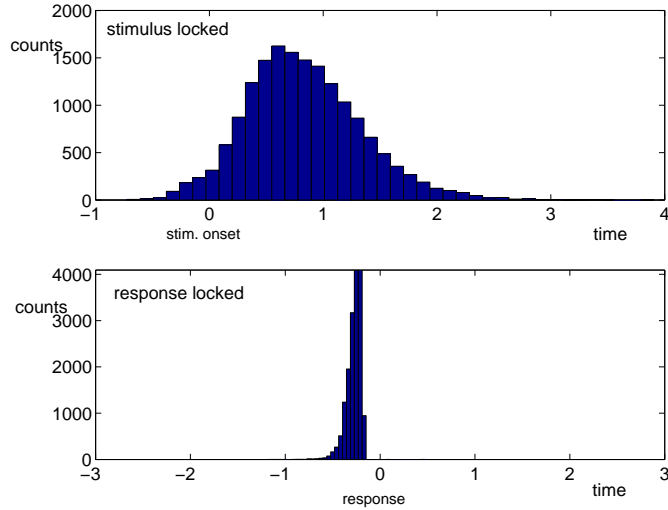


Figure 11: Histograms of latencies between first passage times  $T_y$  –times at which the LC fires its phasic burst in the model– and (top) stimulus onset times  $t_d$  vs. (bottom) response times  $T_z$ , all for the optimized two-layer decision model with dynamic gain.

for the constant gain case. Solving (24)-(25) via numerical optimization, we find that the best reward for the dynamic gain version (24) is  $RR_d=0.339 \text{ sec.}^{-1}$  while  $RR_c=0.337 \text{ sec.}^{-1}$  is achieved for the constant gain version (10): a negligible difference. Therefore the model of LC-mediated adaptive gain has no effect on performance for a one-layer decision maker. The reward values thus obtained by the one-layer model are 97-98% of the way through the  $RR$  range, a significant improvement over all two-layer models. We will revisit this fact in the Discussion of Section 7.

## 5 Two biologically motivated constraints to the standard parameter set

### 5.1 Fixed motor thresholds

Next we define a related optimization problem motivated by a plausible neurobiological constraint and suggested by Josh Gold (University of Pennsylvania [25]). We take a more conservative view of optimization in neural decision networks by assuming that the motor threshold  $h$  is not available for adjustment from task-to-task, but is rather fixed at some sufficiently high value so as to avoid being prematurely crossed in any of a variety of different tasks with a varying signal-to-noise ratios and stimulus magnitudes. Here, we fix  $h = 5$ , roughly 2-3 times typical gain values for nearly-optimized solutions for the problem (19) in which  $h$  was allowed to vary

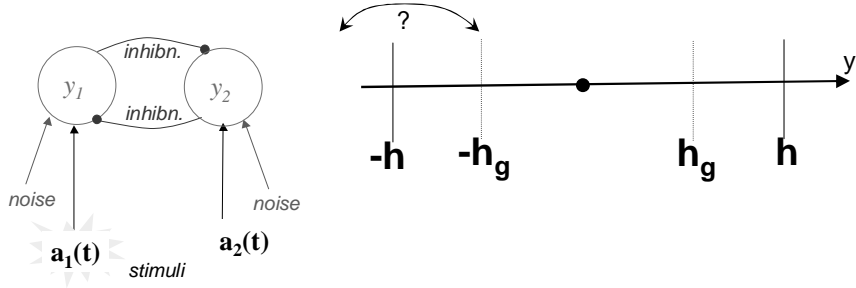


Figure 12: The architecture of the one-layer decision model. Thresholds for both gain adjustments and behavioral responses are collapsed to a single layer. The question mark indicates that the gain thresholds  $h_g$  are not a priori restricted to lie within the response thresholds  $h$ , and may take values outside these thresholds (hence ensuring that gain will be constant) for optimal performance.

freely. That is, we numerically solve the constrained optimization problem

$$RR_d = \max RR(g_y^{pre}, g_z^{pre}, \Delta g, h_g) \text{ under Eqs. (15)-(16) with } h = 5 . \quad (26)$$

for dynamic gain and

$$RR_c = \max RR(g_y^{pre}, g_z^{pre}, h_g) \text{ under Eqs. (15)-(16) with } \Delta g = 0 \text{ and } h = 5 . \quad (27)$$

for the constant gain case.

The results of the optimization problems (26)-(27) in this case are as follows: the best  $RR$  for the dynamic gain problem (26) is  $RR_d=0.299 \text{ sec.}^{-1}$  while  $RR_c=0.247 \text{ sec.}^{-1}$  is achieved for the constant gain version (27). Thus, transient modulation of the decision process enabled by the LC allows the decision maker to enhance expected rewards by 21 percent over any fixed task interval. Furthermore, note that the  $RR$  value  $0.299 \text{ sec.}^{-1}$  obtained for (26) matches that obtained for the problem (19), which allowed  $h$  to vary freely. The conclusion is that fixed motor thresholds have an insignificant effect on the reward rate that a system with dynamic gain can achieve, but are detrimental when gain is constant.

In terms of their position in the dynamic range  $[RR_{floor}, RR_{ceil}]$ , the optimized two-layer constant gain model achieves a  $RR$  level 58.4% of the way through the range; this improves to 80.7% with the addition of dynamic gain.

We also studied versions of the one-layer decision optimization problem (24)-(25) with fixed motor threshold  $h = 5$ , with the same threshold crossing effects as in Fig. 12. In this case the optimal performance for the adaptive-gain model is  $RR_d=0.306 \text{ sec.}^{-1}$ , while  $RR_c=0.281 \text{ sec.}^{-1}$  for constant gain, or positions of 83.7% and 73.0% through  $[RR_{floor}, RR_{ceil}]$ , respectively. Thus, when motor thresholds are fixed, adaptive gain scheduling yields a performance advantage for the one-layer model. This stands in contrast to the situation for one-layer decision models with freely optimized motor thresholds, in which adaptive gain did not improve

performance. Furthermore, unlike for the two-layer model, optimal RR values are reduced by fixing the motor threshold.

## 5.2 Stimulus-dependent noise

Next we allow for depressed noise in inputs to the first layer before the stimulus coherence emerges at time  $t_d$ , presuming that additional fluctuations in afferent inputs accompany the appearance of this coherence. Specifically, we set  $c = \frac{1}{4} * \frac{1}{\sqrt{2}}$  for  $t < t_d$  and  $c = \frac{1}{\sqrt{2}}$  for  $t \geq t_d$ . The  $\frac{1}{4}$  is an arbitrary value chosen to illustrate the present situation.

The optimal reward for the dynamic gain problem (19) in this case is  $RR_d=0.361 \text{ sec.}^{-1}$ . For comparison,  $RR_c=0.311 \text{ sec.}^{-1}$  was obtained achieved for the constant gain problem (10). Here, then, the LC enables reward rate improvements of 16 percent in absolute reward rate over any fixed task interval.

## 6 LC effects on decision performance persist in another model of neural integrators

The papers [26, 16] introduce the leaky integrator *connectionist* model of neural decision processes:

**layer 1:**

$$\tau_c dx_1 = [-x_1 - \beta f_{g(t)}(x_2) + a_1(t)] dt + \frac{c(t)}{\sqrt{2}} dW_t^1 \quad (28)$$

$$\tau_c dx_2 = [-x_2 - \beta f_{g(t)}(x_1) + a_2(t)] dt + \frac{c(t)}{\sqrt{2}} dW_t^2 \quad (29)$$

**layer 2:**

$$\tau_c dv_1 = [-v_1 - \beta f_{g(t)}(v_2) + a_1(t) + w f_{g(t)}(x_1)] dt + \frac{c(t)}{\sqrt{2}} dW_t^1 \quad (30)$$

$$\tau_c dv_2 = [-v_2 - \beta f_{g(t)}(v_1) + a_2(t) + w f_{g(t)}(x_2)] dt + \frac{c(t)}{\sqrt{2}} dW_t^2 \quad (31)$$

where the state variables  $x_j(t)$  and  $v_j(t)$  denote the mean input currents to cell bodies of the  $j$ th neural in the decision and motor layers, respectively. The integration implicit in the differential equations represents temporal summation of dendritic synaptic inputs ([27] and references therein). Population firing rates are given by  $f_{g(t)}(x_j(t))$ , where  $f_{g(t)}(\cdot)$  is as in Eqn. 3. The other terms are as for Eqns.(1-2) and (4-5). The main difference between the “connectionist” ((28-31)and “firing rate” ((1-2), (4-5)) models is whether all of the deterministic inputs, or just the activity of the opposing population, appear inside the function  $f_{g(t)}(\cdot)$ . See [8], cf. [27], for more on the relationship between these models.

Equations for the differences  $x \sim x_1 - x_2$  and  $v \sim v_1 - v_2$  may be derived, as in Section 2.1, via linearization of the activation functions  $f_{g(t)}(\cdot)$ . Following rescaling of  $x, v$ , and the thresholds  $h, h_g$  by  $1/k$ , exactly as in Section 2.4, as well as analogous

redefinitions of the free parameters  $g_x$  and  $g_v$ , these equations are

$$dx = [-x + g_x x + a] dt + cdW_t^2 \quad (32)$$

$$dv = [-v + g_v v + g_x x] dt + cdW_t^1 . \quad (33)$$

We now make an additional simplification. As noted above, study only differences in firing rates, assuming that decisions are determined by these differences alone. Following linearization,  $f_{g(t)}(v_1) - f_{g(t)}(v_2) = g(t)v$ , and it is this quantity that we will require exceeds thresholds  $\pm h$  at the instant that decisions are declared (and similarly for  $g(t)v$  and the triggering of gain variations LC via crossing of  $\pm h_g$ ). Thus, gain transients have two effects: (i) with respect to the variables  $(x, v)$ , the effective thresholds are adjusted and (ii) the linear terms in the equations (32-33) are themselves changed.

In order to assess the extent to which adaptive gain strategies can improve performance in the reduced connectionist model (32-33), we define reward rate optimization problems that allow us to analogous to the those of Section 2 for the firing rate model. First, as in (19), we numerically determine the best achievable reward rate utilizing adaptive gain strategies

$$RR_d^{conn} = \max RR(g_x^{pre}, g_v^{pre}, \Delta g, h, h_g) \text{ under Eqns. (32)-(33)} \quad (34)$$

and compare this with the maximum reward rate under constant gain:

$$RR_c^{conn} = \max RR(g_x^{pre}, g_v^{pre}, \Delta g, h, h_g) \text{ under Eqns. (32)-(33) with } \Delta g = 0 . \quad (35)$$

Using coordinate transformations and rescalings similar to those in 2.4, it may be shown that this constant gain reward rate optimization problem for the connectionist model is actually equivalent to its analogue 10 for the firing rate model.

Therefore, it is not surprising that numerical optimization with the SUBPLEX algorithm yields  $RR_c^{conn} = 0.267 \text{ sec.}^{-1}$ , exactly as for the firing rate model in Section 3.4. However, for the dynamic gain problem (34), the twin effects of  $\Delta g$  mentioned above for the connectionist model yield a slightly higher maximal reward rate than for the firing rate model:  $RR_d^{conn} = 0.316 \text{ sec.}^{-1}$ , an improvement of 18% in raw reward rate over the constant gain case, or of  $88.0\% - 67.0\% = 21\%$  in dynamic range of RR.

We also studied one layer versions of the reduced connectionist model: for dynamic gain,

$$RR_d^{conn, 1 \text{ layer}} = \max RR(g_x^{pre}, \Delta g, h_g, h) \text{ under Eqn. (32)} . \quad (36)$$

for dynamic gain and

$$RR_c^{conn, 1 \text{ layer}} = \max RR(g_x^{pre}, h) \text{ under Eqn. (32)} . \quad (37)$$

for constant gain. We find that  $RR_c^{conn, 1 \text{ layer}} = 0.337 \text{ sec.}^{-1}$ , exactly as for the analogous firing rate problem (37), as expected. Also, numerical optimization shows that  $RR_d^{conn, 1 \text{ layer}} = 0.337 \text{ sec.}^{-1}$ , indicating that, again as for the firing rate

model, adaptive gain does not yield a performance advantage for our one layer decision model.

An implication of the results described in this section is that the qualitative information processing benefits of the adaptive gain strategy are do not depend critically on the details of how gain enters into the reduced equations (15-16) studied in detail in the first sections of the present paper. Therefore, even though it is not yet completely clear what set of effective equations most accurately describes the dynamics of norepinephrine-modulated neural groups in decision tasks, one may expect that this modulation will lead to improved performance in rather general settings.

## 7 Discussion

This paper shows that adaptive gain changes, potentially mediated by the locus coeruleus (LC), can help optimize performance on simulated sensory discrimination tasks, even when no knowledge of stimulus timing is assumed. This type of task is complementary to that studied in [8], where the time course of stimuli was either assumed to be known explicitly or to be expressed implicitly via strong onset effects, which evoke stimulus-locked changes in the processing of inputs.

The primary model considered here has two layers: the first integrates sensory input directly and the second accumulates this filtered input (as well as noise from other brain areas) and translates it into motor responses via threshold crossing events. Gain transients occur after a physiologically motivated delay  $t_{NE}$  following threshold crossings in the first layer – in this sense, gain schedules adapt ‘online’ according to accumulated sensory information. The linearization that we adopt allows a clear understanding of parameter effects, which we use to obtain a simpler set of optimization problems without sacrificing the generality of their solutions. By comparing optimal model reward rates in the presence of simulated LC mediated gain changes with the (separately) optimized reward rates in the absence of such gain changes, we determine the extent to which these gain changes contribute to enhanced task performance, all for a ‘standard parameter set’ derived from fits to experiments. The results, summarized in the bottom rows of Figures 13-14, indicate that a modest but significant improvement in reward is attributable to the LC mediated gain mechanism. Additionally, the optimal gain transients are more tightly correlated with simulated responses than stimuli, in agreement with trends reported in recent experimental studies involving direct recordings from the LC [3] (see Figures 1 and 11). This provides converging evidence for the hypothesis of [1] that the LC plays a part in optimizing the dynamics of simple decision tasks.

The improvement in reward rate that LC-mediated gain transients lend to two-layer models does not carry over to one-layer models (Figure 12), for which the adaptive gain transients considered here do not always yield reward rates higher than those attainable with constant values of gain (top row of Figure 13). Intuitively, this is because the first layer directly integrates incoming sensory inputs, and once sufficient evidence has accumulated in it to determine the presence of a stimulus and

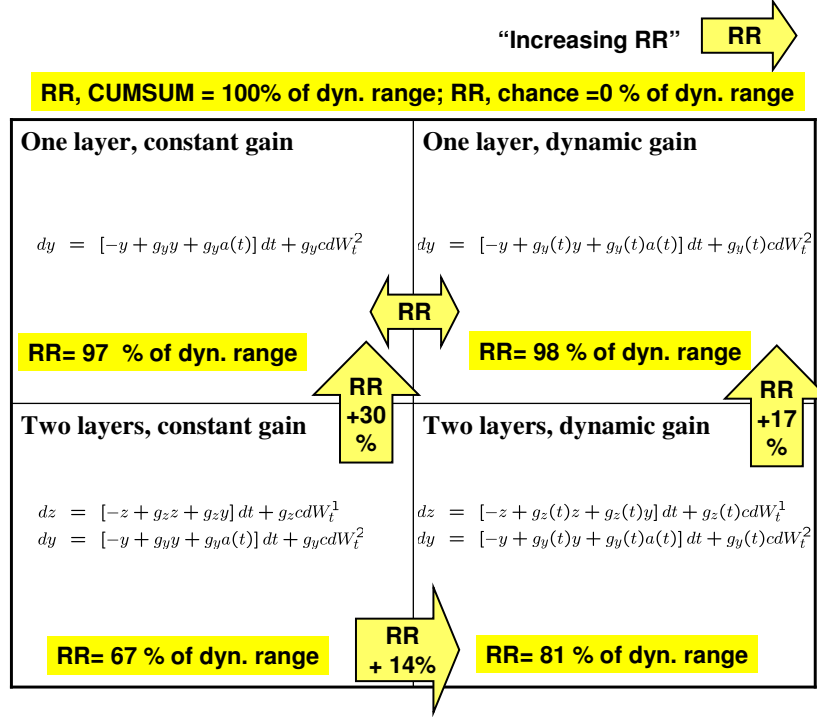


Figure 13: Summary of results for the reward rate ( $RR$ ) optimization problems with motor threshold  $h$  allowed to vary freely. For the two-layer models, the optimization problem with adaptive, dynamic gain (17) gives higher reward rate (bottom right) than its constant gain counterpart (18) (bottom left). However, the optimization problems for the one-layer model yield similar reward rates for both the constant gain ((25), top left) and variable gain ((24), top right) versions. Reward rates are presented as percentages within the  $RR$  ‘dynamic range’  $[RR_{floor}, RR_{ceil}]$  (Sect. 3.3). All results are for the standard parameter set of Sect. 3.2.

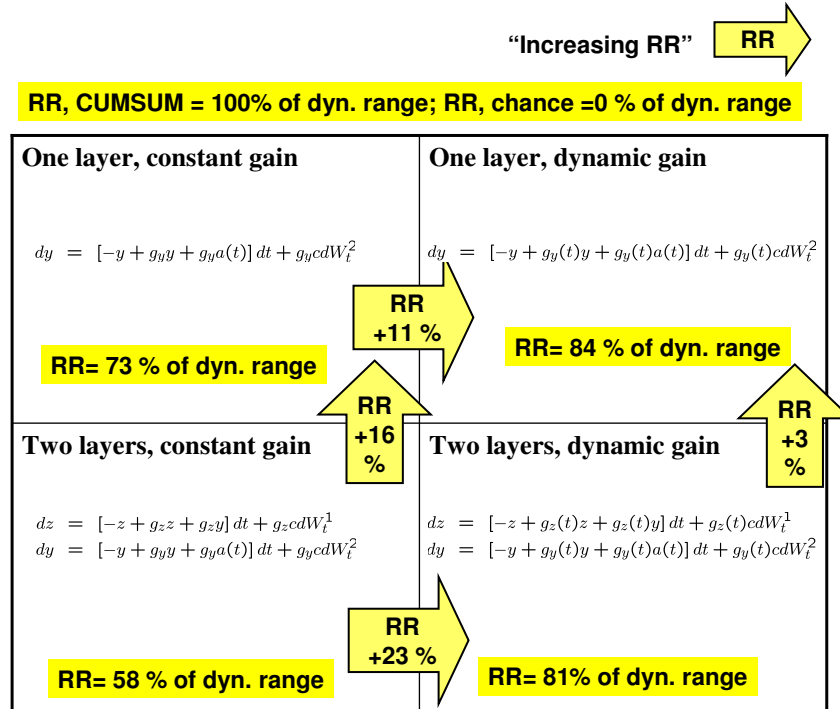


Figure 14: As for Figure 13, but with fixed motor threshold  $h = 5$ . In this case, dynamic, adaptive gain provides a significant increase in optimal  $RR$  for both one- and two-layer models. Further, the difference in  $RR$  for one- and two-layer models with dynamic gain is slight.  $RR$  reported as percentages within  $[RR_{floor}, RR_{ceil}]$ , as in Figure 13.

hence the advantage of a concurrent increase in gain, there is also sufficient evidence to identify which of the two possible stimuli was actually presented. Thus, optimal strategies directly produce a behavioral response at this point, instead of relying on a gain transient.

This strategy of allowing integrated sensory inputs to trigger responses directly is not available for two-layer decision models, in which a second level of (noisy) integration to threshold precedes responses. In this two-layer case, the firing rate trajectories of Figure 9 illustrate how transients in gain can streamline the transmission of sensory evidence through the network. The mechanism exploited here is to increase gain in the second layer to a sufficiently high level so as to cause, via positive ‘unstable’ feedback, a rapid readout of the evidence that has accumulated in the second layer. This gain increase occurs at a delay  $t_{NE}$  following crossing of optimally tuned thresholds  $\pm h_g$  in the first layer (note that evidence, typically in favor of the ‘correct’ alternative, continues to accumulate in *both* layers during this delay period).

Despite their optimal use of transient gain schedules, reward rates for two-layer models can still be significantly worse than those achievable by one-layer models (righthand column of Figure 13), which achieve an effective ‘short-circuit’ from sensory input to response. However, this is only the case if motor thresholds are allowed to vary freely. The righthand column of Figure 14 illustrates that, when thresholds  $\pm h$  for evoking motor responses are fixed, a two-layer model can accomplish a reward rate comparable to that of its one-layer counterpart. In other words, for models with fixed response thresholds, most of the performance apparently lost by adopting a (noisy) two- vs. one-layer architecture can be recovered via LC mediated adaptive gain schedules.

Finally, we reiterate that both of the decision models (16-15) and (32-33) studied here are described completely by differences in activities between the competing neural populations that accumulate evidence for the two alternatives. Our optimization results show that, even when the LC-mediated increases in gain are driven solely by these differences – and hence by the outcome of a discrimination process instead of *directly* by stimuli – there is a benefit to task performance over the null hypothesis of no dynamic LC modulation. Thus, our modelling suggests that recent experimental results ([3], summarized in Fig. 1) that activation of monkey LC neurons in forced choice tasks is “decision-related,” are consistent with a role of the LC in optimizing task performance.

**Acknowledgements** This work was partially funded by the grant NIH P50 MH62196 (Cognitive and Neural Mechanisms of Conflict and Control, Silvio M. Conte Center). E.B. and M.G. were supported under National Science Foundation Graduate Fellowships and E.B. under a Burroughs-Wellcome Training Grant in Biological Dynamics and by the Princeton Graduate School. The authors thank Josh Gold and Rafal Bogacz for useful contributions and discussions, as well as Ed Clayton and Gary Aston-Jones for providing the data of Figure 1 and for their insights into the role of the LC in modulating decisions. The authors are grateful to Philip Holmes for comments on the manuscript and ideas for its improvement.

## References

- [1] M. Usher, J.D. Cohen, D. Servan-Schreiber, J. Rajkowski, and G. Aston-Jones. The role of locus coeruleus in the regulation of cognitive performance. *Science*, 283:549–554, 1999.
- [2] G. Aston-Jones, J. Rajkowski, P. Kubiak, and T. Alexinsky. Locus coeruleus neurons in the monkey are selectively activated by attended stimuli in a vigilance task. *J. Neurosci.*, 14:4467–4480, 1994.
- [3] E. Clayton, J. Rajkowski, J.D. Cohen, and G. Aston-Jones. Decision-related activation of monkey locus coeruleus neurons in a forced choice task. Under review, 2004.
- [4] B. Waterhouse, H. Moises, and D. Woodward. Phasic activation of the locus coeruleus enhances responses of primary sensory cortical neurons to peripheral receptive field stimulation. *Brain Res.*, 790:33–44, 1998.
- [5] D. Servan-Schreiber, H. Printz, and J.D. Cohen. A network model of catecholamine effects: Gain, signal-to-noise ratio, and behavior. *Science*, 249:892–895, 1990.
- [6] M.S. Gilzenrat, B.D. Holmes, J. Rajkowski, G. Aston-Jones, and J.D. Cohen. Simplified dynamics in a model of noradrenergic modulation of cognitive performance. *Neural Networks*, 15:647–663, 2002.
- [7] M. Usher and E.J. Davelaar. Neuromodulation of decision and response selection. *Neural Networks*, 15:635–645, 2002.
- [8] E. Brown, J. Gao, P. Holmes, R. Bogacz, M. Gilzenrat, and J. Cohen. Simple networks that optimize decisions. *Int. J. Bifurcation and Chaos*, to appear, 2004.
- [9] J. Schall. Neural correlates of decision processes: neural and mental chronometry. *Curr. Opin. Neurobiol.*, 2003,.
- [10] B. Reddi. Decision making: two stages of neural judgement. *Curr. Biol.*, 11(15):R603–R606, 2001.
- [11] H. Wilson and J. Cowan. Excitatory and inhibitory interactions in localized populations of model neurons. *Biophys. J.*, 12:1–24, 1972.
- [12] J.J. Hopfield. Neurons with graded response have collective computational properties like those of two-state neurons. *Proc. Natl. Acad. Sci. USA*, 82:3088–3092, 1984.
- [13] L. Abbott. Firing-rate models for neural populations. In O. Benhar, C. Bosio, P. Del Giudice, and E. Tabat, editors, *Neural Networks: from Biology to High-Energy Physics*, pages 179–196. ETS Editrice, Pisa, 1991.

- [14] W. Gerstner and W. Kistler. *Spiking Neuron Models*. Cambridge University Press, 2002.
- [15] R. Bogacz, E. Brown, J. Moehlis, P. Hu, P. Holmes, and J.D. Cohen. The physics of optimal decision making: A formal analysis of models of performance in two alternative forced choice tasks. *Submitted to Psych. Rev.*, 2004.
- [16] M. Usher and J.L. McClelland. On the time course of perceptual choice: The leaky competing accumulator model. *Psych. Rev.*, 108:550–592, 2001.
- [17] E. Brown and P. Holmes. Modeling a simple choice task: stochastic dynamics of mutually inhibitory neural groups. *Stochastics and Dynamics*, 1(2):159–191, 2001.
- [18] J. Gold and M. Shadlen. Banburismus and the brain: Decoding the relationship between sensory stimuli, decisions, and reward. *Neuron*, 36:299–308, 2002.
- [19] E.L. Lehmann. *Testing Statistical Hypotheses*. John Wiley & Sons, New York, 1959.
- [20] T. Rowan. *Functional Stability Analysis of Numerical Algorithms*. PhD thesis, University of Texas at Austin, 1990.
- [21] M. Stone. Models for choice-reaction time. *Psychometrika*, 25:251–260, 1960.
- [22] D.R.J. Laming. *Information Theory of Choice-Reaction Times*. Academic Press, New York, 1968.
- [23] W. Feller. *An Introduction to Probability Theory and its Applications*. Wiley, New York, 1968.
- [24] M. Basseville and I. Nikiforov. *Detection of Abrupt Changes: Theory and Application*. Prentice Hall, New Jersey, 1993.
- [25] Josh Gold. Personal communication. 2004.
- [26] J.L. McClelland. On the time relations of mental processes: An examination of systems of processes in cascade. *Psychological Review*, 86:287–330, 1979.
- [27] S. Grossberg. Nonlinear neural networks: principles, mechanisms, and architectures. *Neural Networks*, 1:17–61, 1988.