

## Optimal decisions: From neural spikes, through stochastic differential equations, to behavior

P. Holmes<sup>1,2</sup>, E. Brown<sup>1</sup>, J. Moehlis<sup>3</sup>, R. Bogacz<sup>4</sup>, J. Gao<sup>2</sup>, P. Hu<sup>5</sup>, G. Aston-Jones<sup>6</sup>, E. Clayton<sup>6</sup>,  
J. Rajkowski<sup>6</sup>, and J.D. Cohen<sup>5</sup>

<sup>1</sup> Program in Applied and Computational Mathematics, Princeton University, Princeton, NJ 08544, USA

<sup>2</sup> Department of Mechanical and Aerospace Engineering, Princeton University, Princeton, NJ 08544, USA

<sup>3</sup> Department of Mechanical Engineering, University of California, Santa Barbara, CA 93106

<sup>4</sup> Department of Computer Science, University of Bristol, Bristol BS8 1UB, UK

<sup>5</sup> Department of Psychology, Princeton University, Princeton, NJ 08544, USA

<sup>6</sup> Department of Psychiatry, University of Pennsylvania, Philadelphia, PA 19104, USA

Email: pholmes@math.princeton.edu

**Abstract**—There is increasing evidence from in vivo recordings in monkeys trained to respond to stimuli by making left- or rightward eye movements, that firing rates in certain groups of ‘visual’ neurons mimic drift-diffusion processes, rising to a (fixed) threshold prior to movement initiation. This supplements earlier observations of psychologists, that human reaction-time and error-rate data can be fitted by random walk and diffusion models, and has renewed interest in optimal decision-making ideas from information theory and statistical decision theory as a clue to neural mechanisms.

We review results from decision theory and stochastic ordinary differential equations, and show how they may be extended and applied to derive explicit parameter dependencies in optimal performance that may be tested on human and animal subjects. We then briefly describe a biophysically-based model of a pool of neurons in locus coeruleus, a brainstem nucleus implicated in widespread norepinephrine release. This neurotransmitter can effect transient gain changes in cortical circuits of the type that the abstract drift-diffusion analysis requires. We argue that a rational account of how neural spikes give rise to simple behaviors is beginning to emerge.

### 1. Introduction: Optimal decisions

The drift diffusion (DD) process, governed by the stochastic differential equation (SDE):

$$dx =adt + \sigma dW, \text{ with thresholds } \pm z, \quad (1)$$

where the constants  $a$  and  $\sigma$  denote the drift rate and standard deviation of the Wiener (white noise) process  $W(t)$ , has been used since the 1960’s to model human reaction time and error statistics in two-alternative forced choice and other tasks [1]. Not only is it the continuum limit of the sequential probability ratio test (SPRT), known to be the optimal decision-maker for two-alternative forced-choice (TAFC) tasks with accumulating noisy data [2, 3], but its threshold-crossing behavior closely matches human

behavioral data [4, 5]. Moreover, direct neural recordings from oculomotor brain areas of monkeys performing choice tasks has recently shown that firing rates of groups of neurons selective for the ‘chosen’ of the two alternatives rise toward a threshold that signals the onset of motor response in a manner that seems to match sample DD paths [6, 7, 8].

In this application  $a$  denotes the net weight of evidence in favor of one alternative vs. the other (the log likelihood ratio). First passage time distributions yielding mean reaction times (RT), and error rates (ER), are readily computed for (1) from the backward Kolmogorov or Fokker-Planck equation associated with it:

$$ER = \frac{1}{1 + \exp\left(\frac{2az}{\sigma^2}\right)}; \quad RT = \frac{z}{a} \tanh\left(\frac{az}{\sigma^2}\right). \quad (2)$$

For fixed signal to noise ratio (SNR), as  $z$  increases, ER decreases but at the expense of longer RTs: this ‘speed-accuracy tradeoff’ is well-known in psychology [1]. However, as suggested by [9], one can explicitly compute thresholds that maximize the average *reward rate*:

$$RR = \frac{1 - ER}{RT + D + D_{\text{pen}} \cdot ER}; \quad (3)$$

here the numerator represents the average fraction correct and the denominator the average time between responses (RT + experimenter-imposed delay  $D$  + possibly an additional penalty delay  $D_{\text{pen}}$  incurred by errors). Substituting (2) into (3) one finds that the unique maximum of RR as a function of threshold for fixed SNR occurs at:

$$\exp\left(\frac{2az}{\sigma^2}\right) - 1 = \frac{2a^2}{\sigma^2} \left(D + D_{\text{pen}} - \frac{z}{a}\right). \quad (4)$$

From Eqns. (2-4) we may derive a unique *optimal performance curve* relating normalised reaction time (RT /  $[D + D_{\text{pen}}]$ ) to ER: see Fig. 1, which also shows behavioral data indicating that all subjects but those with the lowest overall scores follow the optimal curve reasonably closely, albeit with slightly suboptimal (longer) reaction times. See [10] for full details.

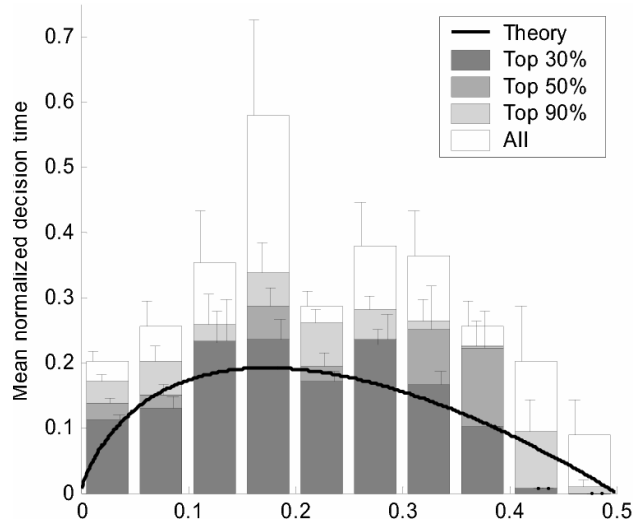


Figure 1: Thick curve shows the optimal performance curve derived from Eqns (2-3), and histogram bars show data collected from 80 human subjects, sorted according to total rewards accrued. White bars: all subjects; light gray bars: lowest 10% excluded; medium gray bars: lowest 50% excluded; dark gray bars: lowest 70% excluded. Error bars indicate standard error.

## 2. A neural model

As shown in [10], the DD process (1) can be derived in suitable limits from connectionist models of neural activity (see [11] and §3 below), which are in turn related to firing rate models that may be derived from biophysically-detailed Hodgkin-Huxley type equations and ‘integrate-and-fire’ simplifications thereof [13, 14]. We have begun studies of specific neural groups involved in the decision process, and, via neurotransmitter release, in control and attention selection. The brainstem nucleus *locus coeruleus* (LC) plays an important role in the latter [12, 15], releasing norepinephrine widely in the cortex when its cells fire action potentials.

We model LC with a heterogeneous set of periodically spiking neurons reduced to planar systems as by Hindmarsh and Rose [16] and further reduced to phase oscillators via the phase response curve (PRC) method [17]. This yields a set of noisy, coupled, SDEs each of the form:

$$d\theta_i = \left[ \omega_i + Z(\theta_i)(I(t) + \sum_j f(\theta_j)) \right] dt + \sigma Z(\theta_i) dW(t) + \mathcal{O}(\sigma^2), \quad (5)$$

where  $I(t)$  and  $f(\theta_j)$  denotes inputs due to external stimuli and from synaptic and electrotonic coupling from other LC cells, and the PRC  $Z(\theta)$  encodes the cell’s sensitivity at different points in its firing cycle or phase  $\theta$ : see [18] for details.

The probability density of phases,  $p(\theta, t)$ , for (5) in the weakly-coupled limit may be found from the associ-

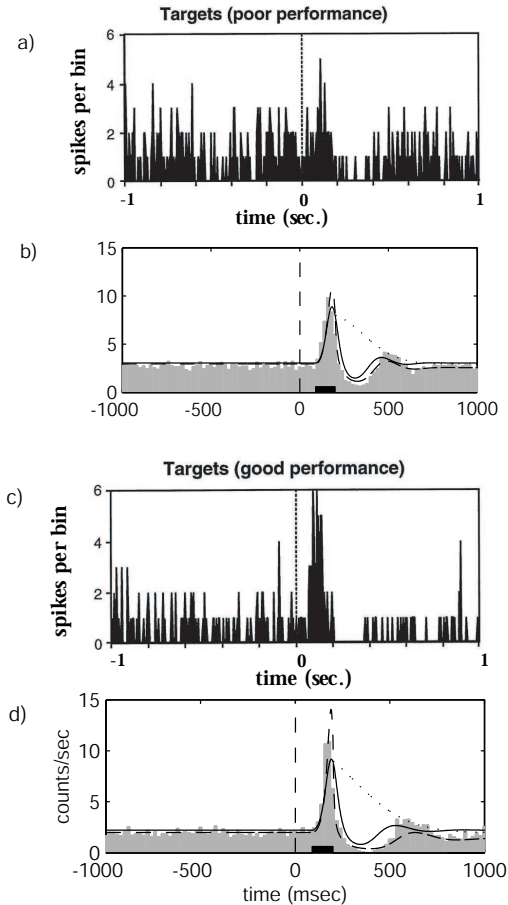


Figure 2: Peri-stimulus time histograms of LC activity for poor (a) and good (c) performance periods during a target identification task, taken from single neurons, averaged over  $\sim 100$  trials, from [15]. (b,d) show corresponding histograms computed by simulating 100 Rose-Hindmarsh neurons (gray bars) and from Fokker-Planck equation for (5) (solid curve), with analytical decay bounds (dashed curve).

ated forward Kolmogorov or Fokker-Planck equation and (semi-) analytical expressions derived for the flux of phases through  $\theta = 0$ , corresponding to the cell firing an action potential. This, in turn allows us to compute average firing rates of (groups of) LC cells in response to stereotyped stimuli representative of simple visual recognition tasks [15, 19]. Fitting noise ( $\sigma$ ) and oscillator frequency distributions  $P(\omega_i)$  to interspike interval data in the absence of stimuli, we compute firing rate histograms for comparison with experimental data: Fig. 2.

This figure illustrates the main result of [18]: that the magnitude of the transient response to stimulus, relative to baseline, is inversely proportional to baseline spiking frequency of LC in the absence of stimuli, and partially explains the correlations between low baseline activity, strong phasic response, and good performance on the one hand, and higher baseline activity, lower response, and poor per-

formance on the other [15].

These results, and others with different stimuli representing more complex decision tasks [18] and different neural models [20], show that analytically-tractable reduced descriptions of neural groups can be derived from biophysically detailed ion-channel models. A major challenge is to assemble such groups into ‘global’ models of interacting brain mechanisms known to be involved in perception and decision-making, e.g. the medio-temporal and lateral intraparietal areas, superior colliculus and frontal eye fields (MT, LIP, CS, FEF), involved in motion-detection and response saccades in monkeys [6, 7, 8], and to integrate brainstem and midbrain areas such as LC and thalamus. At the level of connectionist models, we have begun to study how gain changes such as those due to transient bursts of LC firing of Fig. 2 can affect cortical neurons. We briefly review this before concluding the paper.

### 3. Optimal gain schedules

A firing rate model for TAFC takes the form (cf. [11]):

$$\frac{dy_1}{dt} = -\alpha y_1 + f_{g(t)}(-\beta y_2 + a_1(t)) + g(t) \frac{\sigma(t)}{\sqrt{2}} \eta_1^1, \quad (6)$$

$$\frac{dy_2}{dt} = -\alpha y_2 + f_{g(t)}(-\beta y_1 + a_2(t)) + g(t) \frac{\sigma(t)}{\sqrt{2}} \eta_1^2, \quad (7)$$

where the function  $f_{g(t)}$  relating firing rate to inputs is typically sigmoidal:

$$f_{g(t)}(x) = \frac{1}{1 + \exp(-4g(t)(x - b))} \quad (8)$$

or piecewise-linear, being bounded above (by 1) and below (by 0). Here we allow time-varying stimuli  $a_j(t)$ , noise level  $\sigma(t)$  and gain  $g(t)$  (the maximum slope of  $f_{g(t)}$ ).

If decay (leak)  $\alpha$  and/or inhibition  $\beta$  are large, then (6-7) has a one-dimensional stochastic center manifold [21] that attracts solutions in a probabilistic sense. Moreover, linearizing at the point of maximum slope and subtracting (7) from (6) yields a scalar Ornstein-Uhlenbeck (OU) process for the difference  $x = y_1 - y_2$  in firing rates, and if the system is *balanced* in that leak rate equals inhibition ( $\beta \equiv 1$ ), this OU process reduces to the DD SDE (1) with  $a = a_1 - a_2$ . Hence, for constant SNR, a balanced firing rate, or leaky competing accumulator model, closely approximates the optimal decision-maker [10].

In [22] we address the problem of varying SNR and, using the linearised one-dimensional O-U SDE with time-dependent coefficients, we develop general expressions for optimal gain schedules. These implement the matched filter strategy of signal processing [23]. We compute optimal gains for specific simple cases of stimuli that rise slowly and rapidly and, using a simple linear model of norepinephrine release as a function of LC firing rate, we find that the transient LC firing rates thus predicted are qualitatively similar to experimental PSTH records such as

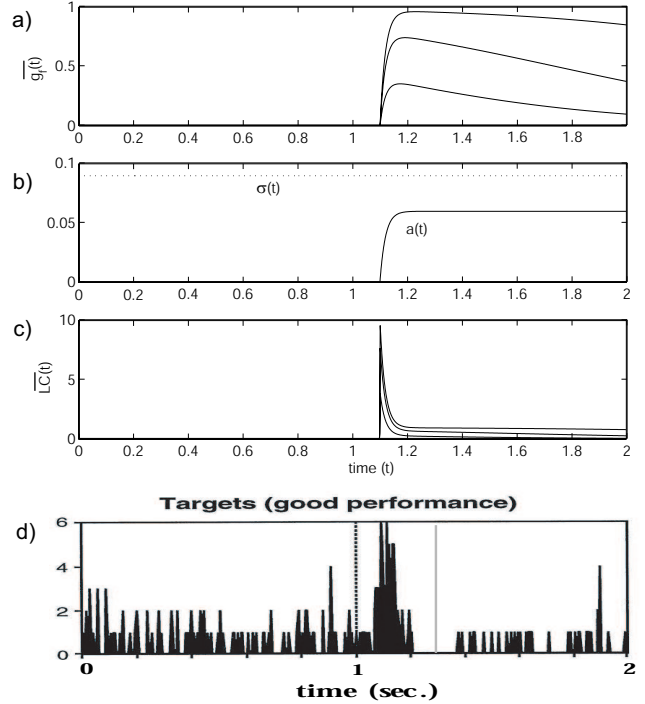


Figure 3: Comparison of optimal gain theory with empirical data for a target detection task. (a) Optimal gain schedules for the firing rate model, with a processing time lag of 0.1 sec following sensory cue, as shown in (b). (c) The corresponding optimal time course of LC firing rate. (d) Histogram of LC firing rates recorded in monkey during good performance, from [15].

those of Fig. 2. See Fig. 3. This lends further support to the hypothesis that LC activity, triggered by the arrival of salient stimuli in cortical decision areas, can ‘tune’ those areas (as well as motor areas) to improve accuracy and speed responses.

### 4. Conclusion

We have reviewed recent work in modeling neural and behavioral response to stimuli at both the level of biophysical detail, beginning with ion channel models of Hodgkin-Huxley type, and at that of abstracted ‘higher level’ connectionist and drift-diffusion SDEs. While numerous gaps remain between models at these disparate spatial and temporal scales, we believe that the general outlines of an integrated theory of neural function in simple decision-making tasks are beginning to emerge.

### Acknowledgments

This work is supported by DoE grant DE-FG02-95ER25238, PHS grants MH58480 and MH62196 (Cognitive and Neural Mechanisms of Conflict and Control, Silvio

M. Conte Center). EB and JM benefitted from National Science Foundation Graduate and Postdoctoral Research Fellowships and Burroughs-Wellcome Training Grant in Biological Dynamics: 1001782.

### References

- [1] D.R.J. Laming, "Information Theory of Choice-Reaction Times," Academic Press, New York, 1968.
- [2] A. Wald, "Sequential Analysis," Wiley, New York, 1947.
- [3] E.L. Lehmann, "Testing Statistical Hypotheses," Wiley, New York, 1959.
- [4] R. Ratcliff, "A theory of memory retrieval," *Psych. Rev.* vol. 85, pp 59-108, 1978.
- [5] R. Ratcliff, T. Van Zandt and G. McKoon "Connectionist and diffusion models of reaction time," *Psych. Rev.* vol. 106(2), pp 261-300, 1999.
- [6] J.D. Schall, "Neural basis of deciding, choosing and acting," *Nature Neurosci.* vol 2, pp 33-42, 2001.
- [7] J.D. Roitman and M.N. Shadlen, "Response of neurons in the lateral intraparietal area during a combined visual discrimination reaction time task," *J. Neurosci.* vol 22(1), pp 9475-9489, 2002.
- [8] R. Ratcliff, A. Cherian and M.A. Segraves, "A comparison of macaque behavior and superior colliculus neuronal activity to predictions from models of two choice decisions," *J. Neurophysiol.* vol 90, pp 1392-1407, 2003.
- [9] J.I. Gold and M.N. Shadlen, "Banburismus and the brain: Decoding the relationship between sensory stimuli, decisions, and reward," *Neuron* vol 36, pp 299-308, 2002.
- [10] R. Bogacz, E. Brown, J. Moehlis, P. Hu, P. Holmes and J.D. Cohen, "The physics of optimal decision making: A formal analysis of performance in two-alternative forced choice tasks," *Submitted*, 2004.
- [11] M. Usher and J.L. McClelland, "On the time course of perceptual choice: The Leaky Competing Accumulator Model," *Psych. Rev.* vol 108, pp 550-592, 2001.
- [12] G. Aston-Jones, J. Rajkowsky, P. Kubiak and T. Alexinsky, "Locus coeruleus neurons in the monkey are selectively activated by attended stimuli in a vigilance task," *J. Neurosci.* vol 14, pp 4467-4480, 1994.
- [13] A. Herrmann and W. Gerstner, "Noise and the PSTH response to current transients: I. General Theory and application to the Integrate-and-Fire Neuron," *J. Comp. Neurosci.* vol 11, pp 135-151, 2001.
- [14] A. Renart, N. Brunel and X-J. Wang, "Mean-field theory of irregularly spiking neuronal populations and working memory in recurrent cortical networks," in *Computational Neuroscience: A Comprehensive Approach*, J. Feng, Ed, CRC Press, Boca Raton, FL, 2004.
- [15] M. Usher, J.D. Cohen, D. Servan-Schreiber, J. Rajkowsky and G. Aston-Jones, "The role of locus coeruleus in the regulation of cognitive performance," *Science* vol 283, pp 549-554, 1999.
- [16] R. Rose and J. Hindmarsh, "The assembly of ionic currents in a thalamic neuron I. The three-dimensional model," *Proc. R. Soc. Lond. B*, vol 237, pp 267-288, 1989.
- [17] G.B. Ermentrout, "Type I membranes, phase resetting curves, and synchrony," *Neural Comp.* vol 8, pp 979-1001, 1996.
- [18] E. Brown, J. Moehlis, P. Holmes, E. Clayton, J. Rajkowski and G. Aston-Jones, "The influence of spike rate and stimulus duration on brainstem noradrenergic neurons," *J. Comp. Neurosci.* vol 17(1), pp 13-29, 2004.
- [19] E. Clayton, J. Rajkowski, J.D. Cohen and G. Aston-Jones, "Decision-related activation of monkey locus coeruleus neurons in a forced choice task," *Submitted*, 2003.
- [20] E. Brown, J. Moehlis and P. Holmes, "On the phase reduction and response dynamics of neural oscillator populations," *Neural Comp.* vol 16(4), pp 673-715, 2004.
- [21] L. Arnold, "Random Dynamical Systems," Springer Verlag, Heidelberg, 1998.
- [22] E. Brown, J. Gao, P. Holmes, R. Bogacz, M. Gilzenrat and J.D. Cohen, "Simple neural networks that optimize decisions," *Int. J. Bifurcation and Chaos* (in press), 2004.
- [23] A. Papoulis, "Signal Analysis," McGraw-Hill, New York, 1977.