

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

P. Holmes^{1,2}, E. Brown¹, J. Moehlis³, R. Bogacz⁴, J. Gao², P. Hu⁵, G. Aston-Jones⁶, E. Clayton⁶,
J. Rajkowski⁶, and J.D. Cohen⁵

¹ Program in Applied and Computational Mathematics, Princeton University, Princeton, NJ 08544, USA

² Department of Mechanical and Aerospace Engineering, Princeton University, Princeton, NJ 08544, USA

³ Department of Mechanical Engineering, University of California, Santa Barbara, CA 93106

⁴ Department of Computer Science, University of Bristol, Bristol BS8 1UB, UK

⁵ Department of Psychology, Princeton University, Princeton, NJ 08544, USA

⁶ 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 σ 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_{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.

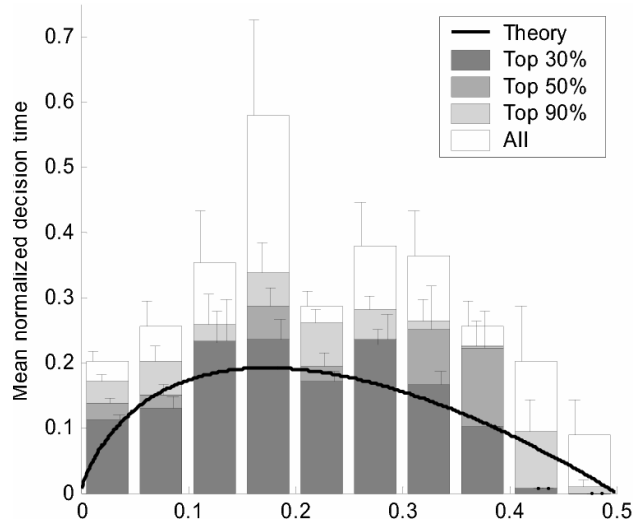


Figure 1: Thick curve shows the optimal performance curve derived from Eqns (2-3), and histogram bars show data collected from 80 human subjects, sorted according to total rewards accrued. White bars: all subjects; light gray bars: lowest 10% excluded; medium gray bars: lowest 50% excluded; dark gray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

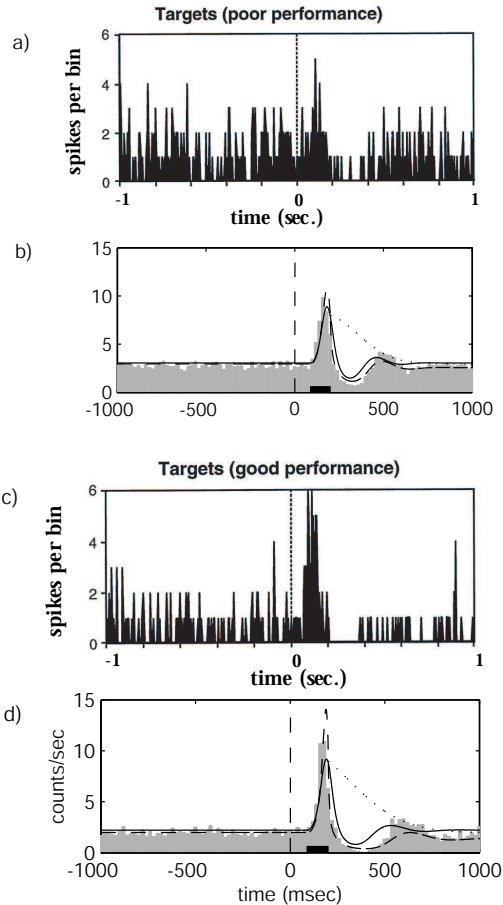


Figure 2: Peri-stimulus time histograms of LC activity for poor (a) and good (c) performance periods during a target identification task, taken from single neurons, averaged over ~ 100 trials, from [15]. (b,d) show corresponding histograms computed by simulating 100 Rose-Hindmarsh neurons (gray bars) 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 (σ) 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-

