Notes on psychometric functions for Ornstein-Uhlenbeck equations

Sam Feng[1] and Philip Holmes[1,2]

[1] Program in Applied and Computational Mathematics,
[2] Department of Mechanical and Aerospace Engineering,
Princeton University, Princeton, NJ 08544, U.S.A.

March 17, 2008

Abstract

We review basic formulae for accuracy as a function of signal to noise ratio (Stimulus coherence), viewing time, and other parameters that characterize a linear stochastic ordinary differential equation that models evidence accumulation in two-alternative forced-choice tasks. We suggest two ways in which reward bias information might enter, and derive the resulting shifted psychometric functions. Finally, we predict shifts that maximize rewards.

1 Introduction

The following notes are motivated by ongoing experiments conducted by Alan Rorie in the Newsome laboratory [16]. Monkeys are trained on a two-alternative forced-choice (2AFC) task with a moving dots motion stimulus [2, 3, 17], but unlike much previous work, in which correct choices of either alternatives are rewarded equally, the experiment is run under four conditions. Rewards may be high for both alternatives A and B, low for both alternatives, high for A and low for B, or low for A and high for B. No rewards are given for incorrect choices, regardless of reward bias.

Each trial begins with a fixation period of 150 ms, followed by a target period of 250 ms and a 250 ms cue period during which the reward values are signaled by colors appearing in the targets. These remain on during the succeeding 500 ms motion period, which is followed by a variable delay period of 300-550 ms before the signal to respond is given. For simplicity, we shall not model dynamics during the fixation, target and delay periods (we comment on this in §1.2 below).

The experiments show that, when rewards are unequal, the psychometric function, which plots the percentage (or fraction) of A choices as a function of stimulus coherence, is shifted so that the 50% mark falls at a nonzero coherence. For example, letting C denote coherence, with positive values indicating alternative A and negative values alternative B, and supposing that A is rewarded more than B, the 50% choice mark is found to lie at a negative value $C_0 < 0$. No significant shift occurs for equal rewards, whether they be high or low.

To investigate how the unequal reward case might bias the decision process, we adopt a simple stochastic choice model: the Ornstein-Uhlenbeck (OU) process. This generalizes the yet simpler drift-diffusion (DD) process, which is a continuum limit of random walk models that have long been used to describe the accumulation of relative evidence in 2AFC tasks [11, 14, 13, 12, 15, 18]. See [1] for an extensive review and discussion of how DD and OU processes emerge from more complex

models, including leaky competing accumulators [19]. A recent paper, from which we draw below, describes fits of behavioral data from monkeys learning the moving dots task, and shows that DD and OU processes can provide good descriptions of psychometric functions [4]. Neural evidence for accumulator models appears in [10, 8, 9, 17].

1.1 The Ornstein-Uhlenbeck process

We begin by stating the OU stochastic differential equation (SDE):

$$dx = (\lambda x + A(t))\,dt + \sigma\,dW, \tag{1}$$

where $x(t)$ represents the firing rate difference between two populations of neurons, each accumulating evidence in favor of one of the two possible choices. The parameter λ represents the leak rate[1] and σ denotes the standard deviation of a white noise (Weiner) process. The stimulus is represented by the drift rate $A(t)$, which may in general vary with time. Without loss of generality we shall take A as positive for stimulus A and negative for B. When $\lambda = 0$ and A is constant, (1) reduces to a DD process. In [1] it is shown that the DD process is the continuum limit of the sequential probability ratio test, which is known to be optimal for 2AFC tasks in the sense that it delivers a decision of guaranteed accuracy in the shortest possible time, or that, given a fixed decision time, it maximizes accuracy [20, 21]. The latter sense is relevant to the fixed viewing time experiments considered here.

The probability distribution of solutions $p(x,t)$ for (1), given suitable distribution of initial data, is governed by the forward Kolmogorov or Fokker-Planck equation [5], cf. [4]:

$$p_t = -p_x\left[(\lambda x + A(t))p\right] + \frac{\sigma^2}{2}p_{xx}. \tag{2}$$

When the distribution of initial conditions is a Gaussian centered about μ_0,

$$p(x,0) = \frac{1}{\sqrt{2\pi\nu_0}}\exp\left[-\frac{(x-\mu_0)^2}{2\nu_0}\right], \tag{3}$$

solutions of (2) remain Gaussian as time evolves:

$$p(x,t) = \frac{1}{\sqrt{2\pi\nu(t)}}\exp\left[-\frac{(x-\mu(t))^2}{2\nu(t)}\right], \quad \text{where} \tag{4}$$

$$\mu(t) = \mu_0 e^{\lambda t} + \int_0^t e^{\lambda(t-s)}A(s)\,ds \quad \text{and} \quad v(t) = \nu_0 e^{2\lambda t} + \frac{\sigma^2}{2\lambda}\left(e^{2\lambda t} - 1\right) \tag{5}$$

If $\lambda = 0$ (the DD limit), $\mu(t)$ and $\nu(t)$ simplify to

$$\mu(t) = \mu_0 + \int_0^t A(s)\,ds \quad \text{and} \quad \nu(t) = \nu_0 + \sigma^2 t. \tag{6}$$

In the calculations below we shall assume that $\nu_0 = 0$, implying that the stream of noise, along with evidence embedded in it and any biases that may enter, begins accumulating at $t = 0$ in each

[1] In reductions from linearized leaky competing accumulators, λ is the difference between the strength of mutual inhibition and the decay rate of an individual accumulator [1].

trial. In this case the initial condition $x(0) = \mu_0$ is identical for all trials (see Eq. (3)). We further assume that the choice made at the end of the stimulus viewing period (or upon presentation of a cue to respond), depends upon the sign of the accumulated net evidence $x(t)$: choice A being made if $x(t) > 0$ and B made if $x(t) < 0$. Thus, the probability that choice A is made can be computed as

$$P(t) = \int_0^\infty p(x,t)\,dx = \frac{1}{2}\left[1 + \mathrm{erf}\left(\frac{\mu(t)}{\sqrt{2\nu(t)}}\right)\right], \tag{7}$$

where $\mathrm{erf}(y) = (2/\sqrt{\pi})\int_0^y \exp(-u^2)\,du$ denotes the error function. Note that $\nu(t) > 0$ regardless of the sign of λ (Eqns. (5-6), so the square roots in (4) and (7) are well-defined. Eqn. (7) defines a *psychometric function* (PMF) analogous to that measured in the Rorie-Newsome experiments [16]. (The range of $P(t)$ runs from 0 to 1 as the argument $(\mu/\sqrt{2\nu})$ of the error function runs from $-\infty$ to $+\infty$, so one must multiply the expression (7) by 100 to obtain % of A choices.)

In addition to its explicit dependence on viewing time T, the PMF also depends on the functional forms of the drift and noise terms embedded in $\mu(t)$ and $\nu(t)$. In particular $\mu(t)$ depends on the coherence or strength of the stimulus via $A(t)$, and upon prior expectations or biases that reward information might introduce. We describe explicit examples of this dependence in §1.2. Along with viewing time, coherence is easily manipulated in experiments, and so we shall sometimes write the PMF in the forms $P(C,t)$ and $P(C,t;\mu_0)$, to denote dependence on C and other parameters such as μ_0. Similarly, we can write the integrated drift as $\mu(C,t)$, etc.

We shall specifically examine two aspects of the PMF as a function of C: the *midpoint shift* and the *slope at the midpoint*. The midpoint shift C_0 is defined as the value of C at which $P(C,t) = 0.5$, or equivalently, the value of C at which $\mu(C,t) = 0$. The slope of $P(C,T)$ at C_0 is given by $\frac{dP(t)}{dC}(C_0)$.

1.2 Models of stimuli and reward biasing

Following the work of [6, 7], we suppose that the part of the drift rate due to the stimulus depends linearly on coherence, so $A = aC$, where a is a scaling or "sensitivity" parameter that allows one to fit the model to different subjects, or to the same subject during different epochs of training [4, Fig. 14]. While power-law dependence on C has been introduced to account for behavior early in training, a linear relationship seems generally adequate for well-trained animals [7]. Here C lies in the range $[-1, 1]$ (between 100% leftward and 100% rightward motion coherence), and is determined by the experimenter, but a is a free parameter for fitting.

An optimality study of 2AFC tasks run under the free response protocol shows that choosing biased initial conditions $x(0) = \mu_0 \neq 0$ is the optimal strategy both for stimuli presented with unequal probabilities, and for unequal rewards [1]. Other biases, such as continuing inputs, will degrade performance. However, this analysis assumes that coherences remain fixed over a block of trials, which is not the case in the Rorie-Newsome experiments, in which coherences are selected from a finite set of values with equal probability on each trial. Moreover, [1] focuses on the free response protocol, so that a speed-accuracy tradeoff must be to solved maximise overall rewards. We shall address the optimality question for the current, mixed coherence experiment with fixed viewing times (interrogation protocol) below, but here we propose two biasing strategies.

(1) The first and simplest is to bias the initial condition at the start of evidence accumulation: $x(0) = \mu_0 \neq 0$, taking $x(0) = \mu_0 > 0$ if alternative A is more lavisly rewarded, and $x(0) = -\mu_0 < 0$ if alternative B is more lavisly rewarded. Since only two discrete reward levels are used in the experiments, the single free parameter μ_0 characterizes the weight that the subject places on the

reward differential. In this case, from (5) the integrated drift rate is simply

$$\mu(C,t) = \mu_0 e^{\lambda t} + \frac{aC}{\lambda}\left(e^{\lambda t} - 1\right), \quad \text{where} \quad t \in [0,T], \tag{8}$$

and the decision is rendered at $t = T$.

(2) Alternatively, one can assume that bias enters throughout the reward cue period and ensuing stimulus period as a separate drift term on which the stimulus is superimposed, as described by the piecewise-constant drift rate:

$$A(C,t) = \begin{cases} b, & t \in [0,\tau) \\ b + aC, & t \in [\tau, T+\tau] \end{cases}, \tag{9}$$

and the corresponding integrated drift during the stimulus period:

$$\mu(C,t) = e^{\lambda t}\left[\int_0^t be^{-\lambda s}\,ds + \int_\tau^t aCe^{-\lambda s}\,ds\right] = \frac{b}{\lambda}\left(e^{\lambda t} - 1\right) + \frac{aC}{\lambda}\left(e^{\lambda(t-\tau)} - 1\right), \quad \text{for } t \in [\tau, T]. \tag{10}$$

Here τ and T are respectively the cue and motion period durations and we take $\mu_0 = 0$, since reward bias is modeled by b. By analogy to the convention for μ_o above, $b > 0$ if choice A (associated with $x > 0$) has higher reward and $b < 0$ if B has higher reward. The decision is rendered at $t = \tau + T$. The experimenter determines τ and T, and b is the free fitting parameter in this case.

More complex drift functions could be defined to cover the fixation, target and delay periods, but analyses of neurophysiological data (LIP firing rates) are necessary before informed modeling can be done. For this reason, and for the sake of simplicity, here we assume that the accumulation process starts at reward cue onset $t = 0$ and ends at motion period offset: $t = T$ in the first case and $t = \tau + T$ in the second. The above models already introduce three bias and two stimulus parameters: μ_0, b (and τ), and a (and C), along with λ and σ. Inclusion of fixation, target and delay periods would introduce many more.

This remark prompts a further observation: that the accuracy defined in Eq. (7) depends upon the single quantity $\mu(t)/\sqrt{2\nu(t)}$ and this function constrains the way in which the underlying parameters interact, implying dependencies among them. Below, we shall see certain characteristic groups of parameters emerge, such as the sensitivity-to-noise ratio a/σ. In this regard we note that in the first case (8) the stimulus and bias parameters a, C, T and μ_0 only enter the numerator $\mu(t)$ at the response time $t = T$, while in the second case the cue duration τ will also enter the noise term $\nu(\tau + T)$.

We also observe that the parameter values selected for graphical examples below are abitrary, and chosen only to illustrate qualitative trends. We have no idea what values might be relevant in fitting real data.

2 Reward bias via initial condition at start of motion period

We consider biasing strategy (1) introduced above, first taking the simplest case of a pure DD process ($\lambda = 0$). We work up from this simple case to the case (2) of continual bias, and $\lambda \neq 0$.

2.1 $\mu_0 \neq 0$, $\lambda = 0$

From Eqs. (6), $\mu(t) = \mu_0 + aCt$ and $\nu(t) = \sigma^2 t$, and at the decision time $t = T$ (7) becomes:

$$P(C,T) = \frac{1}{2}\left[1 + \mathrm{erf}\left(\frac{\mu_0 + aCT}{\sqrt{2\sigma^2 T}}\right)\right]. \tag{11}$$

Setting the argument of the error function equal to zero and differentiating $P(C,T)$ gives the shift and slope:

$$C_0 = -\frac{\mu_0}{aT} \quad \text{and} \quad \frac{dP(T)}{dC}(C_0) = \frac{1}{2}\mathrm{erf}'(0)\frac{a}{\sigma}\sqrt{\frac{T}{2}} = \frac{a}{\sigma}\sqrt{\frac{T}{2\pi}}. \tag{12}$$

Note that the bias-to-stimulus sensitivity ratio μ_0/a appears in the shift, and the sensitivity-to-noise ratio a/σ appears in the slope. Perhaps more significant is the fact that the shift depends inversely on the motion period T, and the slope varies as $\sqrt{T}$.

2.2 $\mu_0 \neq 0$, $\lambda \neq 0$

For the OU process with $\lambda \neq 0$ we use (8) and compute $\nu(T)$ from (5):

$$\mu(C,t) = \mu_0 e^{\lambda T} + \frac{aC}{\lambda}\left(e^{\lambda T} - 1\right),\ \nu(T) = \frac{\sigma^2}{2\lambda}(e^{2\lambda T} - 1)\ \Rightarrow \frac{\mu(T)}{\sqrt{2\nu(T)}} = \frac{\mu_0 \lambda e^{\lambda T} + aC(e^{\lambda T} - 1)}{\sqrt{\sigma^2 \lambda (e^{2\lambda T} - 1)}}. \tag{13}$$

Setting $\mu(T) = 0$ and solving for C we get the shift and slope:

$$C_0 = -\frac{\mu_0 \lambda e^{\lambda T}}{a(e^{\lambda T} - 1)} \quad \text{and} \quad \frac{dP(T)}{dC}(C_0) = \frac{1}{2}\mathrm{erf}'(0)\frac{a(e^{\lambda T} - 1)}{\sigma\sqrt{\lambda(e^{2\lambda T} - 1)}} = \frac{a(e^{\lambda T} - 1)}{\sigma\sqrt{\lambda\pi(e^{2\lambda T} - 1)}}. \tag{14}$$

Note that these formulae apply to both $\lambda > 0$ and $\lambda < 0$, with a change in sign in slope for $\lambda < 0$. To see this, substitute $\lambda = -\bar{\lambda} < 0$ into Eqs. (13-14) and recompute the shift and slope to obtain:

$$C_0 = -\frac{\mu_0 \bar{\lambda} e^{-\bar{\lambda} T}}{a(1 - e^{-\bar{\lambda} T})} \quad \text{and} \quad \frac{dP(T)}{dC}(C_0) = \frac{a(1 - e^{-\bar{\lambda} T})}{\sigma\sqrt{\bar{\lambda}\pi(1 - e^{-2\bar{\lambda} T})}}. \tag{15}$$

As in Eqs. (12), the ratios μ_0/a and a/σ appear, but the motion period T now enters via the exponential terms, implying different behaviors for the stable and unstable OU cases $\lambda < 0$ and $\lambda > 0$. In particular, for long viewing times (or large $|\lambda|$), we have

$$C_0 \approx -\frac{\mu_0 \lambda}{a} \qquad \text{and} \qquad \frac{dP(T)}{dC}(C_0) \approx \frac{a}{\sigma\sqrt{\lambda\pi}}, \quad \text{for} \quad \lambda > 0, \tag{16}$$

$$\text{and} \quad C_0 \approx 0 \qquad \text{and} \qquad \frac{dP(T)}{dC}(C_0) \approx \frac{a}{\sigma\sqrt{-\lambda\pi}}, \quad \text{for} \quad \lambda < 0. \tag{17}$$

Note that not only do the slopes of the PMFs for $\lambda < 0$ and $\lambda > 0$ approach a common limit $a/(\sigma\sqrt{|\lambda|\pi})$ as $T \to \infty$, *they agree for all* T, as one can see by checking that the expressions for dP/dC in Eqns. (14) and (15) are identical for $\bar{\lambda} = \lambda$. In fact for fixed a, σ and $|\lambda|$, the factors multiplying C in the PMF are the same, and in this respect, the stable and unstable OU processes give the *same* accuracy. However, as the expressions for C_0 show, the PMFs are shifted by *different*

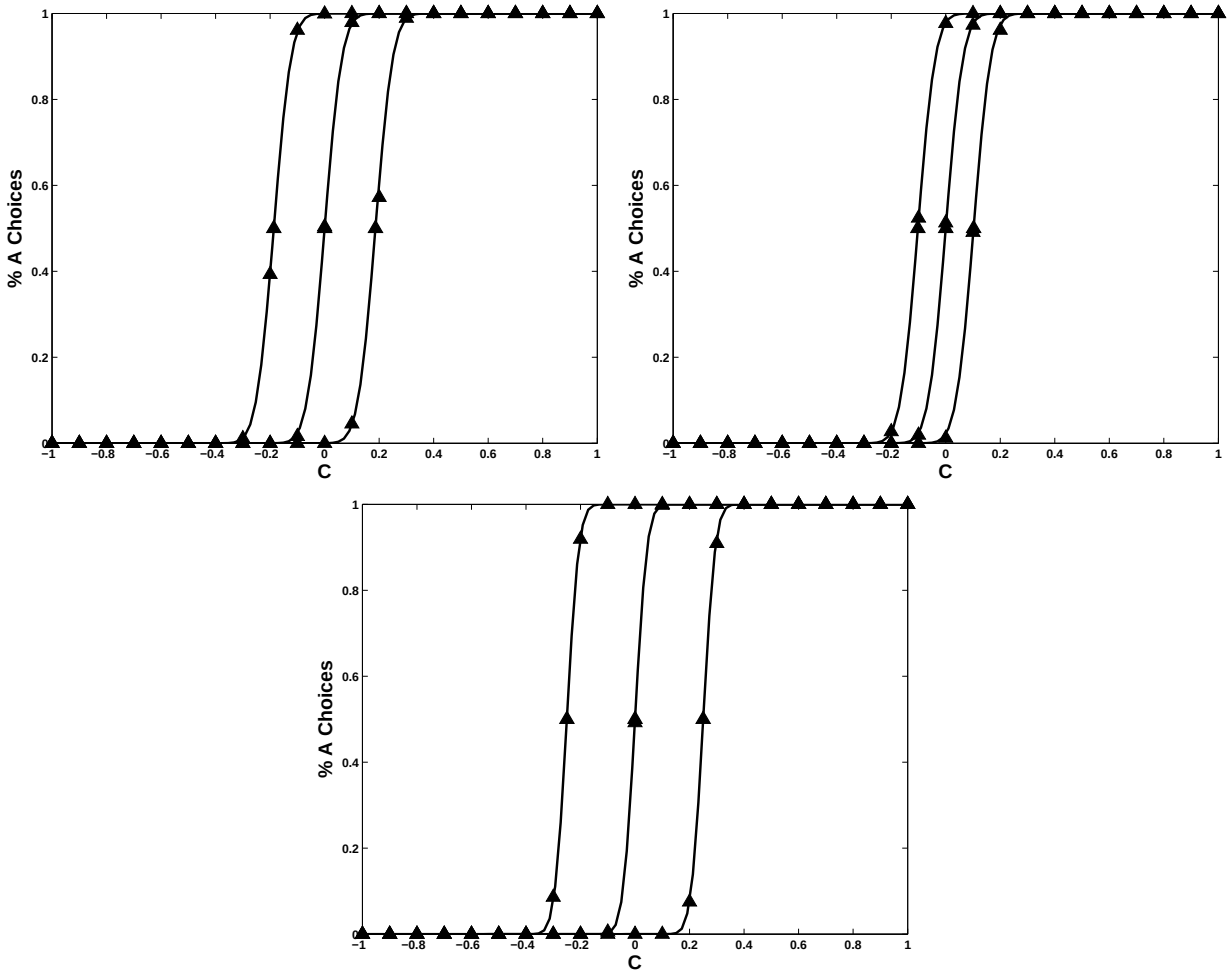

Figure 1: Psychometric functions showing fraction of A choices as a function of coherence C for the case of reward bias applied via initial condition μ_0. Top left: $\lambda = -0.1$, $\mu_0 = 0, \pm 100$; top right: $\lambda = +0.1$, $\mu_0 = 0, \pm 1$; bottom: $\lambda = 0$, $\mu_0 = 0, \pm 10$. Each panel shows curves for $\mu > 0$, 0, < 0 (left to right). Remaining parameters are $a = 1$, $\sigma = 0.2214$ and $T = 40$ (arbitrary time units). Curves show PMFs with shifts and slopes computed from the analysis of §§2.1-2.2, triangles indicate direct simulation results averaged over 1000 trials.

amounts for positive and negative λ: for $\lambda > 0$ the bias μ_0 is remembered, while it is gradually forgotten in the case $\lambda < 0$ that is dominated by decay (cf. Eqns. (16-17)). Since the motion period duration can be varied between blocks of trials, the exponential T-dependence in these shift expressions might allow one to determine which model provides a better fit.

Fig. 1 shows examples of the PMFs computed above. In order to provide well-spaced curves we have selected different values of initial condition μ_0 for each of the three cases $\lambda < 0$, $\lambda = 0$ and $\lambda > 0$.

3 Reward bias applied before and during motion period

We now move to biasing strategy (2), again considering the cases $\lambda = 0$ and $\lambda \neq 0$ separately.

3.1 Bias b, $\lambda = 0$, $(\mu_0 = 0)$

A simple calculation using Eq. (9) and Eqs. (6) yields $\mu(\tau + T) = b(\tau + T) + aCT$ and $\nu(\tau + T) = \sigma^2(\tau + T)$, giving the following shift and slope:

$$C_0 = -\frac{b(\tau + T)}{aT} \quad \text{and} \quad \frac{dP(T)}{dC}(C_0) = \frac{1}{2}\text{erf}'(0)\frac{aT}{\sqrt{2\sigma^2(\tau + T)}} = \frac{aT}{\sigma\sqrt{2\pi(\tau + T)}}. \tag{18}$$

Again the bias-to-stimulus sensitivity and sensitivity-to-noise ratios, b/a and a/σ appear, but now both the reward cue period τ and the motion period T enter.

3.2 $\lambda \neq 0$, $\mu_0 = 0$

Using (10) and computing the integrated noise from (5) at the response time $t = \tau + T$ we have:

$$\mu(C, \tau + T) = \frac{b}{\lambda}\left(e^{\lambda(\tau+T)} - 1\right) + \frac{aC}{\lambda}\left(e^{\lambda T} - 1\right), \ \nu(\tau + T) = \frac{\sigma^2}{2\lambda}\left(e^{2\lambda(\tau+T)} - 1\right). \tag{19}$$

Setting $\mu(\tau + T) = 0$ yields the shift:

$$C_0 = -\frac{b\left(e^{\lambda(\tau+T)} - 1\right)}{a\left(e^{\lambda T} - 1\right)} \tag{20}$$

and since

$$P(\tau + T) = \frac{1}{2}\left[1 + \text{erf}\left(\frac{b(e^{\lambda(\tau+T)} - 1) + aC(e^{\lambda T} - 1)}{\sigma\sqrt{\lambda(e^{2\lambda(\tau+T)} - 1)}}\right)\right]$$

we find the slope at C_0:

$$\frac{dP(T)}{dC}(C_0) = \frac{1}{2}\text{erf}'(0)\frac{a(e^{\lambda T} - 1)}{\sigma\sqrt{\lambda(e^{2\lambda(\tau+T)} - 1)}} = \frac{a(e^{\lambda T} - 1)}{\sigma\sqrt{\pi\lambda(e^{2\lambda(\tau+T)} - 1)}}. \tag{21}$$

As in §2.2, these formulae apply for both $\lambda > 0$ and $\lambda < 0$, with a change in sign for slope (21) when $\lambda < 0$. As T increases, we have

$$C_0 \approx -\frac{be^{\lambda\tau}}{a} \qquad \text{and} \qquad \frac{dP(T)}{dC}(C_0) \approx \frac{ae^{-\lambda\tau}}{\sigma\sqrt{\lambda\pi}}, \quad \text{for} \quad \lambda > 0, \tag{22}$$

$$\text{and} \quad C_0 \approx -\frac{b}{a} \qquad \text{and} \qquad \frac{dP(T)}{dC}(C_0) \approx \frac{a}{\sigma\sqrt{-\lambda\pi}}, \quad \text{for} \quad \lambda < 0. \tag{23}$$

Here the bias b plays a role analogous to μ_0 in Eqs. (16-17), but the pre-stimulus cue period τ enters in the exponential for $\lambda > 0$. Along with the full expression (21), this implies that, for the same $|\lambda|$ and other parameters, the unstable OU ($\lambda > 0$) process provides lower accuracy than the stable ($\lambda < 0$) process. Specifically, the additional factors $\exp(-\lambda\tau)$ in Eqn. (22) and $\sqrt{\exp(2\lambda\tau\ldots)}$ in the denominator of Eqn. (21) are due to the fact that noise accumulates during the cue period, leading to accelerating growth of solutions when $\lambda > 0$ which the stimulus cannot repair.

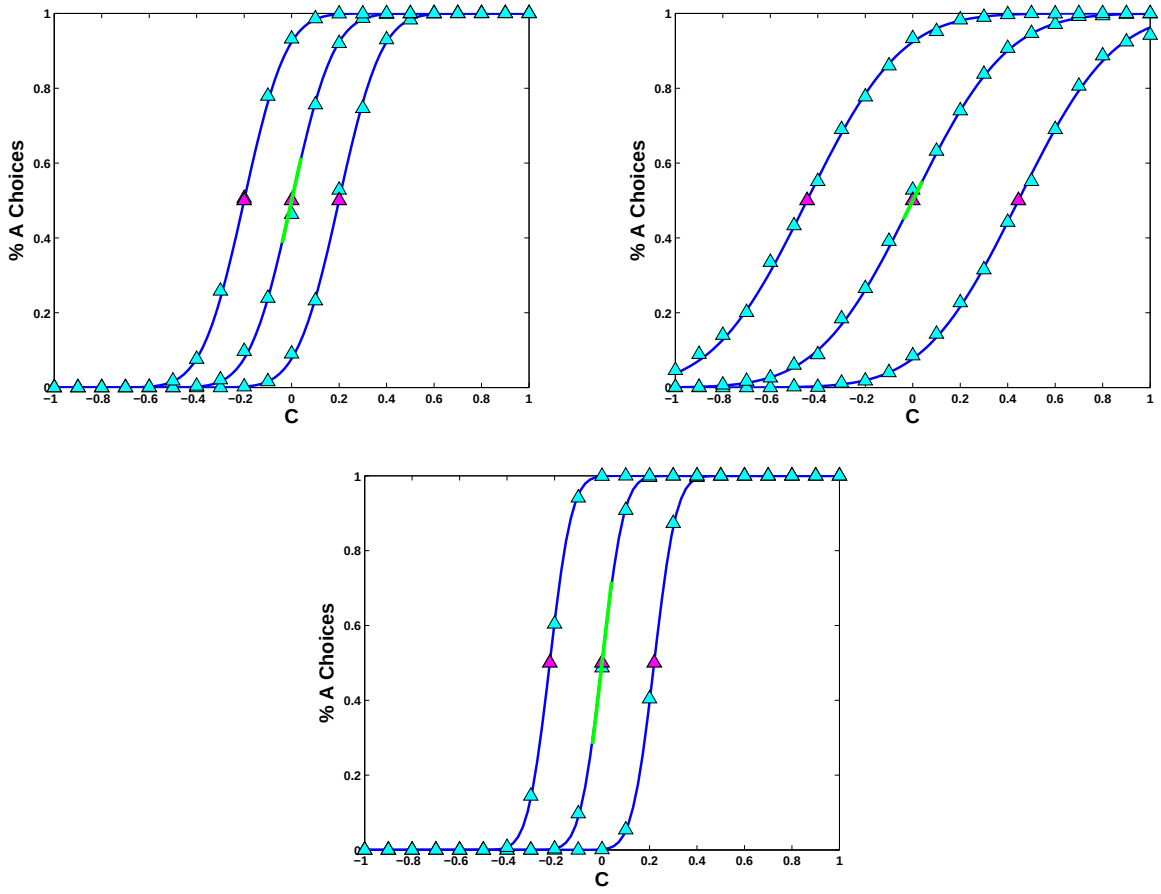

Figure 2: Psychometric functions showing fraction of A choices as a function of coherence C for the case of reward bias applied before and during motion period. Top left: $\lambda = -0.2$; top right: $\lambda = +0.2$; bottom: $\lambda = 0$; each panel shows curves for $b = +0.1, 0, -0.1$ (left to right). Remaining parameters are $a = 0.5$, $\sigma = 0.2214$, $\tau = 4$ and $T = 40$ (arbitrary time units). Blue curves, magenta triangles and green line indicate analytical PMFs, shifts and slopes from §§3.1-3.2, cyan triangles indicate direct simulation results averaged over 1000 trials.

Fig. 2 illustrates the above formulae by showing the PMFs derived in this section for $\lambda = 0$, $\lambda > 0$ and $\lambda < 0$. Note that the overall slopes of the functions are lower for $\lambda < 0$ (top left) and $\lambda > 0$ (top right), than for $\lambda = 0$ (bottom). Thus, accuracy declines as $|\lambda|$ increases, reflecting optimality of the DD process. The functions are rigidly shifted, to the left or right for $b > 0$ and $b < 0$ respectively, by an amount that grows as λ increases from negative to positive. The simulations were computed using ODE45 in Matlab with a basic timestep of $dt = 0.02$, allowing 200 iterations for the reward cue period and 2000 iterations for the motion period. Thus $\tau = 4$ and $T = 40$, in arbitrary time units.

3.3 Bias b, $\lambda \neq 0$, $\mu_0 \neq 0$

For completeness we include results in the case that bias enters both via a drift term *and* a shift in initial condition at the start of the reward cue period. Here $\mu(\tau + T)$ does not change much from §3.2. Repeating the calculations above with $\mu_0 \neq 0$ reveals that

$$\mu(\tau + T) = \mu_0 e^{\lambda(\tau+T)} + \frac{b}{\lambda}(e^{\lambda(\tau+T)} - 1) + \frac{aC}{\lambda}(e^{\lambda T} - 1),$$

and the noise term $\nu(t)$ is unchanged. The shift is thus

$$C_0 = -\frac{[\lambda\mu_0 e^{\lambda(\tau+T)} + b(e^{\lambda(\tau+T)} - 1)]}{a(e^{\lambda T} - 1)},$$

and the slope at C_0 is identical to (21).

4 Optimal biases

Thus far we have computed PMFs for particular models, examining the effect of model parameters on the resulting accuracy functions. We can instead ask what is the optimal PMF: i.e., the accuracy vs. coherence function that maximizes rewards? More precisely, given fixed "system" parameters such as sensitivity-to-noise ratio, what is the shift in the PMF that maximizes expected rewards in the case that the two alternatives are unequally rewarded? It seems reasonable to think of parameters such as a/σ as fixed because it remains stable over relatively long periods for trained animals, or at worst changes slowly [4]. Since the reward cue and motion period durations are fixed by the experimenter, within the context of the models presented here this will determine optimal initial conditions μ_0 or biases b. We are basically asking how much should the subject weight the reward information in order to make optimal use of it.

Let Φ denote the reward obtained on a typical trial, namely, r_1 if alternative A is offered and chosen correctly, and r_2 if B is offered and chosen correctly. The *expected reward* is obtained by multiplying each r_j by the probability that the corresponding alternative is chosen, when it appears in the stimulus. To make this explicit, first suppose that coherence is fixed from trial to trial, and there are only the two possible stimuli A ($C = +\bar{C}$) and B ($C = -\bar{C}$), that each is equally likely, and that biasing is done by selecting the initial state μ_0. In this case we have

$$\mathbb{E}(\Phi) = r_1 \frac{P(+\bar{C}, T; \mu_0)}{2} + r_2 \frac{(1 - P(-\bar{C}, T; \mu_0))}{2}, \tag{24}$$

where we include μ_0 explicitly in the PMF expression. Here we use the fact that $P(C, T; \mu_0)$ is the average proportion of correct A choices, and $1 - P(-C, T; \mu_0)$ is the average proportion of correct B choices, for coherence C and motion period T.

We shall investigate several cases of increasing complexity, much as in §3, but the following is not a complete list.

4.1 Case 1a

We first take $\lambda = 0$ and consider the fixed coherence case. Here, as in (11) we have

$$P(\pm\bar{C}, T; \mu_0) = \frac{1}{2}\left[1 + \mathrm{erf}\left(\frac{\mu_0 \pm a\bar{C}T}{\sqrt{2\sigma^2 T}}\right)\right], \tag{25}$$

so that the derivative with respect to μ_0 is:

$$\frac{\partial}{\partial\mu_0}P(\pm\bar{C},T;\mu_0)=\frac{1}{\sigma\sqrt{2\pi T}}\exp\left[-\frac{(\mu_0\pm a\bar{C}T)^2}{2\sigma^2T}\right]\stackrel{\text{def}}{=}P'_{\pm}. \tag{26}$$

(We let P_+ and P_- denote $P(\ldots)$ with $\pm C$ respectively, and the prime denotes derivative with respect to μ_0.) We then maximize $\mathbb{E}(\Phi)$ over shifts due to the biased initial condition μ_0 by setting its derivative with respect to μ_0 equal to zero. We use the fact that

$$\frac{d}{du}\text{erf}(u)=\frac{2}{\sqrt{\pi}}\exp\left(-u^2\right) \tag{27}$$

to compute

$$\frac{\partial\mathbb{E}(\Phi)}{\partial\mu_0}=\frac{r_1}{2}P'_+-\frac{r_2}{2}P'_-=\frac{2}{\pi}\left\{\frac{r_1}{2}\exp\left[-\frac{(\mu_0+a\bar{C}T)^2}{2\sigma^2T}\right]-\frac{r_2}{2}\exp\left[-\frac{(\mu_0-a\bar{C}T)^2}{2\sigma^2T}\right]\right\}\frac{1}{\sqrt{2\pi\sigma^2T}}=0$$

which implies that

$$\frac{r_1}{r_2}=\frac{\exp\left(-(\mu_0-a\bar{C}T)^2/(2\sigma^2T)\right)}{\exp\left(-(\mu_0+a\bar{C}T)^2/(2\sigma^2T)\right)}=\exp\left(\frac{2a\bar{C}\mu_0}{\sigma^2}\right),\quad\text{and thus}\quad\mu_0^{\text{opt}}=\frac{\sigma^2}{2a\bar{C}}\ln\left(\frac{r_1}{r_2}\right). \tag{28}$$

Note that the initial condition μ_0^{opt} does not depend on viewing time. For equal rewards $r_1=r_2$ we recover $\mu_0^{\text{opt}}=0$: an unbiased PMF with $P(0,T;0)=0.5$, and for a fixed reward ratio, μ_0^{opt} varies inversely with $\bar{C}$, approaching ∞ as $\bar{C}\to 0$. In this limit, of course, the stimulus contains no information and the best one can do is always choose the more lavishly rewarded alternative. Fig. 3 (left, solid blue curves) shows examples of μ_0^{opt} plotted as a function of reward ratio for three different coherence levels.

4.2 Case 1b

Coherences are mixed during blocks of trials in the Rorie-Newsome experiments, so we now consider a continuum idealization of this in which coherences are selected from a uniform distribution over $[C_1,C_2]$ (again positive for stimulus A and negative for B). Instead of summing the weighted probabilites of correct A and B choices for $\pm\bar{C}$, we must now integrate over the range of coherences:

$$\mathbb{E}(\Phi)=\frac{1}{(C_2-C_1)}\int_{C_1}^{C_2}\left[\frac{r_1}{2}P(+C,T;\mu_0)+\frac{r_2}{2}(1-P(-C,T;\mu_0))\right]dC. \tag{29}$$

Computing the derivative (using the Leibniz Rule and noting that the limits of integration do not depend on μ_0), we find:

$$\frac{\partial\mathbb{E}(\Phi)}{\partial\mu_0}=\frac{1}{2(C_2-C_1)}\int_{C_1}^{C_2}\left[r_1P'_+-r_2P'_-\right]\,dC=0\implies\frac{r_1}{r_2}=\frac{\int_{C_1}^{C_2}\exp\left[-\frac{(\mu_0-aCT)^2}{2\sigma^2T}\right]\,dC}{\int_{C_1}^{C_2}\exp\left[-\frac{(\mu_0+aCT)^2}{2\sigma^2T}\right]\,dC},$$

where we have cancelled common terms in the integrands that do not depend upon C, as in §4.1. To turn these expressions into standard error function integrals we change variables by setting $y=(\mu\pm aCT)/\sqrt{2\sigma^2T}$, $dy=\pm\frac{aT}{\sqrt{2\sigma^2T}}dC$. After cancelling more common terms, we obtain

$$\frac{r_1}{r_2}=-\frac{\text{erf}\left(\frac{\mu_0-aC_2T}{\sqrt{2\sigma^2T}}\right)-\text{erf}\left(\frac{\mu_0-aC_1T}{\sqrt{2\sigma^2T}}\right)}{\text{erf}\left(\frac{\mu_0+aC_2T}{\sqrt{2\sigma^2T}}\right)-\text{erf}\left(\frac{\mu_0+aC_1T}{\sqrt{2\sigma^2T}}\right)}. \tag{30}$$

Setting $C_2 = \bar{C} + \epsilon$, $C_1 = \bar{C} - \epsilon$ and expanding (30) in a Taylor series, and letting $\epsilon \to 0$, we recover the single coherence level result (28).

The expression (30) cannot be inverted to solve explicitly for μ_0^{opt} in terms of the the other parameters, but we may use it to plot the reward ratio r_1/r_2 as a function of μ_0 for fixed a, T, σ and coherence limits C_1, C_2. The axes of the resulting graph can then be switched to produce a plot of μ_0 vs. r_1/r_2, allowing us to compare (30) with the single coherence prediction (28). The red curves in Fig. 3 (left) show optimal initial conditions for narrow coherence bands $[\bar{C} - 0.1, \bar{C} + 0.1]$ centered around three fixed coherence levels (solid blue curves). Fig. 3 (right) shows optimal initial conditions for coherence bands of increasing width centered around $\bar{C} = 0.4$, the largest being the range $[0, 0.8]$. The initial condition value, and hence the optimal bias of the PMF, increases as the width of the coherence range increases. This is due to greater weighting of the biased reward information over the low coherence end of the range, for which accuracies approach 50%.

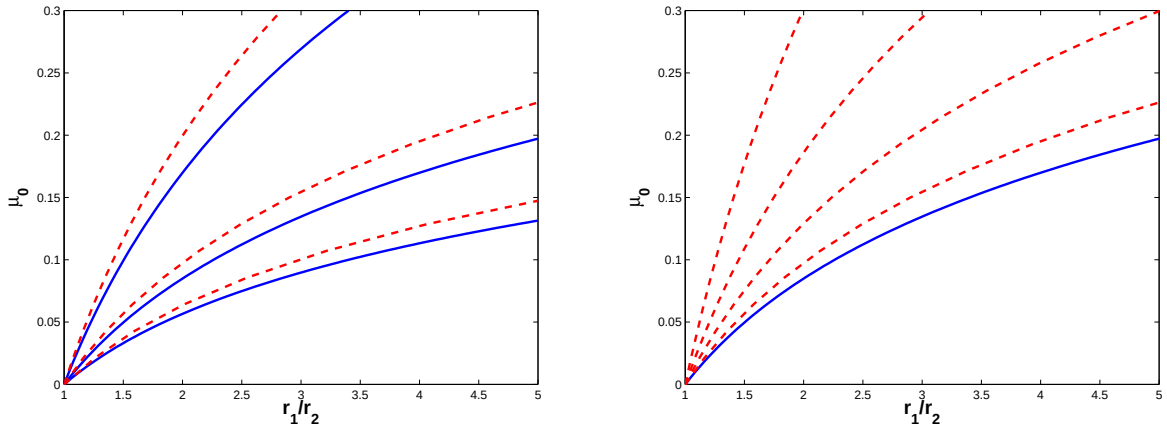

Figure 3: Optimal initial conditions μ_0 as a function of the reward ratio r_1/r_2 for the case of fixed coherence (solid blue curves) and ranges of coherence (dashed red curves). Left: $\bar{C} = 0.2$, 0.4 and 0.6, and $[C_1, C_2] = [0.1, 0.3]$, $[0.3, 0.5]$ and $[0.5, 0.7]$ (top to bottom for each case). Right: Coherence bands centered on $\bar{C} = 0.4$ (solid blue curve) with widths 0.2, 0.4, 0.6 and 0.8 (bottom to top). Other parameters are $a = 0.5$, $\sigma = 0.2214$, and $T = 30$.

In the case that there is a finite set of fixed nonzero coherences $\{\pm C_j,\ j = 1, \ldots N\}$, then the integral in (29) should be replaced by a finite sum. We consider this in the specific case of the Rorie-Newsome data in §5.1 below.

4.3 Case 2a

We now move to the case of continuing bias throughout the reward cue and motion periods, retaining $\lambda = 0$ and starting with fixed coherence $\bar{C}$. Decisions are now made at time $\tau + T$, and from §3.1 the probabilities of correct choices are:

$$P(\pm\bar{C}, T; \mu_0) = \frac{1}{2}\left[1 + \text{erf}\left(\frac{b(\tau + T) \pm a\bar{C}T}{\sqrt{2\sigma^2(\tau + T)}}\right)\right].$$

and the derivatives with respect to the bias parameter are:

$$\frac{\partial P_\pm}{\partial b} = \frac{\tau + T}{\sqrt{2\pi\sigma^2(\tau+T)}} \exp\left[-\frac{(b(\tau+T) \pm a\bar{C}T)^2}{2\sigma^2(\tau+T)}\right].$$

Using (24), a computation analogous to that of §4.1 yields

$$\frac{r_1}{r_2} = \frac{\exp\left(-(b(\tau+T) - a\bar{C}T)^2/(2\sigma^2(\tau+T))\right)}{\exp\left(-(b(\tau+T) + a\bar{C}T)^2/(2\sigma^2(\tau+T))\right)} \implies b^{\text{opt}} = \frac{\sigma^2}{2a\bar{C}T}\ln\left(\frac{r_1}{r_2}\right). \tag{31}$$

Note that bT, the bias integrated over the *motion period*, plays precisely the same role as the initial condition μ_0 in the formula (28).

4.4 Case 2b

As in §4.2 we can include a range of coherences $[C_1, C_2]$ by integrating the expected reward over that range. Proceeding as in §4.2 we obtain the relationship:

$$\frac{r_1}{r_2} = \frac{\int_{C_1}^{C_2} \exp\left[-\frac{(b(\tau+T)-aCT)^2}{2\sigma^2(\tau+T)}\right] dC}{\int_{C_1}^{C_2} \exp\left[-\frac{(b(\tau+T)+aCT)^2}{2\sigma^2(\tau+T)}\right] dC} = -\frac{\text{erf}\left(\frac{b(\tau+T)-aC_2T}{\sqrt{2\sigma^2(\tau+T)}}\right) - \text{erf}\left(\frac{b(\tau+T)-aC_1T}{\sqrt{2\sigma^2(\tau+T)}}\right)}{\text{erf}\left(\frac{b(\tau+T)+aC_2T}{\sqrt{2\sigma^2(\tau+T)}}\right) - \text{erf}\left(\frac{b(\tau+T)+aC_1T}{\sqrt{2\sigma^2(\tau+T)}}\right)}. \tag{32}$$

Note that both T and the total cue plus motion period $\tau + T$ enter this expression. As in §4.3, (32) is a close analogue of (30). Indeed, mapping the initial condition and noise terms in (30) according to

$$\mu_0 \mapsto b(\tau+T) \quad \text{and} \quad \sigma^2 T \mapsto \sigma^2(\tau+T),$$

Eqns. (30) and (32) are *identical*, so plots of optimal biases b^{opt} for the single coherence and coherence band cases would have the same forms as those in Fig. 3. However, the more complex ways in which the cue and motion period durations τ and T enter Eqns. (31-32) in comparison to T in Eqns. (28) and (30) might enable one to design experiments that distinguish which of the two biasing models provides a better description of behavior.

4.5 Case 3

This is the same as Case 2a but with $\lambda \neq 0$. The algebraic manipulations are essentially identical to those of §4.3, but now with

$$P(\pm\bar{C}, T; \mu_0) = \frac{1}{2}\left[1 + \text{erf}\left(\frac{b\left(e^{\lambda(\tau+T)} - 1\right) \pm a\bar{C}\left(e^{\lambda T} - 1\right)}{\sqrt{\sigma^2\lambda(e^{2\lambda(\tau+T)} - 1)}}\right)\right].$$

Setting $\frac{\partial \mathbb{E}(\Phi)}{\partial b} = 0$, and solving for $\frac{r_1}{r_2}$ as above, many terms again cancel, leading to

$$\frac{r_1}{r_2} = \exp\left(\frac{4ab\bar{C}(e^{\lambda T} - 1)}{\sigma^2\lambda(e^{\lambda(\tau+T)} + 1)}\right) \implies b = \frac{\sigma^2\lambda(e^{\lambda(\tau+T)} + 1)}{4a\bar{C}(e^{\lambda T} - 1)}\ln\left(\frac{r_1}{r_2}\right). \tag{33}$$

5 Preliminary fits to data

We now describe fits of PMFs predicted by the theory developed above to accuracy data kindly provided by Alan Rorie. Since all the data was collected in experiments performed with fixed viewing periods (specifically, a 250 ms reward cue period that preceded a 500 ms stimulus viewing period), we can neither distinguish between the two reward biasing models proposed in §1.2, nor can we determine whether the process is best described by a pure DD with constant drift A, an OU process ($\lambda \neq 0$), or whether the drift varies with time. To see this, return to §1.1 and note that the PMF depends only upon the ratio $\mu(T)/\sqrt{2\nu(T)}$, and that in §§2-3 the bias terms appear as additive factors in the numerator $\mu(T)$. Thus, if all parameters other than coherence C are fixed, and coherence appears linearly as assumed in §1.2, all of the various explicit expressions for $\mu(T)/\sqrt{2\nu(T)}$ that appear above can be written in the simple common form

$$\frac{\mu(T)}{\sqrt{2\nu(T)}} = b_1(C + b_2). \tag{34}$$

Here the parameters b_1 and b_2 respectively determine the slope and shift of the PMF, b_2 has the units of coherence, and both parameters depend upon the prameters $a, \sigma, \lambda, \mu_0, T, b$, and τ introduced above. Specifically, from the expressions (11) and (13) in the case of initial condition bias, we have

$$b_1 = \frac{a}{\sigma}\sqrt{\frac{T}{2}}, \quad b_2 = \frac{\mu_0}{aT}, \quad \text{and} \tag{35}$$

$$b_1 = \frac{a(e^{\lambda T} - 1)}{\sigma\sqrt{\lambda(e^{2\lambda T} - 1)}}, \quad b_2 = \frac{\mu_0 \lambda e^{\lambda T}}{a(e^{\lambda T} - 1)}, \tag{36}$$

and from §3.1 and Eqns. (19) the analogous expressions for bias applied during the cue and motion period are:

$$b_1 = \frac{aT}{\sigma\sqrt{2(\tau + T)}}, \quad b_2 = \frac{b(\tau + T)}{aT}, \quad \text{and} \tag{37}$$

$$b_1 = \frac{a(e^{\lambda T} - 1)}{\sigma\sqrt{\lambda(e^{2\lambda(\tau+T)} - 1)}}, \quad b_2 = \frac{b(e^{\lambda(\tau+T)} - 1)}{a(e^{\lambda T} - 1)}. \tag{38}$$

Experiments in which τ and T are varied independently could in principle distinguish among these four cases, but with the present data we can only fit the ratios b_1 and b_2.

5.1 Fits to PMFs

Rorie and Newsome have already performed fits of data collected over a set of discrete coherences, namely $C = 0, \pm 1.5\%, \pm 3\%, \pm 6\%, \pm 12\%, \pm 24\%, \pm 48\%$. They fitted accuracy data to the sigmiodal function

$$\tilde{P}(C) = \frac{1}{1 + \exp[-b_1(C + b_2)]}. \tag{39}$$

Here we repeat that fit under various constraining assumptions, and compare it with fits to the error function predicited by the OU/DD theory:

$$P(C) = \frac{1}{2}\left[1 + \mathrm{erf}\left(\frac{\mu(t)}{\sqrt{2\nu(t)}}\right)\right] = \frac{1}{2}\left[1 + \mathrm{erf}(b_1(C + b_2))\right]. \tag{40}$$

While the resulting values of shift b_2 are directly comparable (both of them being expressed in units of percent coherence), in comparing the slopes it is important to note that the values of b_1 will differ due to the properties of the two functions. Specifically, computing the derivatives of $\tilde{P}(C)$ and $P(C)$ at $C + b_2 = 0$ (50% choices of A), we find

$$\frac{\partial \tilde{P}}{\partial C}(0) = \frac{b_1}{4} \quad \text{and} \quad \frac{\partial P}{\partial C}(0) = \frac{b_1}{\sqrt{\pi}}, \tag{41}$$

so that, if the slopes of the fits are well matched, we can expect the ratio $b_1^{\text{sigmoid}}/b_1^{\text{erf}}$ to approximate $4/\sqrt{\pi} \approx 2.257$.

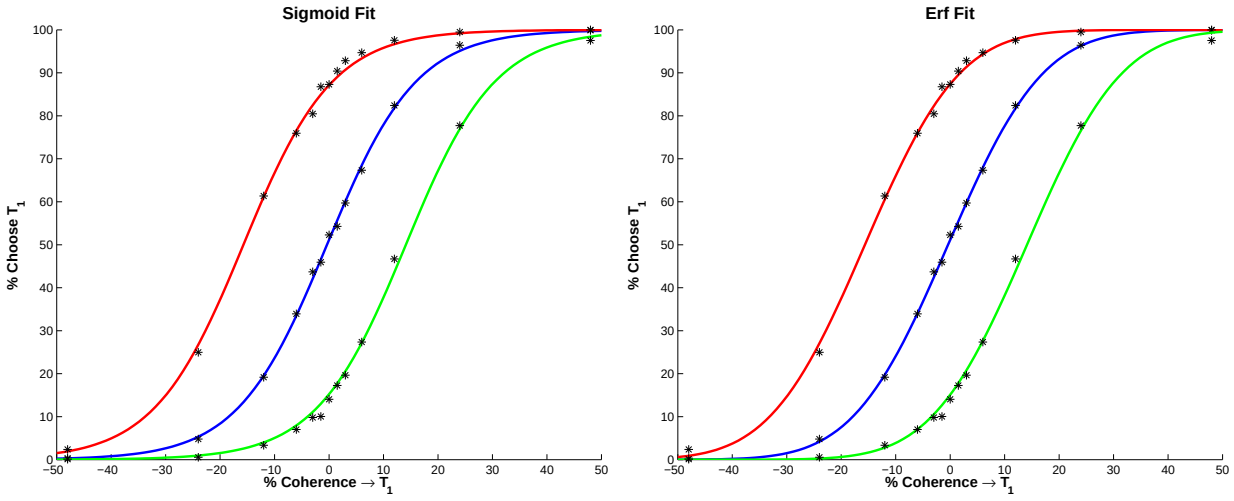

Figure 4: Fits of the Rorie-Newsome data to the sigmoidal function (39) (left) and the error function (40) (right). See text for details.

In Fig. 4 we show fits that were done as follows. Accuracy data from trials with both high and low unbiased rewards ($r_1 = r_2$) were pooled and fitted to determine b_1 and b_2; b_1 was then held fixed, and the biased reward data for $r_1/r_2 = 2$ and for $r_1/r_2 = 0.5$ were fitted separately to determine two values of the bias parameter b_2. Visually, the fits to both functions appear very close, but the residual errors (squared L^2 norm) are slightly smaller for the error function fits, and, using Eqns. (41) we find that the slopes at 50% correct are 0.0304 (sigmoid) and 0.0292 (erf): a difference of $\approx 4\%$. Numerical values are given in the top two rows of Table 1. Note that the values of b_2 are small but nonzero for $r_1 = r_2$ ($< 0.4\%$ coherence), indicating a slight preference for target 1. When this parameter is constrained to the unbiased value $b_2 = 0$ in the fit to equal reward data, the values of b_1 and of b_2 for the biased reward conditions listed in the lower two rows of Table 1 result. With one exception, they are unchanged to four significant figures.

In all cases the b_2 values for biased rewards show a significant asymmetry in that the leftward shift of the PMF for $r_1/r_2 = 2$ exceeds the rightward shift for $r_1/r_2 = 0.5$ by $1.6 - 1.7\%$. This asymmetry is substantially larger than the small effect indicated by the $b_2 \approx 0.4\%$ values for the equally rewarded cases.

Finally we refitted the biased reward data while allowing b_1 as well as b_2 to vary freely. We did this for the $r_1/r_2 = 2$ and $r_1/r_2 = 0.5$ cases separately and together (averaging the leftward

Function	b_1	b_2 for $r_1/r_2 = 1$	b_2 for $r_1/r_2 = 2$	b_2 for $r_1/r_2 = 0.5$
Sigmoid (39)	0.1217	0.3695 (.001)	15.71 (.002)	-14.11 (.0025)
Error Fn. (40)	0.0517	0.3761 (.00068)	15.65 (.0017)	-14.11 (.0023)
Sigmoid (39)	0.1217	0 (.0019)	15.72 (.002)	-14.11 (.0025)
Error Fn. (40)	0.0517	0 (.0016)	15.65 (.0017)	-14.11 (.0023)

Table 1: Parameter values for fits of the Rorie-Newsome data to the sigmoidal and error functions with b_1 held at the value determined for unbiased rewards. Residual fit errors shown in parentheses (squared L^2 norm).

and rightward shifts for the two cases, implicitly assuming no left/right bias). This produced the values shown in Table 2. Note that the values of b_1 change by less than 7.2% from those in Table 1, showing that the predominant effect of reward bias is a lateral shift of the PMF without a significant change in its slope.

Function	b_1, b_2 for $r_1/r_2 = 2$	b_1, b_2 for $r_1/r_2 = 0.5$	b_1, b_2 for both cases
Sigmoid (39)	0.1280, 15.39 (.0014)	0.1304, -13.66 (.0014)	0.1288, $\pm$14.53 (.0072)
Error Fn. (40)	0.0537, 15.39 (.0013)	0.0548, -13.72 (.0015)	0.0541, $\pm$14.56 (.0070)

Table 2: Parameter values for fits of the Rorie-Newsome data to the sigmoidal and error functions with b_1 and b_2 fitted separately for the unbiased and biased reward cases. Residual fit errors shown in parentheses (squared L^2 norm).

5.2 Comparison with optimal shifts

In the Rorie-Newsome experiment there is a finite set of fixed nonzero coherences $\{\pm C_j,\, j = 1, \dots N\}$, each of them, along with zero coherence, being presented with equal probability. Moreover, zero coherence stimuli (for which there is no correct answer), are rewarded with r_1 and r_2 with equal probability. The rewards expected for a full set of each stimulus are therefore:

$$\mathbb{E}(\Phi) = r_1 \sum_{j=1}^{N} P(b_1(+C_j + b_2)) + r_2 \sum_{j=1}^{N} [1 - P(b_1(-C_j + b_2))] + \frac{[r_1 P(b_1 b_2) + r_2(1 - P(b_1 b_2))]}{2}, \quad (42)$$

where we write the argument of P explicitly as $b_1(C + b_2)$ to indicate its dependence on coherence and the slope and bias parameters introduced in §5.1 This expression can be normalised to create a "harvesting efficiency" by dividing by the maximum possible reward per full set $(N + 0.5)(r_1 + r_2)$ [16], but as in §4.2 the optimal shift is determined by seeking zeros of the first derivative of (4.1) with respect to b_2:

$$\frac{\partial \mathbb{E}(\Phi)}{\partial b_2} = r_1 \sum_{j=1}^{N} \frac{\partial P}{\partial b_2}(b_1(+C_j + b_2)) - r_2 \sum_{j=1}^{N} \frac{\partial P}{\partial b_2}(b_1(-C_j + b_2)) + \frac{(r_1 - r_2)}{2} \frac{\partial P}{\partial b_2}(b_1 b_2) = 0. \quad (43)$$

This leads to the expression

$$\frac{r_1}{r_2} = \frac{\sum_{j=1}^{N} \frac{\partial P}{\partial b_2}(b_1(-C_j + b_2)) + \frac{1}{2}\frac{\partial P}{\partial b_2}(b_1 b_2)}{\sum_{j=1}^{N} \frac{\partial P}{\partial b_2}(b_1(+C_j + b_2)) + \frac{1}{2}\frac{\partial P}{\partial b_2}(b_1 b_2)}, \tag{44}$$

in which the partial derivatives may be evaluated for the error function OU/DD theory, using Eqns. (40) and (27):

$$\frac{\partial P}{\partial b_2}(b_1(C_j + b_2)) = \frac{b_1}{\sqrt{\pi}} \exp\left(-[b_1(C_j + b_2)]^2\right). \tag{45}$$

As for Eqn. (30) of §4.2 we cannot solve explicitly for b_2 in terms of the reward ratio and b_1, but we can again plot r_1/r_2 as a function of b_2 for fixed b_1 values, and invert the resulting graph. Taking the error function formula (40) with $b_1 = 0.0522$, the slope fitted to the equal reward data in Table 1, this relationship appears in Fig. 5(left). In this figure we also show the shifts b_2 computed from the biased reward data (triangles). Note that the animal "overshifts" her PMFs by $14-15\%$, compared to the optimal shift of $b_2^{\text{opt}} \approx 11.5\%$.

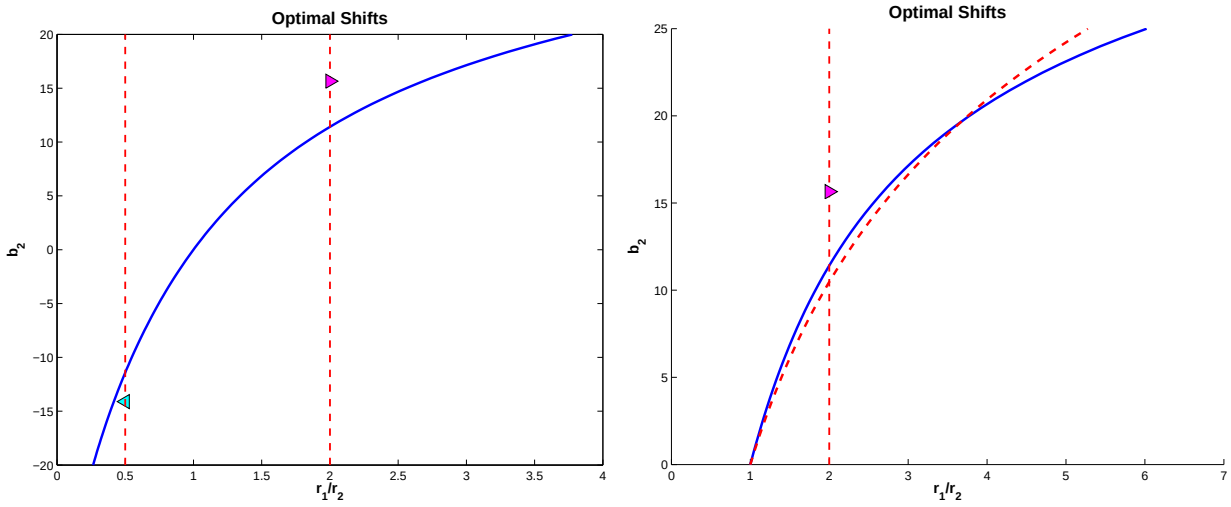

Figure 5: Left: Optimal biases b_2 for a range of reward ratios r_1/r_2 and $b_1 = 0.0517$, corresponding the slope of the error function PMF fitted to the Rorie-Newsome equal rewards data. The vertical lines at $r_1/r_2 = 0.5$ and 2 intersect the curve at the optimal shifts for those reward ratios, which are symmetrically placed ($b_2 \approx \pm 11.5\%$). The blue and pink triangles indicate the shifts determined from the biased rewards data with $r_1/r_2 = 0.5$ and $r_1/r_2 = 2$ respectively (cf. Table 1). Right: Comparison of optimal fits predicted by the error function (solid blue) and sigmoid (dashed red) fits for $r_1/r_2 \geq 1$. Note that the curves increasingly diverge for reward ratios > 4.

5.3 A comment on parameter sensitivity

Comparing the estimates of optimal shifts derived for the error function OU/DD theory with the fit to the sigmoid (39) computed by Rorie and Newsome [16], we note that the latter indicates an optimal shift of $7-8\%$ for reward ratios of $2:1$: substantially smaller than the $\approx 11.5\%$ found above. We repeated the computation of §5.1 with the sigmoid fit of §5.2 (replacing the partial derivative

(45) by the analogous expression for the sigmoid and using the value $b_1 = 0.1217$), and obtained a shift of $b_2^{\text{opt}} \approx 10.5\%$, smaller, but still substantially greater than $7 - 8\%$: see Fig. 5(right). We then examined the expected rewards funtion (42) itself and found that the second derivative at the maximum is small, so that the peak is rather "mild" and fairly small variations in b_1 can lead to substantial changes in b_2^{opt}. Fig. 6 shows non-normalized expected rewards curves for reward ratio $2:1$ and five different slopes b_1: the middle values 0.05 (for error function) and 0.12 (for sigmoid) being close to the fitted values $b_1 = 0.0517$ and $.b_1 = 0.1217$ of Table 1. In both cases, deviations of $\pm 5\%$ from b_2^{opt}, equivalent to increases and reductions in coherence of 43% and 55%, lead to reductions in expected rewards of less than 1.5%.

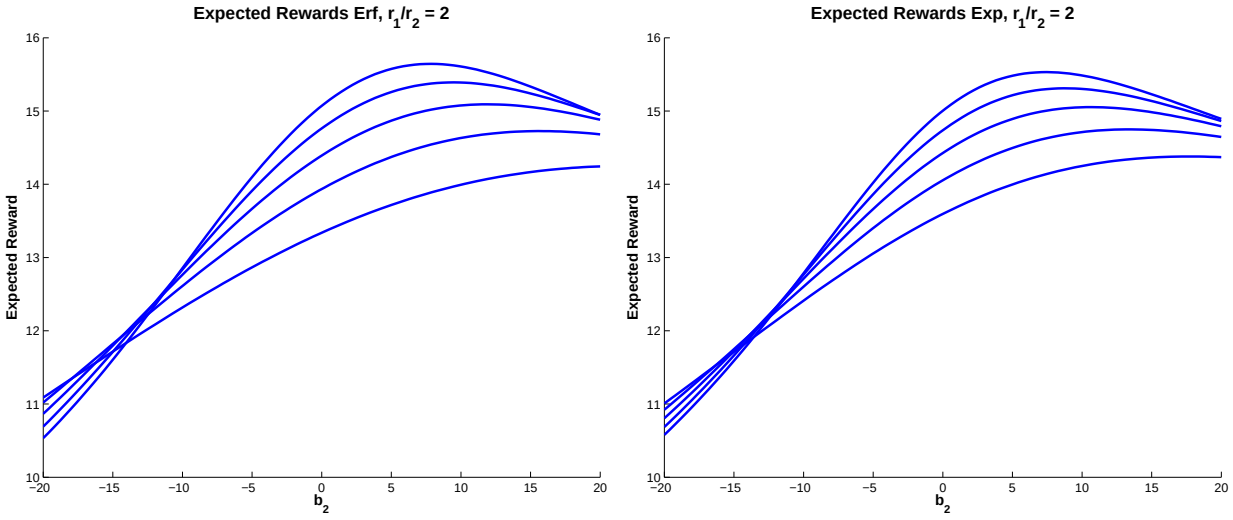

Figure 6: Expected rewards functions for reward ratio $r_1/r_2 = 2$. Left: Error function PMF with $b_1 = 0.03, 0.04, 0.05, 0.06, 0.07$. Right: Sigmoid function PMF with $b_1 = 0.08, 0.10, 0.12, 0.14, 0.16$. Curves become steeper as b_1 increases. Note the mild peaks for lower b_1 values.

As noted by Newsome, for unequal rewards the expected reward values decrease from their maxima more rapidly as b_2 falls below b_2^{opt} than they do for $b_2 > b_2^{\text{opt}}$. (In fact this asymmetric effect becomes stronger as the reward ratio increases. For equal rewards the expected rewards are even functions, symmetric about their maxima at $b_2 = 0$.) This provides a rationale for the animal's overshifting: she suffers a smaller loss than undershifting by the same amount. A similar observation appears in [1, pp 728-729] in connection with reward rate functions in a free response task. However, the above calculations indicate that, if PMF function slopes are relatively small, it may be difficult to determine optimal shifts with sufficient accuracy to judge how close animals are to performing optimally in this task. Taking the values from the error function fit ($b_2^{\text{opt}} \approx 11.5\%$), the animal, at $14 - 15\%$, is working within 1% of the best level of expected rewards, given her ability to discriminate the stimuli.

Finally we can ask if the data allows one to distinguish between the quality of the error function and sigmoid PMF fits. As noted in discussing Fig 4, the fit errors are slightly smaller for the former, and it has the advantage of being derivable from the drift-diffusion model: a (reasonably) well-accepted theory of evidence accumulation and decision making. It has the disadvantage of being a special function, and is therefore less amenable to analytic manipulation than the simple

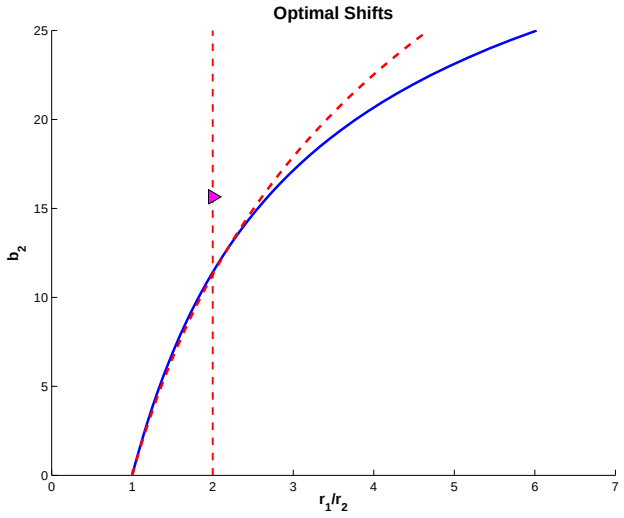

Figure 7: Comparison of optimal shifts for error function (solid blue) and sigmoid (dashed red) PMFs with slopes b_1 chosen to minimize their difference over the reward ratio range $[1, 2]$. Note the increasing divergence as the ratio grows beyond 2.

exponentials of the sigmoid. Given typical error bars on the experimental data, it would be interesting to compute measures of fit reliability. In this respect we note that the optimal shifts predicted by the two PMFs do diverge as the reward ratio increases, as already remarked in Fig. 5(right). Fig 7 indicates that, while one can choose slope parameters such that the optimal shifts are practically indistinguishable for reward ratios between 1 and 2, the increasing difference in their values may allow one to distinguish fit qualities for reward ratios of 4 or 5. Unfortunately, as the reward ratio increases, the maxima in expected rewards become even milder (plots not shown), so the lack of sensitivity to variations in b_2, illustrated in Fig. 6, grows. And here yet another problem rears its head: we have taken the animal's utility function as the (normalised) value of expected rewards, implicitly assuming that two drops of juice are worth twice one drop. The subjective utility may not vary linearly with reward size (it will probably saturate at very high reward ratios).

References

[1] R. Bogacz, E. Brown, J. Moehlis, 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. *Psychological Review*, 113 (4):700–765, 2006.

[2] K.H. Britten, M.N. Shadlen, W.T. Newsome, and J.A. Movshon. The analysis of visual motion: A comparison of neuronal and psychophysical performance. *J. Neurosci.*, 12(12):4745–4765, 1992.

[3] K.H. Britten, M.N. Shadlen, W.T. Newsome, and J.A. Movshon. Responses of neurons in macaque MT to stochastic motion signals. *Visual Neurosci.*, 10:1157–1169, 1993.

[4] P. Eckhoff, P. Holmes, C. Law, P.M. Connolly, and J.I. Gold. On diffusion processes with variable drift rates as models for decision making during learning. *New J. of Physics*, 10:doi:10.1088/1367–2630/10/1/015006, 2008.

[5] C.W. Gardiner. *Handbook of Stochastic Methods, Second Edition*. Springer, New York, 1985.

[6] J.I. Gold and M.N. Shadlen. Representation of a perceptual decision in developing oculomotor commands. *Nature*, 404:390–394, 2000.

[7] J.I. Gold and M.N. Shadlen. The influence of behavioral context on the representation of a perceptual decision in developing oculomotor commands. *J. Neurosci.*, 23(2):632–651, 2003.

[8] G.D. Horwitz and W.T. Newsome. Separate signals for target selection and movement specification in the superior colliculus. *Science*, 284 (5417):1158–1161, 1999.

[9] G.D. Horwitz and W.T. Newsome. Target selection for saccadic eye movements: prelude activity in the superior colliculus during a direction-discrimination task. *J. Neurophysiol.*, 86 (5):2543–2558, 2001.

[10] J.N. Kim and M.N. Shadlen. Neural correlates of a decision in the dorsolateral prefrontal cortex. *Nat. Neurosci.*, 2 (2):176–185, 1999.

[11] D.R.J. Laming. *Information Theory of Choice-Reaction Times*. Academic Press, New York, 1968.

[12] S.W. Link. *The Wave Theory of Difference and Similarity*. Erlbaum, Hillsdale, NJ, 1992.

[13] R.D. Luce. *Response Times: Their Role in Inferring Elementary Mental Organization*. Oxford University Press, New York, 1986.

[14] R. Ratcliff. A theory of memory retrieval. *Psych. Rev.*, 85:59–108, 1978.

[15] R. Ratcliff, T. Van Zandt, and G. McKoon. Connectionist and diffusion models of reaction time. *Psych. Rev.*, 106 (2):261–300, 1999.

[16] A. Rorie and W.T. Newsome. Unpublished notes and presentations. Department of Psychology, Stanford University, 2006.

[17] M.N. Shadlen and W.T. Newsome. Neural basis of a perceptual decision in the parietal cortex (area LIP) of the rhesus monkey. *J. Neurophysiology*, 86:1916–1936, 2001.

[18] P.L. Smith and R. Ratcliff. Psychology and neurobiology of simple decisions. *Trends in Neurosci.*, 27 (3):161–168, 2004.

[19] M. Usher and J.L. McClelland. On the time course of perceptual choice: The leaky competing accumulator model. *Psych. Rev.*, 108:550–592, 2001.

[20] A. Wald. *Sequential Analysis*. John Wiley & Sons, New York, 1947.

[21] A. Wald and J. Wolfowitz. Optimum character of the sequential probability ratio test. *Ann. Math. Statist.*, 19:326–339, 1948.