Published in final edited form as:
Neural Comput. 2008 February ; 20(2): 345–373. doi:10.1162/neco.2007.08-06-313.

A neural network model of the Eriksen task: Reduction, analysis, and data fitting

Yuan Sophie Liu[1], **Philip Holmes**[2], and **Jonathan D. Cohen**[3]

[1]Department of Physics, Princeton University, Princeton, NJ 08544, U.S.A.

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

[3]Department of Psychology, Princeton University, Princeton, NJ 08544, U.S.A.

Abstract

We analyze a neural network model of the Eriksen task, a two-alternative forced choice task in which subjects must correctly identify a central stimulus and disregard flankers that may or may not be compatible with it. We linearize and decouple the model, deriving a reduced drift-diffusion process with variable drift rate that describes the accumulation of net evidence in favor of either alternative, and we use this to analytically describe how accuracy and response time data depend on model parameters. Such analyses both assist parameter tuning in network models, and suggest explanations of changing drift rates in terms of attention. We compare our results with numerical simulations of the full nonlinear model and with empirical data, and show that it provides good fits to both with fewer parameters.

Keywords

connectionist; decoupling; drift-diffusion model; Eriksen task; free-response protocol; linearization; reduction

1 Introduction

There is currently considerable interest in the psychological (Laming, 1968; Ratcliff, 1978; Ratcliff et al., 1999) and neural bases of decision making (Platt and Glimcher, 2001; Schall, 2001; Gold and Shadlen, 2001, 2002). In the simplest, two-alternative forced-choice (2AFC) task, a subject must decide, on each trial, which of two randomly-presented stimuli has actually appeared (at the neural level). The discriminatory process is typically modelled as a competition among different populations of neurons, each preferentially responsive to one of the stimuli (Usher and McClelland, 2001). This is supported by direct recordings in oculo-motor areas of monkeys performing such tasks, which suggest that "decision" neurons (e.g. in the lateral interparietal area (LIP) and frontal eye field (FEF)) accumulate evidence for the stimulus alternatives, and the corresponding behavioral response is initiated when their firing rates cross thresholds, e.g. (Schall, 2001; Gold and Shadlen, 2001; Shadlen and Newsome, 2001; Schall et al., 2002; Roitman and Shadlen, 2002). Moreover, computational simulations and analyses of neural network (connectionist or parallel distributed processing (PDP)) models (Grossberg, 1988; Hopfield, 1982, 1984) show that their solutions can be matched to behavioral data (Cohen et al., 1990, 1992; Usher and McClelland, 2001).

yuanliu@princeton.edu, pholmes@math.princeton.edu, jdc@princeton.edu.

In particular, a simple and analytically-tractible *drift-diffusion model* (DDM) has been extensively fitted in this manner (Ratcliff, 1978; Ratcliff et al., 1999). The DDM is known to be optimal for 2AFC in the sense that, on average, it delivers choices of guaranteed accuracy in the shortest possible time (Laming, 1968), and analytical solutions for DDM error rates and decision times from the DDM can be used to investigate speed-accuracy tradeoffs for optimal performance (Gold and Shadlen, 2002; Bogacz et al., 2006). However, little work has examined the ability of this model to account for performance in more complex and cognitively-interesting tasks, such as those requiring selective attention to a subset of task-relevant stimuli in the presence of distractors. The Eriksen flanker task has been used extensively to study such effects (Eriksen and Eriksen, 1974; Gratton et al., 1988; Cohen et al., 1992).

In this task subjects are asked to respond appropriately to a target letter or arrow (e.g. < or >), visually displayed in the center of a five-symbol stimulus array on a display screen (e.g., by pressing the left button to < and the right button to >). The flanking symbols may be either *compatible* or *incompatible* with the central stimulus. In the compatible conditions, the display reads <<<<< or >>>>>; in the incompatible conditions, it reads >><>> or <<><<, and in each block of trials all four conditions are typically presented with equal probabilities.

Experiments show that subjects are slower and make more errors under the incompatible conditions, as illustrated in the the data of Gratton et al. (1988): see Figure 1. Furthermore, response patterns exhibit an interesting temporal profile: specifically, a dip in accuracy for incompatible trials at short reaction times, and a “crossover time” at which accuracy regains 50%: the chance level for “blind” responses. This dip is thought to reflect the dynamics of an interaction between bottom-up processing of sensory information (which, for the incompatible condition, favors the incorrect response) and the engagement of “top-down” attentional processes which favor processing of the central stimulus and thereby encourage the correct response. Accuracy for compatible trials increases montonically with time.

Trials may be run under a *free-response* paradigm in which decisions are signalled when the subject feels that sufficient evidence in favor of one alternative has accumulated. Since sensory processes are subject to variability, response times vary from trial to trial and performance under the free-response condition is characterised by both reaction time distributions and error rates. In contrast, in a *forced-response* or *deadline* paradigm, subjects must respond at or before a fixed time T following stimulus onset with their best estimate of which alternative was presented. This is how Figure 1 was generated. Reaction times may still vary due to errors in temporal estimation.

Here we shall consider both free response and the “hard limit” of forced response, in which the decision must be rendered when a cue is given: in this limiting case we can ignore RT variability and consider only accuracy as a function of the cue time. To distinguish the latter from deadlining, we call it the *interrogation protocol*. In both cases one can sort the data into response time bins and plot it as in Figure 1, but as we shall see, the two cases lead to somewhat different predictions.

Cohen et al. (1992) proposed a neural network model of the Eriksen task and showed that it can be fitted to the Gratton et al. (1988) data. The model has multiple layers and includes top-down biases applied to perception units associated with the central stimuli. However, like other connectionist models with nonlinear input-output response functions, it is not amenable to analysis, and data fitting and predictive studies must be carried out by numerical simulation, cf. Servan-Schreiber et al. (1998a,b). In this article we derive a simplified, linearized system and show how it can be reduced to a DDM with variable drift rate that models the decision process. In doing so we derive analytical approximations for crossover times and other

characteristics that assist parameter fitting, and reveal how the DDM emerges naturally from more complex multi-layered networks.

The paper is organised as follows. In Section 2 we briefly review connectionist network models, describe the model of Cohen et al. (1992), and analyze a linearised and decoupled version of it, which finally results in a DDM describing the evolution of net evidence in favor of one (or the other) alternative. A key ingredient is the time-varying input to this process from the perception and attention layers of the network, and we compare our analytical predictions of this with numerical simulations of the original nonlinear system. Section 3 contains analyses of the DDM with drift rates of various functional forms derived from simulation data. We compute accuracy vs response time curves for the interrogation protocol explicitly (displaying parameter dependencies), and for the free response protocol numerically, and compare them with simulations of the full Eriksen model. In Section 4 we compare fits of the original model of Cohen et al. (1992) and the reduced DDM to empirical data of Gratton et al. (1988). Our analysis assists parameter tuning in network models, and suggests explanations of variable drift rates in terms of attention, as noted in the summary discussion of Section 5.

2 Connectionist models

Connectionist models are stochastic differential equations (SDEs) or iterated mappings in which the activities (firing rates) of neurons or groups of neurons evolve in a manner determined by their summed inputs transformed via an activation or response function:

$$\text{output}=\psi\left(\sum_{i=1}^{n} w_i x_i + I\right), \tag{1}$$

where ψ is bounded to reflect the resting and maximal firing rates. Here x_i are the activities of other neurons or groups connected via weights w_i, negative values representing inhibition and positive values excitation. The term I models external inputs from stimuli, perhaps modulated via sensory circuits. Typically ψ is taken as the sigmoid

$$\psi(x)=\frac{1}{1+\exp(-4g(x-\beta))}, \tag{2}$$

with parameters g and β specifying gain and bias. Bias sets the input range $x \approx \beta$ in which the unit is maximally responsive, and gain determines its width. Ouside this region, the unit is essentially quiescent (output = 0) or maximally active (output = 1): see Figure 2. Appropriate parameter choices can effectively increase signal-to-noise ratios as information flows though a network, by amplifying larger inputs and suppressing smaller ones (Servan-Schreiber et al., 1990). Eqn. (2) is chosen so that the maximal slope $\partial\psi/\partial x(\beta) = g$, and it has been argued that neural circuits should adjust biases β to work near this point to utilize the resulting sensitivity (Cohen et al., 1990). As in earlier work (Brown et al., 2005), we shall appeal to this in linearizing response functions at their maximal slope regions to yield more tractable models.

Models such as those considered below explain a wide range of behavioral data (Cohen and Servan-Schreiber, 1992; Servan-Schreiber et al., 1998a,b; Aston-Jones and Cohen, 2005), and moreover may be derived from biophysically-based ionic current models of single cells (Abbott, 1991). Since the latter take the form of continuous differential equations rather than iterated maps, we shall focus on SDEs in this paper.

2.1 A connectionist model for the Eriksen task

We consider the architecture proposed by Cohen et al. (1992), shown in Figure 3. The units p_1 through p_6 constitute the perception layer (called the input module in Cohen et al. (1992)), z_1 and z_2 constitute the decision layer (output module), and a_1 through a_3 the attention module. Units within each layer or module are mutually interconnected via inhibitory weights $-w$ (here all assumed equal), implementing competition among representations within that layer. The decision and perception layers receive excitatory inputs of weights $+l$ and $+h$ from the perception and attention layers respectively, and the left, center and right units of the attention layer receive excitatory inputs of weights $+h$ from the corresponding pairs of perception units, as shown. All units are subject to independent additive white noise, to simulate unmodeled inputs. Absent inputs and noise, each unit's activity decays at rate $-k$.

The perception layer contains three pairs of units that receive inputs from the left, central, and right visual fields respectively and it is assumed that in each pair the left unit is preferentially responsive to the symbol < and the right unit to >. Thus, stimuli are modeled as follows: in the compatible condition, either the left (p_3) or right (p_4) central unit has external input $I_j = a$, and the corresponding flanker units (p_1, p_5 or p_2, p_6 resp.) receive input $I_j = b$ (modelling <<<<< or >>>>>); and, in the incompatible condition, either the left (p_3) or right (p_4) central unit has external input $I_j = a$, and the non-corresponding flanker units (p_2, p_6 or p_1, p_5 resp.) recieve input $I_j = b$ (modelling >><>> or <<><<). Since each flanker unit represents two symbols in the stimulus array, we typically assume $b \geq a$. The central unit (a_2) of the attention layer receives an input $A_2 = a_c$ in both conditions. All other inputs I_j, A_j are zero. See Figure 3 for an example.

Under the interrogation paradigm the decision is rendered at a set time t after stimulus onset by taking the larger of the decision unit outputs: i.e. "<" if $z_1(t) > z_2(t)$ and ">" if $z_2(t) > z_1(t)$. Under the free-response paradigm the first of $z_1(t)$ or $z_2(t)$ to cross preset thresholds $z_j = \theta$ determines the choice.

In this article we consider continuously evolving SDE models of these mechanisms, which for the Eriksen model may be written as:

$$\begin{aligned} \dot{z}_1 &= -kz_1 + \psi(-wz_2 + l(p_1 + p_3 + p_5)) + \eta_{a1}, \\ \dot{z}_2 &= -kz_2 + \psi(-wz_1 + l(p_2 + p_4 + p_6)) + \eta_{a2}; \end{aligned} \tag{3}$$

$$\begin{aligned} \dot{p}_1 &= -kp_1 + \psi(-w(p_2 + p_3 + p_4 + p_5 + p_6) + ha_1 + I_1) + \eta_{d1}, \\ \dot{p}_2 &= -kp_2 + \psi(-w(p_1 + p_3 + p_4 + p_5 + p_6) + ha_1 + I_2) + \eta_{d2}, \\ \dot{p}_3 &= -kp_3 + \psi(-w(p_1 + p_2 + p_4 + p_5 + p_6) + ha_2 + I_3) + \eta_{d3}, \\ \dot{p}_4 &= -kp_4 + \psi(-w(p_1 + p_2 + p_3 + p_5 + p_6) + ha_2 + I_4) + \eta_{d4}, \\ \dot{p}_5 &= -kp_5 + \psi(-w(p_1 + p_2 + p_3 + p_4 + p_6) + ha_3 + I_5) + \eta_{d5}, \\ \dot{p}_6 &= -kp_6 + \psi(-w(p_1 + p_2 + p_3 + p_4 + p_5) + ha_3 + I_6) + \eta_{d6}; \end{aligned} \tag{4}$$

$$\begin{aligned} \dot{a}_1 &= -ka_1 + \psi(-w(a_2 + a_3) + h(p_1 + p_2) + A_1) + \eta_{a1}, \\ \dot{a}_2 &= -ka_2 + \psi(-w(a_1 + a_3) + h(p_3 + p_4) + A_2) + \eta_{a2}, \\ \dot{a}_3 &= -ka_3 + \psi(-w(a_1 + a_2) + h(p_5 + p_6) + A_3) + \eta_{a3}, \end{aligned} \tag{5}$$

where the η_j's represent i.i.d. white noise processes.

This 11-dimensional, coupled set of SDEs is effectively insoluble analytically, so we shall employ two strategies that result in more tractible approximations: linearization and decoupling. In all, with the stimulus and attention input choices specified above, 11 parameters are required to specify the system (g and β for the sigmoids; w, l, h and k for leak and connection weights, a, b and a_c for inputs, a threshold value, and an overall noise level. Allowing different values in each layer would significantly increase this number. The reduction done below substantially reduces the number of parameters. Similar analyses of simpler models of the 2AFC task are developed in Brown et al. (2005) and Bogacz et al. (2006).

2.2 Decoupling, linearization, and the drift-diffusion process

Here we use two major ideas to simplify the problem: decoupling and linearization. We decouple the lower layers of the model by assuming that the modulatory output of the attention layer to the perception layer has a predetermined time course that is little-affected by feedback from the perception layer (the decision layer is already decoupled in the version of the model given above, in that it does not feed back to the lower layers). We then appeal to the proposal of Cohen et al. (1990): that biases in the sigmoidal units are adjusted so that they remain near their most sensitive ranges (close to maximum slope) where input modulations have maximal effect on outputs, and we replace the nonlinear functions (2) by their linearizations at $x = b$.

As already noted, the decision layer does not feed back to the perception or attention layers and thus cannot influence their dynamics. It may therefore be analyzed independently, given knowledge of, or assumptions regarding, the inputs $i_1 = l(p_1 + p_3 + p_5)$ and $i_2 = l(p_2 + p_4 + p_6)$ to its two units. Furthermore, assuming that the sigmoid bias parameter β in (2) is selected so that the units remain near their most sensitive range, we linearize (3) about $z_1 = z_2 = \beta$ and let $\bar{z}_j = z_j - \beta$ to obtain:

$$\begin{aligned} \dot{\bar{z}}_1 &= -k\bar{z}_1 - gw\bar{z}_2 + gi_1 + \eta_{a1}, \\ \dot{\bar{z}}_2 &= -k\bar{z}_2 - gw\bar{z}_1 + gi_2 + \eta_{a2}. \end{aligned} \tag{6}$$

Subtracting these equations yields a scalar Ornstein-Uhlenbeck process:

$$\dot{u} = (gw - k)\,u + A + \eta, \tag{7}$$

where $u = \bar{z}_1 - \bar{z}_2$ and $A = g(i_1 - i_2)$ is the difference in the inputs. If $A > 0$ ($i_1 > i_2$) u will tend to increase and if $A < 0$ ($i_1 < i_2$) it will tend to decrease: thus, in its linearised form, the decision layer integrates the net evidence from the perception layer.

When gw-$k = 0$ the first term on the right of equation (7) vanishes, and we say that such a network is *balanced* (Bogacz et al., 2006). In this case (7) is a pure drift-diffusion process, and is particularly simple to analyse, even when the net evidence $A(t)$ varies with time, as it will in the analysis to follow.

We will also formally decouple the perception layer from the attention layer, assuming that the feedback from the latter may be approximated by a specified time-dependent function. We initially neglect noise and inputs from the attention layer, so that after linearization, setting $\bar{p}_j = p_j - \beta$, and writing $\mathrm{p} = (\bar{p}_1, \ldots, \bar{p}_6)$, we have the linear ODE system:

$$\dot{\mathbf{p}} = \mathbf{A}\mathbf{p} + \mathbf{I}, \tag{8}$$

where the $n \times n$ matrix **A** and input vector **I** are:

$$\mathbf{A}=\begin{bmatrix} -k & -gw & -gw & -gw & -gw & -gw \\ -gw & -k & -gw & -gw & -gw & -gw \\ -gw & -gw & -k & -gw & -gw & -gw \\ -gw & -gw & -gw & -k & -gw & -gw \\ -gw & -gw & -gw & -gw & -k & -gw \\ -gw & -gw & -gw & -gw & -gw & -k \end{bmatrix}, \mathbf{I}=\begin{pmatrix} b \\ 0 \\ a \\ 0 \\ b \\ 0 \end{pmatrix} \text{or} \begin{pmatrix} 0 \\ b \\ a \\ 0 \\ 0 \\ b \end{pmatrix}. \tag{9}$$

Here we assume that the central stimulus is "<" and the components of the vector **I** respectively correspond to compatible and incompatible conditions. The ">" case may be derived using symmetry arguments. Also note that, since the inputs to the flanker units p_1, p_5 and p_2, p_6 are equal in both compatible and incompatible cases, solutions remain on the invariant 4-dimensional plane $p_1 = p_5$, $p_2 = p_6$, provided that the corresponding initial conditions are also equal.

The eigenvalues of the symmetric matrix **A** are $\lambda_1 = -(k+5gw)$ and $\lambda_2 = -(k - gw)$ with multiplicities 1 and 5 respectively, so we may diagonalize (8) by an othogonal transformation $\mathbf{p} = \mathbf{T}\mathbf{y}$, solve the resulting decoupled ODEs in the **y** coordinates and transform back to **p**, as detailed in the Appendix. In this way we may compute the sums that form the inputs to the decision layer in (3) and its linearization (7):

$$i_{1,2}=\frac{3y_1}{\sqrt{6}} \pm \left[\frac{2y_2}{\sqrt{2}} - \frac{y_3}{\sqrt{6}} + \frac{2y_4}{\sqrt{12}}\right]; \tag{10}$$

in writing (10) we have also used the symmetry $p_1 = p_5$, $p_2 = p_6$.

If $y_j(0) = 0$, corresponding to unbiased starting points, and a, b = const., the general solution given in the Appendix yields:

$$i_{1,2}=l\left[\frac{(a+2b)\left(e^{\lambda_1 t}-1\right)}{2\lambda_1} \pm \frac{(a \pm 2b)\left(e^{\lambda_2 t}-1\right)}{2\lambda_2}\right]. \tag{11}$$

In (11) the central ± refers to the cases i_1, i_2 and the right-hand ± to the compatible/incompatible conditions. The special case in which decay and inhibition are balanced $\left(k=gw, \lambda_1=-(1+5g)\,k \stackrel{\text{def}}{=} -\bar{\lambda}, \lambda_2=0\right)$ is of particular importance:

$$i_{1,2}=l\left[\frac{(a+2b)\left(1-e^{\bar{\lambda}t}\right)}{2\,\bar{\lambda}} \pm \frac{(a \pm 2b)\,t}{2}\right]. \tag{12}$$

Since the difference between the inputs to the decision layer is

$$i_1 - i_2=\frac{l\,(a \pm 2b)\left(e^{\lambda_2 t}-1\right)}{\lambda_2} \quad (\text{or} \quad l\,(a \pm 2b)\,t \quad \text{when} \quad \lambda_2=0), \tag{13}$$

the flanker inputs dominate in the incompatible case, as one expects (provided $b > a/2$). However, the inputs need not remain constant: in Cohen et al. (1992) and Servan-Schreiber et al. (1998a) the central perception units p_3, p_4 are activated on all trials via the output of the set of attention units shown in Figure 3. For simplicity we shall initially model this effect by boosting the central inputs $I_{3,4}$ to the perception layer by a multiplicative factor that increases linearly with time, replacing a thus:

$$a \mapsto (1+a_c t)\, a. \tag{14}$$

We shall partially justify and improve this simple choice subsequently by comparing with the decision layer inputs and outputs in the nonlinear model (3-5). Replacing a as in (14) and using the solutions given in the Appendix adds the following terms to the individual inputs:

$$i_{1,2}=...+laa_c\left[\frac{\left(e^{\lambda_1 t}-1-\lambda_1 t\right)}{2\lambda_1^2} \pm \frac{\left(e^{\lambda_2 t}-1-\lambda_2 t\right)}{2\lambda_2^2}\right], \tag{15}$$

so that the differences become

$$i_1 - i_2=l\left[\frac{(a \pm 2b)\left(e^{\lambda_2 t}-1\right)}{\lambda_2}+aa_c\frac{\left(e^{\lambda_2 t}-1-\lambda_2 t\right)}{\lambda_2^2}\right], \tag{16}$$

and

$$i_1 - i_2=l\left[(a \pm 2b)\, t+\frac{aa_c t^2}{2}\right] \tag{17}$$

in the balanced case ($\lambda_2 = 0$). The inputs are equal at $t = 0$, and $i_1 > i_2$ for $t > 0$ in the compatible case, but we see that there is now a critical *crossover time* such that $i_1 < i_2$ for $0 < t < t_{ci}$ and $i_1 > i_2$ for $t > t_{ci}$ in the incompatible case. With balanced parameters, we have

$$t_{ci}=\frac{2\,(2b-a)}{aa_c}, \tag{18}$$

and solving the noise-free drift equation (7) with $A = g(i_1 - i_2)$ of (17) and $u(0) = 0$, we find that

$$u\,(t)=l\left[\frac{(a \pm 2b)\, t^2}{2}+\frac{aa_c t^3}{6}\right]. \tag{19}$$

Hence for incompatible stimuli the output of the decision layer is negative for $0 < t < t_{co}$ and positive thereafter, and the crossover time for that output is given by

$$t_{co}=\frac{3\,(2b-a)}{aa_c}. \tag{20}$$

Figure 4 shows examples of decision layer inputs and outputs in the balanced case $\lambda_2 = 0$.

If the effect of attention is modeled by any *additive* term applied equally to both central decision units, so that in place of (14) we have $a \mapsto a + a_c(t)$ applied to p_3 and $0 \mapsto a_c(t)$ applied to p_4, we find that the difference between the inputs to the decision layer is unaffected. A nonlinear (*multiplicative*) interaction is evidently crucial. As we shall see, such an interaction emerges naturally from the nonlinear activation functions ψ of Figure 2.

2.3 Simulations with sigmoidal activation functions

We now return to the the more neurally-realistic network (3-5) with non-linear activation functions, and perform numerical simulations to validate the linearized analysis of §2.2. In order to make direct comparisons we again exclude noise terms. Here and in §3 we set parameters $k = w = l = h = 1$ and $a = b = 0.5$, $a_c = 1$ unless specified otherwise, and take sigmoidal gain and bias $g = 0.55$, $\beta = 0.8$ for the attention and perception units and $g = 1$, $\beta = -0.9$ for the decision units. The biases are selected so that, at rest without stimulus inputs, the units are close the centers of their sensitive ranges where $\phi' = g$, and we allow the units to equilibrate after starting with zero initial conditions before the stimuli are applied at $t = 0$. In §4 we shall derive parameters by fitting to the data of Figure 1.

Examples of solutions analagous to those of Figure 4 are shown in Figure 5, illustrating that the linearized analysis captures the key qualitative effects exhibited by the nonlinear system following stimulus onset, including crossover behavior in incompatible cases. In particular, the inputs from the central attention unit a_2 via the sigmoidal function have the effect of boosting the central stimulus inputs as assumed in our simple analysis. The quantitative predictions of the linear analysis are also adequate: in particular, the ratios between decision layer input and output crossover times fall within 10 - 15% of the value $r = t_{ci}/t_{co} = 2/3$ from (18-20), and the linear dependence of t_{ci} and t_{co} on $(2b - a)$ and inverse dependence on a_c of (18-20) are approximately borne out. Figure 6 shows solutions for incompatible cases for different values of the attention parameter a_c and of the ratio b/a, which measures the relative strength of flanker to center stimuli. Increasing a_c by a factor of 2 reduces t_{ci} from 4.43 to 1.84 and t_{co} from 6.11 to 2.75, and increasing b/a by 2 increases t_{ci} from 1.41 to 5.33 and t_{co} from 2.21 to 7.25, giving 0.42, 0.45 and 3.78, 3.28 respectively, in comparison to the factors 0.5 and 3 predicted by (18-20). (The crossover times are explicitly identified on Figure 5.)

Despite the reasonable match between data from the linearized and fully nonlinear systems, our linear growth assumption (14) for input to the central perception units, which leads to the quadratic and cubic functions of (17) and (19) is evidently too crude. In particular, the difference in inputs $i_1 - i_2$ to the decision layer departs significantly from (17) at large times, due to the limiting effect of the sigmoidal activation functions for large inputs, and it does not account for time delays due to the the fact that the attention units and stimuli are co-activated, and biases to the central perception units take some time to build up. More realistic inputs can be derived by examining noise-free simulation results, as we now show.

Figure 7 (left) shows the net input $i_1 - i_2$ from an incompatible trial in comparison to three analytical approximations: a linear expression

$$g(i_1 - i_2) = d_0 + d_1 t, \tag{21}$$

the quadratic expression (17), which we rewrite in the form

$$g(i_1 - i_2) = q_0 t + q_1 t^2, \tag{22}$$

and the exponential

$$g(i_1 - i_2) = a_0 + a_1 \quad \exp(a_2 t) + a_3 \quad \exp(a_4 t). \tag{23}$$

The former two compare well to the nonlinear system's response in the middle and in the early and middle time ranges respectively, but neither captures its asymptotic approach to a constant as t continues to increase (Figure 7). However, the exponential function (23) provides an excellent fit throughout. The specific parameter values obtained are:

$$\begin{aligned} \text{linear} \quad & : g(i_1 - i_2) \quad = -0.258 + 0.145t, \\ \text{quadratic} \quad & : g(i_1 - i_2) \quad = -0.254t + 0.1420t^2, \\ \text{exponential} \quad & : g(i_1 - i_2) \quad = 0.476 + 6.396 \quad \exp(-0.759t) - 6.906 \quad \exp(-0.659t). \end{aligned} \tag{24}$$

In the compatible case stimuli reinforce rather than compete, and the attention layer again further accentuates the contribution of the center units as time progresses, leading to the monotonically increasing function shown in Figure 7 (right), which can also be well-fitted by an exponential, in this case requiring only three parameters:

$$g(i_1 - i_2) = a_0 + a_1 \quad \exp(a_2 t), \tag{25}$$

the specific values being $a_0 = 0.934$, $a_1 = -0.787$ and $a_2 = -0.960$.

The number of parameters defining the exponential drift rates in the incompatible and compatible cases can each be reduced by one by requiring that the net input $g(i_1 - i_2) = 0$ at an appropriate reference time such as $t = 0$. This was not done for the fits noted above (since the quantities obtained from full simulations did not vanish at $t = 0$), but we use it in fitting empirical data in §4.

3 Analysis of the drift-diffusion decision process

We have already noted that, in case of balanced parameters ($w = gk$) the difference between the activities of the linearized decision layer units follows an Ornstein-Uhlenbeck or drift-diffusion process (cf. equations (6-7)). For convenience we rewrite the latter in Itô form:

$$du = [\lambda u + A(t)]\, dt + c\, dW; \tag{26}$$

here $\lambda = gw - k$, $A(t) = g(i_1 - i_2)$ is the time-varying drift rate, and c denotes the r.m.s. noise strength, assumed to be constant. We discuss psychological interpretations of $A(t)$ in §5.

3.1 The interrogation protocol

We first derive simple analytical expressions for accuracy in terms of response time and the system parameters introduced above, assuming that that responses are delivered under the interrogation protocol. Although this paradigm has not typically been used in empirical studies of performance in the Eriksen task, it provides an approximation of the deadlining procedures used to produce conditional accuracy curves of the sort shown in Figure 1. Such procedures are required to generate an adequate number of responses at short latencies and low accuracy.

To analyze accuracy in terms of response time under the interrogation protocol, we observe that the probability distribution $p(u, t)$ for solutions of (26), derived from the associated Fokker-Planck or forward Kolmogorov equation (Gardiner, 1985):

$$\frac{\partial p(u,t)}{\partial t} = -\frac{\partial}{\partial u}\left[(\lambda u + A(t))\,p(u,t)\right] + \frac{c^2}{2}\frac{\partial^2 p(u,t)}{\partial u^2}, \tag{27}$$

may be written in the form

$$p(u,t) = \frac{1}{\sqrt{2\pi v(t)}}\exp\left[\frac{-(u-\mu(t))^2}{2v(t)}\right], \tag{28}$$

where

$$\mu(t) = \mu_0 e^{\lambda t} + \int_0^t e^{\lambda(t-s)} A(s)\,ds \quad \text{and} \quad v(t) = v_0 e^{2\lambda t} + \frac{c^2}{2\lambda}\left(e^{2\lambda t} - 1\right) \tag{29}$$

denote the evolving mean and variance of $p(u, t)$, and we have assumed that initial conditions $u(0)$ p for (26) are drawn from a Gaussian distribution with mean μ_0 and variance $\sqrt{v_0}$. For a balanced system, $\lambda = 0$ and equations (29) become:

$$\mu(t) = \mu_0 + \int_0^t A(s)\,ds \quad \text{and} \quad v(t) = v_0 + c^2 t. \tag{30}$$

We model the interrogation protocol by assuming that, on a given trial, the subject chooses the alternative that seems more probable at time T. Thus, response > is given if $u(T) > 0$ and < is given if $u(T) < 0$, and if > is the correct alternative, the fraction of correct responses (accuracy) is given by:

$$P_{\text{correct}}(T) = \int_0^\infty p(u,t)\,du = \frac{1}{2}\left[1 + \operatorname{erf}\left(\frac{\mu(T)}{\sqrt{2v(T)}}\right)\right]. \tag{31}$$

To make specific comparisons, we shall assume that $\mu_0 = v_0 = 0$ (initial conditions are reset to $u(0) = 0$ for each trial, corresponding to an unbiased start). We additionally take $\lambda = 0$, corresponding to the optimal drift-diffusion process (Bogacz et al., 2006), and consistent with the parameter choices of Section 2.3 ($w = k = 1$ and $g = 1$ for the decision layer). Substituting the linear (21), quadratic (17), and exponential (23) expressions into the first expression of (30) and appealing to (31), we produce the accuracy vs. mean response time curves of Figure 8. These all exhibit an early dip in accuracy below 50%, as in the data of Gratton et al. (1988) (cf. Figure 1), and we may compute the times t_{min} at which their minima occur and the crossover times t_{50} at which accuracy regains 50%. The latter are given by $\mu(t_{50}) = 0$, leading to the explicit expressions

$$t_{50} = \frac{-2d_0}{d_1} \quad \text{and} \quad t_{50} = \frac{-3q_0}{2q_1}, \tag{32}$$

and the implicit one

$$a_0 t_{50} + \frac{a_1}{a_2}\left(e^{a_2 t_{50}} - 1\right) + \frac{a_3}{a_4}\left(e^{a_4 t_{50}} - 1\right) = 0, \tag{33}$$

in the linear, quadratic and exponential cases respectively. The minima occur at the (negative) turning points of the error function argument $\mu(t)/\sqrt{2}v(t)$, i.e.

$$t_{min}=\frac{-2d_0}{3d_1} \quad \text{and} \quad t_{min}=\frac{-9q_0}{10q_1}, \tag{34}$$

and the solution of

$$\left(a_0+2a_1e^{a_2t_{min}}+2a_3e^{a_4t_{min}}\right)t_{min}=\frac{a_1}{a_2}\left(e^{a_2t_{min}}-1\right)+\frac{a_3}{a_4}\left(e^{a_4t_{min}}-1\right). \tag{35}$$

The fact that the quadratic drift rate leads to the fastest approach to 100% accuracy is due to the rapid growth of that function as t increases. A linear fit, with strong negative drift at early times, leads to the lowest accuracy and latest crossover time. The exponential drift rate gives accuracies between these two cases over the early RT range, only falling below the linear drift case at larger times, when the exponential approaches its finite asymptote.

3.2 The interrogation protocol with bounded domain

While analytical solutions (28) are available for the unbounded DD process (26), this process is unrealistic in that magnitudes $|u(t)|$ can become arbitrarily large. In reality, neural firing rates and differences among them must remain bounded above and below. The original Eriksen model (3-5) respects this via the upper and lower asymptotes of its sigmoidal response functions (Figure 2), and we shall presently compare the above results with direct simulations of this nonlinear model, but one may also incorporate these bounds in a linear context, as we now show.

specifically, to compute the probability distribution with bounded $|u| \leq L$, (27) (with $\lambda = 0$) must be solved subject to no-flux boundary conditions:

$$-A(t)\,p(u,t)+\frac{c^2}{2}\frac{\partial p(u,t)}{\partial u}=0 \quad \text{at} \quad u=-L \quad \text{and} \quad u=L. \tag{36}$$

This can be done analytically via separation of variables for constant drift rates A, but we must resort to Monte Carlo simulations in the present case of time-varying drift. Figure 9 shows examples to illustrate the effect of increasing the boundary L from the small value of $c/3$ to the relatively large one of $5c/3$, illustrating convergence to the unbounded result of Figure 8 as L increases. Note that small L gives *higher* accuracy for mid-range RT responses than large L, but that this effect reverses for slower responses, since the asymptotic accuracy increases toward 1 as L increases. This is due to the fact that more sample paths that would have remained below 0 in the unbounded case (giving errors) are reflected from the lower boundary $u = -L$ and thereafter move above 0 under the influence of positive drift, than are reflected from $+L$ and subsequently move below 0. For example, averages over 10,000 simulated trials show that for $L = 0.1$ about 24% of the sample paths reflected from $+L$ cross and remain below 0, while for $L = 0.3$ and $L = 0.5$ the figures are 3% and 0.3% respectively.

In the limit $t \to \infty$, the exponential drift rate becomes constant $A(t) \to A = a_0$, and at long times the probability distribution $p(u, t)$ approaches the equilibrium solution of (27) with with $\lambda = 0$ and $A(t) \equiv A$ and boundary conditions (36):

$$p(u,t) \to p_{\text{eq}}(u) = \frac{2A\exp\left(\frac{2Au}{c^2}\right)}{c^2\left[\exp\left(\frac{2AL}{c^2}\right) - \exp\left(\frac{-2AL}{c^2}\right)\right]}. \tag{37}$$

The asymptotic accuracy is therefore

$$\lim_{t\to\infty} P_{\text{correct}}(t) = \int_0^L p_{\text{eq}}(u)\,du = \frac{1}{1+\exp\left(\frac{-2AL}{c^2}\right)} < 1. \tag{38}$$

Now $P_{\text{correct}}(t) \to 1$ as $L \to \infty$, but asymptotic accuracy is bounded for all finite L. The $t \to \infty$ limits are indicated for three values of L in Figure 9. In contrast, for the linear and quadratic drift rates of (21-22) that both grow without bound, accuracies approach 100% as t increases (not shown).

3.3 The free-response protocol

For the free-response protocol the appropriate setting is that of a first passage problem, as treated in considerable detail for both DDM and OU processes in Bogacz et al. (2006,Appendix A), cf. (Gardiner, 1985). Unfortunately, to our knowledge, neither explicit solutions nor asymptotic approximations are available for first passage problems with time-varying drift rates, and so we must again use numerical methods. In Figure 10 (left panel) we compare the analytical interrogation results of Section 3.1 with simulations of the full Eriksen model, and in the right panel we compare free response data from the Eriksen model with those of the DDM with exponential drift rate inputs as fitted in Figure 7, Eqns. (25-24). Parameter values for the simulations here and in Figure 11 were: $k = w = l = h = 1$, $a = b = a_c = 1$ and $g = 0.55$, $b = 0.8$ for the attention and perception units and $g = 1$, $b = -0.9$ for the decision units. With the exception of the stimulus strengths a and b, these are identical to those used for our earlier simulations and lead to a pure DDM with $\lambda = 0$. Threshold values are given in the figure captions.

The full model gives higher accuracy at early times and lower accuracy at later times than the unbounded DDM for incompatible stimuli in the interrogation case, much as does the bounded DDM of Figure 9, suggesting that the limiting nature of the sigmoidal response function, ignored in the unbounded linearised analysis, comes into play. The numerical free-response results from the linearized DDM exhibit leftward (time compression) shifts in both compatible and incompatible cases compared with the full simulations. The analytical interrogation results are closer to the full nonlinear simulations for incompatible trials, but a similar shift leads to overestimates of accuracy at early response times.

The free response accuracy results are rather sensitive to the choice of threshold in the DDM, which we based on the difference between the noise-free steady states of the decision layer outputs, as illustrated in Figures 5-6 (note that those figures were computed for parameters differing from those given directly above). specifically, noting that the DDM net evidence variable is $u = z_1 - z_2$ (Equations (6-7)), it follows that if z_1 and z_2 equilibrate to steady state levels $z_1^\infty > z_2^\infty$ as $t \to \infty$ in the noise-free simulation with stimulus 1 (<) applied, the appropriate 1 threshold for u must lie in the range $\left(0, z_1^\infty - z_2^\infty\right)$. Adjustments to match accuracy curves may be made in this range. Figure 11 shows the data from the full Eriksen model again, with DDM data obtained using a modified threshold that provides a better match. We also show reaction-time histograms from the full simulation and from the DDM. Note that, as in the experiments, reaction times are longer under the incompatible condition than the compatible condition, and

that the DDM provides reasonable estimates of RT distributions, especially in the incompatible case.

4 Comparisons with empirical data

In this section we compare results from the Eriksen task model of §2.1 and from the pure $\lambda = 0$ DDM of §3, working under the free response (threshold crossing) paradigm, with the data of Gratton et al. (1988). We reproduce the model fit described in Cohen et al. (1992), and perform a new fit to determine values of the parameters a_j describing the exponential drift rates for both compatible and incompatible stimuli, as in §2.3. Figures 12 and 13 show the resulting accuracy curves and RT distributions in comparison with the experimental data (cf. Figure 1 above). A visual inspection of Figure 12 shows that the DDM fits the accuracy data somewhat better then the full nonlinear model. We quantify and comment further on this below.

Data fits for the DDM were performed using the fmincon() function in MATLAB for comparison with the fits to the original connectionist model of Cohen et al. (1992). As the analyses of §2.3 and §3 predict, different exponentially-varying drift rates $A(t)$ are required for incompatible and compatible cases: the former having 5 and the latter 3 parameters (Equations (23) and (25)). To reduce the number of free parameters we required that $A(0) = 0$, thereby reducing these numbers to 4 and 2 respectively. The noise variance c and threshold θ add 2 more parameters, for a total of 8. These 8 parameters were determined by adjusting them while seeking minima of a fitting error function which averages over all the accuracy and reaction time data for compatible and incompatible trials.

The fitting error utilizes a weighted Euclidean norm. The usual Euclidean (L^2) distance between vectors $\mathbf{u}$ and $\mathbf{v}$ with components u_j and v_j is

$$\|\mathbf{u} - \mathbf{v}\| = \sqrt{(u_1 - v_1)^2 + (u_2 - v_2)^2 + \ldots + (u_n - v_n)^2}. \tag{39}$$

Accuracy and reaction time histogram vectors were first formed from the data ($\mathbf{AC}_d$, $\mathbf{RT}_d$) and model predictions ($\mathbf{AC}_m$, $\mathbf{RT}_m$) (cf. Figures 12-13) and their differences computed by (39). Since the units of accuracy and reaction time differ, each of these was then weighted by dividing it by the mean of the data, indicated by an overbar. This produces the nondimensional quantity:

$$\text{Error} = \sum_{\text{com,incom}} \left[\frac{\|\mathbf{AC}_d - \mathbf{AC}_m\|}{\|\bar{\mathbf{AC}}_d\|} + \frac{\|\mathbf{RT}_d - \mathbf{RT}_m\|}{\|\bar{\mathbf{RT}}_d\|} \right]. \tag{40}$$

This represents the sum of the percentage differences in accuracy and reaction time.

According to the error measure of (40), the DDM provides a 24% improvement over the fit obtained for the full connectionist model of Cohen et al. (1992). Moreover, this is achieved using 8 parameters in comparison with 11 for the connectionist model.

5 Discussion and conclusions

In this article we analyze a linearized version of the connectionist model for the Eriksen two-alternative forced-choice flanker task presented in Cohen et al. (1992) and Servan-Schreiber et al. (1998a). We show that, provided solutions remain within the central domain of the logistic function in which it may be approximated by a linear function that matches its slope g at the bias point β, as proposed by Cohen et al. (1990), analytical solutions of a decoupled, linearized

model modulated by a pre-determined attention signal can provide reasonable estimates of critical times at which evidence in favor of the correct and incorrect alternatives cross over for incompatible trials and hence reproduce the characteristic dip in accuracy for such trials. We also show that the dynamics of the two-unit decision layer can be decoupled and reduced to a drift-diffusion model (DDM) whose drift rate represents the net evidence for one alternative coming from the perception layer.

We then derive estimates of accuracy as a function of response time by interrogating a DDM with variable drift rates that are fitted to outputs from the perception layer of the fully nonlinear model. Collapsing to this model reduces the number of parameters from 11 or more in the connectionist model to 8 in the DDM with exponential drift rates. We compute the evolving probability distribution of solutions to the DDM and integrate it to obtain the psychometric function (% correct) as an explicit function of response time and the parameters defining the drift rate and noise strength. The interrogation protocol assumes that the response delivered reflects the subject's current estimate, and corresponds best to a deadlined task with a cued response.

We also consider a protocol under which subjects respond in their own time, modeled as a first passage problem. The qualitative forms of psychometric functions in the interrogation and free response cases are similar to those of the full nonlinear model for both compatible and incompatible trials, the latter showing the characteristic dip below chance for early responses. The DDM also produces acceptable approximations to accuracy and reaction time distributions derived from simulations of the full nonlinear model, and, more strikingly, it provides a slightly better fit to empirical data than does the full model, while using fewer parameters.

These results show that judicious linearization and decoupling of processing layers in connectionist models can allow analytical studies of how parameters influence the behaviour of such models. They also suggest that parameter tuning based on the explicit formulae available for the DDM interrogation protocol may be generally useful in matching model results with behavioral data. The key linearization step has been justified in model studies of 2AFC tasks (Brown et al., 2005), and extended to multiple alternative decision models (McMillen and Holmes, 2006). The range over which response functions ψ are well-approximated by their linearizations grows with the dynamic range of the neurons involved (the output range is normalized to 1 in Eqn. (2)). Decoupling is more problematic: a simple a priori assumption regarding biases due to attention does not produce realistic inputs from the perception layer to the decision layer or to the DDM, although such inputs can be derived from simulations of the full network, and here they are accurately fitted by simple exponential functions.

The present study also provides a foundation for further theoretical and experimental work. The DDM reduction reveals that interaction of the attention and perception layers produces *variable drift rates*, implying varying signal-to-noise ratios. These modulate the conjectured integration of evidence (in LIP) as attention is progressively engaged by top-down control. This interpretation is consistent with the assumption, within the DDM framework, that attention modulates drift rate.

These findings present an interesting challenge to the hypothesis that human performance in 2AFC tasks reflects the operation of optimal decision making processes. In a stationary environment such as the one modeled here, a constant drift-diffusion process produces optimal behavior (Bogacz et al., 2006). This contrasts with our observation of variable drift, suggesting that human performance in the Eriksen flanker task is not optimal. Such sub-optimality may reflect adaptive biases developed in response to a broader class of experience. For example, a Bayesian approach to analyzing optimality in this task (Yu et al., 2006) has shown that prior

expectations of compatible stimuli, or of correlations between neighboring elements in the visual field, can produce the observed dips in accuracy. Hence, experience of natural environments may have biased perception and decision systems to expect response-compatibility or perceptual similarity of nearby inputs. Work relating such expectations to the dynamics of processing indicate that such biases can be related directly to the variable drift rates incorporated in the present model, and used to account for empirical observations concerning performance (Yu et al., 2006; Liu et al., 2006).

In summary, the work reported here exemplifies how connectionist models of the dynamics of processing in cognitive tasks can be related to the DDM, which then provides an analytic framework for interpreting empirical observations. specifically, we have shown that a DDM reduction of a multi-layer model of the Eriksen task suggests that processing involves a progressive change in the drift rate over the course of a trial, reflecting the influence of top-down attentional mechanisms. This finding can be related, in turn, to an optimality analysis that generates new hypotheses about the factors governing attentional control (e.g., prior expectations of stimulus compatibility). The present work therefore provides links between theoretical analyses of optimal performance and formal specifications of the dynamics of processing mechanisms responsible for actual performance.

Acknowledgments

This work was supported by DoE grant DE-FG02-95ER25238 and PHS grants MH58480 and MH62196 (Cognitive and Neural Mechanisms of Conflict and Control, Silvio M. Conte Center). We thank the referees for their helpful comments.

Appendix

Appendix

We observe that an $n \times n$ matrix **A** with diagonal entries $-k$ and off-diagonal entries gw has two eigenvalues

$$\lambda_1 = -(k + (n-1)gw), \quad \lambda_2 = gw - k, \tag{41}$$

with multiplicities 1 and $n - 1$ respectively, and since **A** is symmetric, $n - 1$ mutually orthogonal eigenvectors belonging to λ_2 can be found. Along with the eigenvector $(1, 1, \ldots, 1)$ of λ_1, these yield an orthonormal transformation **T** with $\mathbf{T}^{-1} = \mathbf{T}^{\mathrm{T}}$ with the latter given explicitly by:

$$y_1 = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} x_i; \quad y_j = \sqrt{\frac{j-1}{j}} \left[\frac{1}{j-1} \sum_{i=1}^{j-1} x_i - x_j \right], \quad 2 \le j \le n. \tag{42}$$

For $n = 6$ the orthonormal eigenvector matrix is:

$$\mathbf{T} = \begin{bmatrix} 1/\sqrt{6} & 1/\sqrt{2} & 1/\sqrt{6} & 1/\sqrt{12} & 1/\sqrt{20} & 1/\sqrt{30} \\ 1/\sqrt{6} & -1/\sqrt{2} & 1/\sqrt{6} & 1/\sqrt{12} & 1/\sqrt{20} & 1/\sqrt{30} \\ 1/\sqrt{6} & 0 & -2/\sqrt{6} & 1/\sqrt{12} & 1/\sqrt{20} & 1/\sqrt{30} \\ 1/\sqrt{6} & 0 & 0 & -3/\sqrt{12} & 1/\sqrt{20} & 1/\sqrt{30} \\ 1/\sqrt{6} & 0 & 0 & 0 & -4/\sqrt{20} & 1/\sqrt{30} \\ 1/\sqrt{6} & 0 & 0 & 0 & 0 & -5/\sqrt{30} \end{bmatrix}, \tag{43}$$

and (8) transforms to the uncoupled system $\dot{\mathbf{y}} = \mathbf{T}^{\mathrm{T}}\mathbf{A}\mathbf{T}\mathbf{y} + \mathbf{T}^{\mathrm{T}}\mathbf{I}$:

$$\begin{aligned}
\dot{y}_1 &= \lambda_1 y_1 + \frac{(a+2b)}{\sqrt{6}},\\
\dot{y}_2 &= \lambda_2 y_2 + \frac{b}{\sqrt{2}} / - \frac{b}{\sqrt{2}},\\
\dot{y}_3 &= \lambda_2 y_3 + \frac{(b-2a)}{\sqrt{6}},\\
\dot{y}_4 &= \lambda_2 y_4 + \frac{(a+b)}{\sqrt{12}},\\
\dot{y}_5 &= \lambda_2 y_5 + \frac{a-3b}{\sqrt{20}} / + \frac{a+b}{\sqrt{20}},\\
\dot{y}_6 &= \lambda_2 y_6 + \frac{(a+2b)}{\sqrt{30}} / + \frac{(a-4b)}{\sqrt{30}},
\end{aligned} \tag{44}$$

where the first of each alternative additive term corresponds to compatible stimuli and the second to incompatible stimuli, and the single terms in components 1,3 and 4 apply to both.

The initial value problem

$$\dot{y} = \lambda y + f(t), y(0) = y_0 \tag{45}$$

has the solution:

$$y(t) = y_0 e^{\lambda t} + \int_0^t e^{\lambda(t-s)} f(s)\, ds, \tag{46}$$

and in case that $y_0 = 0$ and $f(t) = At + B$, we have

$$y(t) = \frac{1}{\lambda^2}\left[A\left(e^{\lambda t} - 1 - \lambda t\right) + \lambda B\left(e^{\lambda t} - 1\right)\right]. \tag{47}$$

In the case of balanced parameters ($k = w \Rightarrow \lambda_2 = 0$), (47) becomes:

$$y(t) = \frac{At^2}{2} + Bt. \tag{48}$$

Equipped with these solutions of (45), we compute $\mathbf{p} = \mathbf{T}\mathbf{y}(t)$ and sum the appropriate components to obtain Equations (11-17) of §2.2. In doing so we also appeal to fact that the 4-dimensional subspace $p_1 = p_5$, $p_2 = p_6$ is invariant, implying, via $\mathbf{y} = \mathbf{T}^{\mathrm{T}}\mathbf{y}$, that solutions started at $y_j(0) = 0$ satisfy

$$\frac{5y_5}{\sqrt{20}} = -\left(\frac{y_2}{\sqrt{2}} + \frac{y_3}{\sqrt{6}} + \frac{y_4}{\sqrt{12}}\right), \frac{6y_6}{\sqrt{30}} = \frac{1}{5}\left(\frac{6y_2}{\sqrt{2}} - \frac{4y_3}{\sqrt{6}} - \frac{4y_4}{\sqrt{12}}\right). \tag{49}$$

References

Abbott, L. Firing-rate models for neural populations. In: Benhar, O.; Bosio, C.; Del Giudice, P.; Tabat, E., editors. Neural Networks: from Biology to High-Energy Physics. ETS Editrice; Pisa: 1991. p. 179-196.

Aston-Jones G, Cohen J. An integrative theory of locus coeruleus-norepinephrine function:Adaptive gain and optimal performance. Ann. Rev. Neurosci 2005;28:403–450. [PubMed: 16022602]

Bogacz R, Brown E, Moehlis J, Holmes P, Cohen J. The physics of optimal decision making: A formal analysis of models of performance in two alternative forced choice tasks. Psychological Review 2006;113(4):700–765. [PubMed: 17014301]

Brown E, Gao J, Holmes P, Bogacz R, Gilzenrat M, Cohen J. Simple neural networks that optimize decisions. Int. J. Bifurcation and Chaos 2005;15:803–826.

Cohen J, Dunbar K, McClelland J. On the control of automatic processes: A parallel distributed processing model of the Stroop effect. Psych. Rev 1990;97(3):332–361.

Cohen J, Servan-Schreiber D. Context, cortex and dopamine:A connectionist approach to behavior and biology in schizophrenia. Psych. Rev 1992;99:45–77.

Cohen J, Servan-Schreiber D, McClelland J. A parallel distributed processing approach to automaticity. American Journal of Psychology 1992;105:239–269. [PubMed: 1621882]

Eriksen B, Eriksen C. Effects of noise letters upon the identification of a target letter in a non-search task. Perception and Psychophysics 1974;16:143–149.

Gardiner, C. Handbook of Stochastic Methods. Vol. Second Edition. Springer; New York: 1985.

Gold J, Shadlen M. Neural computations that underlie decisions about sensory stimuli. Trends in Cognitive Science 2001;5(1):10–16.

Gold J, Shadlen M. Banburismus and the brain: decoding the relationship between sensory stimuli, decisions, and reward. Neuron 2002;36:299–308. [PubMed: 12383783]

Gratton G, Coles M, Sirevaag E, Eriksen C, Donchin E. Pre- and poststimulus activation of response channels: a psychophysiological analysis. J. Exp. Psychol. Hum. Percept. Perform 1988;14:331–344. [PubMed: 2971764]

Grossberg S. Nonlinear neural networks: principles, mechanisms, and architectures. Neural Networks 1988;1:17–61.

Hopfield J. Neural networks and physical systems with emergent collective computational abilities. Proc. Natl. Acad. Sci. USA 1982;79:2554–2558. [PubMed: 6953413]

Hopfield J. Neurons with graded response have collective computational properties like those of two-state neurons. Proc. Natl. Acad. Sci. USA 1984;82:3088–3092. [PubMed: 6587342]

Laming, D. Information Theory of Choice-Reaction Times. Academic Press; New York: 1968.

Liu, Y.; Yu, A.; Holmes, P. Dynamical analysis of Bayesian inference models for the Eriksen task. Center for the Study of Brain, Mind and Behavior, Princeton University; 2006. Preprint

McMillen T, Holmes P. The dynamics of choice among multiple alternatives. J. Math. Psych 2006;50:30–57.

Platt M, Glimcher P. Neural correlates of decision variable in parietal cortex. Nature 2001;400:233–238. [PubMed: 10421364]

Ratcliff R. A theory of memory retrieval. Psych. Rev 1978;85:59–108.

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

Roitman J, Shadlen M. Response of neurons in the lateral intraparietal area during a combined visual discrimination reaction time task. J. Neurosci 2002;22:9475–9489. [PubMed: 12417672]

Schall J. Neural basis of deciding, choosing and acting. Nature Reviews in Neuroscience 2001;2:33–42.

Schall J, Stuphorn V, Brown J. Monitoring and control of action by the frontal lobes. Neuron 2002;36:309–322. [PubMed: 12383784]

Servan-Schreiber D, Bruno R, Carter C, Cohen J. Dopamine and the mechanisms of cognition: Part I. A neural network model predicting dopamine effects on selective attention. Biological Psychiatry 1998a;43:713–722. [PubMed: 9606524]

Servan-Schreiber D, Bruno R, Carter C, Cohen J. Dopamine and the mechanisms of cognition: Part II. D-Amphetamine effects in human subjects performing a selective attention task. Biological Psychiatry 1998b;43:723–729. [PubMed: 9606525]

Servan-Schreiber D, Printz H, Cohen J. A network model of catecholamine effects: Gain, signal-to-noise ratio, and behavior. Science 1990;249:892–895. [PubMed: 2392679]

Shadlen M, Newsome W. Neural basis of a perceptual decision in the parietal cortex (area LIP) of the rhesus monkey. J. Neurophysiology 2001;86:1916–1936.

Usher M, McClelland J. On the time course of perceptual choice: The leaky competing accumulator model. Psych. Rev 2001;108:550–592.

Yu, A.; Cohen, J.; Dayan, P. A Bayesian view of sensory conflicts in decision-making. Center for the Study of Brain, Mind and Behavior, Princeton University; 2006. Preprint

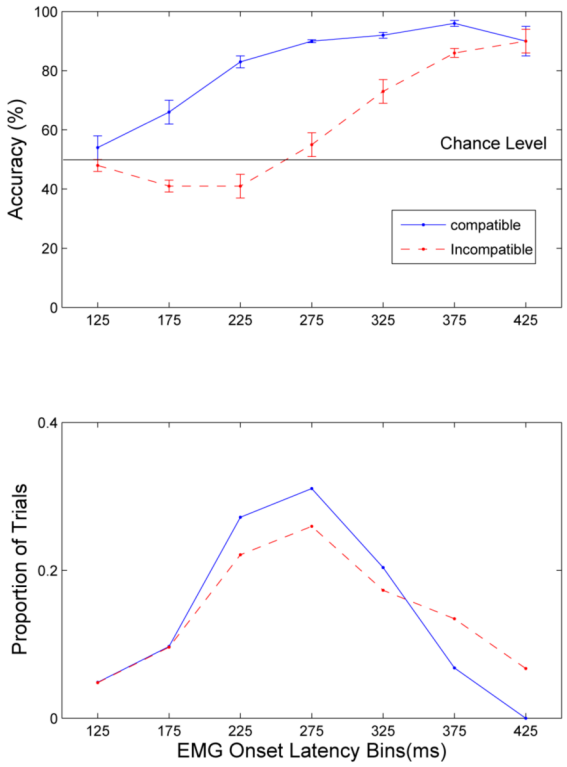

Figure 1.
(Top) Accuracy vs. response time as determined by electromyo-graphic activity (EMG) for compatible (solid) and incompatible (dashed) stimuli; standard errors shown by vertical bars. (Bottom) EMG onset histograms. Data replotted from Gratton et al. (1988, Fig. 1).

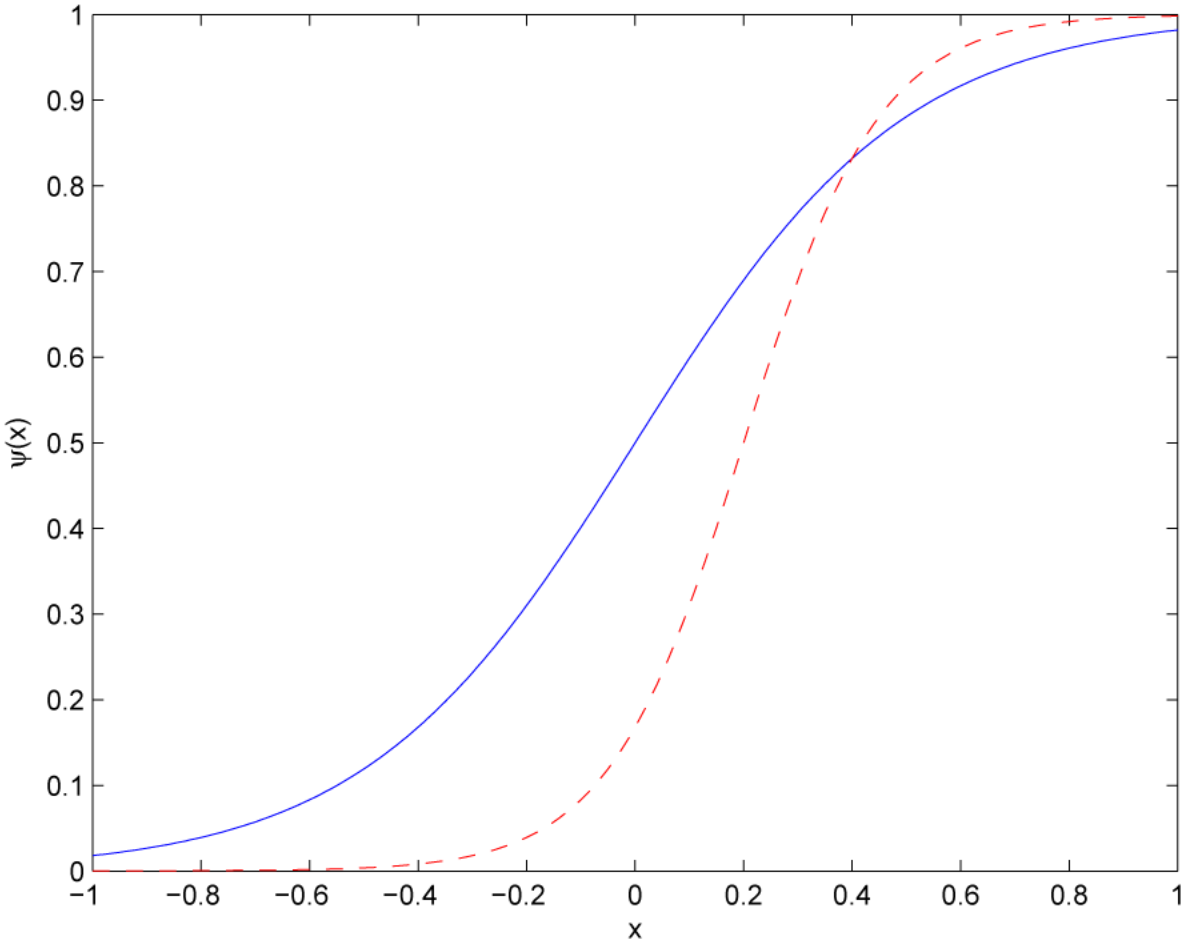

Figure 2.
Logistic activation functions, showing the effects of gain g and bias β. Bias sets the center of the input range over which the response is approximately linear, and gain sets the size of this range. Solid blue curve: $g = 1$, $\beta = 0$; dashed red curve: $g = 2$, $\beta = 0.2$.

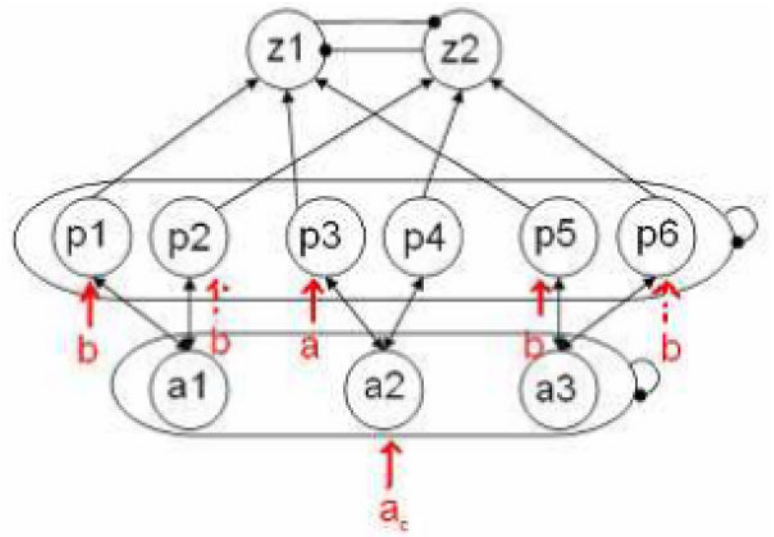

Figure 3.
The architecture of the Eriksen Model, showing inputs corresponding to stimulus < in the center with compatible (solid) and incompatible (dashed) flanking stimuli. From Cohen et al. (1992).

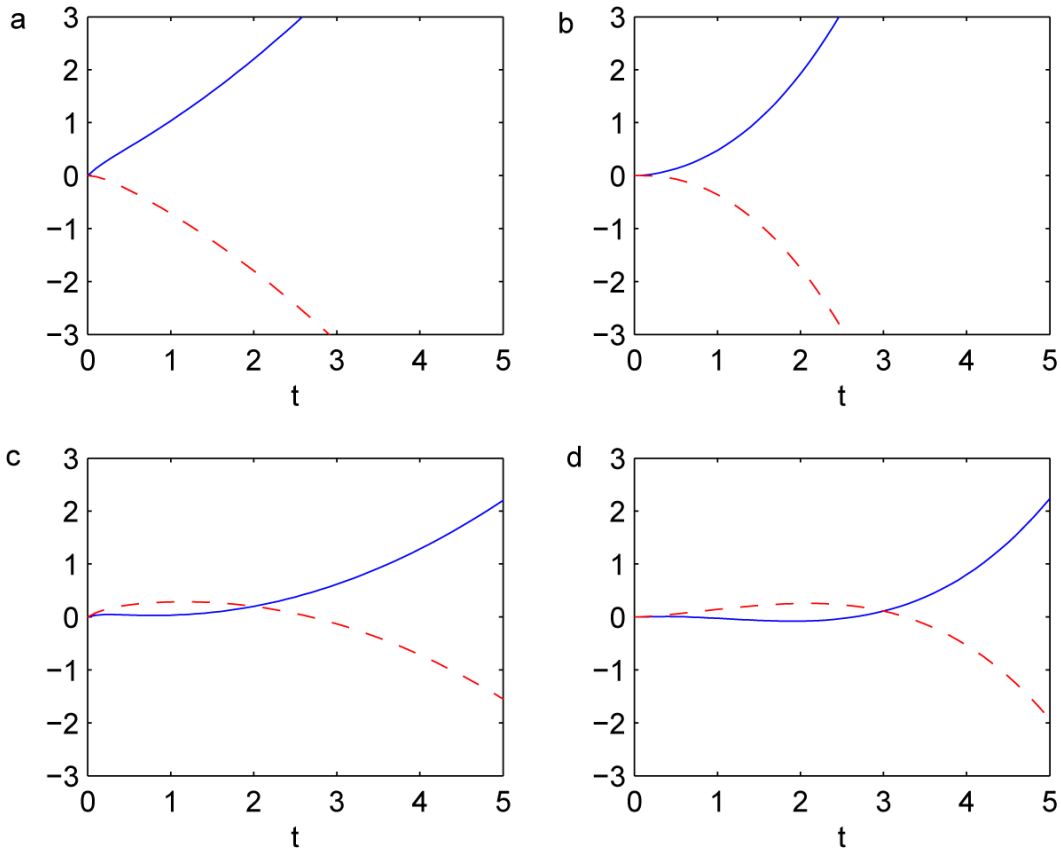

Figure 4.

Analytical solutions of the linearized, noise-free Eriksen model. Decision layer input (a) and output (b) for compatible stimuli, and the same functions for incompatible stimuli (c,d). Solid blue curves indicate i_1 and z_1: correct response; dashed red curves indicate i_2 and z_2: incorrect response. Here $a = b = a_c = 1$ and crossover times $t_{ci} = 2$, $t_{co} = 3$.

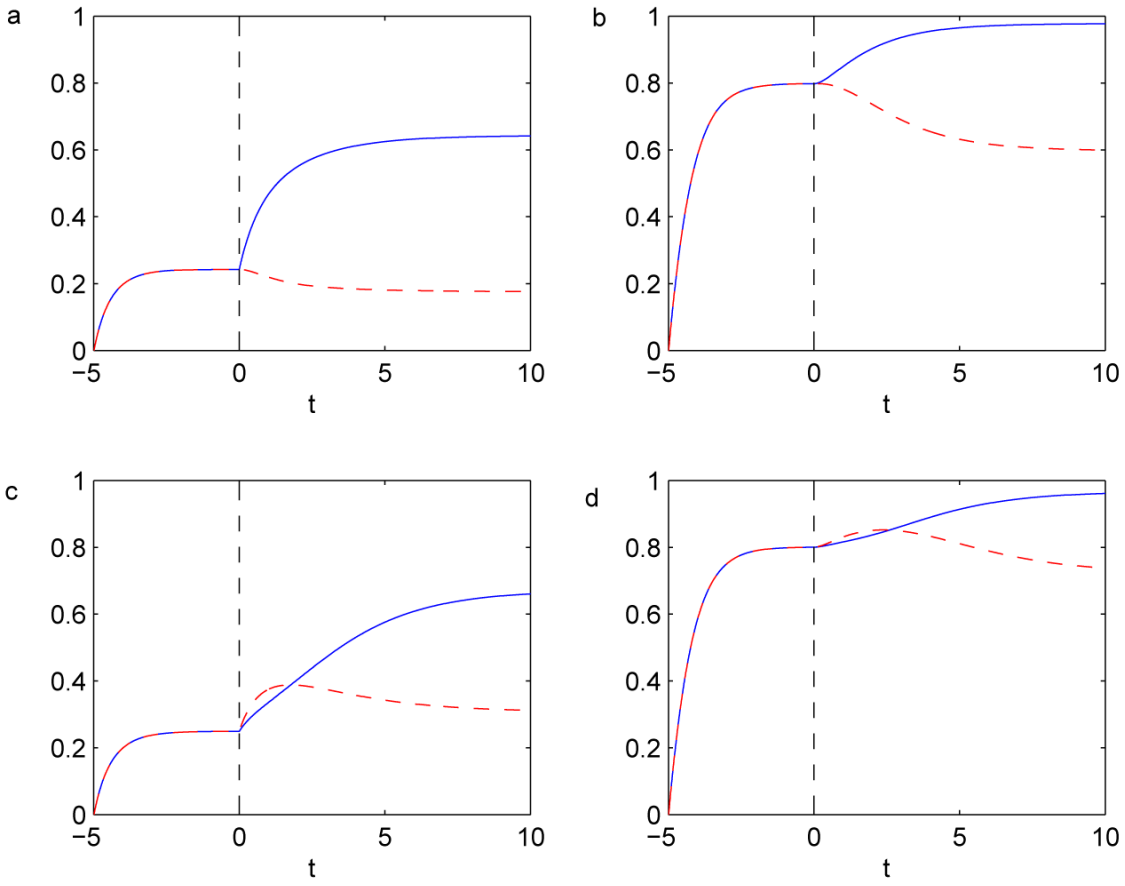

Figure 5.

Simulations of the noise-free Eriksen model with logistic activation functions. Decision layer inputs i_j (a) and outputs z_j (b) for compatible stimuli, and for incompatible stimuli (c,d): crossovers in (c,d) cause the dip in accuracy. Solid blue curves indicate i_1 and z_1: correct response; dashed red curves indicate i_2 and z_2: incorrect response. Stimuli are applied at $t = 0$, after units have settled at resting values. Parameters are specified in text.

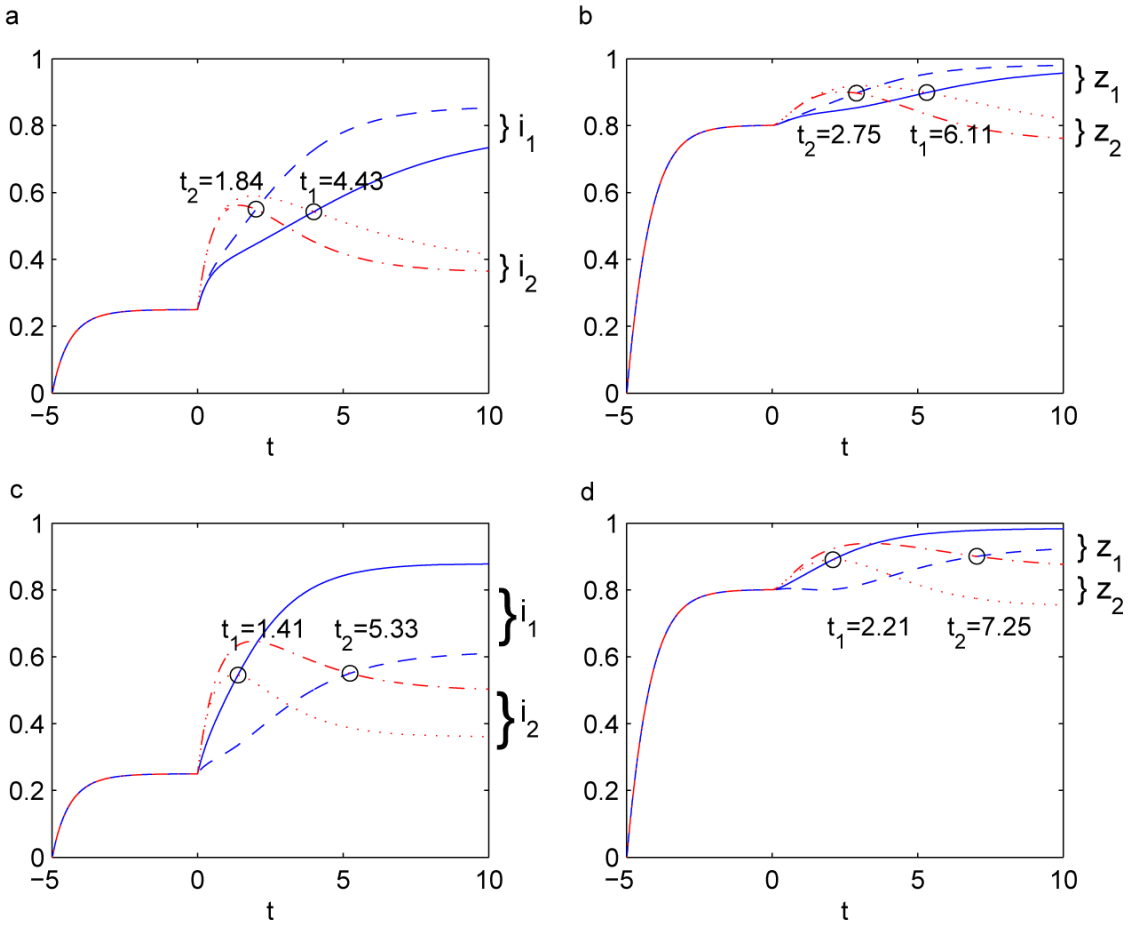

Figure 6.

Simulations of the noise-free Eriksen model with logistic activation functions showing decision layer inputs i_1 and outputs z_1 (solid and dashed blue) and i_2 and z_2 (dash-dotted and dotted red); inputs in panels (a,c) and outputs in (b,d). Panels (a,b) show effect of attention $a_c = 0.5$ (solid and dotted) and $a_c = 1$ (dash-dotted and dotted), and panels (c,d) show effect of the ratio $b/a = 1$ (solid and dotted) and $b/a = 2$ (dashed and dash-dotted). See text for discussion. Other parameters as for Figure 5.

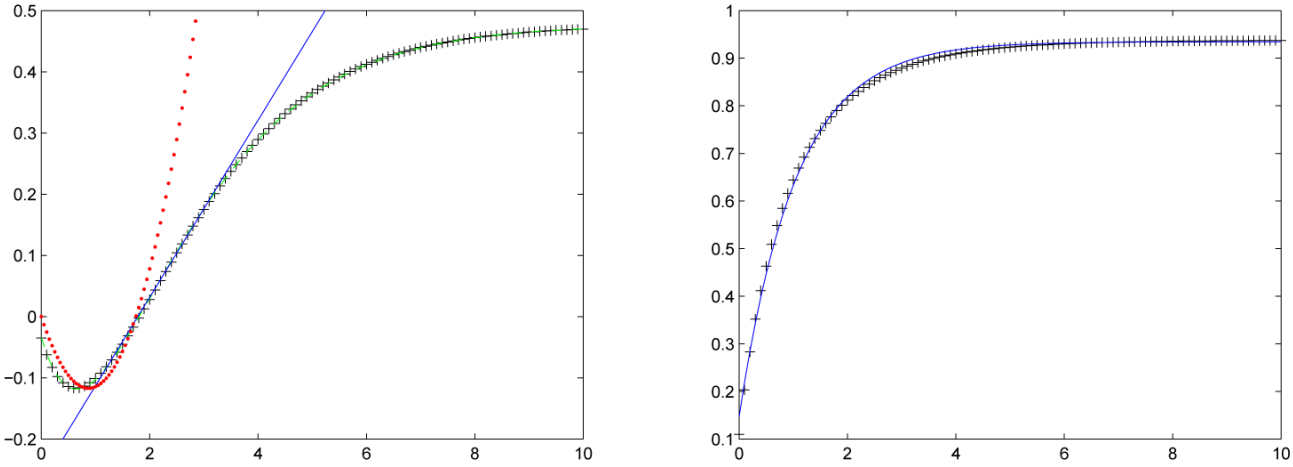

Figure 7.

Inputs to the decision units computed from a simulation of the noise-free Eriksen model logistic activation functions (black +s). Left panel: incompatible stimuli, with linear (solid blue line), quadratic (red dots) and exponential (green dashes) fits to simulation data. Right panel: compatible stimuli with an exponential fit (solid blue curve). See text for discussion.

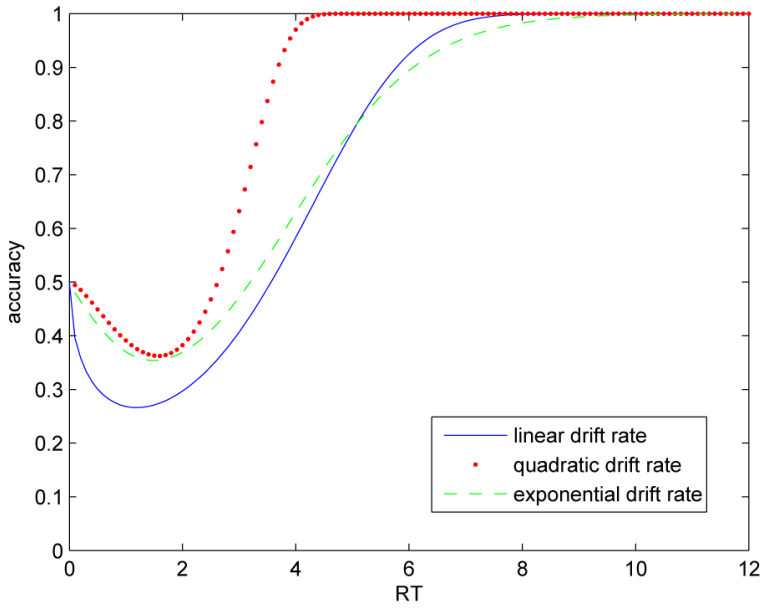

Figure 8.

Analytical predictions of accuracy for the drift-diffusion model under the interrogation protocol with linear (solid, blue), quadratic (dotted, red) and exponential (dashed, green) drift rates, fitted to the nonlinear simulation data of Figure 7 as described in the text. Parameter values for the drift-diffusion process are $\lambda = 0$ (balanced), $c = 0.3$.

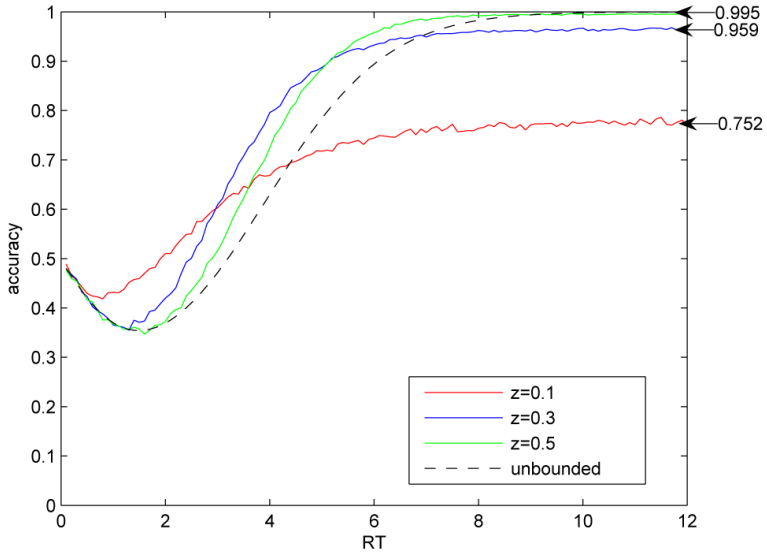

Figure 9.

Accuracy vs. RT for the bounded diffusion process with an exponential drift rate and bounds *L* of 0.1 (red), 0.3 (blue) and 0.5 (green) bottom to top on right of figure, with noise strength *c* = 0.3. Dashed black curve shows accuracy for unbounded process. Arrows at right show theoretical accuracy limits for large RT from (38).

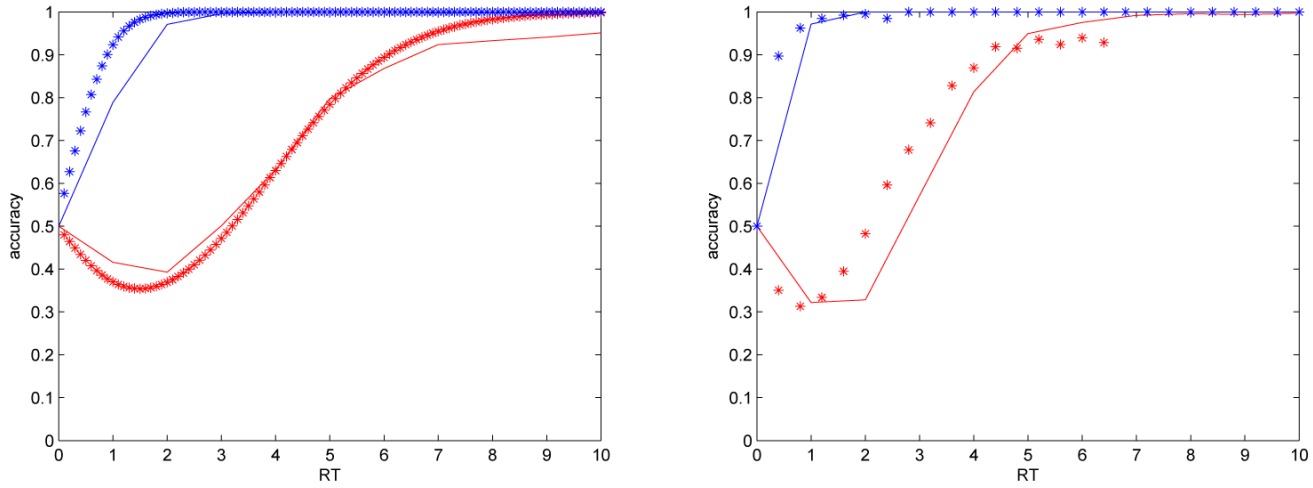

Figure 10.

Left: Accuracy vs. RT for the full Ericksen model under the interrogation protocol (solid line) compared with analytical results for the unbounded DDM (stars). Right: Accuracy vs. RT for the full Ericksen model (solid line) under the free-response protocol compared with numerical results for the unbounded DDM with exponential inputs (stars): thresholds are $\theta = 1.0$ for full model and $\theta = \pm 0.3$ for DDM. Upper curves denote compatible trials and lower curves incompatible trials. Remaining parameters are as specified in text.

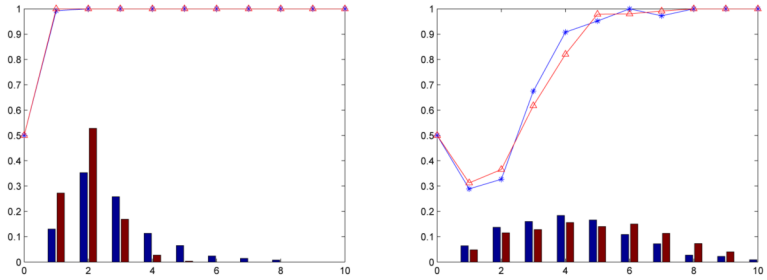

Figure 11.
Accuracy vs. RT for the full Ericksen model (blue lines with solid dots) under the free-response protocol, with RT histograms below (left, blue bars of each pair). Unbounded DDM accuracy results are shown for comparison in solid red with open triangles and RT histograms in right, red bars of each pair. Thresholds are $\theta = 1.1$ for full model and $\theta = \pm 0.3$ for DDM. Left panel shows compatible trials, right panel shows incompatible trials. Remaining parameters are as specified in text.

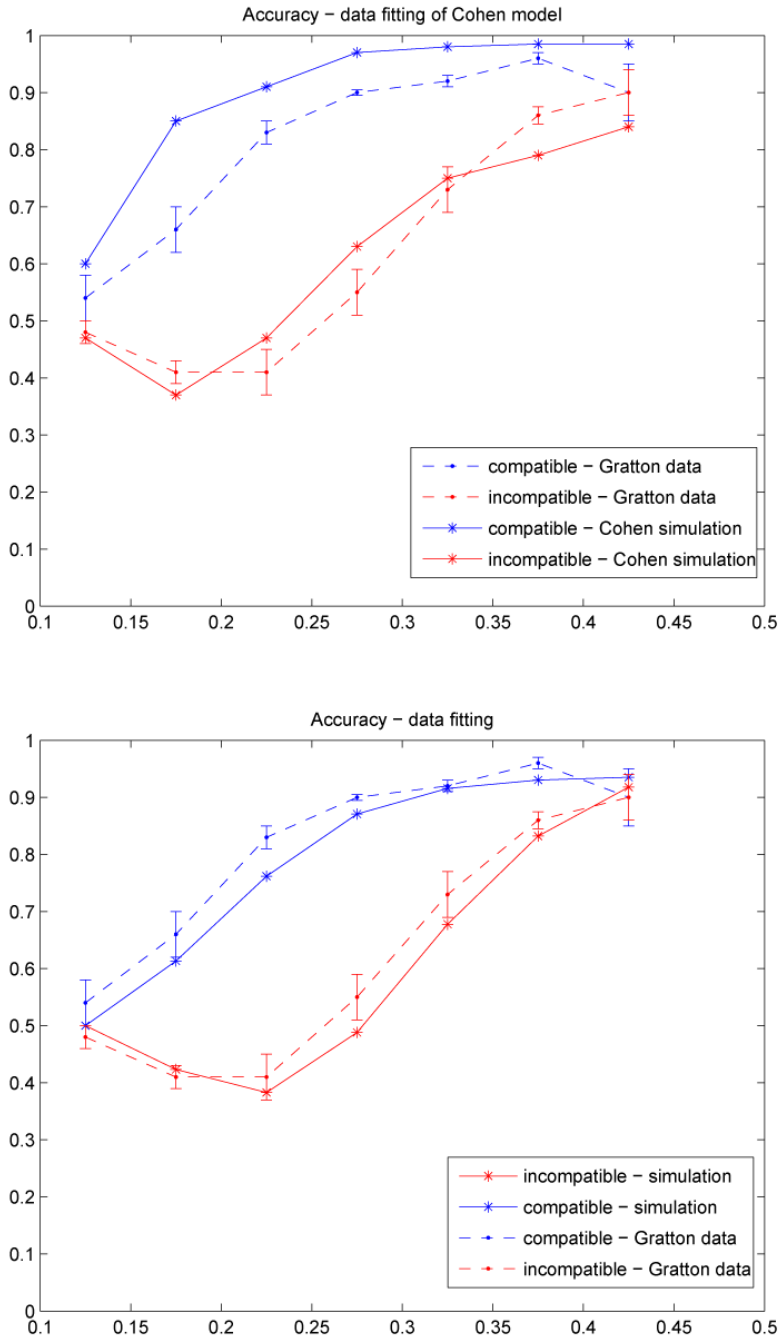

Figure 12.

Accuracy vs. RT for the full model (top, solid) and the DDM (bottom, solid), compared with data from Gratton et al. (1988) (dashed). Standard errors of latter indicated by vertical bars. Curves for compatible stimuli lie above those for incompatible stimuli.

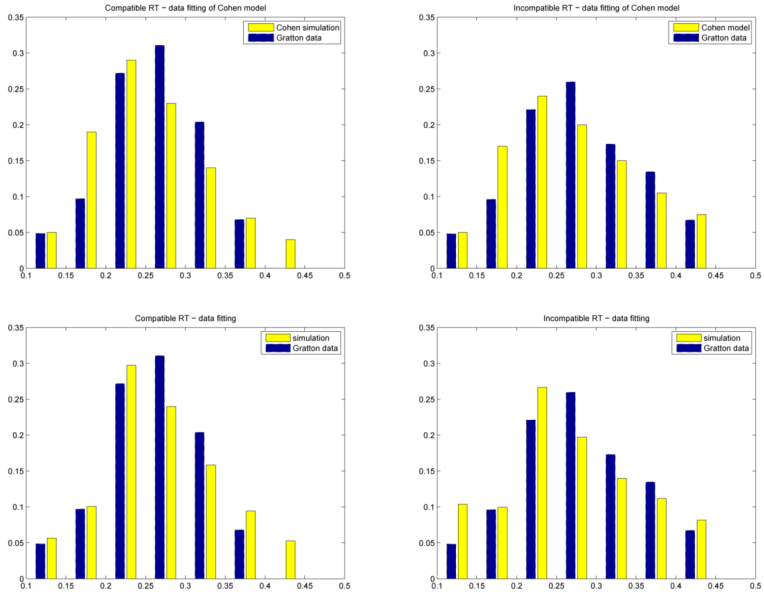

Figure 13.

RT histograms from the full model (top row) and the DDM (bottom row), compared with data from Gratton et al. (1988). Model results are shown in right, yellow bars and empirical data in left, blue bars for each 50 ms bin. Compatible trials appear in left column and incompatible trials in right column.