Journal of Mathematical Psychology 53 (2009) 231–241

Contents lists available at ScienceDirect

Journal of Mathematical Psychology

journal homepage: www.elsevier.com/locate/jmp

A comparison of bounded diffusion models for choice in time controlled tasks

Jiaxiang Zhang [a,b,*], Rafal Bogacz [a], Philip Holmes [c]

[a] *Department of Computer Science, University of Bristol, Bristol BS8 1UB, United Kingdom*
[b] *School of Psychology, University of Birmingham, Birmingham B15 2TT, United Kingdom*
[c] *Program in Applied and Computational Mathematics, Princeton University, United States*

ARTICLE INFO

Article history:
Received 18 June 2008
Received in revised form
3 March 2009
Available online 10 April 2009

Keywords:
Choice
Decision
Diffusion model
Reflecting boundaries
Absorbing boundaries

ABSTRACT

The Wiener diffusion model (WDM) for 2-alternative tasks assumes that sensory information is integrated over time. Recent neurophysiological studies have found neural correlates of this integration process in certain neuronal populations. This paper analyses the properties of the WDM with two different boundary conditions in decision making tasks in which the time of response is indicated by a cue. A dual reflecting boundary mechanism is proposed and its performance is compared with a well-established absorbing boundary in the cases of the WDM, the WDM with extensions, and the WDM with prior probability. The two types of boundary influence the dynamics of the model and introduce differential weighting of evidence. Comparisons with Ornstein–Uhlenbeck models are also done, and it is shown that the WDM with both types of boundary achieves similar performance and produces similar fits to existing behavioural data. Further studies are proposed to distinguish which boundary mechanism is more consistent with experimental data.

1. Introduction

Making choices is a frequent and critical element of human and animal lives. This problem has been studied by psychologists under two types of 2-alternative-forced-choice (2AFC) paradigms, in which subjects must decide between two available alternatives. The *information controlled* (IC) paradigm allows subjects to respond whenever they feel confident (Luce, 1986). Alternatively, subjects can be required to report their choice immediately after a cue to respond (Swensson, 1972; Yellott, 1971), under the *time controlled* (TC) paradigm (sometimes referred to as the response signal paradigm; see Dosher (1984)).

Over the last half century, several sequential sampling models have been proposed to describe experimental results as well as underlying decision making mechanisms in 2AFC tasks (for reviews, see Luce (1986) and Townsend and Ashby (1983)). The Wiener diffusion model (WDM) – the focus of this paper – assumes that subjects integrate partial information representing the relative support for the two alternatives over time (Laming, 1968; Ratcliff, 1978; Stone, 1960), and it has been shown to be the statistically optimal method for choosing between two alternatives on the basis of noisy evidence. For a fixed set of stimulus conditions, the WDM minimizes the reaction time for given accuracy, or maximizes the accuracy for given reaction time (Edwards, 1965; Gold & Shadlen, 2001, 2002; Laming, 1968; Wald, 1947). The WDM successfully describes the reaction time distribution and accuracy in various cognitive decision tasks in the IC paradigm (Laming, 1968; Link, 1975; Link & Heath, 1975; Ratcliff, 1978; Ratcliff & Smith, 2004; Ratcliff, Van Zandt, & McKoon, 1999; Stone, 1960). However, in the TC paradigm, the earlier version of the model (Ratcliff, 1978) allows integrator states to take arbitrary values. This leads to the prediction that the accuracy always increases over time if the drift of the process is in the correct direction. This is contrary to the finding that accuracy in the paradigm grows over time to an asymptote. To avoid this shortcoming, the model needs to assume that drift values are normally distributed across trials (Ratcliff, 1978) or that the integration range is bounded (Ratcliff, 1988). With one of these elaborations, the model correctly predicts the time–accuracy curves found in the IC condition.

This paper introduces a *reflecting boundary* mechanism to model 2AFC tasks in the TC paradigm. We compare the dynamics and performance of the WDM with reflecting and absorbing boundaries (Feller, 1968) and assess their ability to account for published behavioural data from 2AFC experiments. We also compare them with an Ornstein–Uhlenbeck (O–U) model (Busemeyer & Townsend, 1992, 1993; Diederich, 1995, 1997; Ratcliff & Smith, 2004; Smith, 1995, 2000). Analytical and numerical results show that both boundary types lead to similar performance and produce similar fits, but we identify some differences between them. Further studies are proposed that might distinguish which type of boundary better describes the decision making process.

* Corresponding author at: School of Psychology, University of Birmingham, Birmingham B15 2TT, United Kingdom.
E-mail address: J.Zhang.1@bham.ac.uk (J. Zhang).

0022-2496/$ – see front matter © 2009 Elsevier Inc. All rights reserved.
doi:10.1016/j.jmp.2009.03.001

The paper is organized as follows. Section 2 reviews neurophysiological evidence of decision making, and describes the WDM and alternative boundary mechanisms. More detailed reviews are available elsewhere (Bogacz, Brown, Moehlis, Holmes, & Cohen, 2006; Glimcher, 2001; Gold & Shadlen, 2007; Schall, 2001; Smith & Ratcliff, 2004). Section 3 compares the dynamics and the performance of models with the two types of boundary. Section 4 compares fits of bounded models with behavioural data. Finally, Section 5 discusses further experimental studies that could distinguish between the boundary mechanisms. Mathematical details are provided in Appendices A and B.

2. The biology of decision and the sequential sampling models

2.1. The neural basis of 2AFC tasks

Recently, neuronal activity from awake animals has been recorded in choice experiments. For example, in the motion discrimination task, visual stimuli comprise arrays of random moving dots, a proportion of which move coherently to the left or right. Subjects (monkeys) are required to indicate their decision regarding the coherent direction by making a saccade to a left or right target (Roitman & Shadlen, 2002; Shadlen & Newsome, 2001).

The activity of neurons in the middle temporal (MT) area has been shown to correlate with the motion coherence (Britten, Shadlen, Newsome, & Movshon, 1993). However, since these neurons are highly noisy and poorly correlated with choices, the MT area is less likely to be a "decision maker" than to provide temporal information for a further process. It has been suggested that the lateral interparietal (LIP) area may interpret raw information from MT neurons. Shadlen and Newsome (1996) reported that LIP neurons gradually build up or attenuate their activity within a trial, and exhibit persistent activity in the absence of stimuli (Shadlen & Newsome, 2001). The time course of LIP neuronal activity suggests that the LIP area integrates inputs from MT neurons (Hanks, Ditterich, & Shadlen, 2006; Huk & Shadlen, 2005; Roitman & Shadlen, 2002). Similar discharge patterns are also found in the frontal eye field (FEF) (Schall, 2002) and the superior colliculus (SC) (Basso & Wurtz, 1998).

These results indicate a general decision mechanism manifested in different brain regions in which certain neuronal populations integrate sensory information over time to increase the accuracy of selection between alternatives (Gold & Shadlen, 2007; Schall, 2001). Gold and Shadlen (2001, 2002) formalize the decision process in 2AFC tasks as following two processes: two populations of sensory neurons (e.g., in the MT area) generate continuous noisy information streams ($Y_1(t)$ and $Y_2(t)$) for each of two alternatives Y_1 and Y_2 at time t. For simplicity, we assume that $Y_1(t)$ and $Y_2(t)$ have constant means μ_1 and μ_2 during each trial, with the same constant standard deviation, σ. The goal of the second process (reflected in LIP activities) is to successfully identify which input population has higher mean based on sample sequences $Y_1(t)$ and $Y_2(t)$. This framework is the basis for several sequential sampling models in behavioural studies, including the WDM (Ratcliff, 1978), the O–U model (Busemeyer & Townsend, 1992), and the leaky-competing-accumulator (LCA) model (Usher & McClelland, 2001).

2.2. The Wiener diffusion model (WDM)

The Wiener diffusion or Brownian motion is a continuous limit of the random walk (Laming, 1968; Ratcliff, 1978; Stone, 1960). It implies a leak-free integrator that accumulates the *difference* $Y_1(t) - Y_2(t)$ between noisy evidence streams for the two alternatives. Let $X(t)$ denote the accumulated difference at time t: the value of the integrator state, with initial state $X_0 = X(0)$.

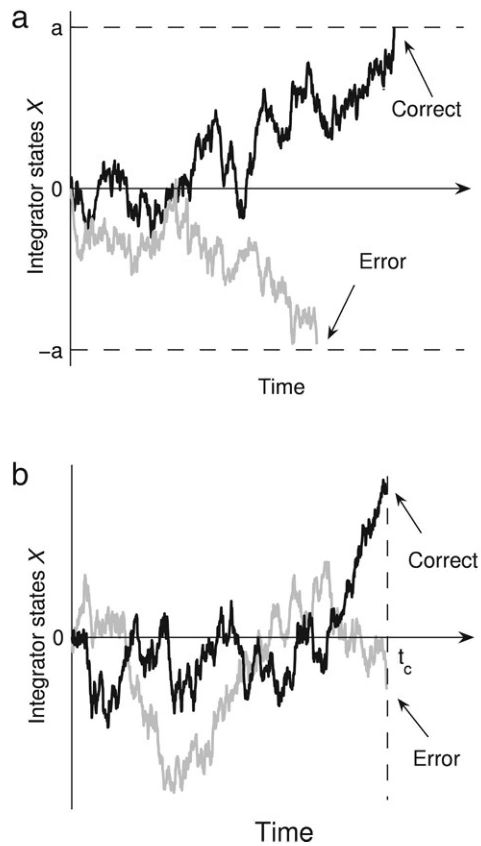

Fig. 1. Evolutions of integrator states after stimulus onset for the WDM, showing correct trials (black) and error trials (grey). The model was simulated with $\mu = \sigma = a = 1$ and time-step $dt = 0.01$ s, and hence $X(t) > 0$ corresponds to the correct alternative. (a) IC paradigm: choices are made on first reaching one of the two thresholds. (b) TC paradigm: choices are determined by the sign of the integrator state $X(t_c)$ at time t_c.

If there is no prior bias towards either choice, the process starts at baseline $X_0 = 0$, and is described by a stochastic differential equation:

$$dX(t) = \mu\, dt + \sigma\, dW(t), \quad \text{with } X_0 = 0, \tag{1}$$

where $dX(t)$ denotes the evidence obtained during time step dt. μ is a constant drift (the accumulation rate), representing the average of evidence difference $\mu_1 - \mu_2$. On a given trial, $\mu > 0$ ($\mu_1 > \mu_2$) implies that Y_1 is the correct choice, while $\mu < 0$ ($\mu_1 < \mu_2$) if Y_2 is correct. For consistency we hereafter set $\mu > 0$ unless indicated specifically, and hence assume that Y_1 is the correct choice. The magnitude of μ reflects the difficulty level of the task: for small μ ($\mu_1 \approx \mu_2$), it is difficult to distinguish which evidence samples have higher mean. The second term, $\sigma\, dW(t)$, denotes Gaussian noise with mean 0 and variance $\sigma^2 dt$. In the absence of noise ($\sigma = 0$), $X(t)$ changes at rate μ and always reaches a correct decision. Noisy inputs cause the $X(t)$ to fluctuate and hence induce incorrect choices on some trials.

Fig. 1 shows the growth of $X(t)$ in the two paradigms. For the IC paradigm, the decision time is unrestricted and two thresholds $\pm a$ are introduced to indicate termination states. Once $X(t)$ reaches a threshold, the corresponding alternative is chosen. For the TC paradigm, the decision process is interrupted by a response cue, and a response is immediately required. We hereafter denote the time delay from stimulus onset to response cue by t_c. The alternatives are selected by locating the final integrator state $X(t_c)$ and selecting Y_1 if $X(t_c) > 0$, and Y_2 if $X(t_c) < 0$.

Performance can be measured by the error rate: the probability of making an incorrect choice in a block of trials,[1] hereafter denoted by P. The error rate is a function of model parameters μ, σ, and threshold setting, a, or response signal, t_c, depending on the paradigm. For unbiased initial conditions $X_0 = 0$, the error rate of the WDM in the IC and TC paradigms is given by Eqs. (2) and (3), respectively (Bogacz, Usher, Zhang, and McClelland (2007); cf. Gardiner (1985) and Ratcliff (1978)):

$$P(a) = \frac{1}{1+e^{\frac{2\mu a}{\sigma^2}}}, \quad \text{and} \tag{2}$$

$$P(t_c) = \int_{-\infty}^{-\frac{\mu}{\sigma}\sqrt{t_c}} \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}} du. \tag{3}$$

Two extensions have been proposed to improve fits to experimental data (Ratcliff et al., 1999). They allow certain parameters to vary randomly across trials. First, the drift rate $\tilde{\mu}$ is assumed to have a Gaussian distribution across trials with mean μ^* and variance σ_μ^2, which might reflect the variability of difficulty between trials, the subject's attention level, or other variable inter-trial factors. On each trial, $\tilde{\mu}$ can take either positive or negative values, promoting accumulation towards different alternatives. The correct choice is determined by the mean drift μ^*, even if the sampled value $\tilde{\mu}$ has opposite sign in some trials. This is motivated by the fact that, in difficult situations, stimulus distributions corresponding to the two alternatives often overlap (Ratcliff et al., 1999). Even after long training, perfect performance in such tasks is impossible. Drift variability is also necessary to ensure that the asymptotic accuracy of the WDM in the IC paradigm is not infinite in the absence of boundaries (Ratcliff, 1978). Second, the theory of premature sampling assumes that subjects start to integrate noise before sensory information is available. Hence the starting point is not at 0 when stimuli onset (Laming, 1968). Instead, on each trial X_0 is chosen from a uniform distribution [2] on the interval $[-\sigma_X, \sigma_X]$.

The extended WDM produces different reaction times on correct and error trials in the IC paradigm (Ratcliff et al., 1999). We assume that the same variability sources also operate in the TC paradigm; their effects on the decision process will be evaluated in the next section.

2.3. Boundary mechanisms

In the TC paradigm, the fact that integrator states $X(t)$ are unbounded implies that the error rate of the WDM with no variability in drift rate diminishes to zero for large t_c (cf. Eq. (3)). To eliminate this contradiction, absorbing boundaries can be introduced at $X = \pm b$. This mechanism was originally used to model tasks in which IC and TC paradigms were intermixed (Ratcliff, 1988). Such tasks promote subjects to respond as quickly as possible before a predetermined deadline. Under this condition the proposed model, also called the internal deadline model, assumes that the decision process is terminated when the accumulated evidence reaches one of the three possible boundaries, whichever comes first. The three boundaries are: top and bottom absorbing boundaries in the state domain, and a deadline boundary in the time domain (Diederich & Busemeyer, 2006; Ratcliff & Rouder, 2000).

In this work, we consider a pure TC paradigm in which subjects are only allowed to respond after the deadline t_c (Roitman & Shadlen, 2002). If the decision process reaches one of the absorbing boundaries before t_c, the accumulation process stops and the activity $X(t) = \pm b$ is maintained until the end of the trial. For sufficiently large t_c, $X(t)$ will almost surely reach one of the boundaries before t_c (cf. Section 3.1). Absorbing boundaries have the same effect as the decision threshold in the IC paradigm. Hence the error rate of the absorbing WDM with infinite t_c can be analytically obtained as (cf. Eq. (2))

$$\lim_{t_c\to\infty} P_{(abs)}(t_c) = \frac{1}{1+e^{\frac{2\mu b}{\sigma^2}}}, \tag{4}$$

where the subscript *abs* stands for the WDM with absorbing boundary. For $b < \infty$, the error rate does not decrease to zero as t_c increases, which is consistent with experimental observations (Meyer, Irwin, Osman, & Kounios, 1988; Usher & McClelland, 2001).

In contrast with the absorbing boundary mechanism, since no time pressure exists in the pure TC paradigm, subjects may use the deadline alone to terminate the decision process. In this case two reflecting boundaries may be more suitable to constrain the accumulation process, because they allow the preferred choice to change even if $X(t)$ reaches a boundary. Here the boundaries restrict the amount of evidence that can be represented (much as a sigmoidal function provides cutoffs at high and low activation). Some previous studies (Diederich, 1995; Diederich & Busemeyer, 2003) use a lower reflecting and an upper absorbing boundary to model the simple reaction time task in which subjects respond immediately after a stimulus is detected. The reflecting boundary in their model and the one proposed here share similar motivations but there is a major difference. In the simple reaction time task, the accumulated sensory information directly represents the absolute evidence to make a response. Since the integrated information cannot drop below a certain baseline, one reflecting boundary is required to model the minimum level of absolute evidence. Single reflecting boundaries have also been used to represent a lower bound on the integration process (see Ratcliff and Smith (2004), Smith and Ratcliff (in press) and Usher and McClelland (2001)).

In the TC paradigm of 2AFC tasks, the integrator in the WDM represents the *relative* evidence supporting the alternatives. The preferred alternative at time t is determined by the sign of $X(t)$. A value of $X(t) > 0$ means that the first alternative is the provisional choice, whereas $X(t) < 0$ means that the second alternative is currently preferred. If we also assume that a minimum activity baseline exists and that decision preferences may switch during the trial, two reflecting boundaries are required to restrict the relative evidence of two alternatives within a certain range. Details of the model are given in Section 3.

2.4. The Ornstein–Uhlenbeck (O–U) model

Another widely applied sequential sampling method is the O–U model (Busemeyer & Diederich, 2002; Busemeyer & Townsend, 1992, 1993). It introduces a new parameter λ to the WDM to represent decay ($\lambda < 0$) or growth ($\lambda > 0$) of accumulated information, its evolution being described by

$$dX(t) = (\lambda X(t) + \mu)\,dt + \sigma dW, \quad X_0 = 0, \tag{5}$$

where the notations are as in Eq. (1). In the O–U model the accumulation rate depends not only on the drift μ but also on the

[1] Here we do not directly measure the probability of choosing certain alternatives (e.g., P_{Y_1} or P_{Y_2}) since in most experiments the correct choice is randomly assigned from the two alternatives across trials (e.g., Roitman and Shadlen (2002) and Shadlen and Newsome (2001)). Note that when we assume Y_1 is correct, then $P = P_{Y_2}$.

[2] The uniform distribution is assumed to prevent X_0 exceeding the thresholds $\pm a$ (by setting $\sigma_X < a$).

current state of the integrator $X(t)$. This allows the error rate of the O–U model to approach a finite asymptote for large t_c even without boundaries.

The O–U model can account for the primacy and recency effects observed in decision making tasks: some subjects pay more attention to initial evidence (primacy) while others focus on later evidence (recency) (Wallsten & Barton, 1982). To illustrate this, note that Eq. (5) has a fixed point $X = -\mu/\lambda$. For $\lambda > 0$, the fixed point is unstable. After $X(t)$ has been driven to one side or other of the fixed point, subsequent evidence has little effect on the final choice due to repulsion, indicating a primacy effect. For $\lambda < 0$, the fixed point is an attractor. Information decays over time so that early evidence is partially lost, indicating a recency effect. A proof is available in Busemeyer and Townsend (1993).

The O–U model has been applied to a variety of tasks (Diederich, 1995, 1997; Smith, 1995) and also been extended to multi-alternative tasks (McMillen & Holmes, 2006; Usher & McClelland, 2001). As it is similar to the WDM, we include the O–U model in our subsequent comparative study.

3. Comparison of properties of bounded WDMs

3.1. The WDM with reflecting boundaries

Motivated by the discussion in Section 2.3, we now propose *reflecting* boundary conditions. In the simplest cases, the boundaries $\pm b$ are still symmetric about X_0 but they differ from absorbing boundaries in that once $X(t)$ reaches $\pm b$, it cannot exceed this value, but it may move back due to noise, as specified by

$$\begin{cases} X(t+dt) = b, & \text{if } X(t) + dX(t) \geq b, \\ X(t+dt) = -b, & \text{if } X(t) + dX(t) \leq -b, \\ X(t+dt) = X(t) + dX, & \text{otherwise}, \end{cases} \tag{6}$$

where $dX(t)$ is given by Eq. (1). In contrast to absorbing boundaries, reflecting boundaries allow information from any time period to contribute to the final choice, which is not determined until the end of t_c. Fig. 2 shows sample paths of integrator states in the WDM with both types of boundary, over one trial. The solid line shows that for reflecting boundaries, even if $X(t)$ reaches the upper boundary, $X(t)$ continues to evolve until t_c and it may "backtrack" to subsequently reach the lower boundary.

Note that, unlike the absorbing WDM in which boundary contact induces a decision, the reflecting WDM requires an external criterion to stop (e.g., a response cue at time t_c), so it cannot be applied to the IC paradigm or to tasks in which IC and TC paradigms are intermixed. Hence we will only consider the TC paradigm in the rest of this paper.

The first step in investigating the bounded WDM is to compute the probability density $p(X, t)$ of Eq. (1) with appropriate boundary conditions, i.e., the probability of finding a sample path at state X at time t. Explicit expressions for $p(X, t)$ of absorbing and reflecting WDMs are given in Eqs. (B.5) and (A.9) of the Appendix. Fig. 3b illustrates the distribution of the reflecting WDM at several time instants. Comparing with that of the absorbing WDM in Fig. 3a, we may distinguish properties of $p(X, t)$ for the two types of boundaries. Initially both are approximately Gaussian (note the curves marked $t = 0.1$ s), with means close to X_0 and shifted towards the upper boundary (the correct choice for $\mu > 0$), but they become significantly different as t grows. For the absorbing WDM, $p(X, t)$ collapses to 0 as $t \to \infty$, since every decision process is absorbed by one of the boundaries after sufficiently large t_c. In contrast, no paths terminate in the reflecting WDM, but $p(X, t)$ remains of unit mass and approaches the equilibrium distribution:

$$\lim_{t\to\infty} p(X, t) = \frac{2\mu e^{\mu(b+X)}}{e^{4\mu b} - 1}. \tag{7}$$

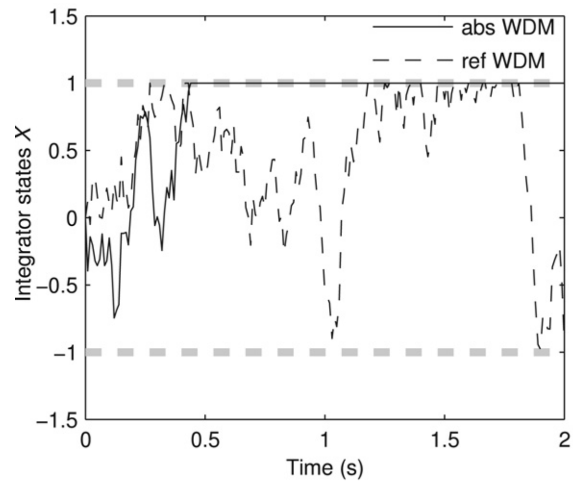

Fig. 2. Examples of trajectories of the WDM with absorbing and reflecting boundaries. Models were simulated with $\mu = 1, \sigma = 1, dt = 0.01$ s, and symmetric boundaries $\pm b = \pm 1$, indicated by thick grey dashed lines. For the absorbing WDM (solid), the choice is determined when the state first hits a boundary (at 0.5 s); for the reflecting WDM (dashed), the preferred choice may change throughout the decision process.

The error rate of the reflecting WDM can be computed by integrating $p(X, t)$ from $-b$ to 0 (cf. Eq. (A.16) of Appendix A), and as t_c increases, it converges to

$$\lim_{t_c\to\infty} P_{(ref)}(t_c) = \frac{1}{1 + e^{\frac{2\mu b}{\sigma^2}}}, \tag{8}$$

where *ref* stands for reflecting WDM. Note that arbitrarily small $P_{(ref)}(t_c)$ can be achieved for $b < \infty$ only if the signal-to-noise ratio $\mu/\sigma \to \infty$, and that the expression of P is identical to Eq. (4).

3.2. Primacy and recency effects

Since the updating rule in Eq. (1) is independent of the current state $X(t)$, the unbounded WDM implicitly assumes that the influence of sensory evidence on the final choice does not depend on the timing of its occurrence. This assumption is contrary to the primacy and recency effects discussed before. Here we show that, by applying the two types of boundary, the WDM can also represent primacy and recency effects.

For the absorbing WDM, once $X(t)$ hits the boundary, the preferred decision is determined and maintained. Only inputs prior to the first boundary hit contribute to the decision process, so the probability that incoming evidence at time t contributes to the final choice (denoted by $P(dX(t) \neq 0)$) is equal to the probability that neither boundary has been reached before t, i.e.,

$$P(dX(t) \neq 0) = P(G(X) > t) = 1 - \Phi_{G(X)}(t), \tag{9}$$

where the random variable $G(X)$ is the time required for $X(t)$ to first reach either boundary (the first-passage time (Feller, 1968)) and $\Phi_{G(X)}$ is the cumulative distribution of $G(X)$. Since $\Phi_{G(X)}$ is a monotonically increasing function, $P(dX(t) \neq 0)$ is monotonically decreasing. Thus on average the decision is more likely based on earlier inputs, indicating a primacy effect.

For the reflecting WDM, each boundary hit results in a loss of information, since the model does not fully integrate the increment that carries it to the boundary. To illustrate this, assume that, at time τ, the incoming evidence $dX(\tau)$ drives the integrator to reach one of the boundaries. Compared with the unbounded WDM, hitting the reflecting boundary leads to a loss of information $|X(\tau) + dX(\tau) - b|$ (cf. Eq. (6)). On average, evidence presented

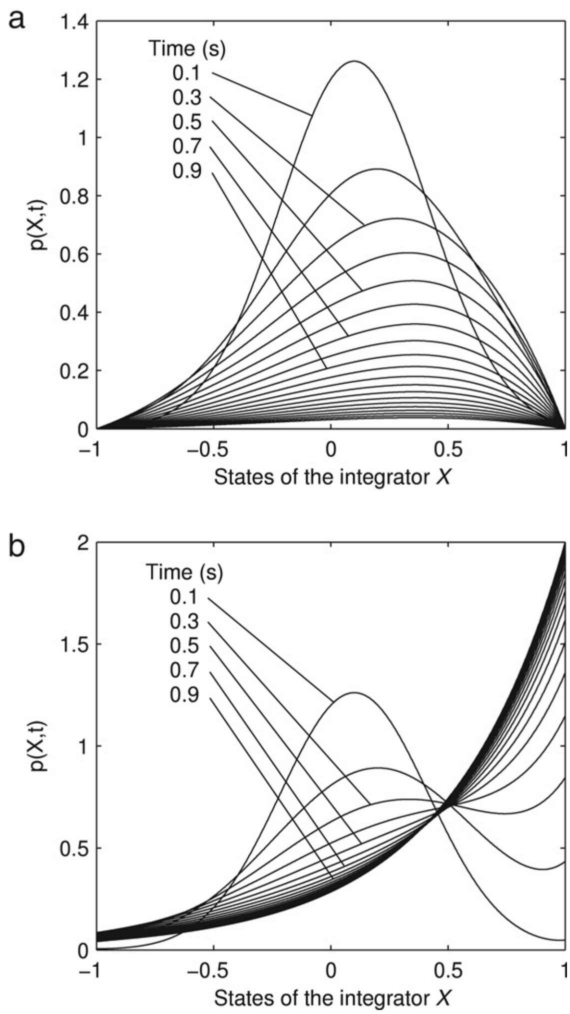

Fig. 3. Probability densities of integrator states in bounded WDMs. State values are shown on the horizontal axes. Densities are shown for parameters $\mu = \sigma = 1$ and boundaries $\pm b = \pm 1$. Different curves show densities at different time instants from 0.1 s to 2 s. The probability density functions $p(X, t)$ are calculated via Eqs. (B.5) and (A.9) in the Appendix. All solutions start at $X_0 = 0$. (a) Absorbing WDM; (b) reflecting WDM.

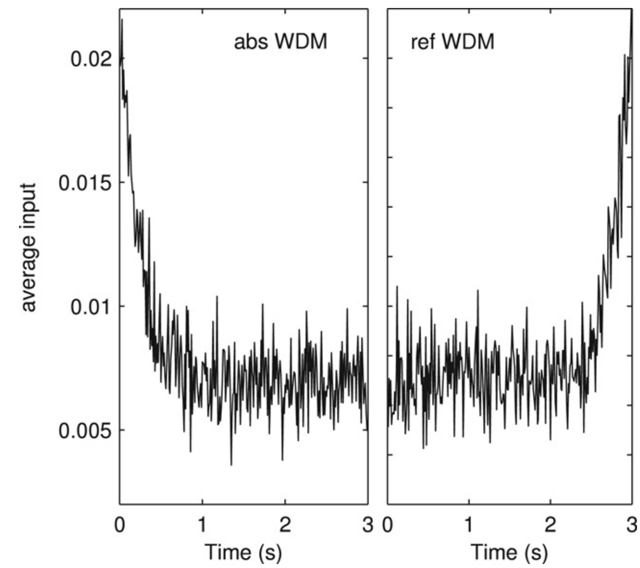

Fig. 4. Primacy and recency effects for the WDM with absorbing (left) and reflecting (right) boundaries. The bounded WDM is simulated for 10,000 trials. For all correct trials the input to the WDM ($\mu dt + \sigma dW$, cf. Eq. (1)) at every time step is recorded and averaged. Curves illustrate averaged inputs plotted against time. Parameter values are $\mu = 0.71, b = 0.47, \sigma = 1$, which are obtained by fitting behavioural data from subject S1 in a 2AFC task (Usher and McClelland (2001), Table 1 and see Section 4) and $t_c = 3$ s.

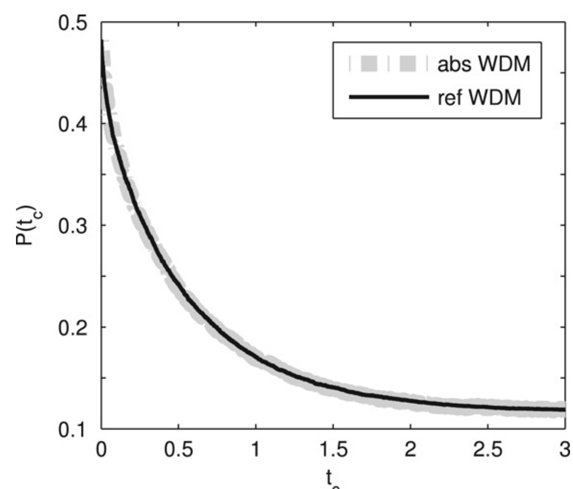

Fig. 5. Error rates of the WDM with reflecting (solid) and absorbing (thick dashed) boundaries at different time t_c. Model parameters are $\mu = \sigma = b = 1$ and data are averaged over 100,000 trials.

earlier in the trial results in more boundary hits. Thus, information arriving earlier is partially lost and the final choice depends to a greater extent on later inputs, producing a recency effect. The reflecting WDM can also (repeatedly) change the preferred alternative within a trial, e.g., in Fig. 2 the state of the reflecting WDM drops below zero, changing its choice in the last 0.2 s.

Primacy and recency effects can be illustrated by showing how the two models weight noisy inputs during the decision process. A sequence of trials was simulated and the noisy inputs recorded (right side of Eq. (1)) only for trials resulting in correct choices. These inputs were then averaged to indicate the correlation between the mean inputs and the final choice, and this averaged input was shown in Fig. 4. For absorbing boundaries, larger inputs at early times contributed more to the final decision (primacy), while for reflecting boundaries, larger inputs before the response led to the correct choice (recency).

3.3. *Performance of bounded WDMs*

Recall that the error rate expressions for the reflecting and absorbing WDM coincide when $t_c \to \infty$ (Eqs. (4) and (8)). In Appendix B we prove that such an equality holds in general: given the same parameter set (μ, σ, b) and no prior bias (i.e., boundaries equidistant from the starting point $X_0 = 0$), the error rates for the two bounded WDMs are identical for any t_c:

$$P_{(abs)}(t_c) = P_{(ref)}(t_c). \tag{10}$$

Hence both bounded models achieve the *same accuracy* under the TC paradigm, as illustrated in Fig. 5: the two curves coincide for any given t_c. Thus, although the different boundaries profoundly influence the dynamics of the decision process, their performances are indistinguishable. For given experimental observations ($P(t_c)$ in the TC paradigm) and the same parameter set, the two types of bounded model can always provide the same fitting ability, as we shall illustrate in Section 4.

3.4. *Performance of bounded WDM with extensions*

We now consider the extended WDM described in Section 2.2.

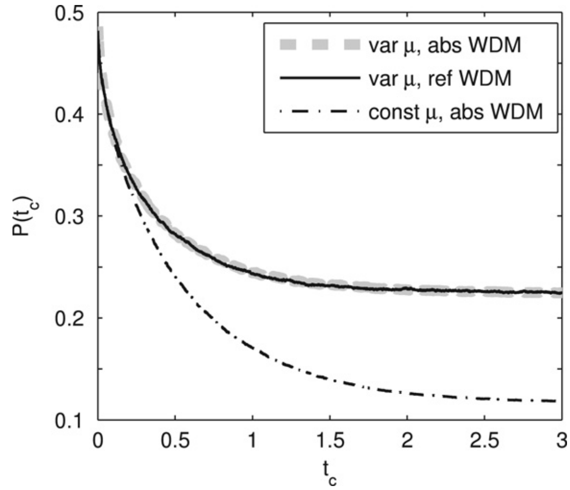

Fig. 6. Error rates of the bounded WDMs at different t_c with variable drift rates across trials sampled from a normal distribution $\mathcal{N}(1, 1)$ (solid and thick dashed), compared with the error rate of the bounded WDM with constant drift rate $\mu = 1$ (dash–dotted). Other model parameters are $\sigma = b = 1$. Data are averaged over 100,000 trials.

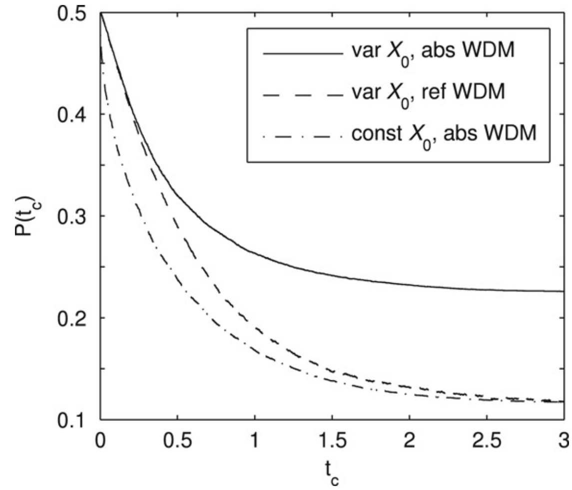

Fig. 7. Error rates of absorbing (solid) and reflecting (dashed) WDMs at different t_c with variable starting point across trials, sampled from a uniform distribution between [1, 1], compared with the error rate for the bounded WDM with starting point $X_0 = 0$ (dash–dotted). Model parameters are $\mu = \sigma = b = 1$. Data are averaged over 100,000 trials.

3.4.1. Variability of drift rate

Fig. 6 compares $P(t_c)$ at different t_c for the bounded WDM with drift rates sampled from a normal distribution, with $P(t_c)$ for the bounded WDM with constant drift rate. The drift variability increases the error rate for all t_c. The figure also shows that reflecting and absorbing boundaries still produce identical error rates for any t_c. This is not surprising, since while the drift differs across trials in the extended model, within each trial the sampled drift $\tilde{\mu}$ remains fixed. Hence both bounded models produce the same error rate, according to Section 3.3.

3.4.2. Variability of starting point

Fig. 7 shows the performance of the bounded WDM with starting points X_0 sampled from a uniform distribution. This produces higher error rates than models with constant X_0. Also, since the starting points are not equidistant from the boundaries, the error rates differ for the two types of boundary. One interesting result is that the reflecting WDM with variable and constant X_0 can achieve the same error rate for large t_c (the dashed and dash–dotted curves converge after 2 s). This follows from the recency effect discussed in Section 3.2: the initial condition variability is attenuated by reflecting boundaries as t_c increases. In contrast, the performance of the absorbing WDM is highly sensitive to the starting points.

3.5. Prior probability and biased starting points

In this section we investigate how boundaries affect the performance if the starting point X_0 depends on the prior probability of the alternatives.

If the subject knows that one of the alternatives is more probable, the performance can be improved by moving the starting point towards that alternative (Edwards, 1965; Link, 1975). Such starting point biases have been observed in behavioural experiments (Laming, 1968; Link, 1975; Ratcliff et al., 1999). We now compare the performance of bounded WDMs in this situation. We denote the probability that the first alternative is correct by p_+, and that the second is correct by $p_- = 1 - p_+$, so that on each trial drift $\mu > 0$ occurs with probability p_+ and drift $\mu < 0$ with probability p_-.

Link (1975) proved that to minimize the error rate under the IC paradigm, the starting point should be set to

$$X_0 = \frac{\sigma^2}{4|\mu|} \ln \frac{p_+}{p_-}. \tag{11}$$

Thus, as the difference in prior probabilities increases, X_0 moves towards the boundary which is more likely to be correct. Recall that in Section 2 we proposed that absorbing boundaries in the TC paradigm act like decision thresholds in the IC paradigm. Hence Eq. (11) also specifies the optimal starting point of the absorbing WDM for large t_c under the TC paradigm. Fig. 8a illustrates this prediction by stimulating bounded models with several starting points. The theoretically optimal X_0 (vertical solid line) of Eq. (11) can minimize the error rate of the absorbing model (solid curve). Note that the correct choice in the simulation is still defined by the sign of the drift rate. Hence a biased starting point close to the error boundary ($X_0 < 0$) induces $P(t_c)$ larger than 0.5. In contrast, the reflecting boundary model achieves the same error rate for different X_0 for large t_c, because the recency effect attenuates the influence of starting point on final choices.

Fig. 8b compares the performance of bounded WDMs with starting points given by Eq. (11). Biased starting points produce lower error rates for all t_c than $X_0 = 0$. As expected from Fig. 8a, for large t_c the reflecting WDM with the starting points of Eq. (11) has similar error rate to the bounded WDM with $X_0 = 0$ because of the recency effect, while the absorbing boundary model with biased starting points of Eq. (11) achieves superior performance, since it can profit from starting point adjustments. For small t_c (<0.3 s), the two bounded models with biased starting points exhibit similar performances, since early in the decision process most integrator states $|X(t)|$ are distant from the boundaries, so the models approximate an unbounded WDM and hence produce similar error rates.

4. Fitting the models to experimental data

An interesting debate has recently taken place between Ratcliff and Usher and McClelland regarding which model best describes data from the TC paradigm. We briefly recall this to place our results in context.

Eq. (3) predicts that the error rate of the unbounded WDM under the TC paradigm converges to zero as $t_c \to \infty$ if the sign of μ represents the correct choice, so that subjects can achieve arbitrarily small error rate even for difficult tasks (small μ). However, it is known that humans cannot produce 100% accuracy even for large t_c. For example, Usher and McClelland (2001) performed an experiment in which subjects were required to decide whether a rectangle is tilted to the left or to the right and to