Journal of Mathematical Psychology 75 (2016) 96–109

Contents lists available at ScienceDirect

Journal of Mathematical Psychology

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

Explicit moments of decision times for single- and double-threshold drift-diffusion processes

V. Srivastava [a,*], P. Holmes [a,b], P. Simen [c]

[a] Department of Mechanical and Aerospace Engineering, Princeton University, NJ 08544, United States

[b] Program in Applied and Computational Mathematics and Princeton Neuroscience Institute, Princeton University, NJ 08544, United States

[c] Department of Neuroscience, Oberlin College, OH 44704, United States

HIGHLIGHTS

- Analytic expressions for first three unconditioned and conditioned moments of decision time for pure drift-diffusion model.
- Semi-analytic expressions for first three unconditioned and conditioned moments of decision time for extended drift-diffusion model.
- Thorough analysis of the behavior of moments of decision time as a function of model parameters.
- Analysis of the effect of non-decision time on moments of reaction time.

ARTICLE INFO

Article history:
Available online 25 April 2016

Keywords:
Decision time
Diffusion model
Conditioned and unconditioned moments

ABSTRACT

We derive expressions for the first three moments of the decision time (DT) distribution produced via first threshold crossings by sample paths of a drift-diffusion equation. The "pure" and "extended" diffusion processes are widely used to model two-alternative forced choice decisions, and, while simple formulae for accuracy, mean DT and coefficient of variation are readily available, third and higher moments and conditioned moments are not generally available. We provide explicit formulae for these, describe their behaviors as drift rates and starting points approach interesting limits, and, with the support of numerical simulations, discuss how trial-to-trial variability of drift rates, starting points, and non-decision times affect these behaviors in the extended diffusion model. Both unconditioned moments and those conditioned on correct and erroneous responses are treated. We argue that the results will assist in exploring mechanisms of evidence accumulation and in fitting parameters to experimental data.

© 2016 Elsevier Inc. All rights reserved.

1. Introduction

In this paper we derive explicit expressions for the mean, variance, coefficient of variation and skewness of decision times (DTs) predicted by the stochastic differential equation (SDE)

$$dx = a\,dt + \sigma\,dW, \quad x(0) = x_0, \tag{1}$$

which models accumulation of the difference $x(t)$ between the streams of evidence in two-alternative forced-choice tasks. An example of such a perceptual decision-making task is one in which a participant determines if the image on the screen has more white or black pixels (e.g., Ratcliff & Rouder, 1998). Here drift rate a and standard deviation σ are constants, dW denotes independent random (Wiener) increments, and dx is the change in evidence during the time interval $(t, t + dt)$. Decision times (DTs) are determined by first passages through upper and lower thresholds $x = +z$ and $-z$ that respectively correspond to correct responses and errors, between which the starting point x_0 is assumed to lie. Thus, without loss of generality we may set $a > 0$, although we will also consider limits $a \to 0$. Predictions of response times (RTs) for comparison to behavioral data are obtained by adding to DTs a non-decision latency, T_{nd}, to account for sensory and motor processes.

SDEs like Eq. (1) are variously called diffusion or drift-diffusion models (DDMs); in Bogacz, Brown, Moehlis, Holmes, and Cohen (2006) Eq. (1) was named the pure DDM to distinguish it from Ratcliff's extended diffusion model (Ratcliff, 1978), which allows trial to trial variability in drift rates and starting points x_0. See Bogacz et al. (2006), Ratcliff (1978) and Ratcliff and Smith (2004) for background on diffusion models, and note that several different variable-naming conventions are used in parameterizing DDMs, e.g. in Ratcliff (1978), Ratcliff and Smith (2004) and Wagenmakers,

* Corresponding author.
E-mail address: vaibhavs@princeton.edu (V. Srivastava).

http://dx.doi.org/10.1016/j.jmp.2016.03.005
0022-2496/© 2016 Elsevier Inc. All rights reserved.

Grasman, and Molenaar (2005) v and s replace a and σ, and thresholds are set at $x = 0$ and $x = a$ with $x_0 \in [0, a]$; in Bogacz et al. (2006) a and σ are named A and c.

Many of the following results have appeared in the stochastic processes literature, or are implicit in it, and some have appeared in the psychological literature (e.g. Grasman, Wagenmakers, & van der Maas, 2009; Ratcliff, 1978; Wagenmakers et al., 2005). However, their dependence on key parameters such as threshold and starting point and behaviors in the limits of low and high drift rates have not been fully explored (see Wagenmakers et al., 2005 for some cases of $a \to 0$). Nor are we aware of explicit derivations of third order moments. Here we provide these, and also prove a Proposition that describes the structure of the coefficient of variation (CV) for DTs predicted by Eq. (1), relating it to the CV for a single-threshold DDM. We end by considering the extended DDM, introduced in Ratcliff (1978), showing how trial-to-trial variability of drift rates and starting points affects the results for the pure DDM and examining the effects of non-decision latency on response times. We summarize the expressions for the unconditioned and conditioned moments of DTs for the pure DDM in Table 1. The MatLab and R implementation of analytic and semi-analytic expressions for the conditioned and unconditioned moments of DTs for the pure and extended DDMs studied here is available at: https://github.com/PrincetonUniversity/higher_moments_ddm.

Notation and units

We start by reviewing definitions and dimensional units, and establishing notation. For a random variable ξ, we define the nth non-central moment by $\mathbb{E}[\xi^n]$ and the nth central moment by $\mathbb{E}[(\xi - \mathbb{E}[\xi])^n]$. The first central moment is zero and the second central moment is the variance. The coefficient of variation (CV) of ξ is defined as the ratio of standard deviation to mean of ξ, i.e., $\mathrm{CV} = \sqrt{\mathbb{E}[(\xi - \mathbb{E}[\xi])^2]}/\mathbb{E}[\xi]$. Similarly, the skewness of ξ is defined as the ratio of the third central moment to the cube of the standard deviation of ξ:

$$\mathrm{skew} = \frac{\mathbb{E}[(\xi - \mathbb{E}[\xi])^3]}{\mathbb{E}[(\xi - \mathbb{E}[\xi])^2]^{3/2}}.$$

The variable $x(t)$ and thresholds $\pm z$ in Eq. (1) are dimensionless, while the parameters a and σ have dimensions $[\mathrm{time}]^{-1}$ and $[\mathrm{time}]^{-\frac{1}{2}}$ respectively. When providing numerical examples we will work in secs. For $a > 0$ we define the normalized threshold k_z and starting point k_x:

$$k_z = \frac{az}{\sigma^2} \geq 0 \quad \text{and} \quad k_x = \frac{ax_0}{\sigma^2} \in (-k_z, k_z); \tag{2}$$

these nondimensional parameters will allow us to give relatively compact expressions.

2. The single-threshold DDM

Eq. (1) with a single upper threshold $z > 0$ necessarily produces only correct responses in decision tasks, but it is of interest because it provides a simple approximation of the double-threshold DDM when accuracy is at ceiling and errors due to passages through the lower threshold are rare. Specifically, for $a > 0$, DTs of this model with starting point x_0 are described by the Wald (inverse-Gaussian) distribution (Borodin & Salminen, 2002, Eq. (2.0.2); Luce, 1986; Wald, 1947):

$$p(t) = \frac{z - x_0}{\sigma}\sqrt{\frac{1}{2\pi t^3}} \exp\left(\frac{-(z - x_0 - at)^2}{2\sigma^2 t}\right). \tag{3}$$

The mean DT, its variance, and CV are:

$$\mathbb{E}[\mathrm{DT}] = \frac{\sigma^2}{a^2}(k_z - k_x), \qquad \mathrm{Var}[\mathrm{DT}] = \frac{\sigma^4}{a^4}(k_z - k_x), \quad \text{and}$$
$$\mathrm{CV} = \frac{\sqrt{\mathrm{Var}[\mathrm{DT}]}}{\mathbb{E}[\mathrm{DT}]} = \frac{1}{\sqrt{k_z - k_x}}, \tag{4}$$

and the skewness is

$$\frac{3}{\sqrt{k_z - k_x}} \quad (= 3\,\mathrm{CV}). \tag{5}$$

In the limit $a \to 0^+$, the distribution (3) converges to the Lévy distribution, and in this limit none of the moments exist. However, as shown below, moments of the double threshold DDM exist in this limit.

The single threshold process has been proposed as a model for interval timing (Balci & Simen, 2014; Luzardo, Ludvig, & Rivest, 2013; Simen, Balci, deSouza, Cohen, & Holmes, 2011; Simen, Vlasov, & Papadakis, 2016). Interval timing, loosely defined, is the capacity either to make a response or judgment at a specific time relative to some event in the environment, or simply to estimate inter-event durations. Classic timing tasks include "production" tasks, such as the Fixed Interval (FI) task, in which a participant receives a reward for any response produced after a delay of a given duration since the last reward was received (Ferster & Skinner, 1957), and discrimination tasks, in which two different stimulus durations are compared to see which is longer (see Creelman, 1962 and Treisman, 1963 for historical reviews of early human timing research). Production tasks can be modeled similarly to decision tasks by a diffusion model: instead of accumulating evidence about a perceptual choice, a timing diffusion model accumulates a steady "clock signal" toward a threshold for responding (Creelman, 1962; Gibbon, Church, & Meck, 1984; Killeen & Fetterman, 1988; Treisman, 1963). The resulting production times, relative to stimulus onset, are then comparable to perceptual decision-making response times, typically yielding a slightly positively skewed Gaussian density (Gibbon & Church, 1990). Simen, Rivest, Ludvig, Balci, and Killeen (2013) show that the single-threshold DDM can fit RT data from a variety of interval timing experiments when the starting point is set to 0, drift is set equal to threshold over duration ($a = z/T$, with $T =$ target duration), and normalized thresholds k_z are set to high values, typically of order 20 (see Simen et al., 2011). In contrast, k_z is usually much lower in fits of typical two-choice decision data, typically of order 1. Noise σ is typically fixed at 0.1 in the literature (Vandekerckhove & Tuerlinckx, 2007) and fitted thresholds typically range from 0.05 to 0.15; see, e.g. Balci et al. (2011), Bogacz, Hu, Holmes, and Cohen (2010), Dutilh, Vandekerckhove, Tuerlinckx, and Wagenmakers (2009) and Ratcliff (2014). Despite this difference, DDM can be fitted to both two-choice decision RTs and timed production RTs in humans with suitably larger thresholds for timing (Simen et al., 2016), suggesting that both tasks may be accomplished by similar accumulation processes.

3. The double-threshold DDM: Unconditioned moments of decision time

We now turn to the double-threshold DDM and derive unconditioned moments of decision time. The DT distribution for the double-threshold DDM may be expressed as a convergent series (Ratcliff, 1978, Appendix), and successive moments of the unconditioned DT (i.e. averaged over correct responses and errors) may be obtained by solving boundary value problems for a sequence of linear ordinary differential equations (ODEs) derived from the backwards Fokker–Planck or Kolmogorov equation (Gardiner, 2009, Chap. 5).

Table 1
Summary of expressions of error rate and moments of decision time.

	1-threshold	2-threshold
Error rate		
ER	NA	$\frac{e^{-2k_x}-e^{-2k_z}}{e^{2k_z}-e^{-2k_z}}$
Mean		
E[DT]	$\frac{\sigma^2}{a^2}(k_z-k_x)$	$\frac{\sigma^2}{a^2}\left[k_z\coth(2k_z)-k_z e^{-2k_x}\operatorname{csch}(2k_z)-k_x\right]$
E[DT]$_+$	NA	$\frac{\sigma^2}{a^2}(2k_z\coth(2k_z)-(k_x+k_z)\coth(k_x+k_z))$
Variance		
Var	$\frac{\sigma^4}{a^4}(k_z-k_x)$	see Eq. (10)
Var$_+$	NA	see Eq. (31)
Coefficient of Variation		
CV	$\frac{1}{\sqrt{k_z-k_x}}$	see Eq. (13)
CV$_+$	NA	see Eq. (33)
Skewness		
Skew $\times$ $Var^{3/2}$	$\frac{3\sigma^6}{a^6}(k_z-k_x)$	see Eq. (21)
Skew$_+$ $\times$ $Var_+^{3/2}$	NA	see Eq. (36)

3.1. Error rate and expected decision time

The expressions for error rate and mean decision time are well known, although the following forms are more compact than those given in Bogacz et al. (2006), for example:

$$\mathrm{ER}=\frac{e^{-2k_x}-e^{-2k_z}}{e^{2k_z}-e^{-2k_z}}, \tag{6}$$

$$\mathbb{E}[\mathrm{DT}]=\frac{\sigma^2}{a^2}\left[k_z\coth(2k_z)-k_z e^{-2k_x}\operatorname{csch}(2k_z)-k_x\right]. \tag{7}$$

In Appendix A we show that these expressions agree with the analogous ones of Bogacz et al. (2006).

For an unbiased starting point $k_x=0$ the mean decision time becomes

$$\mathbb{E}[\mathrm{DT}]=\frac{\sigma^2 k_z}{a^2}\tanh(k_z), \tag{8}$$

and in the limit $a\to 0$ ($k_z\to 0$, $k_x\to 0$) we have

$$\mathrm{ER}=\frac{k_z-k_x}{2k_z}=\frac{z-x_0}{2z}\quad\text{and}\quad \mathbb{E}[\mathrm{DT}]=\frac{\sigma^2(k_z^2-k_x^2)}{a^2}=\frac{z^2-x_0^2}{\sigma^2}. \tag{9}$$

Expressions for the error rate and unconditioned moments of decision time are illustrated in Figs. 1 and 2.

3.2. Variance and coefficient of variation of decision time

We derive the following expression for the unconditioned variance of decision time in Appendix B:

$$\begin{aligned}\mathrm{Var}=\frac{\sigma^4}{a^4}\Big[&3k_z^2\operatorname{csch}^2(2k_z)-2k_z^2e^{-2k_x}\operatorname{csch}(2k_z)\coth(2k_z)\\&-4k_zk_xe^{-2k_x}\operatorname{csch}(2k_z)-k_z^2e^{-4k_x}\operatorname{csch}^2(2k_z)\\&+k_z\coth(2k_z)-k_ze^{-2k_x}\operatorname{csch}(2k_z)-k_x\Big].\end{aligned} \tag{10}$$

For an unbiased starting point $k_x=0$ Eq. (10) reduces to

$$\begin{aligned}\mathrm{Var}&=\frac{\sigma^4}{a^4}\Big[2k_z^2(\operatorname{csch}^2(2k_z)\\&\quad-\operatorname{csch}(2k_z)\coth(2k_z))+k_z(\coth(2k_z)-\operatorname{csch}(2k_z))\Big]\\&=\frac{\sigma^4}{a^4}\left[k_z\tanh(k_z)-k_z^2\operatorname{sech}^2(k_z)\right]\\&=\frac{\sigma^4}{a^4}\left[\frac{k_z(1-4k_ze^{-2k_z}-e^{-4k_z})}{(1+e^{-2k_z})^2}\right]\end{aligned} \tag{11}$$

(cf. Wagenmakers et al., 2005, Eqs. (10–12)), and in the limit $a=0$ we have

$$\mathrm{Var}=\frac{2\sigma^4(k_z^4-k_x^4)}{3a^4}=\frac{2(z^4-x_0^4)}{3\sigma^4}. \tag{12}$$

The coefficient of variation can be determined from Eqs. (10) and (7):

$$\begin{aligned}\mathrm{CV}&=\frac{[\mathrm{Var}]^{\frac12}}{\mathbb{E}[\mathrm{DT}]}\\&=\frac{\left[3k_z^2\operatorname{csch}^2(2k_z)-2k_z^2e^{-2k_x}\operatorname{csch}(2k_z)\coth(2k_z)-\cdots-k_x\right]^{\frac12}}{k_z\coth(2k_z)-k_ze^{-2k_x}\operatorname{csch}(2k_z)-k_x};\end{aligned} \tag{13}$$

the complete numerator appears in brackets in Eq. (10). For $k_x=0$ Eq. (13) reduces to

$$\mathrm{CV}=\sqrt{\frac{1-2k_z\operatorname{csch}(2k_z)}{k_z[\coth(2k_z)-\operatorname{csch}(2k_z)]}}=\sqrt{\frac{1-4k_ze^{-2k_z}-e^{-4k_z}}{k_z(1-e^{-2k_z})^2}}, \tag{14}$$

and in the limit $a\to 0$, from Eqs. (12) and (9) we have

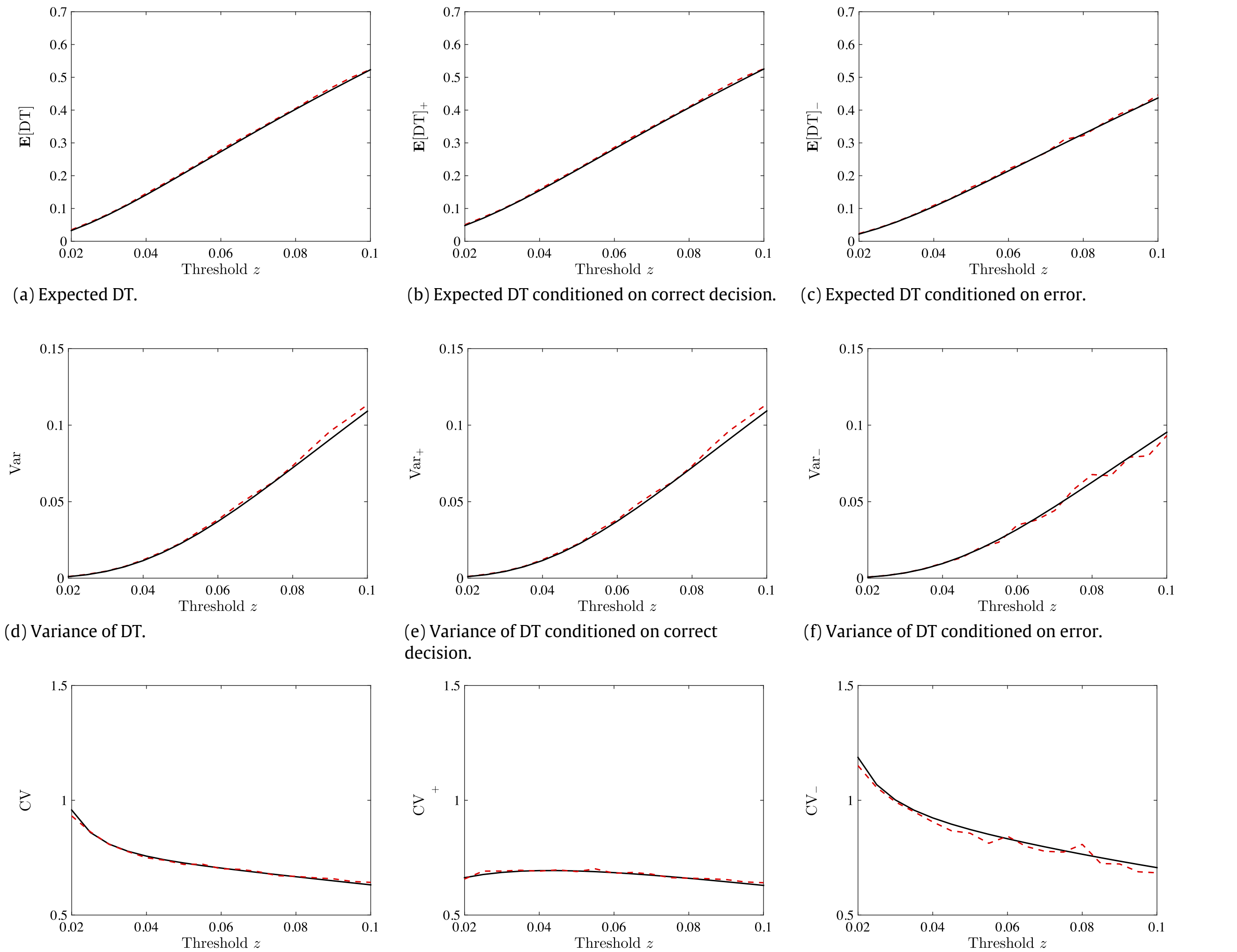
$$\mathrm{CV}=\sqrt{\frac{2(z^2+x_0^2)}{3(z^2-x_0^2)}}\to\sqrt{\frac{2}{3}}\quad\text{as } z\to\infty\text{ or } x_0\to 0. \tag{15}$$

Note that the multiplicative factors σ^2/a^2 cancel and that CV depends only upon the nondimensional threshold and starting point k_z, k_x (or x_0/z in case $a=0$).

If $a>0$, as the threshold z increases, $\mathbb{E}[\mathrm{DT}]$ and Var both increase, but CV decreases, with the following behaviors in the limit $z\to\infty$ ($k_z\to\infty$) for k_x fixed:

$$\frac{\mathbb{E}[\mathrm{DT}]}{k_z}\to\frac{\sigma^2}{a^2},\qquad\frac{\mathrm{Var}}{k_z}\to\frac{\sigma^4}{a^4}\quad\text{and}\quad\mathrm{CV}\to k_z^{-\frac12}; \tag{16}$$

these behaviors follow from the facts that $k_z^m\operatorname{csch}^n(2k_z)\sim k_z^me^{-2nk_z}$ and $\coth(2k_z)\sim 1$ for large k_z. For $a=0$, $\mathbb{E}[\mathrm{DT}]$ and Var also increase with z, as one sees from Eqs. (9) and (12), but CV approaches the limit $\sqrt{2/3}$ (Eq. (15)). In Section 5 we describe the behavior of the CV with unbiased starting point $k_x=0$ throughout the range $k_z\in(0,\infty)$, and show that the CV of the single threshold DDM provides an upper bound for Eq. (14).



(a) Expected DT.
(b) Expected DT conditioned on correct decision.
(c) Expected DT conditioned on error.
(d) Variance of DT.
(e) Variance of DT conditioned on correct decision.
(f) Variance of DT conditioned on error.
(g) CV of DT.
(h) CV of DT conditioned on correct decision.
(i) CV of DT conditioned on error.

Fig. 1. Expected decision times, variances and CVs of decision times for a DDM with $a = 0.2$, $\sigma = 0.1$, and $x_0 = -0.01$, showing dependence on threshold z. Solid curves represent functions derived in Sections 3 and 4; dashed line segments connect point values obtained by 10,000 Monte-Carlo simulations of Eq. (1). Note the non-monotonicity evident in panel h.

3.3. Third moment and skewness of decision time

We end this section by computing the expression for skewness. The third moment of decision time can be computed by solving a boundary value problem analogous to that in Appendix B. However, this computation is very tedious. Instead we obtain skewness from the non-central third moments of DTs conditioned on correct responses and errors derived in Section 4 (this also illustrates the relationships between unconditioned and conditioned moments). Introducing the notation τ for DT, the non-central third moments can be written as

$$\mathbb{E}[\tau^3|x(\tau)=z] = \text{Skew}_+\text{Var}_+^{3/2} + 3\text{Var}_+\,\mathbb{E}[\text{DT}]_+ + \mathbb{E}[\text{DT}]_+^3, \quad \text{and} \tag{17}$$

$$\mathbb{E}[\tau^3|x(\tau)=-z] = \text{Skew}_-\text{Var}_-^{3/2} + 3\text{Var}_-\,\mathbb{E}[\text{DT}]_- + \mathbb{E}[\text{DT}]_-^3, \tag{18}$$

where $\mathbb{E}[\text{DT}]_\pm$, $\text{Var}_\pm$, $\text{Skew}_\pm$ denote expected value, variance, and skewness of DT conditioned on correct responses and errors, respectively. Summing appropriate fractions of these conditioned moments gives the unconditioned third moment

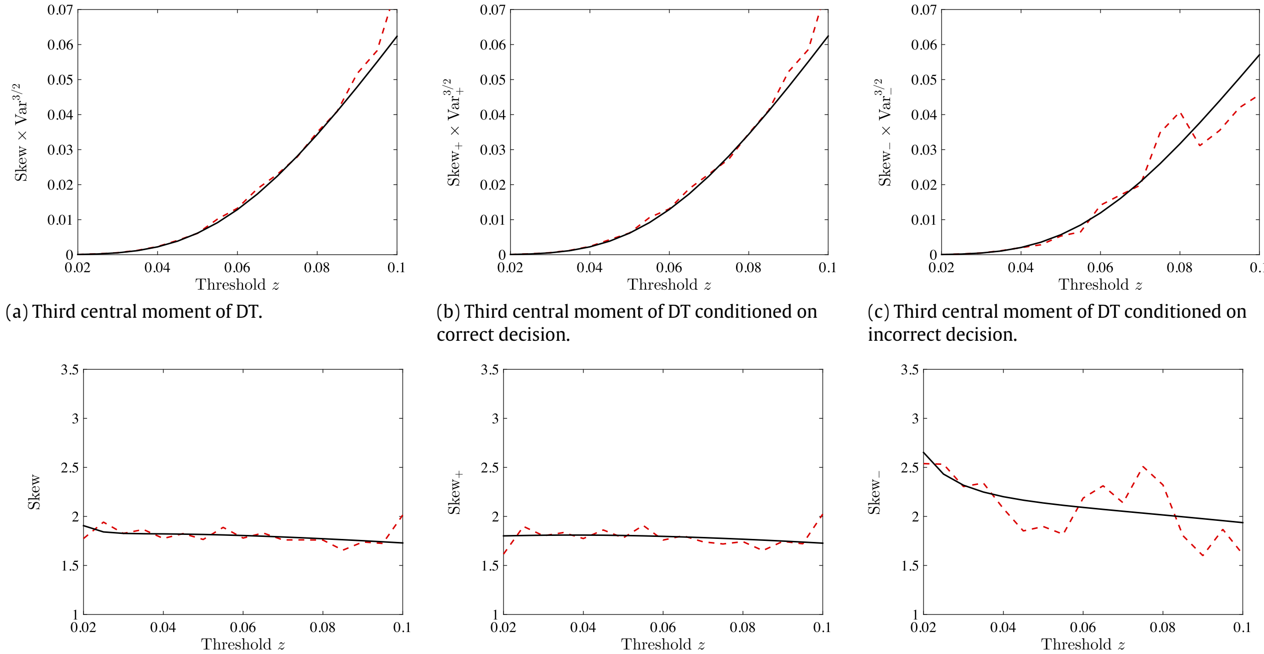
$$\mathbb{E}[\tau^3] = (1-\text{ER}) \times \mathbb{E}[\tau^3|x(\tau)=z] + \text{ER} \times \mathbb{E}[\tau^3|x(\tau)=-z], \tag{19}$$

from which skewness can be derived as follows:

$$\text{Skew} = \mathbb{E}\left[\left(\frac{\tau - \mathbb{E}[\text{DT}]}{\text{Var}^{\frac{1}{2}}}\right)^3\right] = \frac{\mathbb{E}[\tau^3] - 3\text{Var}\,\mathbb{E}[\text{DT}] - \mathbb{E}[\text{DT}]^3}{\text{Var}^{\frac{3}{2}}}. \tag{20}$$

Substituting the expressions (6) for ER and (29), (31) and (36) for conditioned moments from Section 4 into Eqs. (17)–(19), and using the expressions (7) and (10) for the mean and variance of DT, we obtain

$$\begin{aligned}
&\mathbb{E}[\tau^3] - 3\text{Var}\,\mathbb{E}[\text{DT}] - \mathbb{E}[\text{DT}]^3 \\
&= \frac{\sigma^6}{a^6}\Bigg[\Big((24k_xk_z^2 + 6k_z^2 - 12k_z^3)e^{-2k_z-4k_x} \\
&+ (24k_x^2k_z + 24k_xk_z - 16k_z^3 + 6k_z)e^{-2k_x} \\
&- (12k_x^2k_z + 12k_xk_z^2 + 12k_xk_z + 4k_z^3 + 6k_z^2 + 3k_z)e^{4k_z-2k_x} \\
&- (24k_xk_z^2 + 6k_z^2 + 12k_z^3)e^{2k_z-4k_x} \\
&- 8k_z^3e^{-6k_x} - 3k_z\cosh(2k_z) + 3k_z\cosh(6k_z) \\
&+ 9k_x\sinh(2k_z) - 3k_x\sinh(6k_z) + 56k_z^3\cosh(2k_z)
\end{aligned}$$



(a) Third central moment of DT.

(b) Third central moment of DT conditioned on correct decision.

(c) Third central moment of DT conditioned on incorrect decision.

(d) Skewness of DT.

(e) Skewness of DT conditioned on correct decision.

(f) Skewness of DT conditioned on incorrect decision.

Fig. 2. Third central moments and skewnesses of decision times for a DDM with $a = 0.2$, $\sigma = 0.1$, and $x_0 = -0.01$, showing dependence on threshold z. Solid curves represent functions derived in Sections 3 and 4; dashed line segments connect point values obtained by 10,000 Monte-Carlo simulations of Eq. (1).

$$+ 36k_z^2 \sinh(2k_z) - (3k_z - 6k_z^2 + 4k_z^3 + 12k_x k_z$$
$$- 12k_x k_z^2 + 12k_x^2 k_z)e^{-4k_z - 2k_x}\Big) \frac{\operatorname{csch}^3(2k_z)}{4}\Bigg]. \qquad (21)$$

Finally, skewness may be obtained by substituting Eqs. (10) and (21) into Eq. (20). After substitution, the σ^6/a^6 factors cancel out so that, like CV, skewness depends only on k_z and k_x.

For an unbiased starting point $x_0 = k_x = 0$, Eq. (21) can be simplified to

$$\mathbb{E}[\tau^3] - 3\mathrm{Var}\,\mathbb{E}[\mathrm{DT}] - \mathbb{E}[\mathrm{DT}]^3$$
$$= \frac{\sigma^6}{a^6}\Big[3k_z \tanh(k_z) - 3k_z^2 \operatorname{sech}^2(k_z) - 2k_z^3 \tanh(k_z)\operatorname{sech}^2(k_z)\Big]. \qquad (22)$$

We also note that the limits of the double-threshold moments approach those of the single-threshold moments as $k_z \to \infty$ with k_x fixed. Specifically:

$$\frac{\mathbb{E}[\mathrm{DT}]}{k_z} \to \frac{\sigma^2}{a^2}, \qquad \frac{\mathrm{Var}}{k_z} \to \frac{\sigma^4}{a^4},$$
$$\mathrm{CV} \to k_z^{-\frac{1}{2}} \quad \text{and} \quad \mathrm{Skew} \to 3k_z^{-\frac{1}{2}} = 3\,\mathrm{CV}. \qquad (23)$$

In the limit $a \to 0$, we obtain

$$\mathbb{E}[\tau^3] - 3\mathrm{Var}\,\mathbb{E}[\mathrm{DT}] - \mathbb{E}[\mathrm{DT}]^3 = \frac{16(z^6 - x_0^6)}{\sigma^6}, \quad \text{and}$$
$$\mathrm{Skew} = \sqrt{\frac{96}{25}}\frac{(z^6 - x_0^6)}{(z^4 - x_0^4)^{3/2}}, \qquad (24)$$

and the skewness to CV ratio is $12/5$ as $z \to \infty$ or $x_0 \to 0$.

Two further limits are of interest, those in which the starting point approaches either threshold: $k_x \to \pm k_z$ with k_z fixed and finite. In this case ER $\to$ 0 or 1, $\mathbb{E}[\mathrm{DT}] \to 0$, CV $\to \infty$, and Skew $\to \infty$. Letting $k_x = \pm k_z(1 - \epsilon)$ and expanding for small $\epsilon \geq 0$, we have

$$\mathbb{E}[\mathrm{DT}] = \frac{\sigma^2}{a^2}\left[k_z \coth(2k_z) - k_z e^{\mp 2k_z(1-\epsilon)}\operatorname{csch}(2k_z) \mp k_z(1-\epsilon)\right]$$
$$= \frac{\sigma^2}{a^2}\left[\pm 1 - \frac{4k_z}{e^{\pm 4k_z} - 1}\right](k_z \mp k_x) + \mathcal{O}(|k_z \mp k_x|^2) \to 0^+. \qquad (25)$$

Similarly, for the variance and third central moment, we have

$$\mathrm{Var} = \frac{\sigma^2}{a^2}\left[\frac{\mp 8k_z^2(1 + 3e^{\pm 4k_z})}{(e^{\pm 4k_z} - 1)^2} + \frac{4k_z}{e^{\pm 4k_z} - 1} \pm 1\right]$$
$$\times (k_z \mp k_x) + \mathcal{O}(|k_z \mp k_x|^2) \to 0^+, \qquad (26)$$

$$\mathbb{E}[(\tau - \mathbb{E}[\tau])^3] = \mp\frac{\sigma^3}{a^3}\big[18\sinh(2k_z) - 6\sinh(6k_z)$$
$$+ e^{\pm 2k_z}(112k_z^3 - 12k_z) + 24k_z e^{\mp k_z}$$
$$- 12k_z e^{-6k_z} + 256k_z^3 e^{\mp 2k_z} + 16k_z^3 e^{\mp 6k_z}\big]$$
$$\times (k_z \mp k_x) + \mathcal{O}(|k_z \mp k_x|^2) \to 0^+, \qquad (27)$$

so that both CV and skewness diverge like $|k_z \mp k_x|^{-1/2}$. However, the ratio of skewness to CV remains finite as $k_x \to \pm k_z$.

Examples of the functions $\mathbb{E}[\mathrm{DT}]$, Var, CV, skewness and the third central moment of DT are plotted vs. threshold z in the left hand columns of Figs. 1 and 2.

4. The double-threshold DDM: Conditioned moments of decision time

We now turn to moments of DTs conditioned on correct and incorrect responses, deriving them from cumulant and moment generating functions using a method detailed in Appendix C that requires only successive differentiation (see Gut, 2007, Chap 4, Section 6 and Gardiner, 2009, Section 2.6). It suffices to consider only correct decisions, because the moments conditioned on errors can be obtained by replacing x_0 by $-x_0$, or equivalently, k_x by

$$\mathrm{CV}_+ = \frac{\mathrm{Var}_+^{\frac{1}{2}}}{\mathbb{E}[\mathrm{DT}]_+} = \frac{\left[4k_z^2\mathrm{csch}^2(2k_z) + 2k_z\coth(2k_z) - (k_x+k_z)^2\mathrm{csch}^2(k_x+k_z) - (k_x+k_z)\coth(k_x+k_z)\right]^{1/2}}{2k_z\coth(2k_z) - (k_x+k_z)\coth(k_x+k_z)} \tag{33}$$

Box I.

$-k_x$ in the following expressions, as demonstrated by the moment generating functions (58) and (59) in Appendix C. The following expressions for the conditioned moments of decision time are illustrated in Figs. 1 and 2.

4.1. Conditioned cumulant generating function and expected decision time

As derived in Appendix C from Eq. (58), the cumulant-generating function of DTs conditioned on correct decisions is

$$K_+(\alpha) = C(a,\sigma,z,x_0) + \log\sinh\Big(\frac{(z+x_0)\sqrt{a^2-2\alpha\sigma^2}}{\sigma^2}\Big) - \log\sinh\Big(\frac{2z\sqrt{a^2-2\alpha\sigma^2}}{\sigma^2}\Big), \tag{28}$$

where $C(a,\sigma,z,x_0)$ is a function independent of α that will disappear when the cumulants are computed by successive differentiation of $K_+(\alpha)$ with respect to α.

The expected DT conditioned on correct decisions is the first derivative of $K_+(\alpha)$ evaluated at $\alpha = 0$:

$$\begin{aligned}\mathbb{E}[\mathrm{DT}]_+ &= \mathbb{E}[\tau|x(\tau)=z] = \frac{d}{d\alpha}K_+(\alpha)\Big|_{\alpha=0}\\ &= \frac{2z}{a}\coth\Big(\frac{2az}{\sigma^2}\Big) - \frac{z+x_0}{a}\coth\Big(\frac{a(z+x_0)}{\sigma^2}\Big)\\ &= \frac{\sigma^2}{a^2}\Big(2k_z\coth(2k_z) - (k_x+k_z)\coth(k_x+k_z)\Big),\end{aligned} \tag{29}$$

and it can be verified that in the limit $a \to 0^+$

$$\mathbb{E}[\mathrm{DT}]_+ = \frac{4z^2-(z+x_0)^2}{3\sigma^2}. \tag{30}$$

4.2. Conditioned variance and coefficient of variation of decision time

The variance of DT conditioned on correct decisions is the second derivative of $K_+(\alpha)$ at $\alpha = 0$:

$$\begin{aligned}\mathrm{Var}_+ &= \mathrm{Var}[\tau|x(\tau)=z] = \frac{d^2}{d\alpha^2}K_+(\alpha)\Big|_{\alpha=0}\\ &= \frac{4z^2}{a^2}\mathrm{csch}^2\Big(\frac{2za}{\sigma^2}\Big) + \frac{2\sigma^2 z}{a^3}\coth\Big(\frac{2za}{\sigma^2}\Big)\\ &\quad - \frac{(z+x_0)^2}{a^2}\mathrm{csch}^2\Big(\frac{a(z+x_0)}{\sigma^2}\Big)\\ &\quad - \frac{\sigma^2(z+x_0)}{a^3}\coth\Big(\frac{a(z+x_0)}{\sigma^2}\Big)\\ &= \frac{\sigma^4}{a^4}\left[4k_z^2\mathrm{csch}^2(2k_z) + 2k_z\coth(2k_z)\right.\\ &\quad \left. - (k_x+k_z)^2\mathrm{csch}^2(k_x+k_z) - (k_x+k_z)\coth(k_x+k_z)\right];\end{aligned} \tag{31}$$

in the limit $a \to 0^+$:

$$\mathrm{Var}_+ = \frac{32z^4 - 2(z+x_0)^4}{45\sigma^4}. \tag{32}$$

The CV of DT conditioned on correct decisions is therefore CV_+ given in Box I; again, the factors σ^2/a^2 cancel and the conditioned CV depends only on k_z and k_x.

As in Section 3 Eqs. (25)–(26), it can be shown that CV_+ diverges as $k_x \to k_z$ (and hence, by the $k_x \leftrightarrow -k_x$ symmetry, CV_- diverges as $k_x \to -k_z$). However, the behavior as $k_x \to -k_z$ is more interesting and quite subtle, especially as k_z also becomes small. To study this double limit we first set $k_x = \beta k_z$, where $\beta \in (-1, 1)$, and expand the hyperbolic functions in Taylor series for $k_z \ll 1$ (e.g. Abramowitz & Stegun, 1984, Eqs.(4.5.65-66) to obtain

$$\begin{aligned}\mathrm{CV}_+ &= \frac{\left[\frac{2}{45}(\beta^2+2\beta+5)(3-2\beta-\beta^2)k_z^4 + O(k_z^6)\right]^{1/2}}{\frac{1}{3}(3-2\beta-\beta^2)k_z^2 + O(k_z^4)}\\ &= \left[\frac{2(\beta^2+2\beta+5)+O(k_z^2)}{5(3-2\beta-\beta^2)+O(k_z^2)}\right]^{1/2}.\end{aligned} \tag{34}$$

It follows that

$$\mathrm{CV}_+ \to \sqrt{\frac{2(\beta^2+2\beta+5)}{5(3-\beta^2-2\beta)}} \quad \text{as } k_z \to 0^+. \tag{35}$$

In these distinguished limits, CV_+ can approach any value in the range $(\sqrt{2/5}, \infty)$. For $\beta = 0$ ($k_x = 0$) the starting point is unbiased (or $a = 0$), and we obtain the limit $\mathrm{CV}_+ = \sqrt{2/3}$, as for the unconditioned CV; cf. Eq. (15) and see Proposition 5.1. For $\beta \to 1^-$ the starting point lies on the correct threshold and CV_+ diverges as noted above. Aspects of this limiting behavior are illustrated in Fig. 4.

4.3. Conditioned third moment and skewness of decision time

The third central moment of DT conditioned on correct decisions is the third derivative of $K_+(\alpha)$, evaluated at $\alpha = 0$. The skewness of DT is obtained by dividing the third central moment with the cube of standard deviation. Thus, the third central moment of DT is

$$\begin{aligned}\mathrm{Skew}_+\mathrm{Var}_+^{\frac{3}{2}} &= \mathrm{Var}_+^{\frac{3}{2}} \times \mathrm{skewness}[\tau|x(\tau)=z] = \frac{d^3}{d\alpha^3}K_+(\alpha)\Big|_{\alpha=0}\\ &= \frac{12\sigma^2z^2}{a^4}\mathrm{csch}^2\Big(\frac{2az}{\sigma^2}\Big) + \frac{16z^3}{a^3}\coth\Big(\frac{2az}{\sigma^2}\Big)\mathrm{csch}^2\Big(\frac{2az}{\sigma^2}\Big)\\ &\quad + \frac{6\sigma^4 z}{a^5}\coth\Big(\frac{2az}{\sigma^2}\Big) - \frac{3\sigma^2(z+x_0)^2}{a^4}\mathrm{csch}^2\Big(\frac{a(z+x_0)}{\sigma^2}\Big)\\ &\quad - \frac{2(z+x_0)^3}{a^3}\coth\Big(\frac{a(z+x_0)}{\sigma^2}\Big)\mathrm{csch}^2\Big(\frac{a(z+x_0)}{\sigma^2}\Big)\\ &\quad - \frac{3\sigma^4(z+x_0)}{a^5}\coth\Big(\frac{a(z+x_0)}{\sigma^2}\Big)\\ &= \frac{\sigma^6}{a^6}\left[12k_z^2\mathrm{csch}^2(2k_z) + 16k_z^3\coth(2k_z)\mathrm{csch}^2(2k_z)\right.\\ &\quad + 6k_z\coth(2k_z) - 3(k_z+k_x)^2\mathrm{csch}^2(k_z+k_x)\\ &\quad - 2(k_z+k_x)^3\coth(k_z+k_x)\mathrm{csch}^2(k_z+k_x)\\ &\quad \left. - 3(k_z+k_x)\coth(k_z+k_x)\right].\end{aligned} \tag{36}$$

An expression for Skew_+ is obtained by dividing Eq. (36) by the 3/2 power of Eq. (31). In the limit $a \to 0^+$ Eq. (36) becomes

$$\mathrm{Skew}_+\mathrm{Var}_+^{\frac{3}{2}} = \frac{1024z^6 - 16(z+x_0)^6}{945\sigma^6} \quad \text{and} \quad \mathrm{Skew}_+ = \sqrt{\frac{45}{2}}\left[\frac{8(64z^6-(z+x_0)^6)}{21(16z^4-(z+x_0)^4)^{\frac{3}{2}}}\right]. \tag{37}$$

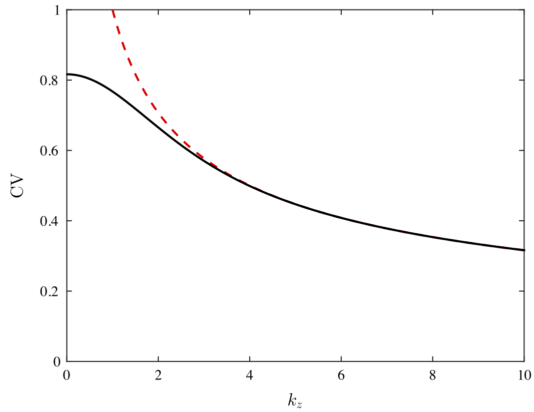


Fig. 3. Coefficients of variation as functions of k_z for the single threshold DDM (dashed) and the DDM with double thresholds and $k_x = 0$ (solid).

Similar to CV_+, Skew_+ diverges as $k_x \to k_z$. For $k_x = \beta k_z$ and $\beta \in (-1, 1)$,

$$\mathrm{Skew}_+ \to \frac{4\sqrt{10}}{7} \frac{(\beta^2+3)(\beta^2+4\beta+7)}{(\beta^2+2\beta+5)^{3/2}(3-2\beta-\beta^2)^{1/2}}, \quad \text{as } k_z \to 0^+. \tag{38}$$

In these distinguished limits Skew_+ can approach any value in the range $(4\sqrt{10}/7, \infty)$.

In Figs. 1 and 2 key expressions derived above are plotted vs. threshold z for the DDM (1) with $a = 0.2, \sigma = 0.1$, and $x_0 = -0.01$. These parameter values were chosen as representative of fits to human data (e.g. Simen et al., 2009), and to illustrate the general forms of the functions. Drift values in this case might be expected to range from -0.4 to 0.4 (e.g. Ratcliff, 2014). See also, among many others, Balci et al. (2011), Balci and Simen (2014), Bogacz et al. (2010) and Dutilh et al. (2009), for similar ranges of fitted parameter values. The results of Monte-Carlo simulations of Eq. (1) using the Euler–Maruyama method (Higham, 2001) with step size 10^{-4} are also shown for comparison. Note that, even with 10,000 sample paths, numerical estimates of the third moment and skewness have not converged very well.

5. Behavior of CVs

We first consider the unconditioned CV with unbiased starting point $x_0 = k_x = 0$, for which we can prove the following result.

Proposition 5.1 (*Behavior of CVs of Decision Times for the DDM*)**.** *The CV for the double-threshold DDM with $k_x = 0$, Eq. (14), is bounded above by the CV for the single-threshold DDM, Eq. (4):*

$$\frac{\sqrt{\frac{1}{k_z}(1-e^{-4k_z}-4k_z e^{-2k_z})}}{1-e^{-2k_z}} \stackrel{\text{def}}{=} F(k_z) < \sqrt{\frac{1}{k_z}}. \tag{39}$$

Moreover, $F(0) = \sqrt{2/3}$ and $F(k_z)$ decays monotonically as k_z increases.

For the proof of the above proposition see Appendix D. Fig. 3 illustrates the proposition by plotting both CV functions over the range $0 \le k_z \le 10$.

It seems difficult to prove a result analogous to Proposition 5.1 for the general CV expressions of Eqs. (10) and (31) due to their complexity. However, plots of the unconditioned and conditioned CVs as functions of the normalized threshold and starting point $k_z = az/\sigma^2$ and $k_x = ax_0/\sigma^2$ shown in Fig. 4 illustrate their behavior over the (k_z, k_x)-plane.

Here, as shown in Proposition 5.1 and Eqs. (34)–(35), for $k_x = 0$ both conditioned and unconditioned CVs converge to $\sqrt{2/3}$ from below as $k_z \to 0^+$ (see right column). However, for $k_x \neq 0$, the behavior is significantly different. In particular, as shown in Section 3, Eqs. (25)–(26), the unconditioned CVs diverge as $k_x \to \pm k_z$ (see left column). CVs for symmetric starting points $\pm k_x$ diverge along different curves as $|k_x| \to k_z$; however, these curves converge to each other as $k_z \to 0^+$ (see left column). Similarly, CVs conditioned on correct responses and errors diverge as $k_x \to k_z$ and $k_x \to -k_z$ respectively. Interestingly, CVs conditioned on correct responses and errors converge to finite limits *smaller than* $\sqrt{2/3}$ as $k_x \to -k_z < 0$ and $k_x \to k_z > 0$ respectively. In Fig. 4(d), as shown in Section 4, CV_+ converges to $\sqrt{2/5}$ as $k_x \to -k_z$ and $k_z \to 0^+$. It is interesting to note that this convergence is not monotone.

The bottom four panels of Fig. 4 illustrate the symmetry of moments conditioned on correct responses and errors with respect to $k_x \mapsto -k_x$, noted at the beginning of Section 4. Unlike the case $k_x = 0$ for which CV is monotonic in k_z, as shown in Proposition 5.1, conditioned CVs are not monotone functions of z or k_z in general. Some instances of non-monotonicity appear in Figs. 1(h), 4(d) and (f).

6. Behavior of moments for the extended DDM

We end by describing some results for the extended DDM introduced by Ratcliff (1978), specifically, the effects of drawing drift rates and starting points for Eq. (1) from Gaussian and uniform distributions $\mathcal{N}(a, \sigma_a)$ and $\mathcal{U}(x_0 - \delta, x_0 + \delta)$ respectively, where $x_0 \pm \delta \in [-z, z]$, and standard deviation σ_a and half-range δ characterize trial-to-trial variability of drift rates and starting points. Complete analytical results on moments for this extended model are not known, and we therefore perform numerical studies. In particular we investigate departures from the analytical results derived above as the variance of the distributions $\mathcal{N}$ and the range of $\mathcal{U}$ increase from zero. We also consider the effects of non-decision time.

6.1. Analytical and semi-analytical expressions

We first discuss how expressions for the moments of decision times and error rate for the pure DDM can be leveraged to efficiently compute analogous explicit expressions for the extended DDM. For clarity, we denote the decision time of the pure DDM for a given drift rate a and starting point x_0 by $\tau(a, x_0)$, and the error rate by $\mathrm{ER}(a, x_0)$. The following expressions for the extended DDM are illustrated in Fig. 5.

The error rate of the extended DDM is the expected value of the error rate of the pure DDM averaged over the distributions of drift rates and starting points:

$$\overline{\mathrm{ER}} = \mathbb{E}_A\big[\mathbb{E}_{X_0}\big[\mathrm{ER}(A, X_0)\big]\big], \tag{40}$$

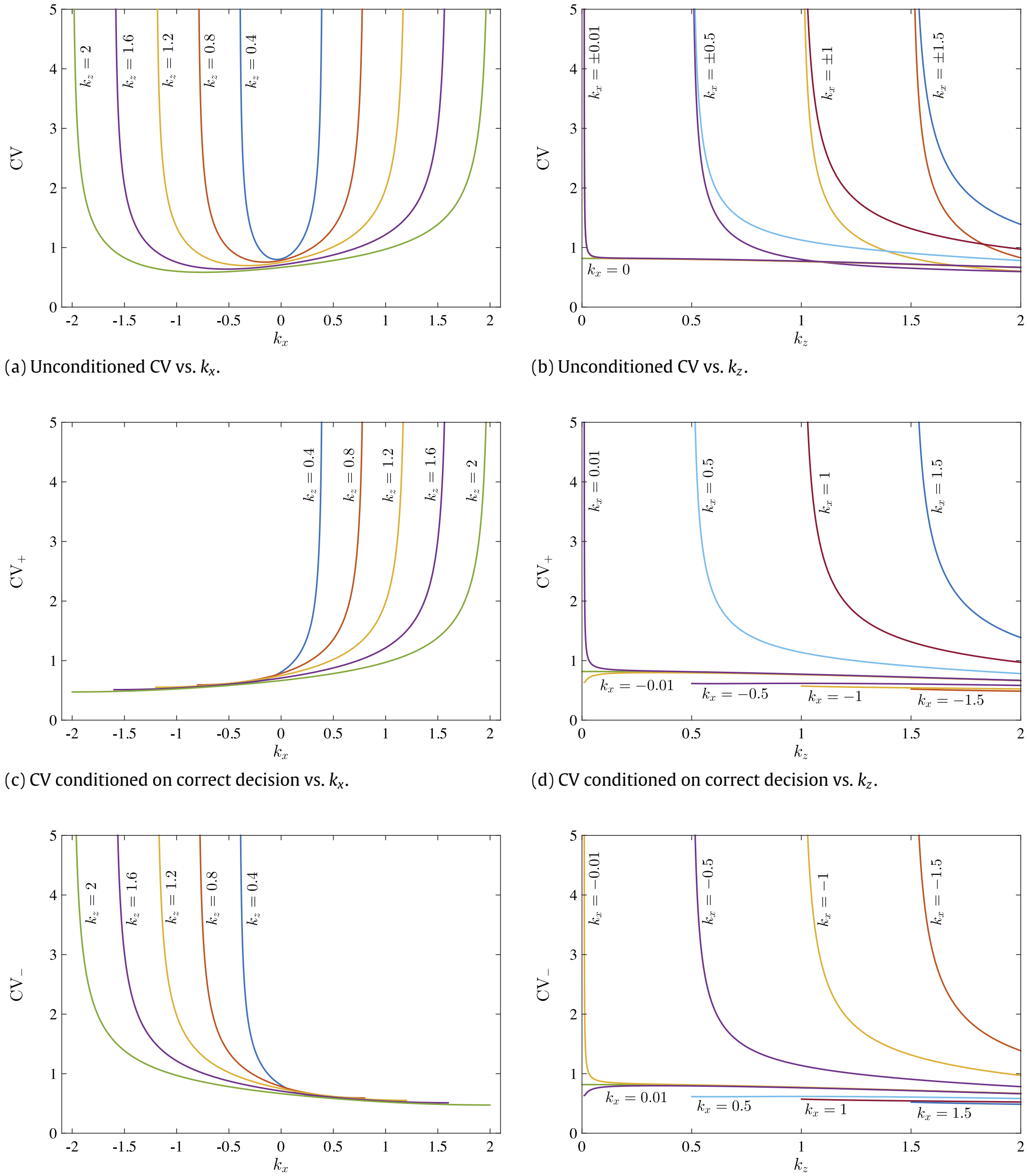
where $\mathbb{E}_Y[\cdot]$ denotes the expected value computed over the distribution of random variable Y. The expectation over the random starting point X_0 in (40) can be computed explicitly as

$$\mathbb{E}_{X_0}[\mathrm{ER}(a, X_0)] = \frac{e^{-2k_x}\mathrm{sinch}(2k_\delta) - e^{-2k_z}}{e^{2k_z} - e^{-2k_z}}, \tag{41}$$

where $k_\delta = a\delta/\sigma^2$ and $\mathrm{sinch}(\cdot) := \sinh(\cdot)/(\cdot)$. Note that this expression reduces to Eq. (6) for $\delta = 0$, using $\mathrm{sinch}(0) = 1$.

The non-central moments of the decision times can be computed similarly. In particular, if $T_n(a, x_0)$ is the non-central nth moment of the decision time for the pure DDM, then the non-central nth moment for the extended DDM is

$$\bar{T}_n = \mathbb{E}_A\big[\mathbb{E}_{X_0}\big[T_n(A, X_0)\big]\big]. \tag{42}$$



(a) Unconditioned CV vs. k_x.

(b) Unconditioned CV vs. k_z.

(c) CV conditioned on correct decision vs. k_x.

(d) CV conditioned on correct decision vs. k_z.

(e) CV conditioned on error vs. k_x.

(f) CV conditioned on error vs. k_z.

Fig. 4. Coefficients of variation of decision time as functions of $k_x = ax_0/\sigma^2$ for several k_z's (left column) and $k_z = az/\sigma^2$ for several k_x's (right column). Unconditioned CVs are shown in top row, conditioned CVs in middle and bottom rows. Observe the symmetry $k_x \mapsto -k_x$ relating the latter, as noted at the beginning of Section 4.

The non-central moments obtained using Eq. (42) can be used with Eqs. (10) and (21) to compute variance and skewness of decision time for the extended DDM. Eq. (42) is valid for both unconditioned and conditioned moments. The above expressions for the error rate and expected decision time for the extended DDM can be found in Bogacz et al. (2006, Appendix, pp 761–763).

For unconditioned moments, the expectation over X_0 in (42) can be computed in closed form for first two moments, which may be written as

$$\mathbb{E}_{X_0}[T_1(a, X_0)] = \frac{\sigma^2}{a^2}\left(k_z \coth(2k_z) - k_z e^{-2k_x}\mathrm{sinch}(2k_\delta)\mathrm{csch}(2k_z) - k_x\right); \quad (43)$$

$$\begin{aligned}\mathbb{E}_{X_0}[T_2(a, X_0)] &= \frac{\sigma^4}{a^4}\Big(k_z^2 + 4k_z^2\mathrm{csch}^2(2k_z) + k_z \coth(2k_z) \\ &\quad - 4k_z^2 e^{-2k_x}\mathrm{sinch}(2k_\delta)\mathrm{csch}(2k_z)\coth(2k_z) \\ &\quad - k_z e^{-2k_x}\mathrm{sinch}(2k_\delta)\mathrm{csch}(2k_z) \\ &\quad - k_x + k_x^2 + \frac{k_\delta^2}{3} - 2k_z k_x \coth(2k_z) \\ &\quad - 2k_z k_x e^{-2k_x}\Big(\mathrm{sinch}(2k_\delta) + \frac{\mathrm{sinch}(2k_\delta) - \cosh(2k_d)}{2k_x}\Big) \\ &\quad \times \mathrm{csch}(2k_z)\Big). \end{aligned} \quad (44)$$

Expected values in Eq. (42), involving integrals over the Gaussian distribution that are not tractable in closed form, can easily be computed numerically, for example, using Simpson's rule.

Fig. 5 illustrates the behavior of the unconditioned moments of the extended DDM, computed as described above. The introduction of variability in starting points results in increase in error rate, decrease in expected decision time, increase in CV, and decrease in skewness to CV ratio. Introduction of variability in drift rate also causes increase in error rate, decrease in expected decision time and increase in CV, but the skewness to CV ratio increases (compare bottom panels). Interestingly, for high values of drift rate variability CV is a monotonically increasing function of k_z, in contrast to the behavior of CV for pure DDM discussed in Section 5. The effect of drift rate variability seems to dominate when both initial condition and drift rate variability are present.

6.2. Effect of non-decision time

Before returning to the extended DDM, we investigate the role of the non-decision part of the reaction time, the sensory–motor latency, on its CV and skewness. Recall that $\mathrm{RT} = \mathrm{DT} + \mathrm{T_{nd}}$, where $\mathrm{T_{nd}}$ is the non-decision time. We define the following coefficients to characterize the dependence of DT and $\mathrm{T_{nd}}$:

$$\rho_{11} = \frac{\mathbb{E}[(\mathrm{DT} - \mathbb{E}[\mathrm{DT}])(\mathrm{T_{nd}} - \mathbb{E}[\mathrm{T_{nd}}])]}{\sqrt{\mathrm{Var}[\mathrm{DT}]\ \mathrm{Var}[\mathrm{T_{nd}}]}}, \tag{45}$$

$$\rho_{12} = \frac{\mathbb{E}[(\mathrm{DT} - \mathbb{E}[\mathrm{DT}])(\mathrm{T_{nd}} - \mathbb{E}[\mathrm{T_{nd}}])^2]}{\sqrt{\mathrm{Var}[\mathrm{DT}]\ \mathbb{E}[(\mathrm{T_{nd}} - \mathbb{E}[\mathrm{T_{nd}}])^4]}}, \tag{46}$$

$$\rho_{21} = \frac{\mathbb{E}[(\mathrm{DT} - \mathbb{E}[\mathrm{DT}])^2(\mathrm{T_{nd}} - \mathbb{E}[\mathrm{T_{nd}}])]}{\sqrt{\mathbb{E}[(\mathrm{DT} - \mathbb{E}[\mathrm{DT}])]^4\ \mathrm{Var}[\mathrm{T_{nd}}]}}. \tag{47}$$

Note that ρ_{11} is the standard correlation coefficient between DT and $\mathrm{T_{nd}}$, and ρ_{12}, ρ_{21} can be interpreted as higher order correlation coefficients. If DT and $\mathrm{T_{nd}}$ are independent, then all these correlation coefficients are zero. In this case, it follows from the definition of RT that

$$\begin{aligned}
\mathbb{E}[\mathrm{RT}] &= \mathbb{E}[\mathrm{DT}] + \mathbb{E}[\mathrm{T_{nd}}] \\
\mathrm{Var}[\mathrm{RT}] &= \mathrm{Var}[\mathrm{DT}] + \mathrm{Var}[\mathrm{T_{nd}}] + 2\rho_{11}\sqrt{\mathrm{Var}[\mathrm{DT}]\ \mathrm{Var}[\mathrm{T_{nd}}]} \\
\mathbb{E}[(\mathrm{RT} - \mathbb{E}[\mathrm{RT}])^3] &= \mathrm{Skew}[\mathrm{DT}]\mathrm{Var}[\mathrm{DT}]^{3/2} + \mathrm{Skew}[\mathrm{T_{nd}}]\mathrm{Var}[\mathrm{T_{nd}}]^{3/2} \\
&\quad + 3\rho_{12}\sqrt{\mathrm{Kur}[T_{\mathrm{nd}}]\mathrm{Var}[\mathrm{DT}]}\mathrm{Var}[\mathrm{T_{nd}}] \\
&\quad + 3\rho_{21}\sqrt{\mathrm{Kur}[\mathrm{DT}]\mathrm{Var}[\mathrm{T_{nd}}]}\mathrm{Var}[\mathrm{DT}],
\end{aligned}$$

where $\mathrm{Kur}[\cdot]$ is the kurtosis.[1] The conditioned mean decision time and variance can be defined similarly by introducing conditioned equivalents of correlation coefficients (45)–(47). However, for simplicity of exposition, in the following we assume that non-decision time and decision time are independent; accordingly, the above correlation coefficients are zero. Formulae for CV and Skew for RT's follow immediately from above expressions. For use below, we assume $\mathrm{T_{nd}}$ is uniformly distributed with mean $\mathbb{E}[\mathrm{T_{nd}}]$ and range s_t.

6.3. Effects of trial-to-trial variability

Seeking to provide a more complete picture, we conducted simulations of the extended and pure DD models. To obtain the following simulation results we used the RTdist package for graphical processing unit (GPU) based simulation of the DDM (Verdonck, Meers, & Tuerlinckx, 2015) to simulate a large subset of the parameter space spanning the range of plausible parameter values. We simulated 1,518,750 parameter combinations in about 5.5 h on a Tesla NVIDIA GPU, with 1 msec timesteps up to 5 secs maximum RT, with 10^5 trials simulated per parameter combination. In Fig. 6, the noise level was fixed at $\sigma = 0.1$ and we varied mean drift a and threshold z over the ranges $[0.1, 1.0]$ and $[0.05, 0.3]$ respectively. Fig. 6 shows accuracy, mean RT, CV, skewness to CV ratio (SCV) and the percentage of trials that failed to cross threshold within 5 secs. (The latter quantity is small except for low drift and high threshold, where it rises to 15%–20%.) Note that the left hand column of Fig. 6 shows results for the pure DDM with $\mathrm{T_{nd}} = 0$, and thus provides standards for comparison with other cases. See Srivastava, Holmes, and Simen (2016, Appendix E) for additional simulation results.

The most profound effect on higher moments of the RT distributions is due to changes in non-decision latency, $\mathrm{T_{nd}}$, as shown in Fig. 6. Specifically, note the dramatic drop in the CV of RTs as $\mathrm{T_{nd}}$ increases from 0 to 0.28 s, and the corresponding increase of skewness to CV ratio (red arrows, row 3).

Fig. 7 shows this phenomenon most clearly, using behaviorally plausible values for the extended DDM. When the correct expected non-decision latency of 0.45 s is subtracted from the RTs, the CV (middle plot) approaches $\sqrt{2/3} \approx 0.8165$ as drift approaches 0. Thus researchers may be able to estimate $\mathrm{T_{nd}}$ at low accuracy levels when behavior is unbiased toward either alternative by progressively subtracting from the RT until the CV approaches $\sqrt{2/3}$ from below (cf. Proposition 5.1 and Fig. 3). In contrast, the SCV ratio grows substantially as drift, and hence accuracy, increase (Fig. 6, red arrows, rows 3 and 4). Researchers may therefore be able to estimate $\mathrm{T_{nd}}$ at high drift levels by subtracting postulated non-decision time from the RT until the SCV ratio declines to 3. These two heuristics for estimating $\mathrm{T_{nd}}$ independently at both low and high levels of drift may provide robust and easily-computable sanity checks for constraining the values of $\mathrm{T_{nd}}$ when using fitting algorithms.

7. Conclusion

We analyzed in detail the first three moments of decision times of the pure and extended DDMs. We derived explicit expressions for unconditioned and conditioned moments and used these expressions to thoroughly investigate the behavior of the CV and skewness of decision times in terms of two useful parameters: the non-dimensional threshold and non-dimensional initial condition (k_z and k_x, Eq. (2)). These expressions are summarized in Table 1, and their MatLab and R implementation is available at: https://github.com/PrincetonUniversity/higher_moments_ddm.

In particular, we computed several limits of interest for the pure DDM. We established that, for an unbiased starting point ($x_0 = 0$), the CV of decision times is a monotonically decreasing function of k_z and that it approaches $\sqrt{2/3}$ as $k_z \to 0$ (Proposition 5.1 and Fig. 3). In the limits of small drift rate and unbiased starting point, we showed that the ratio of skewness to CV approaches 12/5. Furthermore, for non-zero drift rates and in the limit of large thresholds (high accuracy), we showed that skewness to CV ratio approaches 3. We showed that both CV and skewness of decision times diverge as the starting point approaches either threshold; however, the ratio of skewness to CV is a bounded function of non-dimensional threshold. We also showed that in the limit of large thresholds, these moments match those of first passage times for single-threshold drift-diffusion processes, and we established similar results for conditioned CV and skewness of decision times. We established that the decision time distribution for the double-threshold DDM converges to the decision time distribution of the single-threshold DDM for large thresholds (Appendix C).

We then derived analytic and semi-analytic expressions for the moments of decision times of the extended DDM, and numerically investigated the effects of trial-to-trial variability in starting points and drift rates on the DDM's performance. We observed that variability in drift rate appears to dominate these effects, compared to starting point variability.

Finally, we investigated the effect of non-decision times (sensory–motor latencies, $\mathrm{T_{nd}}$) on decision-making performance. We observed that CVs of reaction times ($\mathrm{DT} + \mathrm{T_{nd}}$) decrease and

[1] We consider kurtosis as the ratio of the fourth central moment and the square of the variance. This is in contrast to the convention of subtracting 3 from the above ratio so that the kurtosis of the standard normal random variable is zero.

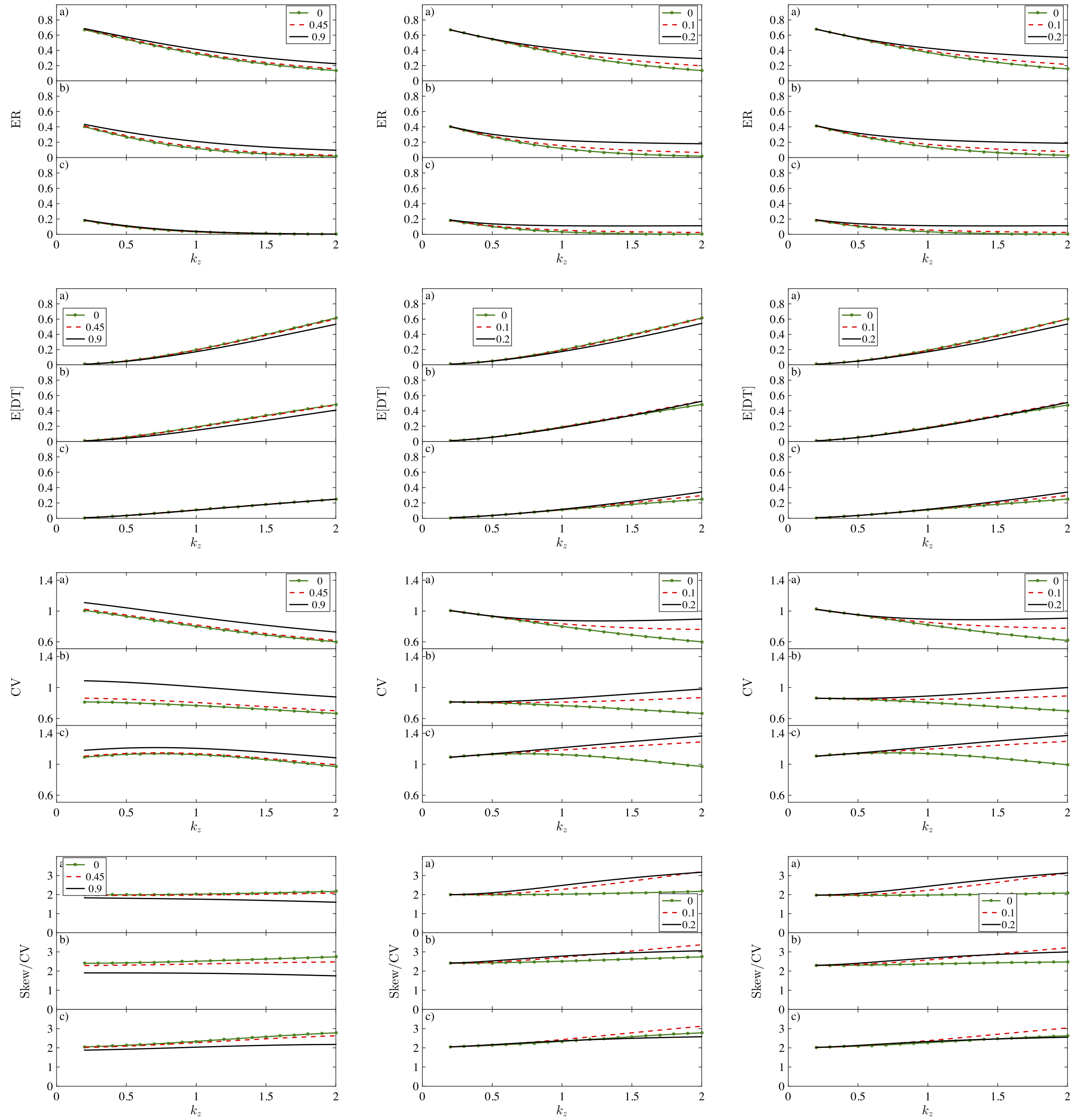


Fig. 5. Behavior of moments for the extended DDM. In all panels $a = 0.2$ and $\sigma = 0.1$. Three sub-panels in each panel correspond to $x_0 = -z/2$, 0 and $z/2$, respectively, from top to bottom. Left panels correspond to $\sigma_a = 0$ and green solid with dots, red dashed, and black solid curves to $\delta = 0, 0.45\min\{z - x_0, x_0 + z\}$ and $0.9\min\{z - x_0, x_0 + z\}$, respectively. Middle panels correspond to $\delta = 0$ and green solid with dots, red dashed, and black solid curves to $\sigma_a = 0$, 0.1 and 0.2, respectively. Right panels are analogous to middle panels and correspond to $\delta = 0.45\min\{z - x_0, x_0 + z\}$.

their skewness-to-CV ratios increase as mean T_{nd}'s increase (Fig. 6). We propose that the decrease in CVs and increase in skewness-to-CV ratios could be used to estimate non-decision times in low and high accuracy regimes respectively (see Fig. 7). The development of rigorous methods using these metrics to estimate non-decision time is a potential avenue for future research.

It should be noted that difficulties in estimating higher moments of empirical RT data have been highlighted in the literature (Luce, 1986; Ratcliff, 1993). However, at least in the context of interval-timing tasks, predictions regarding CV and skewness have proved to be useful in discriminating between different models (Simen et al., 2011, 2016). Furthermore, it is possible that future two-alternative perceptual decision task designs could be found that would yield data amenable to estimation of higher moments, in which case, the expressions we derive here may prove helpful.

More generally, the explicit expressions derived in this paper can be used to quickly identify ranges of parameters that are relevant to fitting specific behavioral data sets, thereby reducing the volumes of multi-dimensional space in which parameter fits need to be run. In principle, the cumulant generating function method outlined in Appendix C can be used to produce formulae for fourth and higher moments, and although the results will be complex, they and their limiting behaviors may also provide guidance for parameter fitting.

Acknowledgments

This work was jointly supported by NIH Brain Initiative grant 1-U01-NS090514-01 (PH), NSF-CRCNS grant DMS-1430077 and the Princeton University Insley-Blair Pyne Fund (VS), and an OKUM

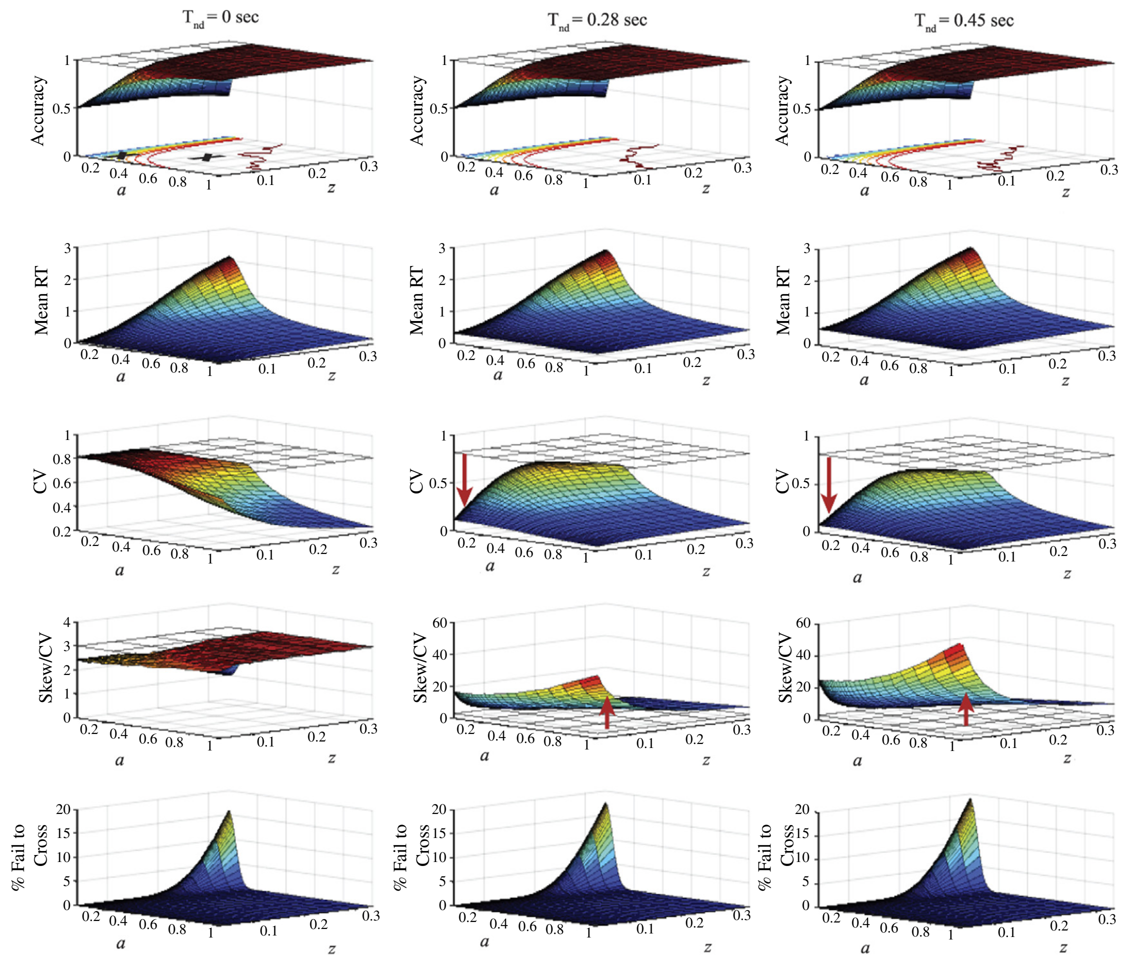


Fig. 6. Effect of deterministic non-decision time T_{nd} in the extended DDM. Left column: $T_{nd} = 0$ s. Middle column: $T_{nd} = 0.28$ s. Right column: $T_{nd} = 0.45$ s. Curves plotted on the drift-threshold plane in top row denote equally spaced contours of the accuracy surface; red arrows show the effect of increasing T_{nd} on CV and SCV.

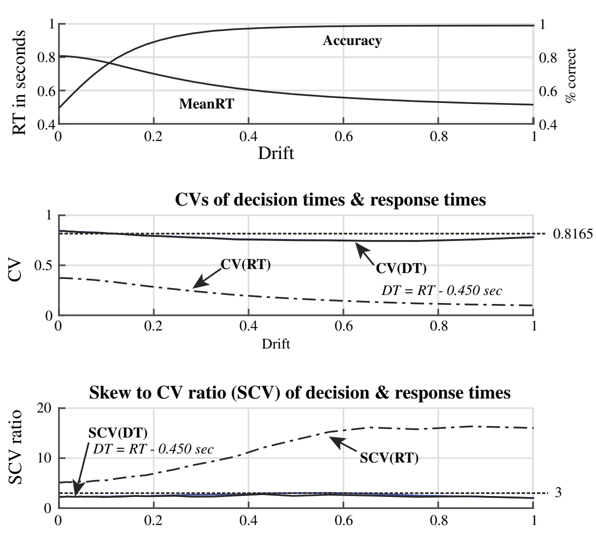


Fig. 7. Comparison of CVs and SCVs, as a function of drift, computed from raw RTs (dot-dashed) and from DTs (solid). DTs have the true, average non-decision latency of 0.45 s subtracted. Representative levels of extended DDM parameter values were used ($z = 0.06$, $x_0 = 0$, $\delta = 0.5\cdot z$, $\sigma_a = 0.25\cdot a$, $\mathbb{E}[T_{nd}] = 0.45$ s, and $s_t = 0.112$ s.)

Fellowship (PS). The authors thank Jonathan Cohen and Michael Shvartsman for helpful suggestions.

Appendix A. Error rate and unconditioned variance of decision time

In this section we show that error rate (6) and expected decision time (7) are equivalent to the expressions given in the subsection "The Drift Diffusion Model" of Bogacz et al. (2006, Appendix, Eqs. (A27-31)). In our notation, the quantities $\tilde{z}$ and $\tilde{a}$ defined in Bogacz et al. (2006) are

$$\tilde{z} = \frac{z}{a}, \quad \text{and} \quad \tilde{a} = \frac{a^2}{\sigma^2}.$$

Define $\tilde{x}_0 = x_0/a$. Note that $k_z = \tilde{z}\tilde{a}$ and $k_x = \tilde{a}\tilde{x}_0$. Also note that $\tilde{x}_0$ and x_0 are referred to as x_0 and y_0, respectively in Bogacz et al. (2006).

The expression (6) for error rate may be rewritten as follows

$$\begin{aligned}
\text{ER} &= \frac{e^{-2k_x} - e^{-2k_z}}{e^{2k_z} - e^{-2k_z}} = \frac{1 - e^{-2k_z}}{e^{2k_z} - e^{-2k_z}} - \frac{1 - e^{-2k_x}}{e^{2k_z} - e^{-2k_z}} \\
&= \frac{e^{2k_z} - 1}{e^{4k_z} - 1} - \frac{1 - e^{-2k_x}}{e^{2k_z} - e^{-2k_z}} \\
&= \frac{e^{2k_z} - 1}{(e^{2k_z} + 1)(e^{2k_z} - 1)} - \frac{1 - e^{-2k_x}}{e^{2k_z} - e^{-2k_z}} \\
&= \frac{1}{1 + e^{2k_z}} - \frac{1 - e^{-2k_x}}{e^{2k_z} - e^{-2k_z}} \\
&= \frac{1}{1 + e^{2\tilde{z}\tilde{a}}} - \left[\frac{1 - e^{-2\tilde{x}_0\tilde{a}}}{e^{2\tilde{z}\tilde{a}} - e^{-2\tilde{z}\tilde{a}}}\right],
\end{aligned} \tag{48}$$

which is identical to the ER expression in Bogacz et al. (2006). Similarly,

$$\begin{aligned}\mathbb{E}[\mathrm{DT}] &= \frac{\sigma^2}{a^2}\left[k_z\coth(2k_z) - k_z e^{-2k_x}\operatorname{csch}(2k_z) - k_x\right]\\ &= \frac{\sigma^2}{a^2}k_z\left[\coth(2k_z) - \operatorname{csch}(2k_z) + (1-e^{-2k_x})\operatorname{csch}(2k_z) - \frac{k_x}{k_z}\right]\\ &= \frac{z}{a}\left[\frac{e^{2k_z}+e^{-2k_z}-2}{e^{2k_z}-e^{-2k_z}} + (1-e^{-2k_x})\operatorname{csch}(2k_z) - \frac{x_0}{z}\right]\\ &= \frac{z}{a}\tanh(k_z) + \frac{2z}{a}\frac{(1-e^{-2k_x})}{e^{2k_z}-e^{-2k_z}} - \frac{x_0}{a}\\ &= \tilde{z}\tanh(\tilde{z}\tilde{a}) + \frac{2\tilde{z}(1-e^{-2\tilde{a}\tilde{x}_0})}{e^{2\tilde{a}\tilde{z}}-e^{-2\tilde{z}\tilde{a}}} - \tilde{x}_0,\end{aligned} \tag{49}$$

which is identical to the expected decision time expression in Bogacz et al. (2006).

Appendix B. Unconditioned variance of decision time

The second moment of the decision time T_2 is the solution of the following linear ODE:

$$a\frac{dT_2}{dx_0} + \frac{\sigma^2}{2}\frac{d^2T_2}{dx_0^2} = -2\mathbb{E}[\mathrm{DT}], \tag{50}$$

with boundary conditions $T_2(\pm z) = 0$ (e.g. Gardiner, 2009, Section 5.5.1; see Eq. (5.5.19) for the general n'th moment ODE). To solve Eq. (50) we first rewrite $\mathbb{E}[\mathrm{DT}]$ to make dependence on the starting point x_0 explicit:

$$\mathbb{E}[\mathrm{DT}] = \alpha_1 - \alpha_2 e^{-2kx_0} - \frac{x_0}{a}.$$

Here $\alpha_1 = \frac{z}{a}\coth(2k_z)$, $\alpha_2 = \frac{z}{a}\operatorname{csch}(2k_z)$ and unlike k_z, k_x defined above, $k = \frac{a}{\sigma^2}$ is independent of z and x_0. A particular solution to (50) is

$$T_2^p = \frac{x_0^2}{a^2} - \alpha_3 x_0 - \frac{2\alpha_2}{a}x_0 e^{-2kx_0},$$

where $\alpha_3 = \frac{2}{a}(\alpha_1 + \frac{1}{2ka})$, and the general solution takes the form

$$T_2(x_0) = c_1 + c_2 e^{-2kx_0} + \frac{x_0^2}{a^2} - \alpha_3 x_0 - \frac{2\alpha_2}{a}x_0 e^{-2kx_0}.$$

Substituting the boundary conditions $T_2(\pm z) = 0$, and solving for c_1 and c_2, we obtain

$$\begin{aligned}c_1 &= \frac{2z^2}{a^2}\coth^2(2k_z) + \frac{z}{ka^2}\coth(2k_z) - \frac{z^2}{a^2} + \frac{2z^2}{a^2}\operatorname{csch}^2(2k_z)\\ &= \frac{z^2}{a^2} + \frac{4z^2}{a^2}\operatorname{csch}^2(2k_z) + \frac{z}{ka^2}\coth(2k_z), \quad \text{and}\\ c_2 &= -\frac{4z^2}{a^2}\operatorname{csch}(2k_z)\coth(2k_z) - \frac{z}{ka^2}\operatorname{csch}(2k_z),\end{aligned}$$

and we therefore find that

$$\begin{aligned}T_2 = {}& \frac{z^2}{a^2} + \frac{4z^2}{a^2}\operatorname{csch}^2(2k_z) + \frac{\sigma^2 z}{a^3}\coth(2k_z)\\ &- \frac{4z^2 e^{-2kx_0}}{a^2}\operatorname{csch}(2k_z)\coth(2k_z) - \frac{\sigma^2 z e^{-2kx_0}}{a^3}\operatorname{csch}(2k_z)\\ &+ \frac{x_0^2}{a^2} - \frac{2zx_0}{a^2}\coth(2k_z) - \frac{\sigma^2 x_0}{a^3} - \frac{2zx_0 e^{-2kx_0}}{a^2}\operatorname{csch}(2k_z).\end{aligned} \tag{51}$$

We can now obtain the expression for the variance of decision time:

$$\begin{aligned}\mathrm{Var} &= T_2 - \mathbb{E}[\mathrm{DT}]^2\\ &= \frac{3z^2}{a^2}\operatorname{csch}^2(2k_z) + \frac{\sigma^2 z}{a^3}\coth(2k_z)\\ &\quad - \frac{\sigma^2 x_0}{a^3} - \frac{2z^2 e^{-2kx_0}}{a^2}\operatorname{csch}(2k_z)\coth(2k_z)\\ &\quad - \frac{\sigma^2 z e^{-2kx_0}}{a^3}\operatorname{csch}(2k_z) - \frac{4zx_0 e^{-2kx_0}}{a^2}\operatorname{csch}(2k_z)\\ &\quad - \frac{z^2 e^{-4kx_0}}{a^2}\operatorname{csch}^2(2k_z).\end{aligned} \tag{52}$$

Equivalently, we may write

$$\begin{aligned}\mathrm{Var} = \frac{\sigma^4}{a^4}\big[&3k_z^2\operatorname{csch}^2(2k_z)\\ &- 2k_z^2 e^{-2k_x}\operatorname{csch}(2k_z)\coth(2k_z) - 4k_z k_x e^{-2k_x}\operatorname{csch}(2k_z)\\ &- k_z^2 e^{-4k_x}\operatorname{csch}^2(2k_z) + k_z\coth(2k_z)\\ &- k_z e^{-2k_x}\operatorname{csch}(2k_z) - k_x\big].\end{aligned} \tag{53}$$

Appendix C. Method for computation of conditioned moments

The *moment generating function* $M_X : \mathcal{H} \to \mathbb{R}_{>0}$ of a random variable X is defined by

$$M_X(\alpha) := \mathbb{E}[e^{\alpha X}],$$

provided the expectation exists for each α in some neighborhood of zero, i.e., for each $\alpha \in \mathcal{H}$, where $\mathcal{H} \subset \mathbb{R}$ is some interval containing zero. The moment generating function is a special case of the *characteristic function* defined on the complex plane (see Grimmett and Stirzaker (2001, Section 5.7, Theorem 12)), and from it the *cumulant generating function* $K_X : \mathcal{H} \to \mathbb{R}$ of X can be obtained by taking the natural logarithm:

$$K_X(\alpha) = \log M_X(\alpha). \tag{54}$$

The nth cumulant κ_n of X is defined as $\kappa_n = \frac{d^n K_X(\alpha)}{d\alpha^n}\big|_{\alpha=0}$, or equivalently $K_X(\alpha) = \sum_{n=1}^{\infty}\frac{\kappa_n\alpha^n}{n!}$. It can then be shown that

$$\kappa_1 = \mu_1, \qquad \kappa_2 = \mu_2^{\mathrm{cen}}, \qquad \kappa_3 = \mu_3^{\mathrm{cen}}, \quad \text{and} \quad \kappa_4 = \mu_4^{\mathrm{cen}} - 3\kappa_2^2,$$

where $\mu_n = \mathbb{E}[X^n]$ and $\mu_n^{\mathrm{cen}} = \mathbb{E}[(X-\mathbb{E}[X])^n]$ denote the nth non-central and central moments. Thus, successive moments of the distribution from which X is drawn can be generated from $M_X(\alpha)$. For further details and derivations of moment generating functions, see Gut (2007, Chap 4, Section 6) and Gardiner (2009, Section 2.6).

We now derive the moment generating function for DTs of the DDM (1). We define $M_{\mathrm{DT}} : \mathcal{A} \to \mathbb{R}_{>0}$, $M_+ : \mathcal{A} \to \mathbb{R}_{>0}$, and $M_- : \mathcal{A} \to \mathbb{R}_{>0}$ by

$$M_{\mathrm{DT}}(\alpha) = \mathbb{E}[e^{\alpha\tau}], \qquad M_+(\alpha) = \mathbb{E}[e^{\alpha\tau}|x(\tau) = z], \quad \text{and}\\ M_-(\alpha) = \mathbb{E}[e^{\alpha\tau}|x(\tau) = -z], \tag{55}$$

where $\mathcal{A} \subset \mathbb{R}$ is some interval containing zero in which the above expectations exist. $M_{\mathrm{DT}}(\alpha)$, $M_+(\alpha)$ and $M_-(\alpha)$ are, respectively, the moment generating functions for unconditioned decision times (for all responses) and for decision times conditioned on correct responses and on errors. Expressions for these functions are well known in the literature (e.g. Borodin & Salminen, 2002). Here, for completeness, we derive them from first principles.

We begin by deriving an expression for $M_+(\alpha)$. We note that for a given set of parameters a, σ, z, and α, $M_+(\alpha)$ depends only on x_0. Let $\tau(x_0)$ denote the decision time (DT) starting from initial condition x_0. Define $g : \mathbb{R} \to \mathbb{R}_{>0}$ as the map from initial condition

x_0 to $M_+(\alpha)\mathbb{P}(x(\tau) = z)$, i.e.,

$$g(x_0) = \mathbb{E}[e^{\alpha\tau(x_0)}\mathbf{1}(x(\tau(x_0)) = z)], \quad (56)$$

where $\mathbf{1}(\cdot)$ is the indicator function.

Consider the evolution of the DDM (1) starting from x_0 at $t = 0$ for an infinitesimal duration $h \in \mathbb{R}_{>0}$. Let $X_h := x(h) = x_0 + ah + \sigma W(h)$. It follows that

$$\begin{aligned} g(x_0) &= \mathbb{E}_{X_h}\mathbb{E}_{\tau(X_h)}[e^{\alpha(h+\tau(X_h))}] \\ &= e^{\alpha h}\mathbb{E}_{X_h}[g(X_h)] \\ &= e^{\alpha h}\Big(g(x_0) + \frac{dg}{dx_0}ah + \frac{1}{2}\frac{d^2g}{dx_0^2}\sigma^2 h\Big) + O(h^2), \end{aligned}$$

where $O(h^2)$ represents terms of order h^2 and higher. Rearranging terms and setting $h \to 0^+$, we obtain the following ODE for g

$$\frac{\sigma^2}{2}\frac{d^2g}{dx_0^2} + a\frac{dg}{dx_0} + \alpha g = 0, \quad (57)$$

with boundary conditions $g(z) = 1$ and $g(-z) = 0$. The solution to (57) is of the form $g(x_0) = \zeta_1 e^{\lambda_1 x_0} + \zeta_2 e^{\lambda_2 x_0}$, where λ_1 and λ_2 are roots of the equation $\sigma^2\lambda^2/2 + a\lambda + \alpha = 0$, i.e.,

$$\lambda_1 = \frac{-a - \sqrt{a^2 - 2\alpha\sigma^2}}{\sigma^2}, \quad \text{and} \quad \lambda_2 = \frac{-a + \sqrt{a^2 - 2\alpha\sigma^2}}{\sigma^2}.$$

Substituting the boundary conditions, we get two simultaneous equations

$$\zeta_1 e^{\lambda_1 z} + \zeta_2 e^{\lambda_2 z} = 1, \quad \text{and} \quad \zeta_1 e^{-\lambda_1 z} + \zeta_2 e^{-\lambda_2 z} = 0,$$

the solution to which is

$$\zeta_1 = \frac{e^{\lambda_1 z}}{e^{2\lambda_1 z} - e^{2\lambda_2 z}}, \quad \text{and} \quad \zeta_2 = -\frac{e^{\lambda_2 z}}{e^{2\lambda_1 z} - e^{2\lambda_2 z}},$$

and consequently,

$$\begin{aligned} g(x_0) &= \frac{e^{\lambda_1(z+x_0)} - e^{\lambda_2(z+x_0)}}{e^{2\lambda_1 z} - e^{2\lambda_2 z}} \\ &= \frac{e^{-a(z+x_0)/\sigma^2}}{e^{-2az/\sigma^2}} \frac{\sinh\left(\frac{(z+x_0)\sqrt{a^2-2\alpha\sigma^2}}{\sigma^2}\right)}{\sinh\left(\frac{2z\sqrt{a^2-2\alpha\sigma^2}}{\sigma^2}\right)} \\ &= e^{\frac{a(z-x_0)}{\sigma^2}} \frac{\sinh\left(\frac{(z+x_0)\sqrt{a^2-2\alpha\sigma^2}}{\sigma^2}\right)}{\sinh\left(\frac{2z\sqrt{a^2-2\alpha\sigma^2}}{\sigma^2}\right)}. \end{aligned}$$

Thus, recalling the definition (56) of $g(x_0)$, the moment-generating function conditioned on correct decisions is

$$\begin{aligned} M_+(\alpha) &= \mathbb{E}[e^{\alpha\tau}|x(\tau) = z] \\ &= \frac{e^{\frac{a(z-x_0)}{\sigma^2}}}{\mathbb{P}(x(\tau) = z)} \frac{\sinh\left(\frac{(z+x_0)\sqrt{a^2-2\alpha\sigma^2}}{\sigma^2}\right)}{\sinh\left(\frac{2z\sqrt{a^2-2\alpha\sigma^2}}{\sigma^2}\right)}, \end{aligned} \quad (58)$$

and substituting this in the definition (54) yields the cumulant generating function (28) used in Section 4.

Similarly, we may obtain analogous expressions for incorrect decisions

$$\begin{aligned} M_-(\alpha) &= \mathbb{E}[e^{\alpha\tau}|x(\tau) = -z] \\ &= \frac{e^{\frac{-a(z+x_0)}{\sigma^2}}}{\mathbb{P}(x(\tau) = -z)} \frac{\sinh\left(\frac{(z-x_0)\sqrt{a^2-2\alpha\sigma^2}}{\sigma^2}\right)}{\sinh\left(\frac{2z\sqrt{a^2-2\alpha\sigma^2}}{\sigma^2}\right)}, \end{aligned} \quad (59)$$

and for all decisions, correct and incorrect:

$$\begin{aligned} M_{\mathrm{DT}}(\alpha) &= \mathbb{E}[e^{\alpha\tau}] \\ &= e^{\frac{-a(z+x_0)}{\sigma^2}} \frac{\sinh\left(\frac{(z-x_0)\sqrt{a^2-2\alpha\sigma^2}}{\sigma^2}\right)}{\sinh\left(\frac{2z\sqrt{a^2-2\alpha\sigma^2}}{\sigma^2}\right)} \\ &\quad + e^{\frac{a(z-x_0)}{\sigma^2}} \frac{\sinh\left(\frac{(z+x_0)\sqrt{a^2-2\alpha\sigma^2}}{\sigma^2}\right)}{\sinh\left(\frac{2z\sqrt{a^2-2\alpha\sigma^2}}{\sigma^2}\right)}. \end{aligned} \quad (60)$$

It should be noted that in the limit $z \to \infty$

$$M_{\mathrm{DT}}(\alpha) = \exp\left(\frac{az}{\sigma^2}\left(1 - \sqrt{1 - \frac{2\alpha\sigma^2}{a^2}}\right)\right),$$

which is the moment generating function of the Wald distribution (Borodin & Salminen, 2002, Eq. 2.0.1), i.e., the decision time distribution of the single-threshold DDM. Consequently, the decision time distribution of the double-threshold DDM converges to the decision time distribution of the single-threshold DDM as $z \to \infty$.

Appendix D. Proof of Proposition 5.1

We first show that the CV for the single-threshold DDM provides an upper bound for the double threshold case. Canceling the $\sqrt{1/k_z}$ terms in the inequality (39), squaring, rearranging and dividing by $2e^{-2k_z}$ shows that this is equivalent to

$$(1 - e^{-2k_z})^2 > (1 - e^{-4k_z} - 4k_z e^{-2k_z}) \Leftrightarrow e^{-2k_z} > 1 - 2k_z, \quad (61)$$

which clearly holds for all $k_z \neq 0$.

We next evaluate the limit of $F(k_z)$ as $k_z \to 0$ by expanding the numerator of Eq. (39) in Taylor series:

$$\begin{aligned} &\sqrt{\frac{1}{k_z}\left[1 - \left(1 - 4k_z + \frac{16k_z^2}{2} - \frac{64k_z^3}{3!}\right) - \left(4k_z(1 - 2k_z + \frac{4k_z^2}{2})\right) + \mathcal{O}(k_z^4)\right]} \\ &= \sqrt{\frac{8}{3}k_z^2 + \mathcal{O}(k_z^3)}. \end{aligned}$$

Expanding the denominator likewise, we have

$$F(k_z) = \frac{\sqrt{\frac{8}{3}k_z^2 + \mathcal{O}(k_z^3)}}{[1 - (1 - 2k_z + \mathcal{O}(k_z^2))]} \to \sqrt{\frac{2}{3}} \quad \text{as } k_z \to 0. \quad (62)$$

The exponentials in the numerator and denominator of $F(k_z)$ decay rapidly, so that it differs from $\sqrt{1/k_z}$ by less than 0.24% for $k_z \geq 4$, implying that the slow monotonic decay $\sim k_z^{\frac{1}{2}}$ dominates for large k_z; see Fig. 3. However, the behavior for smaller k_z is more subtle and requires computation of all terms in the Taylor series.

To prove monotonic decay throughout we use the fact that $F(k_z) > 0$ and show that the derivative of

$$F^2(k_z) = \frac{\left(\frac{1-e^{-4k_z}}{2k_z} - 2e^{-2k_z}\right)}{(1 - e^{-2k_z})^2} \quad (63)$$

is strictly negative for all $k_z > 0$. Henceforth, for convenience, we set $y = 2k_z$ and compute

$$\begin{aligned} &\frac{d}{dy}[F^2(y)] \\ &= \frac{(1 - e^{-y})\left(-\frac{1}{y^2} + \frac{e^{-2y}}{y^2} + \frac{2e^{-2y}}{y} + 2e^{-y}\right) - 2e^{-y}\left(\frac{1-e^{-2y}}{y} - 2e^{-y}\right)}{(1 - e^{-y})^3} \\ &= \frac{\frac{-(1-e^{-y})(1-e^{-2y})}{y^2} - \frac{2e^{-y}(1-e^{-y})}{y} + 2e^{-y}(1 + e^{-y})}{(1 - e^{-y})^3}. \end{aligned} \quad (64)$$

Since $(1-e^{-y})^3 > 0$ it suffices to show that the numerator of Eq. (64) is negative, or, multiplying by y^2e^{3y} and rearranging, that

$$1+e^{3y}+2ye^{2y} \stackrel{\text{def}}{=} L > e^y+e^{2y}+2ye^y+2y^2e^y+2y^2e^{2y} \stackrel{\text{def}}{=} R. \quad (65)$$

We expand both L and R in Taylor series, obtaining

$$\begin{aligned} L &= 1+\left(1+3y+\frac{(3y)^2}{2!}+\frac{(3y)^3}{3!}+\cdots+\frac{(3y)^j}{j!}+\cdots\right) \\ &\quad +2y\left(1+2y+\frac{(2y)^2}{2!}+\cdots+\frac{(2y)^{j-1}}{(j-1)!}+\cdots\right) \\ &= 2+3y+\frac{9y^2}{2}+\frac{27y^3}{6}+\cdots+2y+(2y)^2 \\ &\quad +\frac{(2y)^3}{2!}+\cdots+\frac{(3y)^j}{j!}+\frac{(2y)^j}{(j-1)!}+\cdots \\ &= 2+5y+\frac{17}{2}y^2+\frac{17}{2}y^3+\frac{145}{24}y^4+\frac{403}{120}y^5+\cdots \\ &\quad +\left(\frac{3^j+j2^j}{j!}\right)y^j+\cdots; \quad \text{and} \end{aligned} \quad (66)$$

$$\begin{aligned} R &= 1+y+\frac{y^2}{2!}+\frac{y^3}{3!}+\cdots+\frac{y^j}{j!}+\cdots+1+2y \\ &\quad +\frac{(2y)^2}{2!}+\frac{(2y)^3}{3!}+\cdots+\frac{2^jy^j}{j!}+\cdots \\ &\quad +2y+2y^2+\frac{2y^3}{2!}+\frac{2y^4}{3!}+\cdots+\frac{2y^j}{(j-1)!}+\cdots \\ &\quad +2y^2+2y^3+\frac{2y^4}{2!}+\cdots+\frac{2y^j}{(j-2)!}+\cdots \\ &\quad +2y^2+2^2y^3+\frac{2^3y^4}{2!}+\cdots+\frac{2^{j-1}y^j}{(j-2)!}+\cdots \\ &= 2+5y+\frac{17}{2}y^2+\frac{17}{2}y^3+\frac{145}{24}y^4+\frac{403}{120}y^5+\cdots \\ &\quad +\frac{1+2^j+2j^2+2^{j-1}j(j-1)}{j!}+\cdots. \end{aligned} \quad (67)$$

Note that the first 6 terms of L and R, up to $\mathcal{O}(y^5)$, are identical, and the 4 succeeding coefficients of $L-R$ up to $\mathcal{O}(y^9)$ are strictly positive (specifically, 1/45, 1/30, 11/420 and 1/70). To show that all succeeding coefficients are likewise positive, we make pairwise comparisons of the six terms in the numerator of the general coefficient of $L-R$:

$$\begin{aligned} &3^j+j2^j-[1+2^j+2j^2+2^{j-1}j(j-1)] \\ &\quad =[j2^j-2j^2]+[j2^{j-1}-(1+2^j)]+[3^j-j^22^{j-1}]. \end{aligned} \quad (68)$$

It can be checked that

$$j2^j > 2j^2 \Leftrightarrow 2^j > 2j \quad \text{for } j \geq 3, \quad (69)$$

$$j2^{j-1} > 1+2^j \Leftrightarrow j > 2+\frac{1}{2^{j-1}} \quad \text{for } j \geq 3, \quad (70)$$

$$3^j > j^22^{j-1} \Leftrightarrow \left(\frac{3}{2}\right)^j > \frac{j^2}{2} \quad \text{for } j \geq 10; \quad (71)$$

thus, all coefficients of terms greater than $\mathcal{O}(y^5)$ are strictly positive, completing the proof. □

References

Abramowitz, M., & Stegun, I. (Eds.) (1984). *Handbook of mathematical functions with formulas, graphs, and mathematical tables*. New York: Wiley - Interscience.

Balci, F., & Simen, P. (2014). Decision processes in temporal discrimination. *Acta Psychologica*, *149*, 157–168.

Balci, F., Simen, P., Niyogi, R., Saxe, A., Holmes, P., & Cohen, J. (2011). Acquisition of decision making criteria: Reward rate ultimately beats accuracy. *Attention, Perception & Psychophysics*, *73*(2), 640–657.

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

Bogacz, R., Hu, P., Holmes, P., & Cohen, J. (2010). Do humans produce the speed-accuracy tradeoff that maximizes reward rate? *Quarterly Journal Experimental Psychology*, *63*(5), 863–891.

Borodin, A., & Salminen, P. (2002). *Handbook of Brownian motion: Facts and formulae*. New York: Springer.

Creelman, C. D. (1962). Human discrimination of auditory duration. *The Journal of the Acoustical Society of America*, *34*(5), 582–593.

Dutilh, G., Vandekerckhove, J., Tuerlinckx, F., & Wagenmakers, E.-J. (2009). A diffusion model decomposition of the practice effect. *Psychonomic Bulletin and Review*, *16*(6), 1026–1036.

Ferster, C. B., & Skinner, B. F. (1957). *Schedules of reinforcement*. New York: Appleton-Century-Crofts.

Gardiner, C. (2009). *Stochastic methods: A handbook for the natural and social sciences* (4th Ed.). New York: Springer.

Gibbon, J., & Church, R. M. (1990). Representation of time. *Cognition*, *37*(1), 23–54.

Gibbon, J., Church, R. M., & Meck, W. H. (1984). Scalar timing in memory. *Annals of the New York Academy of Sciences*, *423*(1), 52–77.

Grasman, R. P. P. P., Wagenmakers, E.-J., & van der Maas, H. L. J. (2009). On the mean and variance of response times under the diffusion model with an application to parameter estimation. *Journal of Mathematical Psychology*, *53*(2), 55–68.

Grimmett, G., & Stirzaker, D. (2001). *Probability and random processes*. Oxford, UK: Oxford University Press.

Gut, A. (2007). *Probability: A graduate course* (Corrected 2nd printing). New York: Springer.

Higham, D. (2001). An algorithmic introduction to numerical simulation of stochastic differential equations. *SIAM Review*, *43*(3), 525–546.

Killeen, P. R., & Fetterman, J. G. (1988). A behavioral theory of timing. *Psychological Review*, *95*(2), 274–295.

Luce, R. (1986). *Response times: Their role in inferring elementary mental organization*. New York: Oxford University Press.

Luzardo, A., Ludvig, E. A., & Rivest, F. (2013). An adaptive drift-diffusion model of interval timing dynamics. *Behavioral Processes*, *95*, 90–99.

Ratcliff, R. (1978). A theory of memory retrieval. *Psychological Review*, *85*(2), 59–108.

Ratcliff, R. (1993). Methods for dealing with reaction time outliers. *Psychological Bulletin*, *114*(3), 510–532.

Ratcliff, R. (2014). Measuring psychometric functions with the diffusion model. *Journal of Experimental Psychology: Human Perception and Performance*, *40*(2), 870–888.

Ratcliff, R., & Rouder, J. N. (1998). Modeling response times for two-choice decisions. *Psychological Science*, *9*(5), 347–356.

Ratcliff, R., & Smith, P. (2004). A comparison of sequential sampling models for two-choice reaction time. *Psychological Review*, *111*(2), 333–367.

Simen, P., Balci, F., deSouza, L., Cohen, J., & Holmes, P. (2011). A model of interval timing by neural integration. *Journal of Neuroscience*, *31*(25), 9238–9253.

Simen, P., Contreras, D., Buck, C., Hu, P., Holmes, P., & Cohen, J. (2009). Reward rate optimization in two-alternative decision making: Empirical tests of theoretical predictions. *Journal of Experimental Psychology: Humam Perception Performance*, *35*(6), 1865–1897.

Simen, P., Rivest, F., Ludvig, E. A., Balci, F., & Killeen, P. R. (2013). Timescale invariance in the pacemaker-accumulator family of timing models. *Timing and Time Perception*, *1*(2), 159–188.

Simen, P., Vlasov, K., & Papadakis, S. (2016). Scale (in)variance in a unified diffusion model of decision making and timing. *Psychological Review*, *123*(2), 151–181.

Srivastava, V., Holmes, P., & Simen, P. (2016, January). Explicit moments of decision times for single- and double-threshold drift-diffusion processes. arXiv preprint arXiv:1601.06420.

Treisman, M. (1963). Temporal discrimination and the indifference interval: implications for a model of the 'internal clock'. *Psychological Monographs*, *77*(13), 1–31.

Vandekerckhove, J., & Tuerlinckx, F. (2007). Fitting the Ratcliff diffusion model to experimental data. *Psychonomic Bulletin & Review*, *14*(6), 1011–1026.

Verdonck, S., Meers, K., & Tuerlinckx, F. (2015). Efficient simulation of diffusion-based choice RT models on CPU and GPU. *Behavior Research Methods*, 1–15.

Wagenmakers, E.-J., Grasman, R. P. P. P., & Molenaar, P. C. M. (2005). On the relation between the mean and the variance of a diffusion model response time distribution. *Journal of Mathematical Psychology*, *49*(3), 195–204.

Wald, A. (1947). *Sequential analysis*. New York: John Wiley & Sons.