

# Stochastic Learning Dynamics with Two Applications

## Contents

<b>1</b>	<b>Abstract</b>	<b>1</b>
<b>2</b>	<b>Introduction</b>	<b>2</b>
2.1	Neural . . . . .	2
2.2	Social . . . . .	2
2.3	Extent and Limits of Analogy . . . . .	2
<b>3</b>	<b>Derivation of Model</b>	<b>2</b>
<b>4</b>	<b>Results</b>	<b>3</b>
4.1	Two-Dimensional Decision Making . . . . .	3
4.2	Strategic Decision Making . . . . .	3
4.3	Social Decision Making . . . . .	4
4.4	Comparison to Social Data . . . . .	4
<b>5</b>	<b>Discussion</b>	<b>4</b>
<b>6</b>	<b>Appendix</b>	<b>4</b>
6.1	Derivation of PDEs for waiting time and accuracy . . . . .	4

## 1 Abstract

Learning usually occurs in noisy environments. In order to make a decision, one must accumulate the evidence for different alternatives and choose one only once the uncertainty about the choice is sufficiently reduced. There are two equivalent mathematical formulations of the optimal decision making algorithm: a sequential probability ratio test in which the likelihoods of two hypotheses are evaluated as more data points are collected and a drift-diffusion stochastic differential equation in which the belief about two hypotheses does a random walk until a threshold is reached. In the random walk formulation, the threshold dictating when a decision has been made affects both the accuracy of the decision and the expected time until a decision is reached. If the utility of a decision depends on accuracy and decision time, the threshold can be optimized accordingly. There are two biological systems in which the variables used to make a decision can be described by a random walk: the firing rates of neural populations in the brain of an individual making a decision about a visual stimulus and the beliefs of monkeys about their relative dominance based on the fights they have engaged in. The stochastic differential equations used to model these quite different systems are nearly identical, which suggests that there are universal principles underlying learning through noisy dynamical systems in biology. However, there are (at least) three significant differences between the two models. First, in the social system, there is a strategic component to the system since the optimal threshold of one animal depends on the threshold of the other animal, whereas in the neural system both neural populations are part of the cognitive system of the same individual trying to optimize its decision making. Relatedly, the output of the decision has additional strategic consequences in the social system. Second, whereas in the neural system, the differential equations are often reduced to a one-dimensional model, the social system is inherently two-dimensional. Finally, in the social system, a monkey is not only trying to make a decision about his dominance with respect to one other monkey, but he is additionally trying to make decisions about many other partners at once, whereas in the neural case there is only one decision to be made at a time. As a consequence of these differences... In comparing the output of the model of the social system to data, we found that ...

## 2 Introduction

**2.1 Neural** Stochastic differential equations have been used to model, explain, and predict learning and decision making in both monkeys and humans [6, 2, 7, 1]. Typically, an experimental subject is presented with a visual stimulus that consists of dots, some percentage of which are moving coherently either left or right and the rest of which are moving randomly. The subject is then asked to decide in which direction the dots were moving. If there are two neural populations, each of which fire in response to a particular direction of motion, the firing rates of the two populations will tend to increase or decrease, depending on the strength and direction of the stimulus. Eventually, either the experimenter forces a decision to be made, in which case whichever population has a higher firing rate determines which decision is made, or the firing rate of one population becomes high enough that the subject decides for that direction. To summarize, there are two decision variables— the firing rates of two neural populations responding to left and right motion— that perform a biased random walk until enough evidence has been accumulated to make a decision one way or the other.

**2.2 Social** In groups of pigtailed macaques, different animals have different fighting abilities, and when a pair of animals engages in a fight, one animal may be more likely on average to “win” the fight. Over time, as one animal learns that it is likely to lose fights, it will decide to emit a formal subordination signal, communicating its agreeing to the subordinate role in their relationship [8, 10, 9, 5, 4]. The outcome of a fight is affected by each animal’s allies, the presence of those allies, motivation levels, external environmental variables, and other factors. Additionally, the number of fights that occur in any period of time depends on similarly stochastic factors. Because of this stochasticity, the number of fights won and lost in any period of time is a random variable. Each animal’s opinion about its relative dominance, therefore, will perform a biased random walk until one animal’s opinion of itself is low enough to agree to being subordinate.

Similar models have previously been applied to animal conflicts, although in that case the animals had perfect memory of their past interactions [12]

**2.3 Extent and Limits of Analogy** In Table 1, we list the inputs, decision variables, and outputs of the two systems.

## 3 Derivation of Model

Stochastic differential equations have been phenomenologically applied to learning systems, with some success [6, 2, 7, 1]. However, with two assumptions about the timescales on which probabilistic events occur, a simple mechanistic description of the accumulation of information about random events can be turned into a stochastic differential equation [15]. These assumptions are particularly reasonable in the social system.

$$\begin{aligned} dX_1 &= \left( -\ell X_1(t) + b(2d - 1) \right) dt + b\sqrt{d}dW_1t - b\sqrt{(1-d)}dW_2t \\ dX_2 &= \left( -\ell X_1(t) + b(1 - 2d) \right) dt - b\sqrt{d}dW_1t + b\sqrt{(1-d)}dW_2t, \end{aligned} \tag{1}$$

Each decision variable,  $X_1$  and  $X_2$ , accumulates information about the stimulus. The sign and size of the two variables indicate how much evidence there is for each possible decision outcome. When a lot of evidence for the first possibility has been observed,  $X_1$  will be large and positive, and similarly for the second possibility. If the difference in the variables can be observed, the difference indicates the relative strengths of the evidence for either possibility. If  $Y = X_1 - X_2$ , when  $Y$  is large enough the system decides on  $X_1$  and when  $Y$  is low enough the system decides on  $X_2$ . Specifically, there are two thresholds,  $-T_2$  and  $T_1$  such that if  $Y < -T_2$  the decision is for  $X_2$  and if  $Y > T_1$  the decision is for  $X_1$ . However, it may be the case that the difference in the variables cannot be observed. In this case, the decision depends on whether  $X_1$  or  $X_2$  hits a threshold first. Again, there are two thresholds, now  $-T_2$  and  $-T_1$  and if  $X_1 < -T_1$  the decision is for  $X_1$  and if the  $X_2 < -T_2$  the decision is for  $X_2$ .

In either case, the thresholds affect the expected time until a decision is reached (decision time, DT), the probability that an incorrect decision is made (error rate, ER), and, in the social case, the probability for each animal that he will emit a subordination signal (signaling probability,  $SP_i$ ). These quantities, DT, ER, and  $SP_i$ , satisfy partial differential equations with respect to the starting values of the decision variables. In Appendix 6.1, we derive these partial differential equations. While an analytical solution is known in the one-dimensional case, the two-dimensional case is analytically intractable. We therefore numerically solve the PDEs in the two-dimensional case.

In Table 2, we compare

A drift diffusion model is equivalent to a sequential probability ratio test, a learning algorithm that is optimal with respect to its accuracy and the time required to make a decision [17, 1].

## 4 Results

**4.1 Two-Dimensional Decision Making** In the neural system, the two stochastic equations describing the activity of the two neural populations integrating environmental evidence are often reduced to a one-dimensional equation describing the difference in the activity levels to make the system easier to analyze [2, 1, 7]. However, in the social system we are considering, we cannot make any assumptions about the existence of a third party evaluating the difference in the opinions of the two animals. A “decision,” i.e. the emission of a signal from one of the two animals, is only reached when one of the two animals’ opinions goes below a certain threshold. The social system is therefore inherently two-dimensional, whereas the neural system may be adequately described by one dimension.

We can compare the decision making process in the full two-dimensional system and in a reduced one-dimensional system to see whether decisions are made more accurately or quickly in one or the other. In the neural system, there are two thresholds: if the difference in activity levels reaches the left threshold the subject decides the dots are moving left and analogously for the right threshold. Similarly, in the social system there are two thresholds: each animal chooses a threshold below which he will emit the subordination signal. To make the systems directly comparable, we assume in both cases that the thresholds are symmetric.

We find that the two-dimensional decision is less accurate than the one-dimensional decision. That is, for a given expected time to reach a decision, the two-dimensional decision will reach the correct decision with a lower probability (Figure 1). The difference in probability is very low when memory is perfect (i.e.  $\ell = 0$ ). However, at higher leak rates the difference becomes more pronounced. Further, when the strength of the input ( $d$ ) is lower, the difference becomes greater. Thus, as the decision becomes harder because either memory decreases or the true state of the world becomes harder to detect, the advantage of the one-dimensional system increases.

Without any prior knowledge of the probabilities of the state of the visual stimulus, it is reasonable to assume that the thresholds for deciding for each direction are symmetric. On the other hand, in the social system, the thresholds are chosen by two competing animals and there is less reason to assume that their thresholds are the same. The accuracy in a given amount of time depends on the ratio of the two thresholds and is improved when the weaker animal’s threshold is lower than the stronger animal’s threshold (Figure 2).

**4.2 Strategic Decision Making** Differences in the threshold between animals could just be explained by individual variation. However, there may be a strategic aspect to this variation in the social system. In the neural system, an individual wants to make both a decision that is both quick and accurate. This is also true in the social system. However, it may also be the case that a monkey wants to avoid emitting a subordination signal and wants to receive a signal, regardless of whether or not this is the “correct” outcome, i.e. whether or not he is actually more likely to win a fight. The criterion used to optimize the threshold, therefore, may be quite different in the two systems. Additionally, the speed and outcome of the decision process depend on both animals’ thresholds. Since an animal’s best behavior depends on the behavior of the other animal, optimizing the threshold is strategic in the social system but not in the neural system.

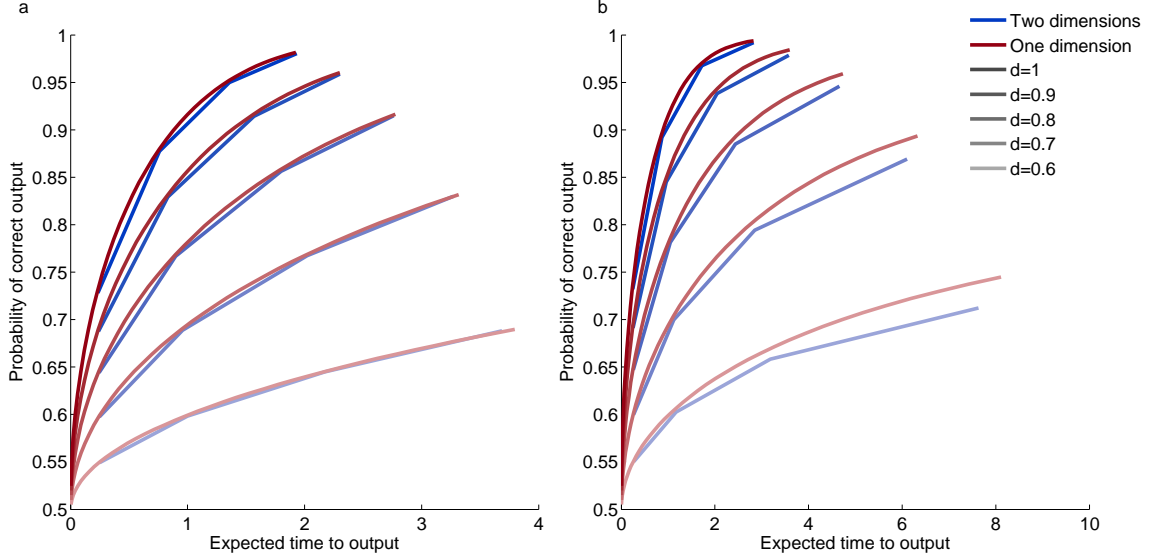


Figure 1: A two-dimensional decision process, in which the decision is made when either of the variables reaches a threshold, is less accurate than a one-dimensional decision process, in which the decision is made when the difference between the variables reaches a threshold. In each panel, the expected time to reach a decision is on the x-axis and the probability of reaching the correct decision in that amount of time is on the y-axis. Red curves represent a one-dimensional decision and red curves represent a two-dimensional decision. Stronger values of the input are represented in higher darker curves. In the left panel, there is no leak. In the right panel, leak rate is set to  $\ell = .5$ . Other parameters are set as follows:

**4.3 Social Decision Making** A monkey can try to decide whether it is more or less likely to win a fight than another monkey using information from the fights the pair has previously engaged in. Further, he can try to make such a decision about every other member of its social group. Since he is trying to use the same algorithm to learn about every other animal, he is constrained to have only one decision threshold for each of these decision processes. The optimal decision threshold in this social context depends on the distribution of fighting abilities in the group and the decision thresholds of the other animals.

When waiting for a decision to be made is more costly, so that minimizing DT is prioritized, the resulting distribution of the number of signals received is more skewed with a heavier right tail (Figure 4).

**4.4 Comparison to Social Data** Empirically, those pairs of monkeys that take the longest to establish a signaling dynamics are those with the most similar fighting abilities...

## 5 Discussion

## 6 Appendix

### 6.1 Derivation of PDEs for waiting time and accuracy

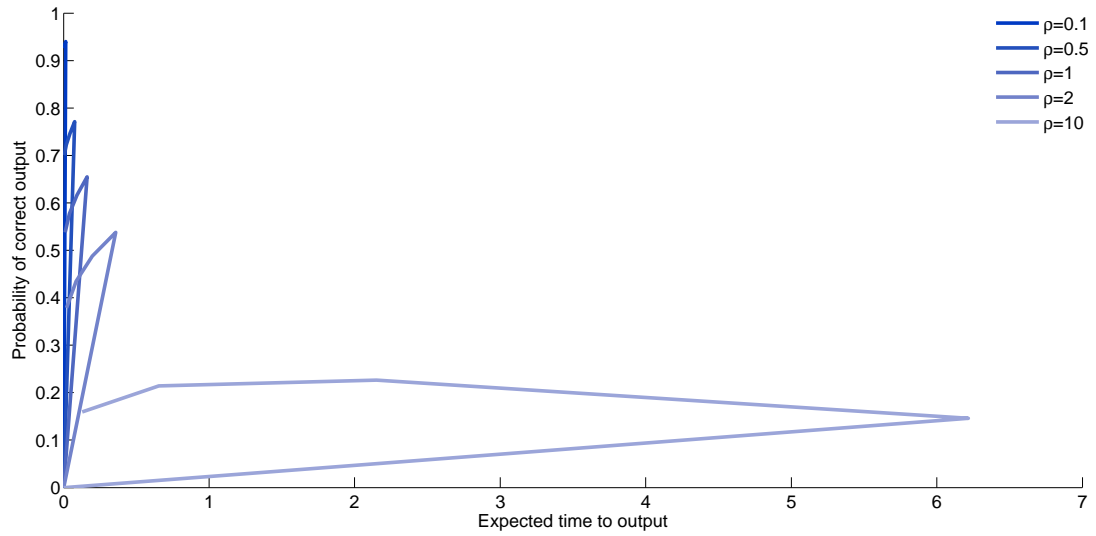


Figure 2: When the two thresholds in the two-dimensional case are allowed to vary freely, the decision process can be arbitrarily accurate in a given amount of time, depending on how much lower the weaker animal’s threshold is than the stronger animal’s threshold. The expected time to reach a decision is on the x-axis and the probability of reaching the correct decision in that amount of time is on the y-axis. The curve depends on the ratio of the weaker animal’s threshold to the stronger animal’s threshold ( $\rho$ ). Lower ratios are represented in higher darker curves.

Figure 3

Figure 4

Table 1: **Analogy between neural and social systems.**

	Neural	Social
input	dots moving left or right	fight won or lost
strength of input	coherence of dots	probability of stronger animal winning
decision variables	firing rates of neural populations	opinions about relative dominance
dist. of states of the world	left / right equally likely	distribution of fighting abilities
decision	subject indicates left or right	one animal emits a subordination signal
correct decision	subject chooses correct direction	weaker animal signals

Table 2: **Comparison of models applied to neural and social systems.** Differences are highlighted in red.

	Neural	Social
dimensionality	1	2
decision	difference hits a threshold	one var. hits a threshold
optimality criterion	reward from being “right”	reward from being “right”
	reward rate	reward rate
	Bayes risk	Bayes risk
		reward from receiving signal
optimization depends on	input strength	input strength
	noise	noise
	leak rate	leak rate
		other animal’s threshold
threshold algorithm	error correction	
	hill climbing	

## References

- [1] Rafal Bogacz, Eric Brown, Jeff Moehlis, Philip Holmes, and Jonathan D. Cohen. The physics of optimal decision making: A formal analysis of models of performamnce in two-alternative forced-choice tasks. *Psychological Review*, 113(4):700–765, 2006.
- [2] Eric Brown, Juan Gao, Philip Holmes, Rafal Bogacz, Mark Gilzenrat, and Jonathan D. Cohen. Simple neural networks that optimize decision. *International Journal of Bifurcation and Chaos*, 15(3):803–826, 2005.
- [3] Eleanor R. Brush, David C. Krakauer, and Jessica C. Flack. A family of algorithms for computing consensus about node state from network data. *PLoS Computational Biology*, 9(7):e1003109, 2013.
- [4] J. O. Caldecott. *An ecological and behavioural study of the pig-tailed macaque*. S. Karger, 1986.
- [5] Frans B. M. de Waal and Lesleigh M. Luttrell. The formal hierarchy of rhesus macaques: An investigation of the bared teeth display. *American Journal of Primatology*, 9:73–85, January 1985.
- [6] P. Eckhoff, P. Holmes, C. Law, P. M. Connolly, and J. I. Gold. On diffusion processes with variable drift rates as models for decision making during learning. *New Journal of Physics*, 10(1):015006, 2008.
- [7] Samuel Feng, Philip Holmes, Alan Rorie, and William T. Newsome. Can monkeys choose optimally when faced with noisy stimuli and unequal rewards? *PLoS Computational Biology*, 5, February 2009.
- [8] Jessica C. Flack and Frans de Waal. Context modulates signal meaning in primate communication. *Proceedings of the National Academy of Sciences*, 104(5):1581–1586, January 2007.
- [9] Jessica C. Flack and Frans B. M. de Waal. Dominance style, social power, and conflict management: a conceptual framework. In Bernard Thierry, Mewa Singh, and Werner Kaumanns, editors, *Macaque Societies: A Model for the Study of Social Organization*, chapter 8, pages 157–181. Cambridge University Press, 2004.
- [10] Jessica C. Flack, Michelle Girvan, Frans B. M. de Waal, and David C. Krakauer. Policing stabilizes construction of social niches in primates. *Nature*, 439, January 2006.

- [11] Jessica C. Flack, David C. Krakauer, and Frans B. M. de Waal. Robustness mechanisms in primate societies. *Proceedings of the Royal Society B: Biological Sciences*, 272(1568):1091–1099, June 2005.
- [12] Adrian Jacques De Froment. *Fighting for information: Decision-making, animal contests and the emergence of social hierarchy*. PhD thesis, Princeton University, 2010.
- [13] Daniel T. Gillespie. A rigorous derivation of the chemical master equation. *Physica A*, 188(404-425), 1992.
- [14] Daniel T. Gillespie. Exact numerical simulation of the ornstein-uhlenbeck process and its integral. *Physical Review E*, 54(2):2084–2091, 1996.
- [15] Daniel T. Gillespie. The chemical langevin equation. *The Journal of Chemical Physics*, 113:297, 2000.
- [16] Daniel T. Gillespie. Stochastic simulation of chemical kinetics. *Annual Review of Physical Chemistry*, 58:35–55, October 2007.
- [17] Jeff Moehlis, Eric Brown, Rafal Bogacz, Philip Holmes, and Jonathan D. Cohen. Optimizing reward rate in two alternative choice tasks: Mathematical formalism. Technical Report 04-01, Center for the Sdy of Brain, Mind, and Behavior, 2004.