

The basal ganglia optimize decision making over general perceptual hypotheses

Nathan F. Lepora¹

Kevin N. Gurney¹

¹Department of Psychology, University of Sheffield, Sheffield S10 2TP, U.K.

Keywords: basal ganglia, perception, decision making, selection, Bayesian inference, sequential analysis, SPRT, MSPRT, reinforcement learning.

Abstract

The basal ganglia are a subcortical group of interconnected nuclei involved in mediating action selection within cortex. A recent proposal is that this selection leads to optimal decision making over multiple alternatives because the basal ganglia anatomy maps onto a network implementation of an optimal statistical method for hypothesis testing, assuming that cortical activity encodes evidence for constrained Gaussian-distributed alternatives. This article demonstrates that this model of the basal ganglia extends naturally to encompass general Bayesian sequential analysis over arbitrary probability distributions, which raises the proposal to a practically-realizable theory over generic perceptual hypotheses. We also show that the evidence in this model can represent either log likelihoods, log likelihood ratios or log odds, which are all leading proposals for the cortical processing of sensory data. For these reasons, we claim that the basal ganglia optimize decision making over *general* perceptual hypotheses represented in cortex. The relation of this theory to cortical encoding, cortico-basal ganglia anatomy and reinforcement learning is discussed.

1 Introduction

Two lines of evidence are converging on an understanding of animal perception as statistically optimal sensory processing by the brain. First, perception is considered as Bayesian inference from noisy and ambiguous sensations [1, 2]. Second, the decisions resulting from this inference are considered optimal in terms of minimizing the costs of making mistakes plus the costs of waiting to gather more sensory data to improve the decision accuracy [3, 4]. For instance, a notable series of experiments considers neuronal activity in lateral intraparietal cortex as monkeys make perceptual judgements of the direction of motion for a group of random dots, and finds individual neurons that noisily ramp-up their firing rates until reaching a threshold when a decision is made [5, 6]. For two alternatives, these processes appear well described by a statistically optimal procedure known as the sequential probability ratio test (SPRT) [3, 4, 7], which accumulates evidence representing a log likelihood ratio for a competing pair of perceptual hypotheses over the time series of sensory data until reaching a preset decision threshold [8].

In a recent proposal, Bogacz and Gurney [9] showed that a key subset of the cortical and basal ganglia anatomy (Figure 1) has the appropriate connectivity to implement an optimal statistical technique for hypothesis testing over multiple perceptual alternatives. They considered a simplified version of the multi-hypothesis sequential probability ratio test (MSPRT) [10] in which the samples of sensory data were Gaussian distributed and the alternatives were such that one was supported if a sequence of sensory data had the highest sample mean and all others had equal lower means. This statistical test has been proved asymptotically optimal for fixed decision thresholds, in that as the error rates tended to zero (with increasing sample numbers) it minimized the overall cost of making mistakes and delaying the decisions [10]. In consequence, Bogacz and Gurney claimed the

basal ganglia and cortex implement optimal decision making [9].

The aim of this article is to demonstrate that the above proposal applies to general distributions of sensory data, thereby raising it to a practically realizable theory of cortico-basal ganglia function over general perceptual hypotheses. Our arguments rest on two new mathematical results. First, that the simplified version of MSPRT derived by Bogacz and Gurney has an identical algebraic form to generic Bayesian sequential analysis over evidence representing arbitrary log likelihood functions of the sensory data. Second, that this decision-making algorithm is agnostic to whether sensory evidence represents log likelihoods or log likelihood ratios. In consequence, the original relation between the basal ganglia and an optimal decision-making algorithm extends to decision making over evidence encoded by either general log likelihoods, log likelihood ratios or log odds, which includes both generic Bayesian sequential analysis and the general form of MSPRT. All of these types of evidence underly leading proposals for the cortical encoding of sensory data (*e.g.* [7, 11] for log likelihood ratios, [12, 13] for log likelihoods and [14, 15] for log odds). For these reasons, we propose that the function of the basal ganglia is to optimize decision making over *general* perceptual hypotheses represented in cortex.

2 Optimal decision making by the basal ganglia for Gaussian-distributed alternatives

2.1 A simplified MSPRT for Gaussian distributions

Bogacz and Gurney related neural computations across the cortico-basal ganglia circuit to sequential decision making over multiple alternatives [9]. They consid-

ered multiple channels of sensory data $\mathbf{s} = (s_1, \dots, s_K)$ sampled simultaneously over time, with the decision to find which channel was activated by a sensory stimulus. The evidence for each alternative was given by a log likelihood ratio between two equal-width Gaussian distributions, $N(\mu_1, \sigma)$ and $N(\mu_2, \sigma)$, for whether a channel was activated or not, which reduced to a gain g and offset c of the channel signal

$$\log \text{LR}_{12}(s_k) = \log \frac{e^{-(s_k - \mu_1)^2 / 2\sigma^2}}{e^{-(s_k - \mu_2)^2 / 2\sigma^2}} = g s_k + c, \quad k = \{1, \dots, K\}, \quad (1)$$

$$g = \frac{\mu_1 - \mu_2}{\sigma^2}, \quad c = \frac{\mu_1^2 - \mu_2^2}{2\sigma^2}, \quad (2)$$

as used originally by Gold and Shadlen for two-choice optimal decision making with SPRT [7]. For application to the basal ganglia, these parameters were assumed identical over all alternatives, so that all channels represented a choice between the same two Gaussian distributions. Then the overall decision was to infer which of the K channels of data $s_k(1), \dots, s_k(T)$ had sample mean μ_1 assuming that all other $K - 1$ channels had mean μ_2 (and the same standard deviation σ). Considering an example of populations of cortical neurons tuned to distinct perceptual features (such as directions of aggregate motion in a visual drifting dots task [16, 17]), perception then becomes the inference of which population fires with the highest mean [7]. An appropriate unit of evidence for each alternative percept is the log likelihood ratio (Eq. 1) that a cortical population is responding versus not responding to the particular perceptual feature to which it is tuned [9].

Analogously to how two-choice sequential decision making over probability ratios relates to an optimal statistical method (SPRT) [8], sequential decision making over more than two alternatives relates to another optimal method, the multi-hypothesis sequential probability ratio test (MSPRT) [10, 18, 19]. The MSPRT is based on applying Bayes' rule to multiple streams of sequentially sampled data,

with the decision of whether hypothesis H_k is supported to a given reliability determined by whether its associated posterior is the first to rise above a threshold

$$O_k(T) = -\log P(H_k | \mathbf{s}(1), \dots, \mathbf{s}(T)) < \Theta_k, \quad (3)$$

where we follow a convention in which the *minus* log posterior (which is positive) falls *below* a minus log probability threshold Θ_k (which is also positive) that is specified for each channel. (Note that Bogacz and Gurney considered a single, common threshold, but their model and optimality arguments extend straightforwardly to multiple, distinct thresholds¹.) The sign convention (3) is consistent with the physiology of the basal ganglia outputs, which influence cortex by a process of focussed disinhibition, corresponding to the channel with $O_k < \Theta_k$ at the level of the output nuclei. Using Bayes' rule and assuming independent samples distributed according to the assumptions in the preceding paragraph, Bogacz and Gurney showed that these log posteriors are given by a simple formula based on accumulating the log likelihood ratio (Eq. 1),

$$O_k(T) = -y_k(T) + \log \sum_{j=1}^K \exp y_j(T), \quad (4)$$

$$y_j(T) = \sum_{t=1}^T x_j(t), \quad x_j(t) = g s_j(t). \quad (5)$$

This simple equation has two key components. First, an evidence accumulation term y_k for each channel that implements a noisy race to threshold [20, 21]. Second, a log-sum-exponential competition term that decreases the activity of all channels when there is no clear winner [9]. Both contributions are necessary for multiple alternatives, as then the decision making is known to be asymptotically

¹In general, our notation is based on Bogacz and Gurney's, but with minor differences such as their x is our s , we use O for their OUT , and we consider a distinct threshold Θ_k for each channel.

optimal for fixed thresholds [10]. Then the cost of making errors plus the cost per sample of delaying the decision is minimized as the sample number approaches infinity, to appropriately balance reaction speed against accuracy [9].

2.2 Map between the basal ganglia and MSPRT

The functional architecture of the basal ganglia has been long understood to influence cortex by a process of focussed disinhibition [22, 23, 24, 25]. In this process, the basal ganglia output nuclei fire tonically to inhibit cortical activity, but strong focused activation of the striatum can locally reduce this tonic firing to release target thalamic areas and their corresponding cortical areas from inhibitory control. Within the basal ganglia, the ‘direct’ pathway projects from striatum to the output nuclei via one population of inhibitory medium spiny neurons, while the ‘indirect’ striatal pathway involves another, physiologically distinct, population of inhibitory medium spiny neurons and passes through an inhibitory relay in the external globus pallidus (GPe). More recently, this architecture has been revealed as an over-simplification, with significant reciprocal connectivity between the various internal nuclei, including a pathway from STN onto GPe (see [26]). Furthermore, the subthalamic nucleus is now considered a major input station of the basal ganglia that receives projections from across cortex and routes through to the basal ganglia output nuclei via the so-called ‘hyper-direct’ pathway [27], which is emphasized in this article by considering the diffuse pathways from the STN to the internal (GPi) and external globus pallidus.

The main point made by Bogacz and Gurney [9] is that their instantiation of the MSPRT in equations (3-5) matches with various aspects of cortico-basal ganglia anatomy and function (Figure 1). First, the cortex could compute the accumulated evidence representing a log likelihood ratio, as proposed by Gold

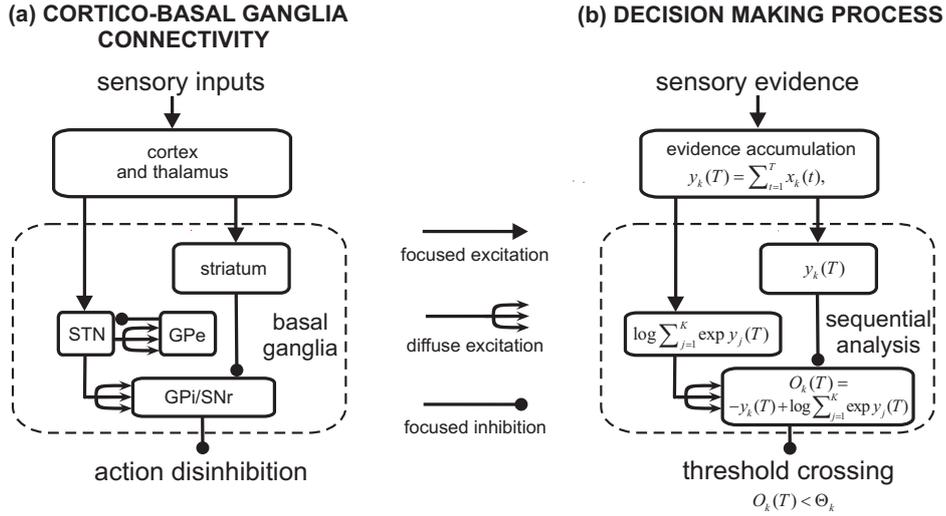


Figure 1: Map between cortico-basal ganglia anatomy and MSPRT. (a) A subset of the connections between the cortex and the basal ganglia nuclei. STN denotes the subthalamic nucleus, SNr the substantia nigra pars reticulata and GPe/GPi the globus pallidus external/internal segments (in primates; other mammals have homologous structures with different names). (b) The network version of the MSPRT decision making algorithm used by Bogacz and Gurney [9]. Actions are selected with downwards threshold crossing in the basal ganglia output nuclei.

and Shadlen [7]. Second, the log-sum-exponential competition term could be computed via the STN-GPe complex in accordance with the STN taking inputs from across cortex and projecting diffusely to the basal ganglia output nuclei (we refer to the original reference for further derivation and discussion of this correspondence [9]). Third, the striatum could then route the accumulated evidence to GPi/SNr for combination with the competition term from STN-GPe complex, giving a centralized selection mechanism similar to that proposed by Redgrave, Prescott and Gurney [25, 28, 29, 30]. Finally, when these basal ganglia outputs fall below a fixed threshold they represent disinhibition of thalamo-cortical motor area

targets responsible for initiating action, as emphasized below equation 3 above. Treating the thalamus and cortex as a single unit does oversimplify the anatomy, for example with respect to subcortical inputs [31], but allows a range of evidence accumulation mechanisms to be treated similarly, such as via thalamo-cortical loops [32] or recurrent cortical networks [33, 34]. In addition, those pathways not included in Figure 1, such as dopaminergic innervation from the substantia nigra pars compacta and ventral tegmental area, have been proposed to give added functionality, such as reinforcement learning of the appropriate action to be elicited by a stimulus [35], and will be described in the discussion.

3 Optimal decision making for arbitrary alternatives

3.1 The basal ganglia and generic Bayesian sequential analysis

A general framework for decision making over multiple alternatives is to use generic sequential analysis in which Bayes' rule is applied sequentially to input data distributed according to arbitrary probability distributions [36]. Mathematically, this approach differs from Bogacz and Gurney's in several important aspects. First, the decision making uses evidence encoded as log likelihoods rather than the log likelihood ratios in MSPRT. Second, the sensory inputs are no longer restricted to Gaussian variables, but can have arbitrary statistics. Third, non-flat priors can be included if desired. Fourth, there are no constraints between the alternatives, so that the losing alternatives are not assumed identically distributed. This latter generalization allows more general perceptual hypotheses to be represented in the model, such as partial activation of multiple channels of evidence leading to hypothesis selection, as may happen in multi-modal perception.

Sequential analysis formalizes the decision making between alternatives as a

competition between K hypotheses based upon sequentially sampling multiple sensory inputs $\mathbf{s} = (s_1, \dots, s_M)$. We emphasize that the number of sensory inputs M does not necessarily equal the number K of evidence channels associated with the hypotheses. When there are more evidence channels than inputs ($K > M$), the formalism allows a choice between an arbitrary number of perceptual hypotheses. For example, tactile sensing could be considered as encoding a time series of contact pressures, but there are multiple types of texture that could be perceived [37]. Conversely, when there are more inputs than evidence channels ($M > K$), the sensory evidence is aggregated over the inputs to make a decision. For example, in two choice (*e.g.* left vs right) motion discrimination experiments, the representation of visual motion is over ensembles of MT neurons, with each neuron having a distinct direction tuning [38].

In sequential analysis, the decision to find the winning hypothesis from its posterior probability is calculated with Bayes' rule from the likelihoods and priors

$$p(H_k | \mathbf{s}(1), \dots, \mathbf{s}(T)) = \frac{p(\mathbf{s}(1), \dots, \mathbf{s}(T) | H_k) p(H_k)}{\sum_{j=1}^K p(\mathbf{s}(1), \dots, \mathbf{s}(T) | H_j) p(H_j)}. \quad (6)$$

The condition for when to stop sampling and make a decision can be expressed as bounds on the posterior probabilities over a time sequence of T samples. Following the arguments from Section 2.1, we consider that the minus log posterior must fall below a minus log probability threshold to trigger a decision

$$O_k(T) = -\log P(H_k | \mathbf{s}(1), \dots, \mathbf{s}(T)) < \Theta_k(T). \quad (7)$$

A distinct threshold is associated with each evidence channel, so that different costs for each hypothesis can be considered. For example, the hypotheses could correspond to actions with differing potential rewards for the agent. Furthermore, these decision thresholds could also vary over time to give optimal decisions

that minimize the cost of making errors plus the cost of delays. For example, in deadline-dependent tasks, the optimal thresholds decrease over time to ensure the decision must be made by the deadline [39]. For general Bayesian sequential analysis it is unlikely that optimal decisions can be obtained with fixed thresholds, and so we employ the threshold-crossing rule with the proviso that optimality may depend on a history-dependent threshold. Some biological considerations of distinct, time-dependent thresholds are considered in the discussion.

The aim of this section is to relate the equations for generic Bayesian sequential analysis (6,7) back to Bogacz and Gurney's decision making algorithm (4,5) for the basal ganglia. To simplify the explanation, we start with samples at a single time, which will be later generalized to a sample sequence. From Bayes rule (6), the log posterior probabilities can be expressed as log likelihoods and log priors, with the denominator expanded using the identity $p_1 p_2 = \exp(\log p_1 + \log p_2)$:

$$\begin{aligned}
O_k &= -\log p(H_k | \mathbf{s}) \\
&= -\log p(\mathbf{s} | H_k) p(H_k) + \log \sum_{j=1}^K p(\mathbf{s} | H_j) p(H_j) \\
\therefore O_k &= -[\log p(H_k) + \log p(\mathbf{s} | H_k)] \\
&\quad + \log \sum_{j=1}^K \exp[\log p(H_j) + \log p(\mathbf{s} | H_j)]. \quad (8)
\end{aligned}$$

Now we replace the samples with an independent, identically distributed sequence over time, giving an accumulation of the log likelihood over these times

$$\log p(\mathbf{s}(1), \dots, \mathbf{s}(T) | H_k) = \sum_{t=1}^T \log p(\mathbf{s}(t) | H_k). \quad (9)$$

Replacing the single samples \mathbf{s} in the expression (8) with the multiple samples $\mathbf{s}(1), \dots, \mathbf{s}(T)$ and substituting the relation (9) for accumulating the log likelihood over these samples, then results in the following algorithm for general

Bayesian sequential analysis:

$$O_k(T) = -y_k(T) + \log \sum_{j=1}^K \exp y_j(T), \quad (10)$$

$$y_j(T) = \sum_{t=0}^T x_j(t), \quad x_j(t) = \begin{cases} \log p(H_j), & t = 0, \\ \log p(\mathbf{s}(t)|H_j), & t > 0. \end{cases} \quad (11)$$

The key point is that this general algorithm for Bayesian sequential analysis (10) is identical to the original Bogacz-Gurney result (4), even though the inputs y_k now represent evidence encoded as general log likelihoods (11) of the sensory data (rather than log likelihood ratios of Gaussians). Therefore, this sequential analysis model of decision making maps onto the basal ganglia architecture in an identical way to the original by Bogacz and Gurney (Figure 1).

As a consequence of the above arguments, we obtain a version of the original Bogacz-Gurney result but now applied to arbitrary probability distributions. Furthermore, the temporal accumulation for each channel $y_k(T)$ now starts from a log prior, extending the results to non-flat priors. The principal difference from the original model of Bogacz and Gurney is that the accumulated evidence sent to the striatum and STN now represents the log likelihoods of sensory data and a log prior, which we assume are together estimated by sensory cortex. Indeed, that cortex can function as a log likelihood estimator has been proposed already for optimal decoding of neural populations [12].

3.2 The basal ganglia and generic MSPRT

A key difference between many models of animal perception and general Bayesian sequential analysis is that the former uses evidence representing log likelihood ratios while the latter uses evidence representing log likelihoods. Given this significant departure in how the evidence is represented, it is surprising that the decision

making algorithms for calculating the log posterior decision variables from the accumulated evidence are identical (*cf.* equations 4 and 10). The aim of this section is to show that there is a simple correspondence between evidence encoded as log likelihoods and log likelihood ratios. As a consequence, we derive a general version of Bogacz and Gurney's original result for evidence representing arbitrary log likelihood ratios in a generic MSPRT.

Following Bogacz and Gurney, we formalize the alternatives for decision making as over K channels $\mathbf{s} = (s_1, \dots, s_K)$ of sensory data, with the decision to infer the winning channel that has samples drawn from a probability distribution, $p_{1,k}(s_k)$ while the losing channels are drawn from distributions $p_{2,j}(s_j)$ with $j \neq k$. Originally, these were the Gaussian distributions reviewed in Section 2.1. Assuming conditional independence, the likelihood that channel k is the winner (denoted as hypothesis H_k) with samples s_k drawn from distribution $p_{1,k}$ is

$$p(\mathbf{s}|H_k) = \prod_{j=1}^K p(s_j|H_k), \quad p(s_j|H_k) = \begin{cases} p_{1,k}(s_k), & j = k, \\ p_{2,j}(s_j), & j \neq k. \end{cases} \quad (12)$$

Using some straightforward algebra, the log likelihood for each hypothesis can be rewritten as a log likelihood ratio and a hypothesis-independent term,

$$\log p(\mathbf{s}|H_k) = \log \left[\frac{p_{1,k}(s_k)}{p_{2,k}(s_k)} \prod_{j=1}^K p_{2,k}(s_j) \right] = \log \text{LR}_{12,k}(s_k) + C(\mathbf{s}), \quad (13)$$

$$\log \text{LR}_{12,k}(s_k) = \log \frac{p_{1,k}(s_k)}{p_{2,k}(s_k)}, \quad C(\mathbf{s}) = \sum_{j=1}^K \log p_{2,j}(s_j). \quad (14)$$

This relation between log-likelihoods and their ratios is sufficient to ensure the identity of the expressions (4,10) in the original Bogacz-Gurney version of MSPRT and general Bayesian sequential analysis. This is because the term $C(\mathbf{s})$ is ignor-

able, due to an invariance of the expressions

$$\begin{aligned}\tilde{O}_k &= -\tilde{y}_k + \log \sum_{j=1}^K \exp \tilde{y}_j = -y_k - b + \log \sum_{j=1}^K \exp(y_j + b) \\ &= -y_k - b + \log(\exp b) + \log \sum_{j=1}^K \exp y_j = O_k,\end{aligned}\tag{15}$$

under shifting the baselines $\tilde{y}_k = y_k + b$ of all channels. Since the term $C(\mathbf{s})$ can be ignored in (13), evidence can be represented interchangeably as a log likelihood or log likelihood ratio for hypotheses pertaining to probabilities of the form (12).

The above results then combine into a general MSPRT over arbitrary log likelihood ratios that has a similar form to the original that Bogacz and Gurney mapped onto the basal ganglia. Substituting the identity (13) between log likelihoods and log likelihood ratios into the decision making algorithm (10) for log likelihoods and using that the outputs are independent of baseline (15) gives

$$O_k(T) = -y_k(T) + \log \sum_{j=1}^K \exp y_j(T),\tag{16}$$

$$y_j(T) = \sum_{t=0}^T x_j(t), \quad x_j(t) = \begin{cases} \log p(H_j), & t = 0, \\ \log \text{LR}_{12,j}(s_j(t)), & t > 0. \end{cases}\tag{17}$$

This relation is the most general that can be derived between input evidence y_k representing log likelihood ratios and outputs O_k that are log posteriors. Decision making via threshold crossing then gives a general MSPRT over arbitrary log likelihood ratios. These decisions are also asymptotically optimal for the same reasons as the original (Gaussian) Bogacz-Gurney model [10].

For an illustrative example of this general MSPRT, we consider the specific case of when the channels represent choices between Gaussian distributions with different means for each channel, rather than being equal for the losing channels

as assumed by Bogacz and Gurney. Then the winning $p_{1,k}$ and losing distributions $p_{2,k}$ for channel k are Gaussian but have channel-dependent means $\mu_{1,k}$ and $\mu_{2,k}$ respectively (for simplicity, we maintain the assumption of equal standard deviations σ), and we assume that only one channel is in a winning state. By the log likelihood relation for two Gaussians (1) considered by Gold and Shadlen [7], the appropriate input evidence for the decision making algorithm (16,17) is then

$$x_j(t) = \log \text{LR}_{12,j}(s_j(t)) = g_j s_j(t), \quad (18)$$

where $g_j = (\mu_{1,j} + \mu_{2,j})/\sigma^2$ and $t > 0$. Note that there is also a channel-dependent prior $x_j(0) = \log p(H_j)$, which when flat (equal) can be dropped from $x_j(t)$ by the identity (15). The decision making algorithm has thus reduced to that considered by Bogacz and Gurney (4,5), but with a channel-dependent gain g_j . Further restricting the winning μ_1 and losing μ_2 means as channel-independent causes this gain g to become equal for all channels and we reobtain the original model of Bogacz and Gurney.

Another special case of the general MSPRT algorithm (16,17) is with application to log odds, defined as the log likelihood ratio between the likelihood of a hypothesis and the likelihood of *not* that hypothesis. The log odds are associated with the human representation of uncertainty [15] and play a prominent role in Bayesian theories of neural computation (*e.g.* [14, 40, 41]). Simply taking $p_{2,j}(s_j) = 1 - p_{1,j}(s_j)$ in equation (12), leads to the input evidence for the decision making algorithm (16,17) becoming

$$x_j(t) = \log \text{LR}_{12,j}(s_j(t)) = \log \frac{p_{1,j}(s_j(t))}{1 - p_{1,j}(s_j(t))}, \quad (19)$$

which is the log odds for the samples s_j being drawn from distribution $p_{1,j}$. This represents an equivalent formulation of the general Bayesian sequential analysis

considered in section 3.1, but with the evidence accumulated over the log odds rather than the log likelihood.

Therefore, Bogacz and Gurney's original mapping between the cortico-basal ganglia anatomy and a network implementation of decision making by a simplified MSPRT [9] is a special case of a more general functional mapping with the generic MSPRT in equations (16,17). In particular, the original assumption that the evidence in all channels represents a common log likelihood ratio of the same two Gaussians can be relaxed to encompass distinct ratios over arbitrary probability distributions, including also the special case of the log odds. For the specific example of where these two distributions are both Gaussian, the input evidence for the decision making algorithm reduces to a gain function of the sensory data (19), as in the original Bogacz-Gurney model. Another aspect of the general case is that the evidence may also begin accumulating from initial log priors. This generality over both the probability distributions and the priors implies that the original proposal that the basal ganglia and cortex implement optimal decision making over multiple alternatives [9] applies not just to artificial sensory data represented by restricted Gaussians, but in principle to arbitrary distributions of statistical stimuli found in practically realizable situations with arbitrary priors representing a preceding history of sensory experience

4 Discussion

This article argues that the basal ganglia architecture appears configured for optimal decision making over multiple channels of sensory evidence under a wide range of assumptions about what the evidence represents. Our arguments rely on extending a previous model of action selection by Bogacz and Gurney, who

mapped the basal ganglia anatomy onto a network implementation of an optimal statistical procedure for testing multiple hypotheses restricted to constrained Gaussian variables compared via a likelihood ratio [9]. We show that their anatomical mapping also holds for this statistical test with generic probability distributions and unconstrained alternatives. In addition, we show that the same mapping holds for general Bayesian sequential analysis with evidence encoding log likelihoods or the log odds rather than log likelihood ratios. All of these sequential tests are optimal in the sense that applying a threshold-crossing decision rule can minimize the total cost of making errors and delaying decisions [8, 10].

In considering generic representations of evidence represented in cortex, the interpretations of basal ganglia function considered here should hold under practically realizable situations, such as with realistic neuronal firing, tuning curves and population coding. Moreover, this functionality applies not just to evidence from sensory cortex but also to any evidence aiding perceptual decision making, including information from the motor and limbic cortices about the state of the animal. Our brains make perceptual judgements over a broad variety of stimuli across a range of task demands, and therefore it seems reasonable that a generic model is necessary to capture the computational processes underlying decision making.

In this discussion section, we describe the relation of these findings to other models of basal ganglia function and discuss some aspects of their biological interpretation and implications.

4.1 Relation to other models of the basal ganglia and cortex

Models of the basal ganglia performing action selection conventionally assume there are multiple parallel channels representing evidence for distinct alterna-

tives [9, 29, 30, 35, 42, 43, 44, 46]. The evidence is processed by the cortico-basal ganglia network to reach a decision about the appropriate action, which is enacted by focused disinhibition of the basal ganglia output nuclei onto the appropriate motor regions of thalamus and cortex [22, 23, 24, 25]. Some of these models place emphasis on two particular pathways from cortex through the basal ganglia to its output nuclei [9, 29, 30]: the direct striatal pathway, where activity in each channel gives a local ‘off-center’ inhibition of the tonically-firing output nuclei; and the STN hyper-direct pathway [24, 27], where conflict between channels gives a global ‘on-surround’ excitation of the output nuclei. The convergence of these two pathways at the basal ganglia output nuclei gives a competitive ‘off-center/on-surround’ mechanism for selection. This mechanism fits parsimoniously with the model of probabilistic selection considered in this article, because then the STN-GPe complex normalizes the probabilities represented at the basal ganglia output nuclei to sum to unity [9, 35]. Thus, hard selection with a single winner-takes-all is obtained for probability thresholds greater than one-half, and soft selection with multiple possible winners is obtained for thresholds less than one-half.

The basal ganglia model in the present article follows an approach of modeling neural decision making that emphasizes optimal sequential tests for deciding between competing perceptual hypotheses with sensory evidence [3, 4, 7, 11, 18, 47, 48]. An underlying assumption in these models is that the samples of data are independent and identically distributed (IID) over time. Although this assumption limits the methods for inference, it also simplifies the decision process to a straightforward evidence accumulation over time. Similar models of evidence accumulation to threshold have been used widely in the neuroscience and psychology of perceptual decision making, for example the race model [49] and diffusion model [50, 51] of two-choice discrimination tasks. However, if neural decision

making does utilize time-dependencies in the sensory data, then these IID model assumptions should be reassessed. For example, one generalization would be to pre-filter the sensory inputs to identify relevant statistical dependencies before applying the optimal sequential test to the filtered data.

Because the present article extends the optimal decision making model of the basal ganglia by Bogacz and Gurney [9] to general representations of sensory evidence, the formalism inherits many of the advantages of their model while having the potential to address further questions (also see later sections). For instance, these decision making models relate to a previous neural network account of basal ganglia function that exhibited appropriate action selection and switching properties [29, 30], and was later extended to include physiological details of thalamo-cortical loops [32], populations of spiking neurons [44] and the role of tonic dopamine [53]. In addition, the diffuse projection of the STN onto the GPi basal ganglia output nucleus relates to proposals that it dynamically modulates the propensity to elicit actions by the degree of conflict [24, 27, 43, 54]. In the present model of optimal decision making, the evidence in every channel is appropriately decreased to delay reaching the decision thresholds so that sufficient evidence can be collected to disambiguate conflicting perceptual hypotheses. Only for the particular log-sum-exponential functional form derived for the STN-GPe network can the decisions be optimal [9], in that the evidence decrement due to ambiguity slows the decision sufficiently to minimize the overall cost of errors and delays.

4.2 Implications for neural encoding

A key aspect of the optimal decision making algorithm considered here is that it is agnostic to whether the input evidence is represented as a log likelihood, log likelihood ratio or log odds, providing the compared inputs are in the same repre-

sentation. All of these types of evidence underly leading proposals for the cortical encoding of sensory data (*e.g.* [7, 11] for log likelihood ratios, [12, 13] for log likelihoods and [14, 15] for log odds). Thus, if it transpires that cortex uses all three methods to encode evidence, we have shown that the basal ganglia can lead to optimal decisions using all three representations. Fundamentally, if the basal ganglia and cortex do make optimal decisions with a procedure based on sequential analysis, then it may be somewhat meaningless to draw a distinction between these probabilistic representations anyway. By the identity (13), a log likelihood ratio (for a channel winning versus it losing) is equivalent to a log likelihood (for a channel winning and all others losing) plus a hypothesis independent term that can be ignored due to an invariance of the decision making algorithms (see below). Similarly, by equation (19) the log odds can be considered a special case of the log likelihood ratio, which by the above identity is equivalent to a log likelihood. Hence, these probabilistic representation can be used interchangeably, and it is just a matter of convention which one is adopted as an interpretation.

If cortex represents evidence as log likelihoods, log likelihood ratios or log odds, then an apparent issue is that these quantities can be negative, whereas neural activity is a positive quantity. However, as observed originally by Bogacz and Gurney [9], the specific form of the decision making algorithm solves this problem: provided all input evidence is shifted up or down (renormalized) by the same amount, then by equation (15) the output decision variables remain unchanged. As a consequence, the input neural activity can be considered as a shifted version of the evidence, such that this renormalized quantity always remains positive. This mechanism also gives added flexibility in how signals are represented in cortex, because the shift in neural activity can change over time. Thus if the neural activity in some channels is becoming small then all channels can be shifted to

renormalize the activity within an appropriate dynamic range.

4.3 Other anatomical mappings

In this article, we concentrated on mapping sequential analysis onto the feed-forward architecture for the basal ganglia considered by Bogacz and Gurney [9]. Sensory inputs are integrated in the cortex to enter the basal ganglia via the striatum and STN, and leave via the basal ganglia output nuclei for action disinhibition (Figure 1). This architecture is here intended as a basic sketch of the actual connectivity to give a simple foundation for constructing anatomically more realistic models. To this end, we now describe two alternative anatomical mappings that are consistent with the arguments presented in this article.

A recurrent basal ganglia architecture has been considered by Bogacz and Larsen [35] based on the cortico-basal ganglia-thalamo-cortical loops found in anatomical tracing studies [55]. Rather than integrating the evidence within cortex, as considered originally [7, 9], they employed the overall loop from the cortex through the basal ganglia back to cortex to implement an iterative update of the basal ganglia outputs up to a decision threshold [35, Figure 3]. They then showed that this recurrent circuit can implement optimal decision making (for Gaussian inputs) and extended their model to include reinforcement learning. Note that their decision making model uses same Gaussian-restricted MSPRT as the feed-forward model by Bogacz and Gurney, for which the general arguments presented here apply. Hence, general Bayesian sequential analysis also maps consistently onto their recurrent basal ganglia architecture, with similar considerations about the types of evidence representation and role of the STN-GPe complex.

A second alternative architecture proposed here is a feed-forward network (Figure 2) that differs from the original Bogacz-Gurney model in the following

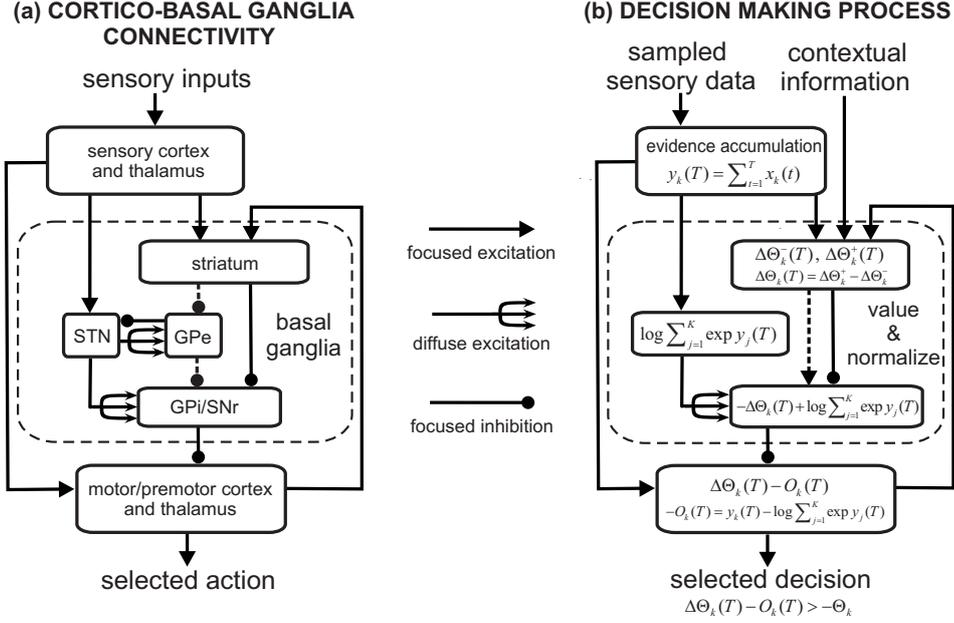


Figure 2: Proposed feed-forward map between cortico-basal ganglia anatomy and sequential analysis. (a) A subset of the connections between the cortex and the basal ganglia nuclei. Labels are as in Figure 1. (b) The network version of Bayesian sequential analysis considered in this article, configured to map onto the cortico-basal ganglia anatomy. The striato-pallidal indirect pathway (dashed lines) is interpreted as conveying negative changes to decision thresholds $-\Delta\Theta_k^-$, which complement positive changes $\Delta\Theta_k^+$ along the direct striato-nigral pathway. Decisions are selected with upwards threshold-crossing in cortex, with the baseline and thresholds shifted consistently to ensure positive neuronal activity.

way. Rather than the evidence being sent via the striatum to combine with the STN-GPe competition term in the basal ganglia output nuclei, the evidence is routed to another area of cortex where this combination takes place. This frees the striatum up for another function, which we interpret as selectively biasing the decisions by shifting the thresholds in response to contextual information about the

task demands. Our motivation for proposing this architecture is that the original model by Bogacz and Gurney lacks a structure for implementing decision thresholds, which are the primary task-dependent parameters for setting system behavior. A striatal pathway that sets decision thresholds would extend the scope of the basal ganglia model to take advantage of the full power of the sequential analysis formalism. A second benefit is that the accumulation mechanisms are now explicitly cortical, similarly to evidence accumulation models of intra-cortical decision making [3, 7]; meanwhile, the biasing of the decision thresholds is striatal, as is consistent with cortico-striatal connections controlling speed and accuracy in perceptual decision making [54, 56] and reward-related signals modulating decision processes by their effect on the striatum [57, 58].

The connectivity of the basal ganglia and cortex in Figure 2 is consistent with a range of anatomical studies. For instance, the intrinsic basal ganglia connectivity is similar to that used throughout this article and is founded on leading studies of the basal ganglia (see section 2.2). Meanwhile, excitatory forward projections from sensory to motor cortex are a standard feature of intra-cortical connectivity [59], and inhibitory projections from the basal ganglia output nuclei to motor thalamus and then cortex are also standard [22, 23]. In addition, the striatum is known to receive excitatory input from across cortex [60], allowing it to process a range of sensory, motor and contextual information.

Overall, Figure 2 models the basal ganglia as functioning together with cortex to ensure that decisions are optimal with respect to the costs for the performed task as represented by the decision thresholds. These thresholds have both cortical and striatal components, with the striatal part recalled from contextual sensory and motor information about the state of the animal and its environment (see section 4.5). Note that the role of the STN-GPe complex is key to ensuring consistent

action selection with distributed decision making in cortex [25, 28]. Continuing the reasoning from section 4.1, the pathway from cortex through STN-GPe to GPi and back to thalamus and cortex ensures that cortical representations of probability are self-consistently normalized to have unit sum. In consequence, this pathway can enforce hard ‘winner-take-all’ selection even when the decision making is distributed across functionally separated regions of cortex.

The basal ganglia model in Figure 2 also contains the direct and indirect striatal pathways [22]. In the present model, the indirect pathway via the GPe is considered distinct from the STN-GPe network. Although there is debate about the anatomy of the globus pallidus, there is scant evidence for this separated connectivity, with instead a more integrative role for the nucleus considered likely. In this work, we follow previous interpretations of striatal modulation via the direct and indirect pathways as relating to a focussed Go and NoGo mechanism for biasing the selection of individual actions [42]. However, we do not rule out more elaborate mechanisms for setting decision thresholds, such as a striatal-GPe-STN global NoGO pathway for striatal suppression of all actions.

4.4 Integration with reinforcement learning

Decision making and reinforcement learning can be viewed as complementary processes underlying action selection [35]. Reinforcement learning theories describe how to learn which action to select for a given stimulus to maximize reward [61, 62], whereas decision making theories describe how to select the action corresponding to the stimulus best supported by the incoming sensory evidence [25, 29]. We now discuss some implications of integrating reinforcement learning and decision making theories of the basal ganglia.

The recurrent model of optimal decision making (for Gaussian inputs) consid-

ered by Bogacz and Larsen has been integrated with reinforcement learning [35]. To summarize, their model utilizes an actor-critic formalism [63] mapped on the matrix/patch physiology of the striatum [64]. The actor implements an iterative update of the basal ganglia outputs via recurrent cortico-basal ganglia-thalamo-cortical loops, while the critic reinforces cortico-striatal weights to connect a sensory channel with a decision to act. Then the decision making is optimal when all weights reach zero or one. For more details we refer to their article, noting that their recurrent architecture is compatible with the general formalism presented here on optimal decision making (see previous section).

The feed-forward architecture proposed in Figure 2 can also be integrated with reinforcement learning, but with a new interpretation of the cortico-striatal weights as connecting the task context with the decision thresholds. Reinforcement learning of these weights then biases the selection of one action over another, while the decision making remains optimal throughout learning. In sequential analysis, these decision thresholds are the primary parameters for setting system behavior, and thus it is appropriate that they be learnt by reinforcement from the agent's environment. The direct striatal pathway is then interpreted as transmitting negative changes to the individual decision thresholds for the perceptual hypotheses, which learns biases towards appropriate actions; meanwhile, the indirect striatal pathway is interpreted as transmitting only positive changes, which learns biases against inappropriate actions.

This model of biasing the decision thresholds has much in common with a Go/NoGo interpretation of the direct/indirect pathways and its relation to reward [42, 43, 45, 46]. The Go population of striatal neurons facilitates the selection of a particular response by inhibiting the associated channels in the basal ganglia output nuclei, thereby releasing the inhibition on the thalamus; meanwhile,

the NoGo striatal population suppresses the selection of particular responses by a double inhibitory pathway via the GPe onto the basal ganglia output nuclei, thereby preventing the release of inhibition on the thalamus. Hence, we expect that reinforcement learning in the present basal ganglia model of optimal decision making proceeds in a qualitatively similar way to existing studies of phasic dopamine in cortico-striatal plasticity [42, 43, 65], although it remains an open question how these mechanisms should be applied to achieve decision optimality under general task demands.

4.5 Decision thresholds and optimality

The decision thresholds are the primary task-dependent parameters for setting the system behavior, in terms of specifying the balance between decision speed and accuracy in sequential analysis models of decision making [3, 4, 37]. Setting low thresholds gives quick but inaccurate decisions, for example in urgent situations such as evading predators; conversely, high thresholds give slow decisions with fewer mistakes, for example in dangerous situations such as climbing.

Optimal decision making models minimize the overall cost of the outcome, such as the total cost of delaying decisions and making mistakes. However, there are some caveats to this optimality. For two alternatives, both SPRT and Bayesian sequential analysis (which are then equivalent) give optimal decisions for constant thresholds and no deadline [8], while with a deadline the optimal decision threshold decreases over time to ensure a decision is made sufficiently early [36, 39]. For multiple alternatives, MSPRT with fixed thresholds gives optimal decisions as the error rate tends to zero at infinite decision times [10]. However, humans and animals make decisions from limited durations of sensory data and are error prone, for which it is unknown how close MSPRT with fixed thresholds is to op-

tinality (and what would then be the form of the thresholds that give optimality). Also, to the best of our knowledge, there is no analogous result for the thresholds leading to optimal Bayesian sequential analysis. If animals do make near-optimal decisions over general perceptual hypotheses, this raises the question of how such history-dependent optimal thresholds could be implemented in the brain.

One possibility is that history-dependent thresholds occur via feedback mechanisms in the underlying biology. For example, reafferent cortico-basal ganglia-thalamo-cortical loops could naturally give feedback processes that lead to varying thresholds, while the sensory evidence itself could give a time-varying context for setting the threshold throughout the decision process. It is an open theoretical question how the thresholds should vary to achieve optimal decision making and we conjecture that the diverse inputs to striatum convey the necessary sensory, motor and other contextual information (shown by the three cortico-striatal connections on Figure 2). That the striatum utilizes many distinct sources of information is consistent with recordings from the caudate nucleus of multiple signals for influencing and assessing perceptual decisions [66] and the existence of diverse striatal inputs from across cortex [67]. Understanding how these sources of information influence perceptual decision making and their relation to achieving decision optimality could shed light on the normal and pathological function of the basal ganglia.

Acknowledgments

We thank the anonymous referees and are grateful to Tony Prescott, Javier Caballero, Mark Humphries and other members of the Adaptive Behavior Research Group (ABRG) and Active Touch Laboratory at Sheffield (ATLAS) for discussions. This work was supported by FP7 grants BIOTACT (ICT-215910), IM-

References

- [1] D.C. Knill and A. Pouget. The Bayesian brain: the role of uncertainty in neural coding and computation. *TRENDS in Neurosciences*, 27(12):712–719, 2004.
- [2] D. Kersten, P. Mamassian, and A. Yuille. Object perception as Bayesian inference. *Annu. Rev. Psychol.*, 55:271–304, 2004.
- [3] J.I. Gold and M.N. Shadlen. The neural basis of decision making. *Annu. Rev. Neurosci.*, 30:535–574, 2007.
- [4] R. Bogacz, E. Brown, J. Moehlis, P. Holmes, and J.D. Cohen. The physics of optimal decision making: A formal analysis of models of performance in two-alternative forced-choice tasks. *Psychological Review*, 113(4):700, 2006.
- [5] A.C. Huk and M.N. Shadlen. Neural activity in macaque parietal cortex reflects temporal integration of visual motion signals during perceptual decision making. *Journal of Neuroscience*, 25(45):10420, 2005.
- [6] M.L. Platt and P.W. Glimcher. Neural correlates of decision variables in parietal cortex. *Nature*, 400(6741):233–238, 1999.
- [7] J.I. Gold and M.N. Shadlen. Neural computations that underlie decisions about sensory stimuli. *Trends in Cognitive Sciences*, 5(1):10–16, 2001.
- [8] A. Wald and J. Wolfowitz. Optimum character of the sequential probability ratio test. *The Annals of Mathematical Statistics*, 19(3):326–339, 1948.

- [9] R. Bogacz and K. Gurney. The basal ganglia and cortex implement optimal decision making between alternative actions. *Neural computation*, 19(2):442–477, 2007.
- [10] V.P. Dragalin, A.G. Tartakovsky, and V.V. Veeravalli. Multihypothesis sequential probability ratio tests. I. asymptotic optimality. *Information Theory, IEEE Transactions on*, 45(7):2448–2461, 1999.
- [11] T. Yang and M.N. Shadlen. Probabilistic reasoning by neurons. *Nature*, 447(7148):1075–1080, 2007.
- [12] M. Jazayeri and J.A. Movshon. Optimal representation of sensory information by neural populations. *Nature Neuroscience*, 9(5):690–696, 2006.
- [13] A.B.A. Graf, A. Kohn, M. Jazayeri, and J.A. Movshon. Decoding the activity of neuronal populations in macaque primary visual cortex. *Nature neuroscience*, 2011.
- [14] S. Deneve. Bayesian spiking neurons I: inference. *Neural computation*, 20(1):91–117, 2008.
- [15] H. Zhang and L.T. Maloney. Ubiquitous log odds: a common representation of probability and frequency distortion in perception, action, and cognition. *Frontiers in Neuroscience*, 6, 2012.
- [16] K.H. Britten, M.N. Shadlen, W.T. Newsome, and J.A. Movshon. Responses of neurons in macaque MT to stochastic motion signals. *Visual Neuroscience*, 10:1157–1157, 1993.

- [17] J.N. Kim, M.N. Shadlen, et al. Neural correlates of a decision in the dorso-lateral prefrontal cortex of the macaque. *Nature neuroscience*, 2:176–185, 1999.
- [18] T. McMillen and P. Holmes. The dynamics of choice among multiple alternatives. *Journal of Mathematical Psychology*, 50(1):30–57, 2006.
- [19] C.W. Baum and V.V. Veeravalli. A sequential procedure for multihypothesis testing. *Information Theory, IEEE Transactions on*, 40(6), 1994.
- [20] M. Stone. Models for choice reaction time. *Psychometrika*, (25):251–260, 1960.
- [21] D. Vickers. Evidence for an accumulator model of psychophysical discrimination. *Ergonomics*, (13):37–58, 1970.
- [22] R.L. Albin, A.B. Young, and J.B. Penney. The functional anatomy of basal ganglia disorders. *Trends in neurosciences*, 12(10):366–375, 1989.
- [23] G. Chevalier and J.M. Deniau. Disinhibition as a basic process in the expression of striatal functions. *Trends in Neurosciences*, 13(7):277–280, 1990.
- [24] J.W. Mink. The basal ganglia: focused selection and inhibition of competing motor programs. *Progress in neurobiology*, 50(4):381–425, 1996.
- [25] P. Redgrave, T.J. Prescott, and K. Gurney. The basal ganglia: a vertebrate solution to the selection problem. *Neuroscience*, 89:1009–1023, 1999.
- [26] P. Redgrave, M. Rodriguez, Y. Smith, M.C. Rodriguez-Oroz, S. Lehericy, H. Bergman, Y. Agid, M.R. DeLong, and J.A. Obeso. Goal-directed and habitual control in the basal ganglia: implications for parkinson’s disease. *Nature Reviews Neuroscience*, 11(11):760–772, 2010.

- [27] A. Nambu, H. Tokuno, and M. Takada. Functional significance of the cortico-subthalamo-pallidal ‘hyperdirect’ pathway. *Neuroscience Research*, 43(2):111–117, 2002.
- [28] T.J. Prescott, P. Redgrave, and K. Gurney. Layered control architectures in robots and vertebrates. *Adaptive Behavior*, 7(1):99, 1999.
- [29] K. Gurney, T.J. Prescott, and P. Redgrave. A computational model of action selection in the basal ganglia. I. A new functional anatomy. *Biological cybernetics*, 84(6):401–410, 2001.
- [30] K. Gurney, T.J. Prescott, and P. Redgrave. A computational model of action selection in the basal ganglia. II. Analysis and simulation of behaviour. *Biological cybernetics*, 84(6):411–423, 2001.
- [31] J.G. McHaffie, T.R. Stanford, B.E. Stein, V. Coizet, and P. Redgrave. Subcortical loops through the basal ganglia. *TRENDS in Neurosciences*, 28(8):401–407, 2005.
- [32] M.D. Humphries and K.N. Gurney. The role of intra-thalamic and thalamocortical circuits in action selection. *Network: Computation in Neural Systems*, 13(1):131–156, 2002.
- [33] C.C. Lo and X.J. Wang. Cortico–basal ganglia circuit mechanism for a decision threshold in reaction time tasks. *Nature neuroscience*, 9(7):956–963, 2006.
- [34] X.J. Wang. Decision making in recurrent neuronal circuits. *Neuron*, 60(2):215–234, 2008.

- [35] R. Bogacz and T. Larsen. Integration of reinforcement learning and optimal decision-making theories of the basal ganglia. *Neural computation*, pages 1–35, 2011.
- [36] D. Siegmund. *Sequential analysis: tests and confidence intervals*. Springer, 1985.
- [37] N.F. Lepora, C.W. Fox, M.H. Evans, M.E. Diamond, K. Gurney, and T.J. Prescott. Optimal decision-making in mammals: insights from a robot study of rodent texture discrimination. *Journal of The Royal Society Interface*, published online 25 Jan, 2012.
- [38] M.E. Mazurek, J.D. Roitman, J. Ditterich, and M.N. Shadlen. A role for neural integrators in perceptual decision making. *Cerebral cortex*, 13(11):1257–1269, 2003.
- [39] P. Frazier and A.J. Yu. Sequential hypothesis testing under stochastic deadlines. *Advances in neural information processing systems*, 20:465–472, 2008.
- [40] A. Soltani and X.J. Wang. Synaptic computation underlying probabilistic inference, *Nature neuroscience*, 13(1):112–119, 2009.
- [41] L. Buesing, J. Bill, B. Nessler, and W. Maass. Neural dynamics as sampling: A model for stochastic computation in recurrent networks of spiking neurons. *PLoS Computational Biology*, 7(11):e1002211, 2011.
- [42] M.J. Frank. Dynamic dopamine modulation in the basal ganglia: a neuro-computational account of cognitive deficits in medicated and nonmedicated parkinsonism. *Journal of cognitive neuroscience*, 17(1):51–72, 2005.

- [43] M.J. Frank. Hold your horses: a dynamic computational role for the subthalamic nucleus in decision making. *Neural Networks*, 19(8):1120–1136, 2006.
- [44] M.D. Humphries, R.D. Stewart, and K.N. Gurney. A physiologically plausible model of action selection and oscillatory activity in the basal ganglia. *The Journal of neuroscience*, 26(50):12921, 2006.
- [45] P. Simen, J.D. Cohen, and P. Holmes. Rapid decision threshold modulation by reward rate in a neural network. *Neural Networks*, 19(8):1013–1026, 2006.
- [46] R. Ratcliff and M.J. Frank. Reinforcement-based decision making in corticostriatal circuits: mutual constraints by neurocomputational and diffusion models. *Neural Computation*, pages 1–44, 2012.
- [47] A.K. Churchland, R. Kiani, and M.N. Shadlen. Decision-making with multiple alternatives. *Nature neuroscience*, 11(6):693–702, 2008.
- [48] J.M. Beck, W.J. Ma, R. Kiani, T. Hanks, A.K. Churchland, J. Roitman, M.N. Shadlen, P.E. Latham, and A. Pouget. Probabilistic population codes for bayesian decision making. *Neuron*, 60(6):1142–1152, 2008.
- [49] P.L. Smith and D. Vickers. The accumulator model of two-choice discrimination. *Journal of Mathematical Psychology*, 32(2):135–168, 1988.
- [50] P.L. Smith and R. Ratcliff. Psychology and neurobiology of simple decisions. *Trends in Neurosciences*, 27(3):161–168, 2004.
- [51] R. Ratcliff. A theory of memory retrieval. *Psychological review*, 85(2):59, 1978.

- [52] R. Ratcliff and G. McKoon. The diffusion decision model: Theory and data for two-choice decision tasks. *Neural computation*, 20(4):873–922, 2008.
- [53] M.D. Humphries, M. Khamassi, and K. Gurney. Dopaminergic control of the exploration-exploitation trade-off via the basal ganglia. *Frontiers in neuroscience*, 6, 2012.
- [54] B.U. Forstmann, A. Anwander, A. Schäfer, J. Neumann, S. Brown, E.J. Wagenmakers, R. Bogacz, and R. Turner. Cortico-striatal connections predict control over speed and accuracy in perceptual decision making. *Proceedings of the National Academy of Sciences*, 107(36):15916–15920, 2010.
- [55] A. Parent and L.N. Hazrati. Functional anatomy of the basal ganglia. I. the cortico-basal ganglia-thalamo-cortical loop. *Brain Research Reviews*, 20(1):91–127, 1995.
- [56] B.U. Forstmann, G. Dutilh, S. Brown, J. Neumann, D.Y. Von Cramon, K.R. Ridderinkhof and E.J. Wagenmakers, Striatum and pre-SMA facilitate decision-making under time pressure, *Proceedings of the National Academy of Sciences*, 105(45):17538–17542, 2008.
- [57] J.N.J. Reynolds, B.I. Hyland, and J.R. Wickens. A cellular mechanism of reward-related learning. *Nature*, 413(6851):67–70, 2001.
- [58] W. Schultz. Getting formal with dopamine and reward. *Neuron*, 36(2):241–263, 2002.
- [59] E.G. Jones, J.D. Coulter, and S.H.C. Hendry. Intracortical connectivity of architectonic fields in the somatic sensory, motor and parietal cortex of monkeys. *The Journal of comparative neurology*, 181(2):291–347, 1978.

- [60] C.R. Gerfen and C.J. Wilson. Chapter II: The basal ganglia. *Handbook of chemical neuroanatomy*, 12:371–468, 1996.
- [61] P.R. Montague, P. Dayan, and T.J. Sejnowski. A framework for mesencephalic dopamine systems based on predictive hebbian learning. *The Journal of Neuroscience*, 16(5):1936–1947, 1996.
- [62] W. Schultz, P. Dayan, and P.R. Montague. A neural substrate of prediction and reward. *Science*, 275(5306):1593–1599, 1997.
- [63] R.S. Sutton and A.G. Barto. *Reinforcement learning: An introduction*, volume 1. Cambridge Univ Press, 1998.
- [64] K. Doya. Complementary roles of basal ganglia and cerebellum in learning and motor control. *Current Opinion in Neurobiology*, 10(6):732–739, 2000.
- [65] M.J. Frank, B.B. Doll, J. Oas-Terpstra, and F. Moreno. Prefrontal and striatal dopaminergic genes predict individual differences in exploration and exploitation. *Nature neuroscience*, 12(8):1062–1068, 2009.
- [66] L. Ding and J.I. Gold. Caudate encodes multiple computations for perceptual decisions. *The Journal of Neuroscience*, 30(47):15747–15759, 2010.
- [67] E.G. Jones, J.D. Coulter, H. Burton, and R. Porter. Cells of origin and terminal distribution of corticostriatal fibers arising in the sensory-motor cortex of monkeys. *The Journal of comparative neurology*, 173(1):53–80, 1977.