

Stochastic Models of Evidence Accumulation in Changing Environments*

Alan Veliz-Cuba[†]
Zachary P. Kilpatrick[‡]
Krešimir Josić[§]

Abstract. Organisms and ecological groups accumulate evidence to make decisions. Classic experiments and theoretical studies have explored this process when the correct choice is fixed during each trial. However, we live in a constantly changing world. What effect does such impermanence have on classical results about decision making? To address this question we use sequential analysis to derive a tractable model of evidence accumulation when the correct option changes in time. Our analysis shows that ideal observers discount prior evidence at a rate determined by the volatility of the environment, and the dynamics of evidence accumulation is governed by the information gained over an average environmental epoch. A plausible neural implementation of an optimal observer in a changing environment shows that, in contrast to previous models, neural populations representing alternate choices are coupled through excitation. Our work builds a bridge between statistical decision making in volatile environments and stochastic nonlinear dynamics.

Key words. mathematical neuroscience, decision making, dynamic environment, Bayesian inference, recursive Bayesian estimation, sequential probability ratio test, drift-diffusion model

AMS subject classifications. 62C10, 60H10, 62L10, 91E30, 92C20

DOI. 10.1137/15M1028443

I. Introduction. To navigate a constantly changing world, we intuitively use the most recent and pertinent information. For instance, when planning a route between home and work we use recent reports of accidents and weather. We discount older information, as our environment is in constant flux: The clouds threatening rain last night may have dissipated, and an accident reported an hour ago has likely been cleared. The optimal strategy is therefore to weight recent evidence more strongly.

How to make decisions in the face of uncertainty and impermanence is a question that recurs in fields ranging from economics to ecology and neuroscience. Mammals [8, 11, 12, 24], insects [13, 38], single cells [3], and animal collectives [28] gather

*Received by the editors June 30, 2015; accepted for publication (in revised form) October 5, 2015; published electronically May 5, 2016. This work was supported by grants NSF-DMS-1311755 (ZPK), NSF-DMS-1517629 (KJ and ZPK), NSF/NIGMS-R01GM104974 (AV-C and KJ), and NSF-DMS-1122094 (KJ). The second and third authors contributed equally to this work.

<http://www.siam.org/journals/sirev/58-2/M102844.html>

[†]Department of Mathematics, University of Dayton, Dayton, OH 45469 (avelizcuba1@udayton.edu).

[‡]Department of Mathematics, University of Houston, Houston, TX 77204 (zpkilpat@math.uh.edu).

[§]Department of Mathematics, University of Houston, Houston, TX 77204, and Department of Biology and Biochemistry, University of Houston, Houston, TX 77204 (josic@math.uh.edu).

evidence to make decisions. However, information about the state of the world is typically incomplete, and perception is noisy. Therefore, animals make choices based on uncertain evidence. The case of an observer deciding between two alternatives based on a series of noisy measurements has been studied extensively when the environment is static [8, 23, 32, 45]. In this case, humans [36] and other mammals [11, 24] can accumulate incoming evidence near optimally to reach a decision.

Stochastic accumulator models provide a plausible neural implementation of decision making between two or more alternatives [4, 43]. These models are analytically tractable [8], and can implement optimal decision strategies [10]. Remarkably, there is also a parallel between these models and experimentally observed neural activity. Recordings in animals during a decision task suggest that neural activity reflects the weight of evidence for one of the choices [24].

A key assumption in many models is that the correct choice is fixed in time; i.e., decisions are made in a static environment. This assumption may hold in the laboratory, but natural environments are seldom static [17, 34]. Recent experimental evidence suggests that human observers integrate noisy measurements near optimally even when the state of the environment changes. For instance, when observers need to decide between two options and the corresponding reward changes in a history-dependent manner, human behavior approximates that of a Bayes optimal observer [5]. An important feature of evidence accumulation in volatile environments is an increase in learning rate when recent observations do not support a current estimate [31]. Both behavioral and fMRI data show that human subjects employ this strategy when they must predict the position of a stochastically moving target [29]. Experimental work thus suggests that humans adjust evidence valuation to account for environmental variability.

However, the dynamics of decision making in changing environments has not been fully investigated. To address this question we extend optimal stochastic accumulator models to a changing environment. These extensions are amenable to analysis and reveal that an optimal observer discounts old information at a rate adapted to the frequency of environmental changes. As a result, the certainty that can be attained about any of the choices is limited. Our approach frames the decision-making process in terms of a first passage problem for a doubly stochastic nonlinear model that can be examined using techniques of nonlinear dynamics. Extending previous work, we also identify accurate piecewise linear approximations to the nonlinear model. This model also suggests a biophysical neural implementation for evidence integrators consisting of neural populations whose activity represents the evidence in favor of a particular choice. When the environment is not static, optimal evidence discounting can be performed exactly by populations coupled through excitation. We also show that the computation can be well approximated by appropriately tuned classical linear population models [10, 30, 41, 44].

2. Optimal Decisions in a Static Environment. We develop our model in a way that parallels the case of a static environment with two possible states. We therefore start with the derivation of the recursive equation for the log-likelihood ratio of the two states, and the approximating stochastic differential equation (SDE), when the underlying state is fixed in time.

To make a decision, an optimal observer integrates a stream of measurements to infer the present environmental state. In the static case, this can be done using sequential analysis [12, 45]: An observer makes a stream of independent, noisy measurements, $\xi_{1:n} = (\xi_1, \xi_2, \dots, \xi_n)$, at equally spaced times, $t_{1:n} = (t_1, t_2, \dots, t_n)$. The

probability of each measurement, $f_+(\xi_n) := \Pr(\xi_n|H_+)$ and $f_-(\xi_n) := \Pr(\xi_n|H_-)$, depends on the environmental state. Combined with the prior probability, $\Pr(H_\pm)$, of the states, this gives the ratio of probabilities,

$$R_n = \frac{\Pr(H_+|\xi_{1:n})}{\Pr(H_-|\xi_{1:n})} = \frac{f_+(\xi_1)f_+(\xi_2)\cdots f_+(\xi_n)\Pr(H_+)}{f_-(\xi_1)f_-(\xi_2)\cdots f_-(\xi_n)\Pr(H_-)},$$

which can also be written recursively [45]:

$$(2.1) \quad R_n = \left(\frac{f_+(\xi_n)}{f_-(\xi_n)} \right) \cdot R_{n-1},$$

where $R_0 = \Pr(H_+)/\Pr(H_-)$ can describe an observer's prior belief about the probability of the two choices.

With a fixed number of observations, the ratio in (2.1) can be used to make a choice that minimizes the total error rate [32] or maximizes reward [23]. Equation (2.1) gives a recursive relation for the log-likelihood ratio, $y_n = \ln R_n$,

$$(2.2) \quad y_n = y_{n-1} + \ln \frac{f_+(\xi_n)}{f_-(\xi_n)}.$$

When the time between observations, $\Delta t = t_j - t_{j-1}$, is small, we can use the functional central limit theorem (p. 357 in [6]) to approximate this stochastic process by the SDE [8, 35]

$$(2.3) \quad dy = g_\pm dt + \rho_\pm dW_t,$$

where W_t is a Wiener process, and the constants $g_\pm = \frac{1}{\Delta t} E_\xi[\ln \frac{f_+(\xi)}{f_-(\xi)} | H_\pm]$ and $\rho_\pm^2 = \frac{1}{\Delta t} \text{Var}_\xi[\ln \frac{f_+(\xi)}{f_-(\xi)} | H_\pm]$ depend on the environmental state. Below we approximate other discrete-time processes, such as that given by (2.2), with SDEs. Details of these derivations are provided in the appendices.

In state H_+ we have $g_+ \Delta t = \int_{-\infty}^{\infty} f_+(\xi) \ln \frac{f_+(\xi)}{f_-(\xi)} d\xi$. The drift between two observations thus equals the Kullback–Leibler divergence between f_+ and f_- , i.e., the strength of the observed evidence from a measurement in favor of H_+ [14]. An equivalent interpretation holds for g_- . Hence g_+ and g_- are the rates at which an optimal observer accumulates information. We will use this observation to interpret the parameters of the model in a changing environment.

3. Two Alternatives in a Changing Environment. We use the same assumptions to derive a recursive equation for the log-likelihood ratio between two alternatives in a changing environment. The state of the environment, $H(t)$, is H_+ or H_- , but can now change in time (see Figure 1A, B). When the environment is in one of these two possible states, the statistics of the observations are fixed. Observation statistics are therefore piecewise stationary in time. An observer infers the present state from a sequence of observations, $\xi_{1:n}$, made at equally spaced times, $t_{1:n}$ with $\Delta t = t_j - t_{j-1}$ and characterized by probabilities $f_\pm(\xi_n) := \Pr(\xi_n|H_\pm)$. The state of the environment changes according to a telegraph process (e.g., p. 77 in [21]), and the probability of a change between two observations is $\epsilon_\pm \Delta t := \Pr(H(t_n) = H_\mp | H(t_{n-1}) = H_\pm)$. We assume that ϵ_+ and ϵ_- are known to the observer.

The probabilities, $L_{n,\pm} = \Pr(H(t_n) = H_\pm | \xi_{1:n})$, then satisfy (see Appendix A)

$$(3.1) \quad L_{n,\pm} \propto f_\pm(\xi_n) ((1 - \Delta t \epsilon_\pm) L_{n-1,\pm} + \Delta t \epsilon_\mp L_{n-1,\mp}),$$

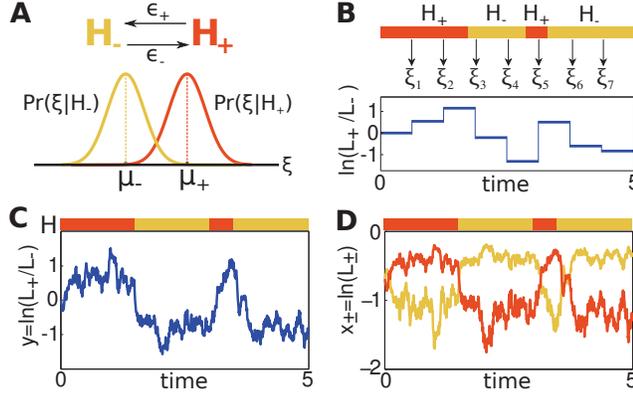


Fig. 1 Evidence accumulation in a changing environment. (A) The environmental state transitions from state H_+ to H_- and back with rates ϵ_+ and ϵ_- , respectively. Observations follow state-dependent probabilities, $f_{\pm}(\xi) = \Pr(\xi|H_{\pm})$. (B) The distributions of the measurements, ξ_n , change with the environmental state. Each individual observation changes the log-likelihood ratio, $\ln(L_{n,+}/L_{n,-})$. A single realization is shown. (C, D) The evolution of the continuous approximation of the log-likelihood ratio, $y(t)$ (panel C), and the log probabilities $x_{\pm}(t)$ (panel D). At time t , evidence favors the environmental state H_+ if $y(t) > 0$, or, equivalently, if $x_+(t) > x_-(t)$.

with proportionality constant $\Pr(\xi_{1:n-1})/\Pr(\xi_{1:n})$. As in the static case, the ratio of the probabilities of the two environmental states at time t_n can be determined recursively (see Appendix A) and equals

$$(3.2) \quad R_n = \frac{L_{n,+}}{L_{n,-}} = \frac{f_+(\xi_n)}{f_-(\xi_n)} \frac{(1 - \Delta t \epsilon_+) R_{n-1} + \Delta t \epsilon_-}{\Delta t \epsilon_+ R_{n-1} + 1 - \Delta t \epsilon_-}.$$

In this expression, the ratio of probabilities at the time of the previous observations, R_{n-1} , is discounted in a way that depends on the frequency of environmental changes, ϵ_{\pm} . This equation and the continuum limits we discuss below have been derived previously [16, 50], but their dynamics were not analyzed.

Equation (3.2) describes a variety of cases of evidence accumulation studied previously (see Figure 2): If the environment is fixed ($\epsilon_{\pm} = 0$), we recover (2.1). If the environment starts in state H_- , changes to H_+ , but cannot change back ($\epsilon_- > 0, \epsilon_+ = 0$), we obtain

$$R_n = \frac{f_+(\xi_n)}{f_-(\xi_n)} \frac{R_{n-1} + \Delta t \epsilon_-}{1 - \Delta t \epsilon_-},$$

a model used in change point detection [1, 39, 49].

We can again approximate evolution of $y_n = \ln R_n$, i.e., the stochastic process describing the evolution of the log of the likelihood ratio in (3.2), by an SDE:

$$(3.3) \quad dy = [g(t) + \underbrace{\epsilon_-(e^{-y} + 1) - \epsilon_+(e^y + 1)}_{\text{nonlinearity}}] dt + \rho(t) dW_t,$$

where $g(t) = \frac{1}{\Delta t} E_{\xi} \left[\ln \frac{f_+(\xi)}{f_-(\xi)} | H(t) \right]$ and $\rho^2(t) = \frac{1}{\Delta t} \text{Var}_{\xi} \left[\ln \frac{f_+(\xi)}{f_-(\xi)} | H(t) \right]$. Note that the drift and variance are no longer constant, but depend on the state of the environment $H(t)$ at time t . We use $E_{\xi} [F(\xi) | H(t)]$ to denote the expectation of $F(\xi)$ when ξ is drawn from the distribution associated with the current state $H(t)$, i.e., $f_{\pm}(\xi)$ when

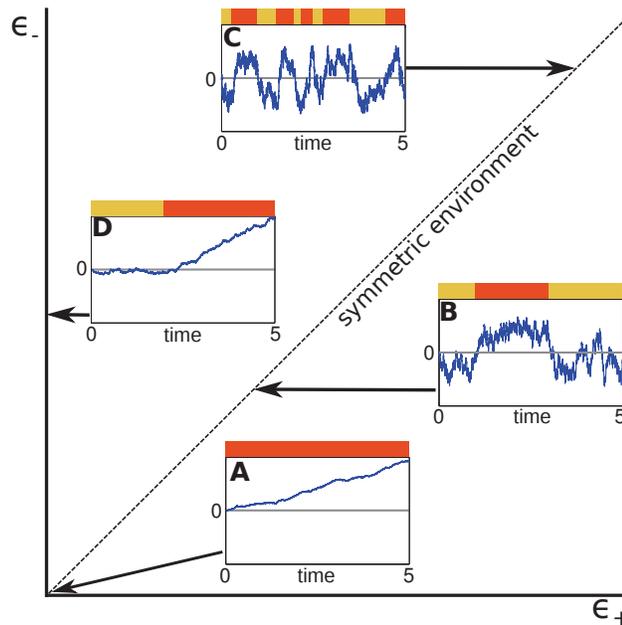


Fig. 2 In a dynamic environment, the dynamics of the log-likelihood ratio, y , depends on the rates of switching between states. (A) When $\epsilon_{\pm} = 0$, the environment is static, and the model reduces to the one derived in section 2. (B) When the environment changes slowly, $|\epsilon_{\pm}| \ll 1$, the log-likelihood ratio, y , can saturate. (C) In a rapidly changing environment, y tends not to equilibrate. (D) When $\epsilon_+ = 0$ and $\epsilon_- > 0$, the task becomes a change detection problem.

$H(t) = H_{\pm}$. In Appendix B we derive (3.3) as the continuum limit of the discrete process y_n .

As a consequence of the nonlinearity of (3.3) the state variable $y(t)$ will not drift indefinitely when $g(t)$ is fixed for some time interval $t \in [a, b]$. Rather, trajectories will tend to accumulate about the single fixed point of the noise-free system (the case $\rho(t) \equiv 0$). Importantly, more volatile environments (larger ϵ_{\pm}) correspond to fixed points that are closer to the midline $y = 0$, allowing for more rapid changes in $\text{sign}[y(t)]$. The observer's belief about the environmental state is encoded by the log-likelihood ratio and changes at a rate related to the frequency of environmental changes.

The nonlinear term in (3.3) does not appear in (2.3). It serves to discount older evidence by a factor determined by environmental volatility, i.e., the frequencies of changes in environmental states, ϵ_{\pm} . In previous work such discounting was modeled heuristically by a linear term [41–43]; however, our derivation shows that the resulting Ornstein–Uhlenbeck (OU) process is only an approximation of an optimal observer's evidence accumulation process.

3.1. Equal Switching Rates between Two States. When $\epsilon := \epsilon_+ = \epsilon_-$, the frequencies of switches between states are equal. Equation (3.3) then becomes

$$(3.4) \quad dy = g(t)dt - 2\epsilon \sinh(y)dt + \rho(t)dW_t.$$

The steepness of the function $\sinh(y)$ at large values of y ensures that evidence is discounted more rapidly for large log-likelihood ratios than for small ones (see Fig-

ure 4F, below). As a result, evidence builds up faster when y is closer to zero; i.e., the observer is more uncertain. If we rescale time using $\tau = \epsilon t$, the rate of switches between environmental states is unity. We obtain an equation for $y_\tau := y(\tau/\epsilon)$:

$$(3.5) \quad dy_\tau = [\tilde{g}(\tau)] d\tau - 2 \sinh(y_\tau) d\tau + [\tilde{\rho}(\tau)] dW_\tau,$$

where $\tilde{g}(\tau) := g(t)/\epsilon = g(\tau/\epsilon)/\epsilon$ and $\tilde{\rho}(\tau) := \rho(\tau/\epsilon)/\sqrt{\epsilon}$. Recall that $g(t)$ is the rate of evidence accumulation in the present state, and ϵ^{-1} is the average time spent in each state. Hence, $\tilde{g}(\tau) = g(t)/\epsilon$ can be interpreted as the *information gained over an average duration of the present environmental state*.

When observations follow Gaussian distributions, $f_\pm \sim \mathcal{N}(\pm\mu, \sigma^2)$, then $g(t) = \pm 2\mu^2/\sigma^2$, $\rho = 2\mu/\sigma$, and

$$(3.6) \quad dy_\tau = \text{sign}[\tilde{g}(\tau)] m d\tau - 2 \sinh(y_\tau) d\tau + \sqrt{2m} dW_\tau,$$

where $m = 2\mu^2/(\sigma^2\epsilon)$. Thus, the behavior of this system is completely determined by the single parameter m , the information gain over an average environmental epoch.

We now analyze the results of two decision-making processes that utilize the log-likelihood ratio. Under the *interrogation protocol*, the observer waits until a given time $\tau = T$ and reports $\text{sign}[y_\tau(T)] = \pm 1$. Under the *free response protocol*, we assume that the observer uses a predetermined threshold, θ , waits until time τ^* , at which the decision variable meets this threshold, $|y_\tau(\tau^*)| = \theta$, and then reports $\text{sign}[y_\tau(\tau^*)]$. For (3.6), the probabilities of a correct response (accuracy) under both interrogation (Figure 3A) and free response (Figure 3B) protocols increase with m . When an optimal observer is interrogated about the state of the environment at time T , the answer is determined by the sign of the log-likelihood ratio, y_τ . Since observers discount old evidence at a rate increasing with $1/m$, decisions are effectively based on a fixed amount of evidence, and accuracy saturates at a value smaller than 1 (Figure 3A). On the other hand, accuracy arbitrarily close to 1 can be obtained in the free response protocol by increasing the threshold θ (Figure 3B). Equations for the case of multiple alternatives ($N > 2$) are provided in section 4, and increasing N decreases accuracy for a fixed decision time (Figure 3C, D).

If the environment in (3.6) remains in a single state for a long time, the log-likelihood ratio, y_τ , approaches a stationary distribution,

$$(3.7) \quad \mathcal{S}_\pm(y_\tau) = K \exp\left(\pm y_\tau - \frac{2 \cosh(y_\tau)}{m}\right), \quad \tilde{H}(\tau) = H_\pm,$$

where $\tilde{H}(\tau) := H(\tau/\epsilon)$ and K is a normalization constant. Details on finding the stationary density of the Fokker–Planck equation associated with a nonlinear SDE such as (3.6) can be found in Chapter 5 of [21]. The distribution, (3.7), is concentrated around $\bar{y}_{\tau\pm} = \pm \sinh^{-1} \frac{m}{2}$, the fixed points of the deterministic counterpart of (3.6) obtained by setting $W_\tau \equiv 0$. Since old evidence is continuously discounted, the belief of an optimal observer tends to saturate. In contrast, no stationary distribution exists when $\epsilon = 0$, and the environment is static: Aggregating new evidence then always tends to increase an optimal observer’s belief in one of the choices.

Since $\mathcal{S}_\pm(y)$ is obtained by assuming that the environment is trapped in a single state for an extended time, $\int_0^\infty \mathcal{S}_+(y) dy = \int_{-\infty}^0 \mathcal{S}_-(y) dy$ provides an upper bound on the accuracy (Figure 3A). To achieve accuracy a in the free response protocol (Figure 3B), we require $|y_\tau| \geq \ln \frac{a}{1-a}$ [8]. While the threshold $\theta = \ln \frac{a}{1-a}$ that $y_\tau(\tau)$ must cross to obtain a specific accuracy, a , does not change with m , the time to reach

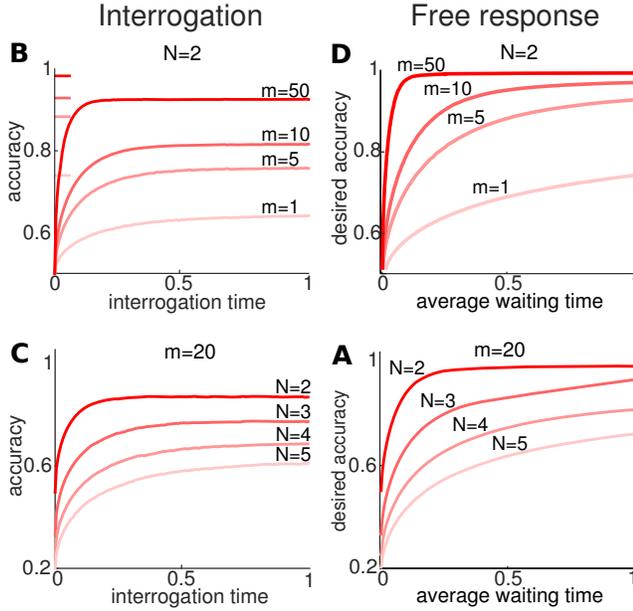


Fig. 3 *Dependence of the probability of the correct response (accuracy) on normalized information gain, m , in a symmetric environment. (A) Accuracy in an interrogation protocol increases with m and interrogation time, t , but saturates. Horizontal bars on the left indicate the accuracy when the environment is in a single state for a long time, as in (3.7). (B) When the observer responds freely, accuracy is similar but saturates at 1. The increase in accuracy with waiting time is exceedingly slow for low m . We fix $\epsilon_{ij} \equiv \epsilon$ for all $i \neq j$, $g_i \equiv g$, and set $m = g/\epsilon \equiv 20$. (C) Accuracy in an interrogation protocol decreases with the number of alternatives N (see section 4 for $N > 2$), saturating at ever lower levels. (D) The free response protocol results in similar behavior, but the accuracy saturates at 1. The increase in accuracy with waiting time is exceedingly slow for higher numbers of alternatives N .*

this threshold increases steeply with a and decreases with m . This is partly due to the fact that for smaller m , environmental switches are rapid, causing frequent changes in the drift of $y_\tau(\tau)$ and keeping it close to the midline $y_\tau = 0$.

3.2. Linear Approximation of the SDE. An advantage of (3.3) is that it is amenable to standard methods of stochastic analysis. We can find an accurate piecewise linear approximation to (3.3), although, for simplicity, we focus on (3.6). The piecewise OU process that models an observer that linearly discounts evidence has the form

$$(3.8) \quad dy_\tau = b(\text{sign}[\tilde{g}(\tau)]m d\tau + \sqrt{2m}dW_\tau) + \lambda y_\tau d\tau.$$

For (3.8) to be the continuum limit of a linear log-likelihood update process, the drift and diffusion need to be co-scaled by the common parameter b . We begin by focusing on a linear approximation of (3.6) with the same equilibria and local stability, obtained by setting $\lambda = -\sqrt{m^2 + 4}$ and $b = \sqrt{1 + \frac{4}{m^2}} \sinh^{-1} \frac{m}{2}$. Individual realizations of (3.8) and (3.6) agree in quickly changing environments (Figure 4A, $m = 1$), but are less similar in slowly changing environments (Figure 4B, $m = 10$; see also panel C). Thus, as observer performance improves, the nonlinear term in (3.6) becomes more important. Note that the corresponding drift-diffusion model, $dy_\tau = \text{sign}[\tilde{g}(\tau)]m d\tau +$

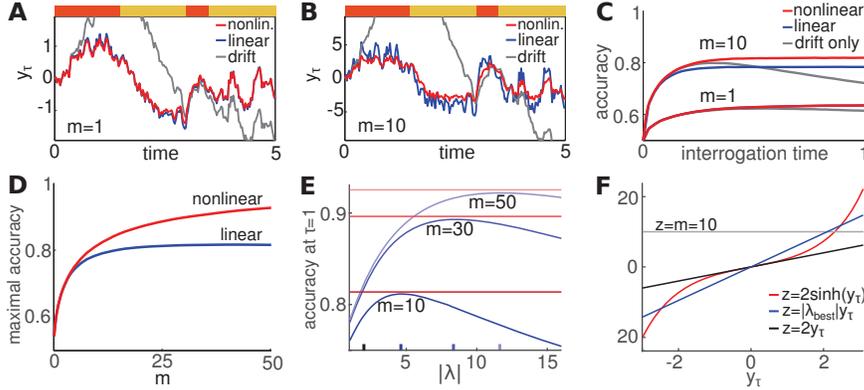


Fig. 4 Closest linear approximations of the nonlinear SDE (3.6). (A, B) Single realizations of the nonlinear (3.6), linear approximation (3.8), and corresponding drift-diffusion model $dy_\tau = \text{sign}[\tilde{g}(\tau)]m d\tau + \sqrt{2m}dW_\tau$, in (A) a quickly changing environment ($m = 1$), and in (B) a slowly changing environment ($m = 10$). We used the same realizations of drift $\tilde{g}(\tau)$ and noise W_τ for all models. (C) In the interrogation protocol, accuracy increases faster in the nonlinear (3.6) than in the linear approximation (3.8). Accuracy eventually decreases in the drift model since all evidence is weighted equally across time. (D) In the limit $t \rightarrow \infty$, accuracy saturates below unity in both the nonlinear model and the linear approximation. The linear model discounts evidence suboptimally and hence performs worse. (E) Accuracy under the interrogation protocol with stopping time $\tau = 1$ for the linear model, (3.8), with leak λ (blue ticks: λ_{best} ; black tick: $\lambda = -2$). Optimal λ -values for the linear approximation (blue curves) result in accuracy that is very close to that of the optimal nonlinear model, (3.6) (red lines). (F) Plot of $2 \sinh(y_\tau)$ demonstrating two possible linear approximations: most accurate linear approximation from panel E (blue), and linearization of $2 \sinh(y_\tau)$ at the origin (black).

$\sqrt{2m}dW_\tau$, is qualitatively different as it lacks a restorative leak term. This difference becomes more pronounced as m increases (Figure 4C).

Equation (3.8) can be integrated explicitly using standard methods in stochastic calculus [21]. Furthermore, the accuracies of both systems saturate to a value smaller than 1 in the interrogation protocol as the interrogation time increases (Figure 4C).

This linearized approximation can differ considerably from the full nonlinear model. For instance, in the interrogation protocol the performance of an ideal observer modeled by (3.6) increases with interrogation time (Figure 4D), and accuracy approaches 1 as m diverges. In contrast, the accuracy of an observer that discounts evidence linearly limits to a value below 1 as m diverges. Indeed, this can be seen by employing the quasi-steady state approximation (fixing $\text{sign}[\tilde{g}(\tau)] = 1$) and computing $\int_0^\infty \mathcal{T}(y)dy$, where $\mathcal{T}(y)$ is the steady state distribution of the OU process given in (3.8) with $\lambda = -\sqrt{m^2 + 4}$ and $b = \sqrt{1 + \frac{4}{m^2}} \sinh^{-1} \frac{m}{2}$, to obtain

$$U_m := \int_0^\infty \mathcal{T}(y)dy = \frac{1}{2} + \frac{1}{2} \text{erf} \left(\sqrt{\frac{m}{2\sqrt{m^2 + 4}}} \right),$$

and $\lim_{m \rightarrow \infty} U_m = \frac{1}{2} + \frac{1}{2} \text{erf} \frac{1}{\sqrt{2}} \approx 0.84 < 1$.

In general, there is a family of linear approximations to (3.6) given by (3.8), where $\lambda \in (-\infty, 0]$. However, the choice of λ depends on the way we measure the quality of the approximation. For example, we may want to maximize decision accuracy under the interrogation policy with a specific stopping time, or maximize accuracy under the

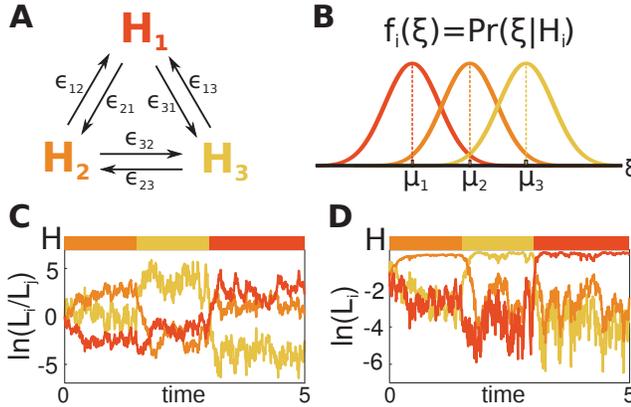


Fig. 5 Evidence accumulation with multiple choices in a changing environment. (A) The environment switches between N states (here $N = 3$). (B) Distributions $f_i(\xi) = \Pr(\xi|H_i)$ describing the probability of observation ξ in environmental state H_i (here $N = 3$). (C, D) Realization of the log-likelihood ratios $\ln(L_1/L_2)$, $\ln(L_2/L_3)$, $\ln(L_3/L_1)$ (panel C); and log-likelihoods $\ln L_1$, $\ln L_2$, $\ln L_3$ (panel D).

free response policy. In general, we need numerical optimization methods to identify the λ that provides the best linear approximation. Without loss of generality, we can fix $b \equiv 1$ in (3.8), since the rescaling $z_\tau = y_\tau/b$ preserves $\text{sign}[z_\tau] = \text{sign}[y_\tau]$ and eliminates b . Thus, we need only study the system $dy_\tau = \text{sign}[\tilde{g}(\tau)] m d\tau + \sqrt{2m} dW_\tau + \lambda y_\tau d\tau$. For a given m , there is a single value, $\lambda = \lambda_{\text{best}}$, that maximizes the accuracy of decisions after an interrogation at $\tau = 1$ (Figure 4E). However, there is a different λ_{best} for each value of m . Interestingly, the best linear approximation has accuracy close to that of the nonlinear system. We note that the linear approximation at the origin ($\lambda = -2$; see also Figure 4F) did not perform well. Since the accuracy has saturated at $\tau = 1$ (Figure 4C), the optimal value of λ will not change significantly for larger interrogation times.

Similarly, λ_{best} will change with different thresholds under the free response protocol, or with other measures of performance, such as reward rate [8]. In contrast, the nonlinear model given by (3.6) reflects the log-likelihood ratio exactly. Therefore, we can use this single model for any decision that can be made optimally using the log-likelihood ratio.

4. Multiple Alternatives in a Changing Environment. We next extend our analysis of evidence accumulation in changing environments to the case of multiple alternatives. With multiple environmental states, H_i ($i = 1, \dots, N$), the optimal observer computes the present probability of each state (Figure 5A) from a sequence of measurements, $\xi_{1:n}$. Measurements have probability $f_i(\xi_n) := \Pr(\xi_n|H_i)$ dependent on the states H_i [4, 9] (Figure 5B). We assume that the state of the environment, $H(t)$, changes as a memoryless process. A change from state j to i between two measurements occurs with probability $\epsilon_{ij}\Delta t = \Pr(H(t_n) = H_i|H(t_{n-1}) = H_j)$ for $i \neq j$, and $\Pr(H(t_n) = H_i|H(t_{n-1}) = H_i) = 1 - \sum_{j \neq i} \Delta t \epsilon_{ji}$ (Figure 5A).

We again use sequential analysis to obtain the probabilities $L_{n,i} = \Pr(H(t) = H_i|\xi_{1:n})$ that the environment is in state H_i given observations $\xi_{1:n}$. The index that maximizes the posterior probability, $\hat{i} = \text{argmax}_i L_{n,i}$, corresponds to the most probable state, given the observations $\xi_{1:n}$. Following the approach above, we obtain (see

Appendix D)

$$(4.1) \quad L_{n,i} = \frac{\Pr(\xi_{1:n-1})}{\Pr(\xi_{1:n})} f_i(\xi_n) \left(\left(1 - \sum_{j \neq i} \Delta t \epsilon_{ji} \right) L_{n-1,i} + \sum_{j \neq i} \Delta t \epsilon_{ij} L_{n-1,j} \right).$$

Again after taking logarithms, $x_{n,i} = \ln L_{n,i}$, we can approximate the discrete stochastic process in (4.1) with an SDE:

$$(4.2) \quad d\mathbf{x} = \mathbf{g}(t)dt + \Lambda(t)d\mathbf{W}_t + K(\mathbf{x})dt,$$

where the drift has components $g_i(t) = \frac{1}{\Delta t} E_\xi [\ln f_i(\xi) | H(t)]$, $\Lambda(t)\Lambda(t)^T = \Sigma(t)$ with entries $\Sigma_{ij} = \frac{1}{\Delta t} \text{Cov}_\xi [\ln f_i(\xi), \ln f_j(\xi) | H(t)]$, components of \mathbf{W}_t are independent Wiener processes, and $K_i(\mathbf{x}) = \sum_{j \neq i} (\epsilon_{ij} e^{x_j - x_i} - \epsilon_{ji})$. The drift g_i is maximized in environmental state H_i (Figure 5C, D).

We can recover the case of two alternatives by setting $N = 2$ and exchanging the numbers in (4.2) with \pm to obtain the approximating SDEs

$$(4.3) \quad dx_\pm = [g_\pm(t) + (\epsilon_\mp e^{x_\mp - x_\pm} - \epsilon_\pm)]dt + dW_\pm,$$

where $\langle W_i W_j \rangle = \Sigma_{ij}(t) \cdot t$ for $i, j \in \{+, -\}$. We obtain (3.3) by setting $y := x_+ - x_-$. Note that since $x_\pm = \ln(L_\pm)$ are the log-likelihoods, $y := x_+ - x_- = \ln(L_+/L_-)$ is the log-likelihood ratio. Analogous expressions for the log-likelihood ratios $y_{ij} = \ln(L_i/L_j)$ are derived in Appendix E. The matrix of these log-likelihood ratios quantifies how much more likely one alternative is compared to others (e.g., Figure 5C) [18].

5. A Continuum of States in a Changing Environment. Lastly, we consider the case of a continuum of possible environmental states. This provides a tractable model for recent experiments with observers who infer the location of a hidden, intermittently moving target from noisy observations. Evidence suggests that humans update their beliefs quickly and near optimally when observations indicate that the target has moved [29].

Suppose the environmental state, $H(t)$, intermittently switches between a continuum of possible states, H_θ , where $\theta \in [a, b]$. An observer again computes the probabilities of each state from observations, $\xi_{1:n}$, with distributions $f_\theta(\xi_n) := \Pr(\xi_n | H_\theta)$. The environment switches from state θ' to state θ between observations with transition probabilities $\epsilon_{\theta\theta'} d\theta \Delta t := \Pr(H(t_n) = H_\theta | H(t_{n-1}) = H_{\theta'})$ for $\theta \neq \theta'$, and $\Pr(H(t_n) = H_\theta | H(t_{n-1}) = H_\theta) = 1 - \int_a^b \Delta t \epsilon_{\theta\theta'} d\theta'$ (see Appendix F for details). From (4.1) the expression for the probabilities $L_{n,\theta} = \Pr(H(t_n) = H_\theta | \xi_{1:n})$ is derived in Appendix F, yielding

$$L_{n,\theta} = \frac{\Pr(\xi_{1:n-1})}{\Pr(\xi_{1:n})} f_\theta(\xi_n) \left(\left(1 - \int_a^b \Delta t \epsilon_{\theta\theta'} d\theta' \right) L_{n-1,\theta} + \int_a^b \Delta t \epsilon_{\theta\theta'} L_{n-1,\theta'} d\theta' \right).$$

We again approximate the logarithms of the probabilities, $\ln L_{n,\theta}$, by a temporally continuous process,

$$(5.1) \quad dx_\theta(t) = g_\theta(t)dt + d\widehat{W}_\theta(t) + K_\theta(x(t))dt,$$

where $x = (x_\theta)_{\theta \in [a,b]}$, $g_\theta(t) = \frac{1}{\Delta t} E_\xi [\ln f_\theta(\xi) | H(t)]$, \widehat{W}_θ is a spatiotemporal noise term with mean zero and covariance function given by

$$\Sigma_{\theta\theta'}(t) = \frac{1}{\Delta t} \text{Cov}_\xi [\ln f_\theta(\xi), \ln f_{\theta'}(\xi) | H(t)],$$

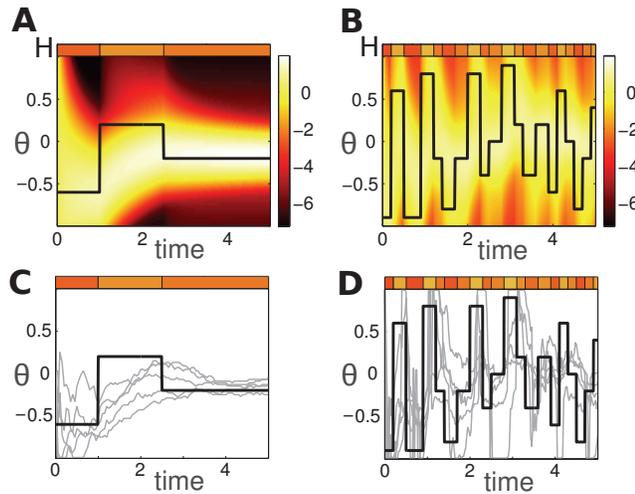


Fig. 6 Evidence accumulation with a continuum of choices. The observer infers the state of the environment, H_θ , where $\theta \in [-1, 1]$, and the state changes at discrete points in time. (A) In slowly changing environments, the distribution of the log probabilities, x_θ , can nearly equilibrate between switches. (Solid line represents the true state of the environment at time t . For clarity, we show results of simulations without noise, $\bar{W}_\theta \equiv 0$.) (B) In quickly changing environments, the distribution does not have time to equilibrate between switches. (C) In slowly changing environments, the most probable state of the environment, $\hat{\theta}(t) = \operatorname{argmax}_\theta x_\theta(t)$ (thin lines), fluctuates around the true value (thick line). (D) In quickly changing environments, $\hat{\theta}(t)$ fluctuates more widely, as it is in a transient state much of the time.

and $K_\theta(x) = \int_a^b (\epsilon_{\theta\theta'} e^{x_{\theta'} - x_\theta} - \epsilon_{\theta'\theta}) d\theta'$ is an interaction term describing the discounting process.

The drift $g_\theta(t)$ is maximal when θ agrees with the present environmental state. The most likely state, given observations up to time t , is $\hat{\theta}(t) = \operatorname{argmax}_\theta x_\theta(t)$.

In slowly changing environments, the log probability $x_\theta(t)$ *nearly* equilibrates to a distribution with a well-defined peak between environmental switches (Figure 6A). This does not occur in quickly changing environments (Figure 6B). However, each logarithm $x_\theta(t)$ approaches a stationary distribution if the environmental state remains fixed for a long time. The term $K_\theta(x)$ in (5.1) causes rapid departure from this quasi-stationary density when the environment changes, a mechanism proposed in [29].

Even when the environment is stationary for a long time, noise in the observations stochastically perturbs the log probabilities, $x_\theta(t)$, over the environmental states. This leads to fluctuations in the estimate $\hat{\theta}(t)$ of the most probable alternative (Figure 6C, D). Thus, as opposed to the case of a discrete space of N alternatives, the observer's estimate of the most probable choice will change continuously, fluctuating about the continuum of possible alternatives. Unless changes are too rapid, the peak of the log probability distribution, $\hat{\theta}(t)$, fluctuates around the true environmental state and tracks abrupt changes in $H_\theta(t)$. This is in line with recent observations in human behavioral data [22, 29].

6. A Neural Implementation of an Optimal Observer. Previous neural models of decision making typically relied on mutually inhibitory neural networks [10, 30, 43, 46], with each population representing one alternative. These models match the

recorded neural activity and responses of monkeys performing two-alternative forced-choice decision tasks, where single trial stimuli have stationary statistics [24]. Even when reward rates are varied across trials, animals can adjust their behavior near optimally from trial to trial in ways that are well captured by mutually inhibitory models [20]. Interestingly, these networks also provide a plausible model of decision making in house-hunting honeybee swarms [33]. In previous studies, it has been shown that a single fixed point can be stabilized in linear population models, as long as the strength of mutual inhibition is weaker than the leak of individual populations [8, 10, 43]. As we will show, a complementary approach in linear population models is to consider a mutually excitatory network, with arbitrary leak in individual populations. As with the linear approximations discussed above, such models perform suboptimal inference in changing environments, but can approach the performance of the ideal nonlinear discounting process given by (3.3).

Optimal inference in dynamic environments with two states, H_+ and H_- , can be performed by mutually excitatory nonlinear neural populations with activities (firing rates) r_+ and r_- evolving according to

$$(6.1a) \quad dr_+ = [I_+(t) - \alpha r_+ + F_+(r_- - r_+)] dt + dW_+,$$

$$(6.1b) \quad dr_- = [I_-(t) - \alpha r_- + F_-(r_+ - r_-)] dt + dW_-,$$

where the transfer functions are $F_{\pm}(x) = -\alpha x/2 + \epsilon_{\mp} e^x - \epsilon_{\pm}$, the mean input $I_{\pm}(t) = I_{\pm}^0$ when $H(t) = H_{\pm}$ and vanishes otherwise, and W_{\pm} are Wiener processes representing the variability in the input signal with covariance defined as in (4.3) (see Appendix D). Thus, $I_{\pm}(t)dt + dW_{\pm}$ represents the total input to population r_{\pm} . When $\alpha > 0$ and is sufficiently small, population activities are modulated by self-inhibition/leak, and mutual excitation (Figure 7A). The parameter α determines the leak in each individual population, which depends on both the time constants and recurrent architecture of the local network [46]. The difference $y = r_+ - r_-$ evolves according to the SDE for the log-likelihood ratio, (3.3). In the limit of a stationary environment, $\epsilon_{\pm} \rightarrow 0$, we obtain a linear integrator $dr_{\pm} = [I_{\pm}dt + dW_{\pm}] - \alpha(r_+ + r_-)dt/2$, as in previous studies [8, 30].

To show that the populations mutually excite each other, we set $W_+ = W_- = 0$ and study the dynamics in the vicinity of the fixed points of (6.1). When the environment has not changed for a long time, (6.1) approaches a fixed point (\bar{r}_+, \bar{r}_-) with

$$(\bar{r}_+, \bar{r}_-) = \left(\frac{I_+^0 + \epsilon_- e^{-\bar{y}_+} - \epsilon_+}{\alpha} + \frac{\bar{y}_+}{2}, \frac{\epsilon_+ e^{\bar{y}_+} - \epsilon_-}{\alpha} - \frac{\bar{y}_+}{2} \right),$$

when $I_+(t) = I_+^0$ and $I_-(t) = 0$, and

$$(\bar{r}_+, \bar{r}_-) = \left(\frac{\epsilon_- e^{-\bar{y}_-} - \epsilon_+}{\alpha} + \frac{\bar{y}_-}{2}, \frac{I_-^0 + \epsilon_+ e^{\bar{y}_-} - \epsilon_-}{\alpha} - \frac{\bar{y}_-}{2} \right),$$

when $I_+(t) = 0$ and $I_-(t) \equiv I_-^0$, where

$$\bar{y}_{\pm} = \ln \left[\frac{\pm I_{\pm}^0 + \epsilon_- - \epsilon_+}{2\epsilon_+} + \sqrt{\frac{(\pm I_{\pm}^0 + \epsilon_- - \epsilon_+)^2}{4\epsilon_+^2} + \frac{\epsilon_-}{\epsilon_+}} \right].$$

Note that by increasing (decreasing) α , the fixed points (\bar{r}_+, \bar{r}_-) move closer to (farther from) the origin $(0, 0)$. To determine the sign of the coupling near these fixed points,

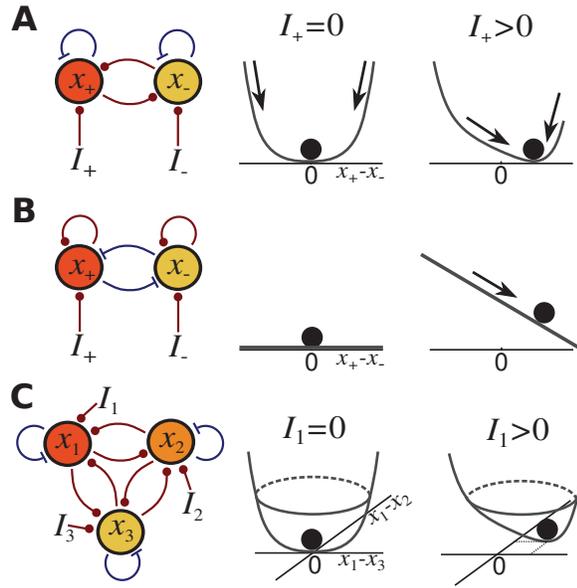


Fig. 7 *Neural population models of evidence accumulation. (A) Two populations u_{\pm} receive a fluctuating stimulus with mean I_{\pm} ; they are mutually coupled by excitation (circles) and locally affected by inhibition/leak (flat ends). When $I_{+} > 0$, the fixed point of the system has coordinates satisfying $x_{+} > x_{-}$, as shown in the plots of the associated potentials. (B) Taking $\epsilon_{\pm} \rightarrow 0$ in (6.1) generates a mutually inhibitory network that perfectly integrates inputs I_{\pm} and has a flat potential function. (C) With $N = 3$ alternatives, three populations coupled by mutual excitation can still optimally integrate the inputs $I_{1,2,3}$, rapidly switching between the fixed point of the system in response to environmental changes.*

note that the Jacobian matrix of (F_{+}, F_{-}) has the form

$$(6.2) \quad J(r_{+}, r_{-}) = \begin{bmatrix} \alpha/2 - \epsilon_{-}e^{r_{-}-r_{+}} & -\alpha/2 + \epsilon_{-}e^{r_{-}-r_{+}} \\ -\alpha/2 + \epsilon_{+}e^{r_{+}-r_{-}} & \alpha/2 - \epsilon_{+}e^{r_{+}-r_{-}} \end{bmatrix}.$$

For $\epsilon_{\pm} > 0$, taking $\alpha < 2 \min\{\epsilon_{-}e^{-\bar{v}}, \epsilon_{+}e^{\bar{v}}\}$ will guarantee that the sign of the Jacobian matrix is $\begin{bmatrix} - & + \\ + & - \end{bmatrix}$ in a region that contains the fixed point. This corresponds to a neural network with self-inhibition/leak and mutual excitation, as illustrated in Figure 7A.

We can compare our results to previous studies of linear connectionist models [8, 10, 43] by deriving a linear rate model that best accumulates evidence in changing environments. To do so, we focus on the best linear approximation of the log-likelihood ratio, given by (3.8). We have shown that when the coefficients of the linear models are tuned appropriately, their accuracy is remarkably close to that of the full nonlinear model (Figure 4E). Assuming symmetric switching ($\epsilon_{\pm} \equiv \epsilon$), the following system describes a linear rate model that can be mapped to the linear (3.8):

$$(6.3a) \quad dr_{+} = [I_{+}(t) - \kappa r_{+} + \gamma r_{-}] dt + dW_{+},$$

$$(6.3b) \quad dr_{-} = [I_{-}(t) - \kappa r_{-} + \gamma r_{+}] dt + dW_{-}.$$

Here $\kappa > 0$ denotes the leak in each population’s activity, and $\gamma > 0$ is the strength of mutual excitation between populations. Selecting $I_{\pm}(t) = I_0$ when $H(t) = H_{\pm}$ and zero otherwise, it can be shown that the system will tend to the quasi-equilibria

$(r_+, r_-) = \frac{I_0}{\kappa^2 - \gamma^2} \cdot (\kappa, \gamma)$ and $\frac{I_0}{\kappa^2 - \gamma^2} \cdot (\gamma, \kappa)$ in either case. Stability of these fixed points is given by the nonzero eigenvalue $\lambda = -(\kappa + \gamma) < 0$, so these quasi-equilibria are always attractive. Note also that the reduced SDE for the difference $y = r_+ - r_-$ will take the form $dy = [I_d(t) - (\kappa + \gamma)y]dt + dW_d$, where $I_d(t) = I_+(t) - I_-(t)$ and $W_d = W_+ - W_-$, which matches the form of (3.8). Thus, in addition to the large leak in mutually inhibitory networks [8, 10, 43], linear population networks with mutual excitation possess a stable fixed point for arbitrary leak κ and mutual excitation γ . Any particular decision task has an optimal λ in (3.8). Thus, a linear neural network could be trained to learn this best evidence discounting rate if supplemented with a plasticity rule that properly tunes the excitation strength γ .

Returning to the nonlinear model given by (6.1), the dynamics is matched to the timescale of the environment determined by ϵ_{\pm} , and solutions approach stationary distributions if input is constant. The network's dynamics is very sensitive to changes in inputs, a feature absent in population models with winner-take-all dynamics [48]. Even when ϵ is small, (6.1) has a single attracting state determined by the mean inputs I_{\pm}^0 . We illustrate the response of the model to inputs using potentials (Figure 7A). In contrast to the single attractor of (6.1), mutually inhibitory models can possess a neutrally stable line attractor that integrates inputs ($\epsilon_{\pm} \equiv 0$, Figure 7B) [27].

We can extend our results to $N > 2$. In [4], the reliability of motion information was assumed to vary during a trial, and the optimal model encoded the posterior probability distribution over the possible stimulus space. Here, we assume the true hypothesis, $H(t)$, changes in time. For an arbitrary but finite number of possible alternatives, $\{H_1, \dots, H_N\}$, decisions can be performed optimally by neural populations (r_1, \dots, r_N) coupled by mutual excitation,

$$(6.4) \quad dr_i = \left[I_i(t) - \alpha r_i + \sum_{j \neq i} F_{ij}(r_j - r_i) \right] dt + d\widehat{W}_i(t),$$

where the mean input is $I_i(t) = I_i^0$ when $H(t) = H_i$ and 0 otherwise, and the noise vector $(d\widehat{W}_1(t), \dots, d\widehat{W}_N(t))^T = \Lambda(t)d\mathbf{W}_t$ describes input noise with $\Lambda(t)$ defined as in (4.2). Population firing rates are again determined by inhibition/leak within each population and excitation between populations as described by the arguments of the firing rate function

$$F_{ij}(r) = -\alpha r/N + \epsilon_{ij}e^r - \epsilon_{ji}.$$

In this case coupling between populations is again excitatory (Figure 7C).

Note that, as in the case of $N = 2$ alternatives, taking the limit of (6.4) as $\epsilon_{ij} \rightarrow 0$, we obtain linear integrators [30]

$$dr_i = \left[I_i(t) - \alpha \sum_{j=1}^N r_j/N \right] dt + d\widehat{W}_i(t).$$

The nonlinear population rate models described by (6.1) and (6.4) react rapidly, but not instantaneously, to changes in their inputs. Recent evidence suggests that in monkeys the activity of single neurons in the parietal cortex exhibits jumps, rather than a gradual increase as previously suggested [26]. Furthermore, the performance of rats and humans discriminating the direction of auditory click sequences can be optimally fit by a pulse-accumulating mechanism with zero noise [11]. However, the activity of a population of cells encoding behavior may still ramp upward or downward.

7. Discussion. We have derived a nonlinear stochastic model of optimal evidence accumulation in changing environments. Importantly, the resulting SDE is not an OU process, as suggested by previous heuristic models [35, 41, 43]. Rather, an exponential nonlinearity allows for optimal discounting of old evidence, and rapid adjustment of decision variables following environmental changes. As a result, the certainty of an optimal observer tends to saturate, even if the environment happens to be stuck in a single state for long periods of time.

We have made several assumptions about the model to simplify the derivations. Our ideal observer is assumed to be aware both of the uncertainty of their own measurements and of the frequency with which the environment changes. A more realistic model would require that a naive observer learn the underlying volatility of the environment. Modeling the case of initially unknown transition rates leads to hierarchical models that identify the location of change points [47]. However, this approach quickly grows in computational complexity, since the probability of change points is determined by accounting for all possible transition histories [1]. We also assumed that changes in the environment follow a memoryless process. In more general cases, we would not be able to obtain a recursive equation for the probability of a state. An ideal observer would have to use all previous observations at each step, rather than integrating the present observation with the posterior probability obtained with the previous observation. This process cannot be approximated by an SDE.

Sequential sampling in dynamic environments with two states has been studied previously in special cases, such as adapting spiking models, capable of responding to environmental changes [16]. Likelihood update procedures have also been proposed for multiple alternative tasks in the limit $\epsilon_{ij} \rightarrow 0$ [2, 18], but their dynamics was not analyzed. A related case of a temporally changing context has also been examined [40]. One important conclusion of our work is that $m = g/\epsilon$, the information gain over the characteristic environmental timescale, is the key parameter determining the model's dynamics and accuracy. It is easy to show that equivalent parameters govern the dynamics of likelihoods of multiple choices. This allows for a straightforward approximation of the nonlinear model by a linear SDE, which can be analyzed fully.

Models of evidence accumulation are of interest in disciplines ranging from neuroscience and robotics to psychology and economics. They can help us understand how decisions are made in cells, animals, ecological groups, and social networks. We presented a principled derivation of a series of nonlinear stochastic models amenable to stochastic analysis and have used quasi-static approximations, first passage techniques, and dimensional analysis to examine their dynamics. Thus we have built a bridge between classic models in signal detection theory and nonlinear stochastic processes. Continuous stochastic models have been very useful in interpreting human decision making in static environments [8, 24]. Dynamic environments offer a promising future direction for theory and experiments to probe the biophysical mechanisms that underlie decisions.

Appendices. In these appendices, we present the derivations for the probability update formulas and their approximations discussed in the main text. We begin by deriving the update expression for the probability ratio, R_n , in the case of two alternatives in a changing environment. The result is a nonlinear recursive equation. Subsequently, we show how to approximate the log-likelihood ratio, $y_n = \ln R_n$, using an SDE. To make the approximation precise, it is key to view the discrete equation for y_n as a family of equations parameterized by the time interval, Δt , over which each observation, ξ_n , is made [8]. Furthermore, we extend our derivations to multiple

($N > 2$) alternatives and show that the log probability updates can be approximated by a nonlinear system of SDEs in the continuum limit. With the appropriate scaling of the probabilities, $f_i(\xi) = \Pr(\xi|H_i)$, we can make precise the correspondence between the discrete and continuum models of posterior probability evolution. Lastly, we present a derivation for the stochastic integrodifferential equation that represents the log probability for a continuum of possible environmental states, $\theta \in [a, b]$.

Note that throughout the appendices, we use notation involving a subscript Δt . This helps us define a family of stochastic processes indexed by the spacing between observations $\Delta t = t_n - t_{n-1}$. For instance, $f_{\Delta t, \pm}(\xi)$ represents the probability of an observation, ξ , in environmental state H_{\pm} (or, in the language of statistics, when hypothesis H_{\pm} holds). This probability changes with the timestep Δt . This approach allows us to properly take the continuum limit $\Delta t \rightarrow 0$. However, for simplicity we refrain from using this notation in the main text. Rather, we treat the limiting SDEs as approximations of discrete update processes. Also, we slightly abuse notation and write $f_i(\xi) = \Pr(\xi|H_i)$, even when ξ is a continuous random variable.

Appendix A. Likelihood Ratio for Two Alternatives. We begin by deriving the recursive update equation for the probabilities $L_{n, \pm} := \Pr(H(t_n) = H_{\pm} | \xi_{1:n})$ associated with each alternative H_{\pm} , where each observation (measurement), ξ_i , is made at time t_i . This is the probability that alternative H_{\pm} is true at time t_n , given that the series of observations $\xi_{1:n}$ has been made. Importantly, the underlying truth changes stochastically, and in a memoryless way, with transition probabilities given by $\epsilon_{\Delta t, \pm} := \Pr(H(t_n) = H_{\mp} | H(t_{n-1}) = H_{\pm})$, so that $\Pr(H(t_n) = H_{\pm} | H(t_{n-1}) = H_{\pm}) = 1 - \epsilon_{\Delta t, \pm}$. We begin by examining the probability $L_{n,+}$ associated with the alternative H_+ . Using Bayes' rule and the law of total probability (Chapter 3 in [37]) we can relate the current probability, $L_{n,+}$, to the conditional probabilities at the time of the previous observation, t_{n-1} ,

$$L_{n,+} = \frac{1}{\Pr(\xi_{1:n})} \sum_{s=\pm} \Pr(\xi_{1:n} | H(t_n) = H_+, H(t_{n-1}) = H_s) \\ \times \Pr(H(t_n) = H_+, H(t_{n-1}) = H_s),$$

marginalizing over the joint distribution for the current, $H(t_n)$, and previous, $H(t_{n-1})$, environmental states. Next, we can apply the definition of the conditional probability $\Pr(H(t_n) = H_+ | H(t_{n-1}) = H_s)$ to write

$$L_{n,+} = \frac{1}{\Pr(\xi_{1:n})} \sum_{s=\pm} \Pr(\xi_{1:n} | H(t_n) = H_+, H(t_{n-1}) = H_s) \\ \times \Pr(H(t_n) = H_+ | H(t_{n-1}) = H_s) \Pr(H(t_{n-1}) = H_s).$$

Furthermore, we can split the joint condition on the first term by using the fact that the probability of making an observation ξ_n is independent of $H(t_{n-1})$ when we condition on the present state $H(t_n) = H_+$ of the environment, so $\Pr(\xi_{1:n} | H(t_n) = H_+, H(t_{n-1}) = H_s) = \Pr(\xi_n | H(t_n) = H_+) \Pr(\xi_{1:n-1} | H(t_{n-1}) = H_s)$ and

$$L_{n,+} = \frac{1}{\Pr(\xi_{1:n})} \sum_{s=\pm} \Pr(\xi_n | H(t_n) = H_+) \Pr(\xi_{1:n-1} | H(t_{n-1}) = H_s) \\ \times \Pr(H(t_n) = H_+ | H(t_{n-1}) = H_s) \Pr(H(t_{n-1}) = H_s).$$

Lastly, we apply Bayes' rule to switch the order of $\Pr(\xi_{1:n-1} | H(t_{n-1}) = H_s)$, yielding terms involving $L_{n-1,s} = \Pr(H(t_{n-1}) = H_s | \xi_{1:n-1})$. In addition, we use $\Pr(H(t_n) =$

$H_+|H(t_{n-1}) = H_+) = 1 - \epsilon_{\Delta t,+}$ and $\Pr(H(t_n) = H_+|H(t_{n-1}) = H_-) = \epsilon_{\Delta t,-}$, so that

$$(A.1) \quad L_{n,+} = \frac{\Pr(\xi_{1:n-1})\Pr(\xi_n|H(t_n) = H_+)}{\Pr(\xi_{1:n})} ((1 - \epsilon_{\Delta t,+})L_{n-1,+} + \epsilon_{\Delta t,-}L_{n-1,-}),$$

where $L_{0,+} = \Pr(H(t_0) = H_+)$.

Similarly, we can obtain an update equation for the probability $L_{n,-}$ of the alternative H_- at time t_n :

$$(A.2) \quad L_{n,-} = \frac{\Pr(\xi_{1:n-1})\Pr(\xi_n|H(t_n) = H_-)}{\Pr(\xi_{1:n})} (\epsilon_{\Delta t,+}L_{n-1,+} + (1 - \epsilon_{\Delta t,-})L_{n-1,-}),$$

where $L_{0,-} = \Pr(H(t_0) = H_-)$.

From (A.1) and (A.2), the ratio $R_n = L_{n,+}/L_{n,-}$ is readily seen to satisfy the recursive equation

$$(A.3) \quad R_n = \frac{f_{\Delta t,+}(\xi_n)}{f_{\Delta t,-}(\xi_n)} \frac{(1 - \epsilon_{\Delta t,+})R_{n-1} + \epsilon_{\Delta t,-}}{\epsilon_{\Delta t,+}R_{n-1} + 1 - \epsilon_{\Delta t,-}},$$

where $f_{\Delta t,\pm}(\xi_n) = \Pr(\xi_n|H(t_n) = H_\pm)$ is the distribution for each choice parameterized by the timestep $\Delta t = t_n - t_{n-1}$, and $R_0 = \frac{\Pr(H_+,t_0)}{\Pr(H_-,t_0)}$.

Appendix B. The Continuum Limit for the Log-Likelihood Ratio of Two Alternatives. In this section, we derive a continuum equation for the log-likelihood ratio $y_n := \ln R_n$. We will proceed by first defining a family of stochastic difference equations for y_n , which are parameterized by the timestep, $\Delta t = t_n - t_{n-1}$, between pairs of observations. By choosing an appropriate parameterization, we obtain a continuum limit that is an SDE. To begin, we divide both sides of (A.3) by R_{n-1} and take logarithms to yield

$$(B.1) \quad y_n - y_{n-1} = \ln \frac{f_{\Delta t,+}(\xi_n)}{f_{\Delta t,-}(\xi_n)} + \ln \frac{1 - \epsilon_{\Delta t,+} + \epsilon_{\Delta t,-}e^{-y_{n-1}}}{1 - \epsilon_{\Delta t,-} + \epsilon_{\Delta t,+}e^{y_{n-1}}}.$$

Following [7, 8], we assume that the time interval between individual observations, Δt , is small. Denote by $\Delta y_n = y_n - y_{n-1}$ the change in the log-likelihood ratio due to the observation at time t_n . By assumption, the probability that the environment changes between two observations scales linearly with Δt up to higher order terms, so that $\epsilon_{\Delta t,\pm} := \Delta t\epsilon_\pm + o(\Delta t)$. Omitting higher order terms in Δt , (B.1) can then be rewritten as

$$\Delta y_n = \ln \frac{f_{\Delta t,+}(\xi_n)}{f_{\Delta t,-}(\xi_n)} + \ln(1 + \Delta t(-\epsilon_+ + \epsilon_-e^{-y_{n-1}})) - \ln(1 + \Delta t(-\epsilon_- + \epsilon_+e^{y_{n-1}})).$$

Since we assumed $\Delta t \ll 1$, we can use the approximation $\ln(1 + a) \approx a$, which is valid to linear order in $|a| \ll 1$. We also assume that the change in the log-likelihood ratio, Δy_n , is small over the time interval Δt , so y_{n-1} can be replaced by y_n on the right-hand side of the equation. We obtain

$$(B.2) \quad \begin{aligned} \Delta y_n &\approx \ln \frac{f_{\Delta t,+}(\xi_n)}{f_{\Delta t,-}(\xi_n)} + \Delta t(\epsilon_-(e^{-y_n} + 1) - \epsilon_+(1 + e^{y_n})) \\ &= \mathbb{E}_\xi \left[\ln \frac{f_{\Delta t,+}(\xi_n)}{f_{\Delta t,-}(\xi_n)} \middle| H(t_n) \right] + \left(\ln \frac{f_{\Delta t,+}(\xi_n)}{f_{\Delta t,-}(\xi_n)} - \mathbb{E}_\xi \left[\ln \frac{f_{\Delta t,+}(\xi_n)}{f_{\Delta t,-}(\xi_n)} \middle| H(t_n) \right] \right) \\ &\quad + \Delta t(\epsilon_-(e^{-y_n} + 1) - \epsilon_+(1 + e^{y_n})), \end{aligned}$$

where we have conditioned on the state of the environment, $H(t_n) = H_{\pm}$ at time t_n . Replacing the index n with the time t , we can therefore write

$$(B.3) \quad \Delta y_t \approx \Delta t g_{\Delta t}(t) + \sqrt{\Delta t} \rho_{\Delta t}(t) \eta + \Delta t (\epsilon_- (e^{-y_t} + 1) - \epsilon_+ (1 + e^{y_t})),$$

where η is random variable with standard normal distribution, and

$$(B.4) \quad \begin{aligned} g_{\Delta t}(t) &:= \frac{1}{\Delta t} \mathbb{E}_{\xi} \left[\ln \frac{f_{\Delta t,+}(\xi)}{f_{\Delta t,-}(\xi)} \middle| H(t) \right], \\ \rho_{\Delta t}^2(t) &:= \frac{1}{\Delta t} \text{Var}_{\xi} \left[\ln \frac{f_{\Delta t,+}(\xi)}{f_{\Delta t,-}(\xi)} \middle| H(t) \right]. \end{aligned}$$

As before, $\mathbb{E}_{\xi}[F(\xi)|H(t)]$ is the expectation of $F(\xi)$ when ξ is drawn from the distribution $f_{\pm}(\xi)$ associated with the current state $H(t) = H_{\pm}$. Clearly, the drift $g_{\Delta t}$ and variance $\rho_{\Delta t}^2$ will diverge or vanish unless $f_{\Delta t,\pm}(\xi)$ are scaled appropriately in the $\Delta t \rightarrow 0$ limit. We discuss different ways of introducing such a scaling in the next section.

Assuming that we have well-defined limits $g(t) := \lim_{\Delta t \rightarrow 0} g_{\Delta t}(t)$ and $\rho^2(t) := \lim_{\Delta t \rightarrow 0} \rho_{\Delta t}^2(t)$, the discrete-time stochastic process, (B.3), approaches the SDE

$$(B.5) \quad dy = g(t)dt + \rho(t)dW_t + (\epsilon_- (e^{-y} + 1) - \epsilon_+ (1 + e^y))dt,$$

where W_t is a standard Wiener process. This limit holds in the sense of distributions. Roughly, the smaller Δt is, the closer the distributions of the random variables y_n and $y(t_n)$, whose evolutions are described by (B.1) and (B.5), respectively. This correspondence can be made precise using the Donsker invariance principle (p. 520 in [6]).

In sum, (B.5) can be viewed as an approximation of the logarithm of the likelihood ratio whose evolution is given exactly by (A.3). For a fixed interval Δt , the parameters of the two equations are related via (B.4), and $\epsilon_{\Delta t,\pm}/\Delta t = \epsilon_{\pm}$.

Appendix C. Precise Correspondence. We now discuss two approaches in which the correspondence between (B.1) and (B.5) can be made exact. We choose a specific scaling for the drift and variance arising from each observation, ξ_n . Suppose that over the time interval Δt , an observation, ξ_n , is a result of $r\Delta t$ separate observations—for example, the measurement of the direction of $r\Delta t$ different moving dots [24]. In this case the estimate of the average of the individual measurements—e.g., the average of the velocities of dots in a display—will have both a mean and a variance that increase linearly with Δt .

As a concrete example we can compute $g(t)$ and $\rho(t)$ in SDE (B.5) when observations, ξ_n , follow normal distributions with mean and variance scaled by Δt ,

$$f_{\Delta t,\pm}(\xi) = \frac{1}{\sqrt{2\pi\Delta t\sigma^2}} e^{-(\xi - \Delta t\mu_{\pm})^2/(2\Delta t\sigma^2)}.$$

Using (B.4), it is then straightforward to compute [7, 8]

$$\begin{aligned} g_{\Delta t}(t) &= \pm \frac{(\mu_+ - \mu_-)^2}{2\sigma^2} = g_{\pm}, \\ \rho_{\Delta t}^2(t) &= \frac{(\mu_+ - \mu_-)^2}{\sigma^2} = \rho^2, \end{aligned}$$

and note that $g(t) \in \{g_+, g_-\}$ is a telegraph process (e.g., p. 77 in [21]), with the

probability masses $P(g_{\pm}, t)$ evolving according to the master equation $P_t(g_{\pm}, t) = \mp \epsilon_+ P(g_+, t) \pm \epsilon_- P(g_-, t)$. In this case $\rho^2(t) = \rho^2$ remains constant.

More generally, we can obtain an identical result by considering that each observation made on a time interval consists of a number of subobservations, each with statistics that scale with the length of the interval and the number of subobservations. We define a family of stochastic processes parameterized by k , the number of subobservations made in an interval of length Δt . As above, when $k = 1$, we assume that an observation ξ_n is the result of $r\Delta t$ separate observations. Assuming r is large, note that for $k > 1$ each of the k subobservations contains roughly $r_k = \lfloor r\Delta t/k \rfloor$ observations with mean and variance that scale linearly with $r_k \propto \Delta t/k$. We can achieve this by approximating $\ln \frac{f_{\Delta t, +}(\xi_n)}{f_{\Delta t, -}(\xi_n)}$ in (B.2) by the family of stochastic processes parameterized by k [8]:

$$\sum_{l=1}^k \frac{\Delta t}{k} \ln \frac{f_+(\xi_l)}{f_-(\xi_l)} + \sum_{l=1}^k \frac{\sqrt{\Delta t}}{\sqrt{k}} \left(\ln \frac{f_+(\xi_l)}{f_-(\xi_l)} - E_{\xi} \left[\ln \frac{f_+(\xi)}{f_-(\xi)} \middle| H(t) \right] \right).$$

The scaling in this approximation guarantees that the drift is given by the limit

$$g(t) = \lim_{\Delta t \rightarrow 0} g_{\Delta t}(t) = \lim_{\Delta t \rightarrow 0} \frac{1}{\Delta t} E_{\xi} \left[\ln \frac{f_{\Delta t, +}(\xi)}{f_{\Delta t, -}(\xi)} \middle| H(t) \right] = E_{\xi} \left[\ln \frac{f_+(\xi)}{f_-(\xi)} \middle| H(t) \right]$$

and the variance

$$\rho^2(t) = \lim_{\Delta t \rightarrow 0} \rho_{\Delta t}^2 = \lim_{\Delta t \rightarrow 0} \frac{1}{\Delta t} \text{Var}_{\xi} \left[\ln \frac{f_{\Delta t, +}(\xi)}{f_{\Delta t, -}(\xi)} \middle| H(t) \right] = \text{Var}_{\xi} \left[\ln \frac{f_+(\xi)}{f_-(\xi)} \middle| H(t) \right].$$

Furthermore, as $k \rightarrow \infty$, by the central limit theorem,

$$\begin{aligned} \Delta y_t \approx & \sum_{l=1}^k \frac{\Delta t}{k} \ln \frac{f_+(\xi_l)}{f_-(\xi_l)} + \sum_{l=1}^k \frac{\sqrt{\Delta t}}{\sqrt{k}} \left(\ln \frac{f_+(\xi_l)}{f_-(\xi_l)} - E_{\xi} \left[\ln \frac{f_+(\xi)}{f_-(\xi)} \middle| H(t) \right] \right) \\ & + \Delta t (\epsilon_-(e^{-y_t} + 1) - \epsilon_+(1 + e^{y_t})) \end{aligned}$$

converges in distribution to

$$\Delta y_t \approx \Delta t g(t) + \sqrt{\Delta t} \rho(t) \eta + \Delta t (\epsilon_-(e^{-y_t} + 1) - \epsilon_+(1 + e^{y_t})),$$

where η is a standard normal random variable. Taking the limit $\Delta t \rightarrow 0$ yields (B.5). When observations follow Gaussian distributions, $f_{\pm} \sim \mathcal{N}(\pm\mu, \sigma^2)$, then $g(t) = \pm 2\mu^2/\sigma^2$, $\rho = 2\mu/\sigma$, and

$$dy = [g(t) + (\epsilon_-(e^{-y} + 1) - \epsilon_+(1 + e^y))] dt + \rho dW,$$

where dW is a standard white noise process.

Appendix D. Continuum Limit for Log Probabilities with Multiple Alternatives. We now describe the calculation of the continuum limit of the recursive system defining the evolution of the probabilities $L_{n,i} = \Pr(H(t_n) = H_i | \xi_{1:n})$ of one among multiple alternatives (environmental states), H_i , $i = 1, \dots, N$. The state of the environment and, equivalently, the correct choice at time t again change stochastically. We assume that the transitions between the alternatives are memoryless, with transition rates $\epsilon_{\Delta t, ij} := \Pr(H(t_n) = H_i | H(t_{n-1}) = H_j)$. Using Bayes' rule and rearranging

terms (analogous to the derivation of (A.1) and (A.2)), we can express each probability $L_{n,i}$ in terms the probability at the time of the previous observation, $L_{n-1,j}$:

$$L_{n,i} = \frac{\Pr(\xi_{1:n-1})}{\Pr(\xi_{1:n})} \Pr(\xi_n | H_i, t_n) \sum_{j=1}^N \epsilon_{\Delta t, ij} L_{n-1, j}.$$

Since we are only interested in comparing the magnitude of the probabilities, we can drop the common prefactor $\frac{\Pr(\xi_{1:n-1})}{\Pr(\xi_{1:n})}$ and use the fact that $\sum_{j=1}^N \epsilon_{\Delta t, ji} = 1$ (since $\epsilon_{\Delta t, ij}$ is a left stochastic matrix) to write $\epsilon_{\Delta t, ii} = 1 - \sum_{j \neq i} \epsilon_{\Delta t, ji}$ and obtain

$$(D.1) \quad L_{n,i} = f_{\Delta t, i}(\xi_n) \left(\left[1 - \sum_{j \neq i} \epsilon_{\Delta t, ji} \right] L_{n-1, i} + \sum_{j \neq i} \epsilon_{\Delta t, ij} L_{n-1, j} \right),$$

where $f_{\Delta t, i}(\xi_n) = \Pr(\xi_n | H_i, t_n)$. From (D.1), it follows that the log of the rescaled probabilities, $x_i := \ln L_i$, satisfies the recursive relation

$$x_{n,i} - x_{n-1,i} = \ln f_{\Delta t, i}(\xi_n) + \ln \left(1 - \sum_{j \neq i} \epsilon_{\Delta t, ji} + \sum_{j \neq i} \epsilon_{\Delta t, ij} e^{x_{n-1, j} - x_{n-1, i}} \right).$$

To derive an approximating SDE, we denote by $\Delta x_{n,i} = x_{n,i} - x_{n-1,i}$ the change in the log probability due to an observation at time t_n . As before, we assume $\epsilon_{\Delta t, ij} := \Delta t \epsilon_{ij} + o(\Delta t)$ for $i \neq j$ and drop the higher order terms, giving

$$\Delta x_{n,i} = \ln f_{\Delta t, i}(\xi_n) + \ln \left(1 - \sum_{j \neq i} \Delta t \epsilon_{ji} + \sum_{j \neq i} \Delta t \epsilon_{ij} e^{x_{n-1, j} - x_{n-1, i}} \right).$$

Assuming $\Delta t \ll 1$, we again use the approximation $\ln(1 + a) \approx a$ for $|a| \ll 1$. We also assume that the change in the log probability, $|\Delta x_{n,i}| \ll 1$, is small over the time interval, Δt so that

$$(D.2) \quad \begin{aligned} \Delta x_{n,i} &\approx \ln f_{\Delta t, i}(\xi_n) + \Delta t \sum_{j \neq i} (\epsilon_{ij} e^{x_{n-1, j} - x_{n-1, i}} - \epsilon_{ji}) \\ &= \mathbb{E}_\xi [\ln f_{\Delta t, i}(\xi) | H(t_n)] + (\ln f_{\Delta t, i}(\xi_n) - \mathbb{E}_\xi [\ln f_{\Delta t, i}(\xi) | H(t_n)]) \\ &\quad + \Delta t \sum_{j \neq i} (\epsilon_{ij} e^{x_{n-1, j} - x_{n-1, i}} - \epsilon_{ji}), \end{aligned}$$

where we condition on the current state of the environment $H(t_n) \in \{H_1, \dots, H_N\}$.

Replacing the index n by the time t , we can therefore write

$$(D.3) \quad \Delta x_{t,i} \approx \Delta t g_{\Delta t, i}(t) + \sqrt{\Delta t} \rho_{\Delta t, i}(t) \eta_i + \Delta t \sum_{j \neq i} (\epsilon_{ij} e^{x_{t, j} - x_{t, i}} - \epsilon_{ji}),$$

where η_i 's are correlated random variables with standard normal distribution

$$\begin{aligned} g_{\Delta t, i}(t) &:= \frac{1}{\Delta t} \mathbb{E}_\xi [\ln f_{\Delta t, i}(\xi) | H(t)], \\ \rho_{\Delta t, i}^2(t) &:= \frac{1}{\Delta t} \text{Var}_\xi [\ln f_{\Delta t, i}(\xi) | H(t)]. \end{aligned}$$

The correlation of η_i 's is given by

$$\text{Corr}_\xi[\eta_i, \eta_j] := \text{Corr}_\xi[\ln f_{\Delta t, i}(\xi), \ln f_{\Delta t, j}(\xi) | H(t)].$$

Note that (D.3) is the multiple-alternative version of (B.3). Equivalently, we can write (D.3) as

$$\Delta x_{t, i} \approx \Delta t g_{\Delta t, i}(t) + \sqrt{\Delta t} \widehat{W}_{\Delta t, i} + \Delta t \sum_{j \neq i} (\epsilon_{ij} e^{x_{t, j} - x_{t, i}} - \epsilon_{ji}),$$

where $\widehat{W}_{\Delta t} := (\widehat{W}_{\Delta t, 1}, \dots, \widehat{W}_{\Delta t, N})$ follows a multivariate Gaussian distribution with mean zero and covariance matrix $\Sigma_{\Delta t}$ given by

$$\Sigma_{\Delta t, ij} = \frac{1}{\Delta t} \text{Cov}_\xi[\ln f_{\Delta t, i}(\xi), \ln f_{\Delta t, j}(\xi) | H(t)].$$

Finally, taking the limit $\Delta t \rightarrow 0$, and assuming that the limits

$$(D.4) \quad g_i(t) := \lim_{\Delta t \rightarrow 0} g_{\Delta t, i}(t) \quad \text{and} \quad \Sigma_{ij}(t) := \lim_{\Delta t \rightarrow 0} \Sigma_{\Delta t, ij}(t)$$

are well defined, we obtain the system of SDEs

$$(D.5) \quad dx_i = g_i(t)dt + d\widehat{W}_i(t) + \sum_{j \neq i} (\epsilon_{ij} e^{x_j - x_i} - \epsilon_{ji}) dt,$$

or equivalently the vector system

$$d\mathbf{x} = \mathbf{g}(t)dt + \Lambda(t)d\mathbf{W}_t + K(\mathbf{x})dt,$$

where $\mathbf{g}(t) = (g_1(t), \dots, g_N(t))^T$ and $\Lambda(t)\Lambda(t)^T = \Sigma(t)$ are defined using the limits in (D.4), $K_i(\mathbf{x}) = \sum_{j \neq i} (\epsilon_{ij} e^{x_j - x_i} - \epsilon_{ji})$, and the components of \mathbf{W}_t are independent Wiener processes. We can recover (B.5) by taking $N = 2$, letting $y = x_1 - x_2$, and exchanging the indices 1 and 2 with + and -, respectively.

As in the case of two alternatives, (D.5) can be viewed as an approximation of the logarithm of the probability whose evolution is given exactly by (D.1). For a fixed interval Δt , the parameters of these equations are related via (D.5), and $\epsilon_{\Delta t, ij} / \Delta t = \epsilon_{ij}$.

The limits $g_i(t) := \lim_{\Delta t \rightarrow 0} g_{\Delta t, i}(t)$ and $\Sigma_{ij}(t) := \lim_{\Delta t \rightarrow 0} \Sigma_{\Delta t, ij}(t)$ are defined when the statistics of the observations scale with Δt . As we argued above, this can be obtained by considering observations drawn from a normal distribution with mean and variance scaled by Δt :

$$f_{\Delta t, i}(\xi) = \frac{1}{\sqrt{2\pi\Delta t\sigma^2}} e^{-(\xi - \Delta t\mu_i)^2 / (2\Delta t\sigma^2)}.$$

Alternatively, the required scaling can also be obtained when each observation made on a time interval consists of a number of subobservations, (ξ_1, \dots, ξ_k) , with mean and variance scaled by $\frac{\Delta t}{k}$. To do so we approximate $\ln f_{\Delta t, i}(\xi_n)$ in (D.2) by

$$\sum_{l=1}^k \frac{\Delta t}{k} \ln f_i(\xi_l) + \sum_{l=1}^k \frac{\sqrt{\Delta t}}{\sqrt{k}} (\ln f_i(\xi_l) - E_\xi[\ln f_i(\xi) | H(t)]).$$

Appendix E. Log-Likelihood Ratio for Multiple Alternatives. We can also derive a continuum limit for the log-likelihood ratio for any two choices $i, j \in \{1, 2, \dots, N\}$. From (D.1), the likelihood ratio $R_{n,ij} = L_{n,i}/L_{n,j}$. We note that this will provide us with a matrix of stochastic processes. We start with the recursive equation

$$R_{n,ij} = \frac{f_{\Delta t,i}(\xi_n) \left(1 - \sum_{k \neq i} \epsilon_{\Delta t,ki}\right) R_{n-1,ij} + \sum_{k \neq i} \epsilon_{\Delta t,ik} R_{n-1,kj}}{f_{\Delta t,j}(\xi_n) \left(1 - \sum_{k \neq j} \epsilon_{\Delta t,kj} + \sum_{k \neq j} \epsilon_{\Delta t,jk} R_{n-1,kj}\right)}.$$

We can thus derive the continuum equation for the log-likelihood ratio $y_{n,ij} := \ln R_{n,ij}$, as we did in the case of two alternatives. Since $y_{ij}(t)$ is the difference $y_{ij}(t) = x_i(t) - x_j(t)$, from (D.5) we obtain

$$dy_{ij} = (g_i(t) - g_j(t))dt + d\widehat{W}_i(t) - d\widehat{W}_j(t) + \sum_{k \neq i} (\epsilon_{ik} e^{y_{ki}} - \epsilon_{ki}) dt - \sum_{k \neq j} (\epsilon_{jk} e^{y_{kj}} - \epsilon_{kj}) dt,$$

or

$$(E.1) \quad dy_{ij} = \left[g_{ij}(t) + \left(\sum_{k \neq j} \epsilon_{kj} - \sum_{k \neq i} \epsilon_{ki} + \sum_{k \neq i} \epsilon_{ik} e^{y_{ki}} - \sum_{k \neq j} \epsilon_{jk} e^{y_{kj}} \right) \right] dt + d\widehat{W}_{ij},$$

where $g_{ij}(t) = E_\xi \left[\ln \frac{f_i(\xi)}{f_j(\xi)} | H(t) \right]$ and \widehat{W} is a Wiener process with covariance matrix given by $\text{Cov}_\xi [\widehat{W}_{ij}, \widehat{W}_{i'j'} | H(t)] = \text{Cov}_\xi \left[\ln \frac{f_i(\xi)}{f_j(\xi)}, \ln \frac{f_{i'}(\xi)}{f_{j'}(\xi)} | H(t) \right]$. We can also write (E.1) in vector form,

$$dy = gdt + \Lambda(t)d\mathbf{W}_t + \mathbf{K}(y)dt,$$

where $\mathbf{K}_{ij}(y) = \sum_{k \neq j} \epsilon_{kj} - \sum_{k \neq i} \epsilon_{ki} + \sum_{k \neq i} \epsilon_{ik} e^{y_{ki}} - \sum_{k \neq j} \epsilon_{jk} e^{y_{kj}}$, $\Lambda(t)\Lambda(t)^T = \Sigma(t)$ is the covariance matrix, and the components of \mathbf{W}_t are independent Wiener processes.

Appendix F. Log Probabilities for a Continuum of Alternatives. Finally, we examine the case where an observer must choose between a continuum of hypotheses H_θ , where $\theta \in [a, b]$. Thus, we will first derive a discrete recursive equation for the evolution of the probabilities $L_{n,\theta} = \Pr(H(t_n) = H_\theta | \xi_{1:n})$. The state of the environment, the correct choice at time t , again changes according to a continuous time Markov process. We define this stochastically switching process through its transition rate function $\epsilon_{\Delta t, \theta\theta'}$, which is given for $\theta' \neq \theta$ as

$$\int_{\theta_1}^{\theta_2} \epsilon_{\Delta t, \theta\theta'} d\theta := \Pr(H(t_n) \in H_{[\theta_1, \theta_2]} | H(t_{n-1}) = H_{\theta'}),$$

where $H_{[\theta_1, \theta_2]}$ is the set of all states H_θ with θ in the interval $[\theta_1, \theta_2]$. Thus, $\epsilon_{\Delta t, \theta\theta'}$ describes the probability of a transition over a timestep, Δt , from state $H_{\theta'}$ to some state H_θ , with $\theta \in [\theta_1, \theta_2]$. This means that $\Pr(H(t_n) = H_\theta | H(t_{n-1}) = H_\theta) = 1 - \int_a^b \epsilon_{\Delta t, \theta\theta'} d\theta'$. As in the derivation of the multiple alternative $2 \leq N < \infty$ case, we can express each probability $L_{n,\theta}$ at time t_n in terms of the probabilities $L_{n-1,\theta'}$ at time t_{n-1} , so

$$L_{n,\theta} = \frac{\Pr(\xi_{1:n-1})}{\Pr(\xi_{1:n})} \Pr(\xi_n | H(t_n) = H_\theta) \times \left(\Pr(H(t_n) = H_\theta | H(t_{n-1}) = H_\theta) L_{n-1,\theta} + \int_a^b \epsilon_{\Delta t, \theta\theta'} L_{n-1,\theta'} d\theta' \right).$$

Notice that the sum from the $N < \infty$ case, as in (D.1), has been replaced with an integral over all possible hypotheses $H_{\theta'}$, $\theta' \in [a, b]$, and a term corresponding to the probability of the environment not changing. Again we drop the common factor $\frac{\Pr(\xi_{1:n-1})}{\Pr(\xi_{1:n})}$, since we wish to compare the magnitudes of the probabilities. We obtain

$$(F.1) \quad L_{n,\theta} = f_{\Delta t,\theta}(\xi_n) \left(\left[1 - \int_a^b \epsilon_{\Delta t,\theta'\theta} d\theta' \right] L_{n-1,\theta} + \int_a^b \epsilon_{\Delta t,\theta\theta'} L_{n-1,\theta'} d\theta' \right),$$

where $f_{\Delta t,\theta}(\xi_n) = \Pr(\xi_n | H(t_n) = H_\theta)$. From (F.1), we can thus derive a recursive relation for the log of the rescaled probabilities $x_{n,\theta} := \ln L_{n,\theta}$ in terms of $x_{n-1,\theta}$, so

$$x_{n,\theta} - x_{n-1,\theta} = \ln f_{\Delta t,\theta}(\xi_n) + \ln \left(1 - \int_a^b \epsilon_{\Delta t,\theta'\theta} d\theta' + \int_a^b \epsilon_{\Delta t,\theta\theta'} e^{x_{n-1,\theta'} - x_{n-1,\theta}} d\theta' \right).$$

To approximate this discrete-time stochastic process with an SDE, we denote by $\Delta x_{n,\theta} = x_{n,\theta} - x_{n-1,\theta}$, the change in log probability due to the observation at time t_n . Furthermore, we assume $\epsilon_{\Delta t,\theta\theta'} := \Delta t \epsilon_{\theta\theta'} + o(\Delta t)$ and drop higher order terms,

$$\Delta x_{n,\theta} = \ln f_{\Delta t,\theta}(\xi_n) + \ln \left(1 - \int_a^b \Delta t \epsilon_{\theta'\theta} d\theta' + \int_a^b \Delta t \epsilon_{\theta\theta'} e^{x_{n-1,\theta'} - x_{n-1,\theta}} d\theta' \right).$$

Assuming $\Delta t \ll 1$, we again use the approximation $\ln(1+a) \approx a$ for $|a| \ll 1$. Assuming $|\Delta x_{n,\theta}| \ll 1$,

$$\begin{aligned} \Delta x_{n,\theta} &\approx \ln f_{\Delta t,\theta}(\xi_n) + \Delta t \int_a^b (\epsilon_{\theta\theta'} e^{x_{n-1,\theta'} - x_{n-1,\theta}} - \epsilon_{\theta'\theta}) d\theta' \\ &= \mathbb{E}_\xi [f_{\Delta t,\theta}(\xi) | H(t_n)] + (\ln f_{\Delta t,\theta}(\xi_n) - \mathbb{E}_\xi [\ln f_{\Delta t,\theta}(\xi) | H(t_n)]) \\ &\quad + \Delta t \int_a^b (\epsilon_{\theta\theta'} e^{x_{n-1,\theta'} - x_{n-1,\theta}} - \epsilon_{\theta'\theta}) d\theta', \end{aligned}$$

conditioned on the current state of the environment $H(t_n) = H_\varphi$, where $\varphi \in [a, b]$.

Exchanging the index n with the time, t , we can therefore write

$$(F.2) \quad \Delta x_{t,\theta} \approx \Delta t g_{\Delta t,\theta}(t) + \sqrt{\Delta t} \rho_{\Delta t,\theta}(t) \eta_\theta + \Delta t \int_a^b (\epsilon_{\theta\theta'} e^{x_{t,\theta'} - x_{t,\theta}} - \epsilon_{\theta'\theta}) d\theta',$$

where η_θ 's are correlated random variables which marginally follow a standard normal distribution, and

$$\begin{aligned} g_{\Delta t,\theta}(t) &:= \frac{1}{\Delta t} \mathbb{E}_\xi [\ln f_{\Delta t,\theta}(\xi) | H(t)], \\ \rho_{\Delta t,\theta}^2(t) &:= \frac{1}{\Delta t} \text{Var}_\xi [\ln f_{\Delta t,\theta}(\xi) | H(t)]. \end{aligned}$$

The correlation of η_i 's is given by

$$\text{Corr}_\xi [\eta_\theta, \eta_{\theta'}] := \text{Corr}_\xi [\ln f_{\Delta t,\theta}(\xi), \ln f_{\Delta t,\theta'}(\xi) | H(t)].$$

Equivalently, we can write (F.2) as

$$\Delta x_{t,\theta} \approx \Delta t g_{\Delta t,\theta}(t) + \sqrt{\Delta t} \widehat{W}_{\Delta t,\theta} + \Delta t \int_a^b (\epsilon_{\theta\theta'} e^{x_{t,\theta'} - x_{t,\theta}} - \epsilon_{\theta'\theta}) d\theta',$$

where $\widehat{W}_{\Delta t} := (\widehat{W}_{\Delta t, \theta})_{\theta \in [a, b]}$. For $\theta \in [a, b]$, $\widehat{W}_{\Delta t, \theta}$ is a Gaussian process in the sense that any finite subset of points $\{\theta_1, \dots, \theta_n\} \in [a, b]$ have a multivariate Gaussian distribution with mean zero and covariance, $\Sigma_{\Delta t, \theta\theta}$, given by

$$\Sigma_{\Delta t, \theta\theta'} = \frac{1}{\Delta t} \text{Cov}_{\xi} [\ln f_{\Delta t, \theta}(\xi), \ln f_{\Delta t, \theta'}(\xi) | H(t)].$$

Finally, taking the limit $\Delta t \rightarrow 0$, and assuming that the limits

$$(F.3) \quad g_{\theta}(t) := \lim_{\Delta t \rightarrow 0} g_{\Delta t, \theta}(t) \quad \text{and} \quad \Sigma_{\theta\theta'}(t) := \lim_{\Delta t \rightarrow 0} \Sigma_{\Delta t, \theta\theta'}(t)$$

are well defined, we obtain the system of SDEs

$$(F.4) \quad dx_{\theta} = g_{\theta}(t)dt + d\widehat{W}_{\theta}(t) + \int_a^b (\epsilon_{\theta\theta'} e^{x_{\theta'} - x_{\theta}} - \epsilon_{\theta'\theta}) d\theta' dt,$$

or equivalently the system of SDEs

$$d\mathbf{x} = \mathbf{g}(t)dt + \Lambda(t)d\mathbf{W}_t + K(\mathbf{x})dt,$$

where $\mathbf{g}(t) = (g_{\theta}(t))_{\theta \in [a, b]}$ and $\Lambda(t)\Lambda(t)^T = \Sigma(t)$ are defined using the limits in (F.3), $K(\mathbf{x}) = \int_a^b (\epsilon_{\theta\theta'} e^{x_{\theta'} - x_{\theta}} - \epsilon_{\theta'\theta}) d\theta'$, and the components of \mathbf{W}_t are independent Wiener processes.

While we have formally taken the limit of the discrete (F.2), it is important to note that establishing the well-posedness of stochastic integrodifferential equations is not straightforward. Conditions for the existence and uniqueness of solutions to certain nonlinear stochastic partial differential equations (SPDEs) are demonstrated in Chapter 7 of [15]. This approach considers the solutions to SPDEs to be random processes that take their values in a Hilbert space of functions. Recently, this concept has been extended to provide general conditions on the constituent functions of stochastic neural fields to ensure the existence of solutions [19, 25]. The form of stochastic neural fields is closely related to (F.4), since both types of equation possess a linear drift and a convolution defining a nonlocal coupling between their state variables. It may be possible to utilize these previous approaches to establish the existence and uniqueness of solutions to (F.4) in future studies.

Acknowledgments. We thank Eric Shea-Brown, Shawn Gu, and the anonymous reviewers for helpful comments.

REFERENCES

- [1] R. P. ADAMS AND D. J. MACKEY, *Bayesian Online Changepoint Detection*, arXiv preprint, arXiv:0710.3742, 2007.
- [2] P. ARMITAGE, *Sequential analysis with more than two alternative hypotheses, and its relation to discriminant function analysis*, J. R. Statist. Soc. Ser. B, 12 (1950), pp. 137–144.
- [3] G. BALÁZSI, A. VAN OUDENAARDEN, AND J. J. COLLINS, *Cellular decision making and biological noise: From microbes to mammals*, Cell, 144 (2011), pp. 910–925.
- [4] 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 (2008), pp. 1142–1152.
- [5] T. E. BEHRENS, M. W. WOOLRICH, M. E. WALTON, AND M. F. RUSHWORTH, *Learning the value of information in an uncertain world*, Nat. Neurosci., 10 (2007), pp. 1214–1221.
- [6] P. BILLINGSLEY, *Probability and Measure*, Wiley, New York, 1995.

- [7] S. BITZER, H. PARK, F. BLANKENBURG, AND S. J. KIEBEL, *Perceptual decision making: Drift-diffusion model is equivalent to a Bayesian model*, *Front. Hum. Neurosci.*, 8 (2014), article 102.
- [8] 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*, *Psychol. Rev.*, 113 (2006), pp. 700–765.
- [9] R. BOGACZ AND K. GURNEY, *The basal ganglia and cortex implement optimal decision making between alternative actions*, *Neural Comput.*, 19 (2007), pp. 442–477.
- [10] R. BOGACZ, M. USHER, J. ZHANG, AND J. L. MCCLELLAND, *Extending a biologically inspired model of choice: Multi-alternatives, nonlinearity and value-based multidimensional choice*, *Philos. Trans. R. Soc. Lond. B*, 362 (2007), pp. 1655–1670.
- [11] B. W. BRUNTON, M. M. BOTVINICK, AND C. D. BRODY, *Rats and humans can optimally accumulate evidence for decision-making*, *Science*, 340 (2013), pp. 95–98.
- [12] J. R. BUSEMEYER AND J. T. TOWNSEND, *Decision field theory: A dynamic-cognitive approach to decision making in an uncertain environment*, *Psychol. Rev.*, 100 (1993), pp. 432–59.
- [13] I. D. COUZIN, J. KRAUSE, N. R. FRANKS, AND S. A. LEVIN, *Effective leadership and decision-making in animal groups on the move*, *Nature*, 433 (2005), pp. 513–516.
- [14] T. M. COVER AND J. A. THOMAS, *Elements of Information Theory*, John Wiley & Sons, 2012.
- [15] G. DA PRATO AND J. ZABCZYK, *Stochastic Equations in Infinite Dimensions*, 2nd ed., *Encyclopedia Math. Appl.* 152, Cambridge University Press, Cambridge, UK, 2014.
- [16] S. DENEVE, *Bayesian spiking neurons I: Inference*, *Neural Comput.*, 20 (2008), pp. 91–117.
- [17] P. DOMENICI AND R. W. BLAKE, *The kinematics and performance of fish fast-start swimming*, *J. Exp. Biol.*, 200 (1997), pp. 1165–1178.
- [18] V. DRAGLIA, A. G. TARTAKOVSKY, AND V. V. VEERAVALLI, *Multihypothesis sequential probability ratio tests. I. Asymptotic optimality*, *IEEE Trans. Inform. Theory*, 45 (1999), pp. 2448–2461.
- [19] O. FAUGERAS AND J. INGLIS, *Stochastic neural field equations: A rigorous footing*, *J. Math. Biol.*, (2013), pp. 1–42.
- [20] S. FENG, P. HOLMES, A. RORIE, AND W. T. NEWSOME, *Can monkeys choose optimally when faced with noisy stimuli and unequal rewards?*, *PLoS Comput. Biol.*, 5 (2009), article e1000284.
- [21] C. W. GARDINER, *Handbook of Stochastic Methods for Physics, Chemistry, and the Natural Sciences*, 3rd ed., Springer-Verlag, Berlin, 2004.
- [22] C. GLAZE, J. KABLE, AND J. GOLD, *Normative Evidence Accumulation in an Unpredictable Environment*, Poster at Computational and Systems Neuroscience (CoSyne 2015), Salt Lake City, UT, 2015.
- [23] J. I. GOLD AND M. N. SHADLEN, *Banburismus and the brain: Decoding the relationship between sensory stimuli, decisions, and reward*, *Neuron*, 36 (2002), pp. 299–308.
- [24] J. I. GOLD AND M. N. SHADLEN, *The neural basis of decision making*, *Ann. Rev. Neurosci.*, 30 (2007), pp. 535–74.
- [25] C. KUEHN AND M. RIEDLER, *Large deviations for nonlocal stochastic neural fields*, *J. Math. Neurosci.*, 4 (2014), pp. 1–33.
- [26] K. W. LATIMER, J. L. YATES, M. L. R. MEISTER, A. C. HUK, AND J. W. PILLOW, *Single-trial spike trains in parietal cortex reveal discrete steps during decision-making*, *Science*, 349 (2015), pp. 184–187.
- [27] C. K. MACHENS, R. ROMO, AND C. D. BRODY, *Flexible control of mutual inhibition: A neural model of two-interval discrimination*, *Science*, 307 (2005), pp. 1121–1124.
- [28] J. A. R. MARSHALL, R. BOGACZ, A. DORNHAUS, R. PLANQUÉ, T. KOVACS, AND N. R. FRANKS, *On optimal decision-making in brains and social insect colonies.*, *J. R. Soc. Interface*, 6 (2009), pp. 1065–1074.
- [29] J. T. MCGUIRE, M. R. NASSAR, J. I. GOLD, AND J. W. KABLE, *Functionally dissociable influences on learning rate in a dynamic environment*, *Neuron*, 84 (2014), pp. 870–881.
- [30] T. McMILLEN AND P. HOLMES, *The dynamics of choice among multiple alternatives*, *J. Math. Psychol.*, 50 (2006), pp. 30–57.
- [31] M. R. NASSAR, K. M. RUMSEY, R. C. WILSON, K. PARIKH, B. HEASLY, AND J. I. GOLD, *Rational regulation of learning dynamics by pupil-linked arousal systems*, *Nat. Neurosci.*, 15 (2012), pp. 1040–1046.
- [32] J. NEYMAN AND E. PEARSON, *On the problem of the most efficient tests of statistical hypotheses*, *Philos. Trans. R. Soc. Lond. A*, 231 (1933), pp. 289–337.
- [33] D. PAIS, P. M. HOGAN, T. SCHLEGEL, N. R. FRANKS, AND N. E. LEONARD, *A mechanism for value-sensitive decision-making*, *PLoS ONE*, 8 (2013), article e73216.
- [34] K. M. PASSINO, T. D. SEELEY, AND P. K. VISSCHER, *Swarm cognition in honey bees*, *Behav. Ecol. Sociobiol.*, 62 (2008), pp. 401–414.

- [35] R. RATCLIFF, *A theory of memory retrieval*, Psychol. Rev., 85 (1978), pp. 59–108.
- [36] R. RATCLIFF AND P. L. SMITH, *A comparison of sequential sampling models for two-choice reaction time*, Psychol. Rev., 111 (2004), pp. 333–367.
- [37] Y. A. ROZANOV, *Probability Theory: A Concise Course*, Dover, New York, 1977.
- [38] T. D. SEELEY, S. CAMAZINE, AND J. SNEYD, *Collective decision-making in honey bees: How colonies choose among nectar sources*, Behav. Ecol. Sociobiol., 28 (1991), pp. 277–290.
- [39] A. N. SHIRYAEV, *Optimal Stopping Rules*, Springer, Berlin, 2007.
- [40] M. SHVARTSMAN, V. SRIVASTAVA, AND J. COHEN, *A theory of decision making under dynamic context*, Adv. Neural Inf. Process. Syst., 28 (2015), pp. 2476–2484.
- [41] P. L. SMITH AND R. RATCLIFF, *Psychology and neurobiology of simple decisions*, Trends Neurosci., 27 (2004), pp. 161–168.
- [42] L. P. SUGRUE, G. S. CORRADO, AND W. T. NEWSOME, *Matching behavior and the representation of value in the parietal cortex*, Science, 304 (2004), pp. 1782–1787.
- [43] M. USHER AND J. L. MCCLELLAND, *The time course of perceptual choice: The leaky, competing accumulator model*, Psychol. Rev., 108 (2001), pp. 550–592.
- [44] M. USHER AND J. L. MCCLELLAND, *Loss aversion and inhibition in dynamical models of multi-alternative choice*, Psychol. Rev., 111 (2004), pp. 757–769.
- [45] A. WALD AND J. WOLFOWITZ, *Optimum character of the sequential probability ratio test*, Ann. Math. Statist., 19 (1948), pp. 326–339.
- [46] X.-J. WANG, *Probabilistic decision making by slow reverberation in cortical circuits*, Neuron, 36 (2002), pp. 955–968.
- [47] R. C. WILSON, M. R. NASSAR, AND J. I. GOLD, *Bayesian online learning of the hazard rate in change-point problems*, Neural Comput., 22 (2010), pp. 2452–76.
- [48] K.-F. WONG, A. C. HUK, M. N. SHADLEN, AND X.-J. WANG, *Neural circuit dynamics underlying accumulation of time-varying evidence during perceptual decision making*, Front. Comput. Neurosci., 1 (2007), article 6.
- [49] A. J. YU, *Optimal change-detection and spiking neurons*, in Advances in Neural Information Processing Systems 19 (NIPS 2006), 2007, pp. 1545–1552.
- [50] Q. ZHAO AND J. YE, *Quickest detection in multiple on-off processes*, IEEE Trans. Signal Process., 58 (2010), pp. 5994–6006.