Stochastic models of evidence accumulation in changing environments

Alan Veliz-Cuba* and Zachary P. Kilpatrick❃†

Department of Mathematics, University of Houston, Houston, TX 77204

Krešimir Josić❃‡

Department of Mathematics, University of Houston, Houston, TX 77204 and
Department of Biology and Biochemistry, University of Houston, Houston, TX 77204

(Dated: June 12, 2015)

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, the natural world constantly changes. Using sequential analysis we 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 that 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.

Popular Summary. 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. How to make decisions in face of uncertainty and impermanence is a question that recurs in fields ranging from economics to ecology and neuroscience. Here, we explore this problem in a general setting where an observer evaluates multiple options based on a series of noisy observations. We assume that the best option is not fixed in time. The optimal strategy is therefore to sequentially update the probability of each alternative, weighting recent evidence more strongly. Our work builds a bridge between statistical decision making in volatile environments and stochastic nonlinear dynamics.

I. INTRODUCTION

Mammals [1–4], insects [5, 6], and single cells [7] gather 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 both theoretically and experimentally [2, 8–10]. In this case humans [11], and other mammals [3, 4] 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 [12, 13]. These models are analytically tractable [2], and can implement optimal decision strategies [14]. 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 [3].

However, 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 [15, 16]. Recent experimental evidence suggests that human observers integrate noisy measurements near optimally even when the underlying truth 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 [17]. An important feature of evidence accumulation in volatile environments is an increase in learning rate when recent observations do not support a current estimate [18]. Both behavioral and fMRI data show that human subjects employ this strategy when they must predict the position of a stochastically moving target [19]. Experimental work thus suggests that humans adjust evidence valuation to account for environmental variability.

Here we show that optimal stochastic accumulator models can be extended 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. This model also suggests a biophysical neural implementation

───────

* alanavc@math.uh.edu
† zpkilpat@math.uh.edu
‡ josic@math.uh.edu
❃ equal contribution

for evidence integrators consisting of neural populations whose activity represents the evidence in favor of a particular choice. Surprisingly, when the environment is not static, these populations are coupled through excitation. This is in contrast to optimal integrators in a static environment [14, 20], and other non-optimal integrators [21, 22] where the populations representing different choices inhibit each other.

II. 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 [1, 9]: An observer makes a stream of independent, noisy measurements, $\xi_{1:n} = (\xi_1, \xi_2, ..., \xi_n)$, at equally spaced times, $t_{1:n} = (t_1, t_2, ..., t_n)$. The probability of each measurement, $f_+(\xi_j) := \Pr(\xi_j|H_+)$, and $f_-(\xi_j) := \Pr(\xi_j|H_-)$, depend on the environmental state. Combined with the prior probability, $\Pr(H_\pm)$, of the states, this gives the ratio of probabilities,

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

which can also be written recursively [9]:

$$R_n = \left(\frac{f_+(\xi_n)}{f_-(\xi_n)}\right)\cdot R_{n-1}, \quad \text{with} \quad R_0 = \frac{\Pr(H_+)}{\Pr(H_-)}. \quad (1)$$

With a fixed number of observations, this ratio can be used to make a choice that minimizes the total error rate [8], or maximizes reward [10]. Eq. (1) gives a recursive relation for the log likelihood ratio, $y_n = \ln R_n$,

$$y_n = y_{n-1} + \ln\frac{f_+(\xi_n)}{f_-(\xi_n)}. \quad (2)$$

When the time between observations, $\Delta t = t_j - t_{j-1}$, is small, we can use the Functional Central Limit Theorem [23] to approximate this stochastic process by the stochastic differential equation (SDE) [2, 24],

$$dy = g_\pm dt + \rho_\pm dW_t, \quad (3)$$

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 process, like Eq. (2), with SDEs. Details of these derivations are provided in the Appendix.

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
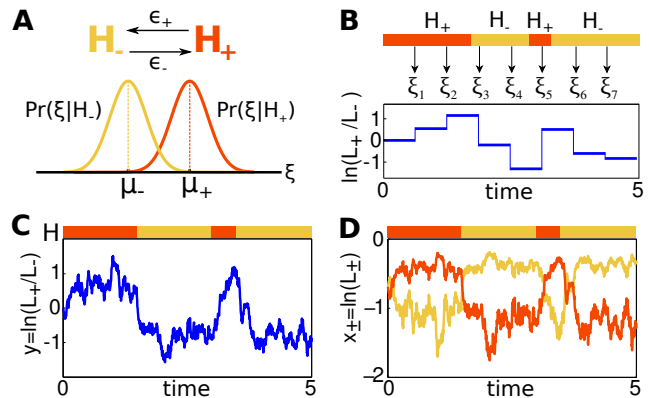

FIG. 1. Evidence accumulation in a changing environment. (**A**) The environmental state transitions from state H_+ to H_- and back with rates ϵ_+, and ϵ_-, respectively. Observations follow state dependent probabilities, $f_\pm(\xi) = \Pr(\xi|H_\pm)$. (**B**) The distributions of the measurements, ξ_j, 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 $x_+(t) > x_-(t)$.

Kullback–Leibler divergence between f_+ and f_-, i.e. the strength of the observed evidence from a measurement in favor of H_+. 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.

III. TWO ALTERNATIVES IN A CHANGING ENVIRONMENT

We use the same assumptions to derive a recursive equation for the log likelihood ratio between to alternatives in a changing environment. The state of the environment, $H(t)$, is H_+ or H_-, but can now change in time. An observer infers the present state from a sequence of observations, $\xi_{1:n}$, made at equally spaced times, $t_{1:n}$, and characterized by probabilities $f_\pm(\xi_n) := \Pr(\xi_n|H_\pm)$. The state of the environment changes according to a telegraph process [25], 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):

$$L_{n,\pm} \propto f_\pm(\xi_n)\left((1 - \Delta t\epsilon_\pm)L_{n-1,\pm} + \Delta t\epsilon_\mp L_{n-1,\mp}\right), \quad (4)$$

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 re-
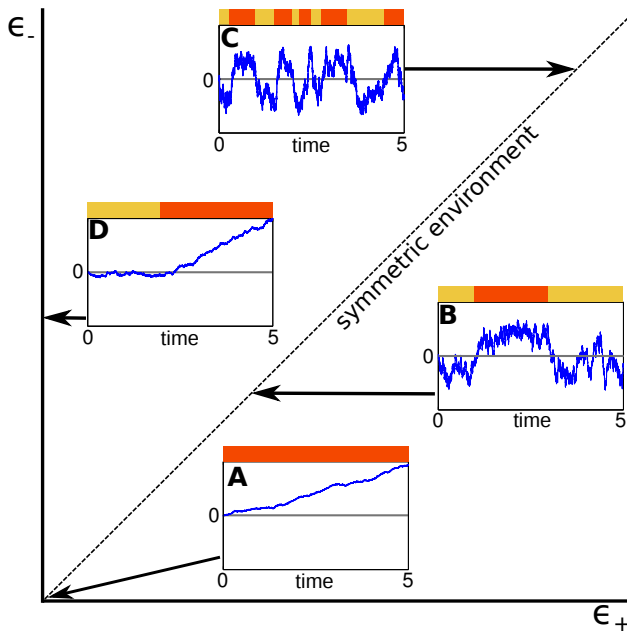
FIG. 2. In a dynamic environment, the dynamics of the log likelihood ratio, y, depends on the rates of switching between states. (**A**) When $\epsilon_\pm = 0$, the environment is static, and the model reduces to the one derived in Section II (**B**) When then environment changes slowly, $|\epsilon_\pm| \ll 1$, 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.
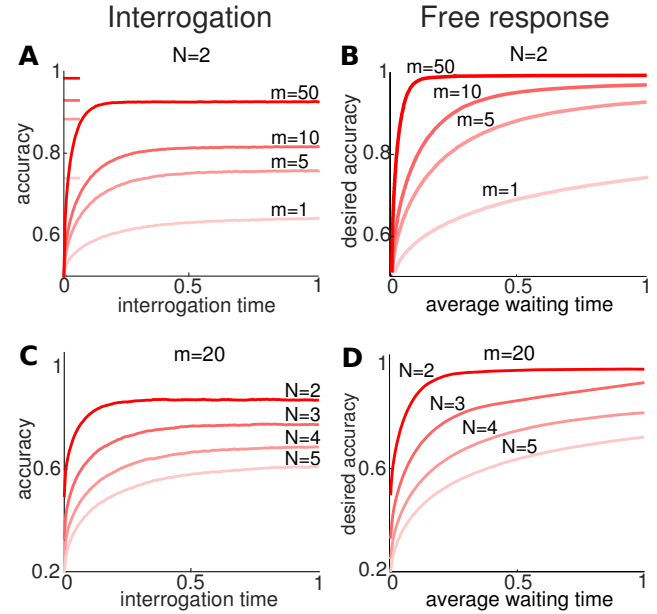

FIG. 3. Dependence of the probability of the correct response (accuracy) on normalized information gain, m, in a symmetric environment. (**A**) Accuracy in an interrogation protocol increases with m and interrogation time, t, but saturates. Horizontal bars on left indicate the accuracy when the environment is in a single state for a long time, as in Eq. (10). (**B**) When the observer responds freely accuracy is similar, but saturates at 1. The increase in accuracy in 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, saturating at ever lower levels. (**D**) The free response protocol results in similar behavior, but the accuracy saturates at 1. The increase in accuracy in time is exceedingly slow for higher numbers of alternatives N.

cursively (See Appendix A), and equals

$$R_n = \frac{L_{n,+}}{L_{n,-}} = \frac{f_+(\xi_n)}{f_-(\xi_n)} \frac{(1 - \Delta t \epsilon_+) R_{n-1} + \Delta t \epsilon_-}{\Delta t \epsilon_+ R_{n-1} + 1 - \Delta t \epsilon_-}. \quad (5)$$

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, $\epsilon_\pm$.

Eq. (5) describes a variety of cases of evidence accumulation studied previously (See Fig. 2): If the environment is fixed ($\epsilon_\pm = 0$), we recover Eq. (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 [26–28].

We can again approximate the stochastic process describing the evolution of the log likelihood ratio, $y_n = \ln R_n$, by an SDE:

$$dy = [g(t) + \underbrace{\epsilon_-(e^{-y} + 1) - \epsilon_+(e^y + 1)}_{\text{nonlinearity}}]dt + \rho(t)dW_t, \quad (6)$$

where the drift $g(t) = \frac{1}{\Delta t} E_\xi \left[\ln \frac{f_+(\xi)}{f_-(\xi)} \Big| H(t) \right]$, and variance $\rho^2(t) = \frac{1}{\Delta t} \text{Var}_\xi \left[\ln \frac{f_+(\xi)}{f_-(\xi)} \Big| H(t) \right]$ are no longer constant,

but depend on the state of the environment at time t. We derive Eq. (6) as the continuum limit of the discrete process y_n in Appendix B.

The nonlinear term in Eq. (6) does not appear in Eq. (3). It serves to discount older evidence by a factor determined by environmental volatility, i.e. the frequency of changes in environmental states. In previous work such discounting was modeled by a linear term [12, 22, 29], however our derivation shows that the resulting Ornstein-Uhlenbeck (OU) process does not model an optimal observer.

A. Equal switching rates between two states

When $\epsilon := \epsilon_+ = \epsilon_-$, the frequencies of switches between states are equal. Eq. (6) then becomes

$$dy = g(t)dt - 2\epsilon \sinh(y)dt + \rho(t)dW_t. \quad (7)$$

If we rescale time using $\tau = \epsilon t$, the rate of switches between environmental states is unity. We obtain

$$\mathrm{d}y_\tau = [\tilde{g}(\tau)]\,\mathrm{d}\tau - 2\sinh(y)\mathrm{d}\tau + [\tilde{\rho}(\tau)]\,\mathrm{d}W_\tau, \qquad (8)$$

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

$$\mathrm{d}y_\tau = \mathrm{sign}[\tilde{g}(\tau)]m\mathrm{d}\tau - 2\sinh(y)\mathrm{d}\tau + \sqrt{2m}\,\mathrm{d}W_\tau, \quad (9)$$

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.

The probabilities of a correct response (accuracy) in both interrogation (Fig. 3A) and free response (Fig. 3B) protocols increase with m. When an optimal observer is interrogated about the state of the environment at time τ, the answer is determined by the sign of the log likelihood ratio, y. 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 (Fig. 3A).

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

$$\mathcal{S}_\pm(y_\tau) = Ke^{\pm y_\tau - \frac{2\cosh(y_\tau)}{m}}, \quad \tilde{H}(\tau) = H_\pm, \qquad (10)$$

where $\tilde{H}(\tau) := H(\tau/\epsilon)$ and K is a normalization constant. This distribution is concentrated around $\bar{y}_\pm = \pm\sinh^{-1}\frac{m}{2}$, the fixed points of the deterministic counterpart of Eq. (9) 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, $\int_0^\infty \mathcal{S}_+(y)\mathrm{d}y = \int_{-\infty}^0 \mathcal{S}_-(y)\mathrm{d}y$ provides an upper bound for accuracy (Fig. 3A). To achieve accuracy a in the free response protocol (Fig. 3B), we require $|y| \geq \ln\frac{a}{1-a}$ [2]. The waiting time for this accuracy steeply increases with a and decreases with m.

B. Linear approximation of the SDE

An advantage of Eq. (6) is that it is amenable to standard methods of stochastic analysis. We can find an accurate piecewise linear approximation to Eq. (6), although, for simplicity, we focus on Eq. (9). The piecewise OU
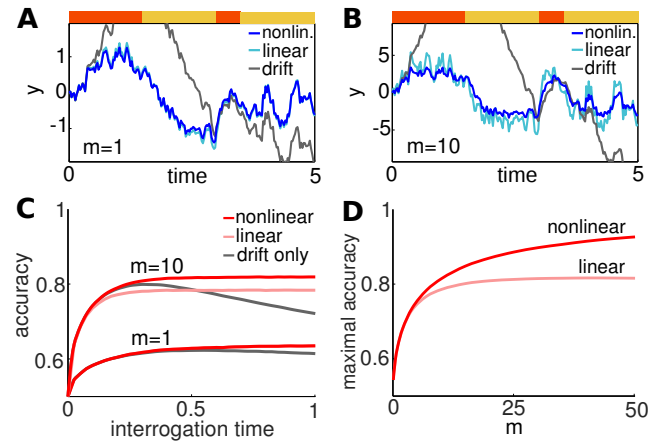

FIG. 4. Closest linear approximations of the nonlinear SDE, Eq. (9). (A,B) Single realizations of the nonlinear Eq. (9), linear approximation Eq. (11), and corresponding drift-diffusion model $\mathrm{d}y_\tau = \mathrm{sign}[\tilde{g}(\tau)]m\mathrm{d}\tau + \sqrt{2m}\mathrm{d}W_\tau$, in (A) a quickly changing environment ($m = 1$), and (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 Eq. (9) than in the linear approximation Eq. (11). Accuracy eventually decreases in the drift model since all evidence is weighted equally across time. (D) In the limit $t \to \infty$, accuracy saturates below unity in both the nonlinear model and linear approximation. The linear model employs the discounting strategy of a non-ideal observer, and hence performs worse.

process that models an observer that linearly discounts evidence has the form

$$\mathrm{d}y_\tau = b(\mathrm{sign}[\tilde{g}(\tau)]m\mathrm{d}\tau + \sqrt{2m}\mathrm{d}W_\tau) + \lambda y_\tau \mathrm{d}\tau, \qquad (11)$$

since both drift and diffusion need to be co-scaled. A linear approximation of Eq. (9), with the same equilibria and local stability is obtained by setting $\lambda = -\sqrt{m^2 + 4}$ and $b = \sqrt{1 + \frac{4}{m^2}}\sinh^{-1}\frac{m}{2}$. Individual realizations of Eq. (11) and Eq. (9) agree in quickly changing environments (Fig. 4A, $m = 1$), but are less similar in slowly changing environments (Fig. 4B, $m = 10$; see also panel C). Thus, as observer performance improves, the nonlinear term in Eq. (9) becomes more important. Note that the corresponding drift-diffusion model, $\mathrm{d}y_\tau = \mathrm{sign}[\tilde{g}(\tau)]m\mathrm{d}\tau + \sqrt{2m}\mathrm{d}W_\tau$, is qualitatively different as it lacks a restorative leak term. This difference becomes more pronounced as m increases (Fig. 4C).

Eq. (11) can be integrated explicitly using standard methods in stochastic calculus [30]. Furthermore, the accuracy of both systems saturate to a value smaller than 1 in the interrogation protocol as the interrogation time increases (Fig. 4C).

The linearized approximation can differ considerably from the full nonlinear model. For instance, in the interrogation protocol the performance of an ideal observer modeled by Eq. (9) increases with interrogation time (Fig. 4D), and accuracy approaches 1 as m diverges.
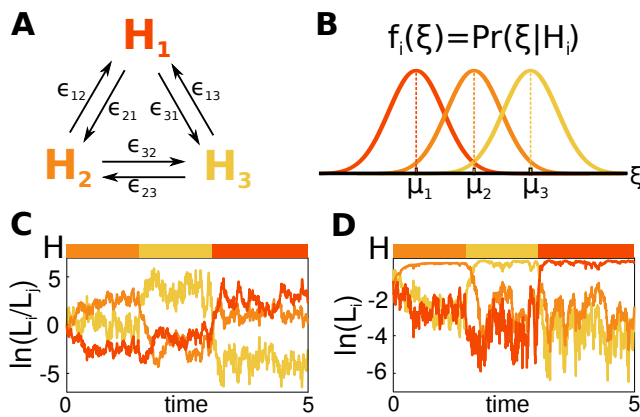
FIG. 5. Evidence accumulation with multiple choices in a changing environment. (**A**) The environment switches between N states (here $N = 3$). (**B**) Distributions $f_i(\xi) = \Pr(\xi|H_i)$ describing the probability of observation ξ in environmental state H_i (here $N = 3$). (**C,D**) Realization of the log likelihood ratios (panel **C**: $\ln(L_1/L_2)$, $\ln(L_2/L_3)$, $\ln(L_3/L_1)$), and $\ln L_i$ (panel **D**).

In contrast, the accuracy of an observer that discounts evidence linearly limits under 1 as m diverges. Indeed, this can be seen by employing the quasi-steady state approximation (fixing $\mathrm{sign}[\tilde{g}(\tau)] = 1$), and computing $\int_0^\infty \mathcal{T}(y)dy$, where $\mathcal{T}(y)$ is the steady state distribution of the OU process given in Eq. (11), to obtain

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

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

IV. 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 $(i = 1, ..., N)$, the optimal observer computes the present probability of each state 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 [13, 31]. 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 $\Pr(H(t_n) = H_i|H(t_{n-1}) = H_i) = 1 - \sum_{j\neq i}\Delta t\epsilon_{ji}$ (Fig 5**A**).

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, $\hat{i} = \mathrm{argmax}_i L_{n,i}$, corresponds to the most probable state, given the observations $\xi_{1:n}$. Following the approach above, we obtain

(See Appendix D):

$$L_{n,i} = \frac{\Pr(\xi_{1:n-1})}{\Pr(\xi_{1:n})} f_i(\xi_n) \times \qquad (12)$$

$$\left(\left(1 - \sum_{j\neq i}\Delta t\epsilon_{ji}\right)L_{n-1,i} + \sum_{j\neq i}\Delta t\epsilon_{ij}L_{n-1,j}\right).$$

Again after taking logarithms, $x_{n,i} = \ln L_{n,i}$, we can approximate the discrete stochastic process in Eq. (12), with an SDE:

$$d\mathbf{x} = \mathbf{g}(t)dt + \Lambda(t)d\mathbf{W}_t + K(\mathbf{x})dt, \qquad (13)$$

where the drift has components $g_i(t) = \frac{1}{\Delta t}\mathbb{E}_\xi[\ln f_i(\xi)|H(t)]$, $\Lambda(t)\Lambda(t)^T = \Sigma(t)$ with entries $\Sigma_{ij} = \frac{1}{\Delta t}\mathrm{Cov}_\xi[\ln f_i(\xi), \ln f_j(\xi)|H(t)]$, 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.

We can recover the case of two alternatives, $N = 2$, after exchanging the numbers in Eq. (13) with $\pm$ to obtain the approximating SDEs:

$$dx_\pm = [g_\pm(t) + (\epsilon_\mp e^{x_\mp - x_\pm} - \epsilon_\pm)]dt + dW_\pm, \qquad (14)$$

where $\langle W_i W_j\rangle = \Sigma_{ij}(t) \cdot t$ for $i, j \in \{+, -\}$. We obtain Eq. (6) by setting $y = x_+ - x_-$. Analogous expressions for the 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., Fig. 5**C**) [32].

V. 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 [19].

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 Eq. (12) the expression for the probabilities $L_{n,\theta} = \Pr(H(t_n) = H_\theta|\xi_{1:n})$ is derived in Appendix F, yielding:

$$L_{n,\theta} = \frac{\Pr(\xi_{1:n-1})}{\Pr(\xi_{1:n})} f_\theta(\xi_n) \times$$

$$\left(\left(1 - \int_a^b \Delta t\epsilon_{\theta'\theta}d\theta'\right)L_{n-1,\theta} + \int_a^b \Delta t\epsilon_{\theta\theta'}L_{n-1,\theta'}d\theta'\right).$$
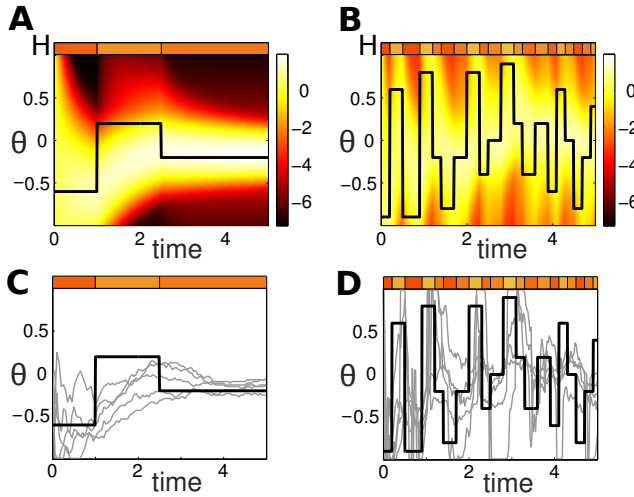
FIG. 6. Evidence accumulation with a continuum of choices. The observer infers the state of the environment, H_θ, where $\theta \in [-1, 1]$, and the state changes at discrete points in time. (**A**) In slowly changing environments, the distribution of the log probabilities, x_θ, can nearly equilibrate between switches (Solid line represents the true state of the environment at time t. For clarity, we show results of simulations without noise). (**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, $\widehat{\theta}(t) = \mathrm{argmax}_\theta x_\theta(t)$ (thin lines), fluctuates around the true value (thick line). (**D**) In quickly changing environments, $\widehat{\theta}(t)$ fluctuates more widely, as it is in a transient state much of the time.

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

$$\mathrm{d}x_\theta(t) = g_\theta(t)\mathrm{d}t + \mathrm{d}\widehat{W}_\theta(t) + K_\theta(x(t))\mathrm{d}t, \qquad (15)$$

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

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

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 $\widehat{\theta}(t) = \mathrm{argmax}_\theta x_\theta(t)$.

In slowly changing environments, the log probability $x_\theta(t)$ *nearly* equilibrates to a distribution with a well-defined peak between environmental switches (Fig. 6**A**). This does not occur in quickly changing environments (Fig. 6**B**). 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 Eq. (15) causes rapid departure from this quasi-stationary density when the environment changes, a mechanism proposed in [19].

Even when the environment is stationary for a long time, noise in the observations stochastically perturbs
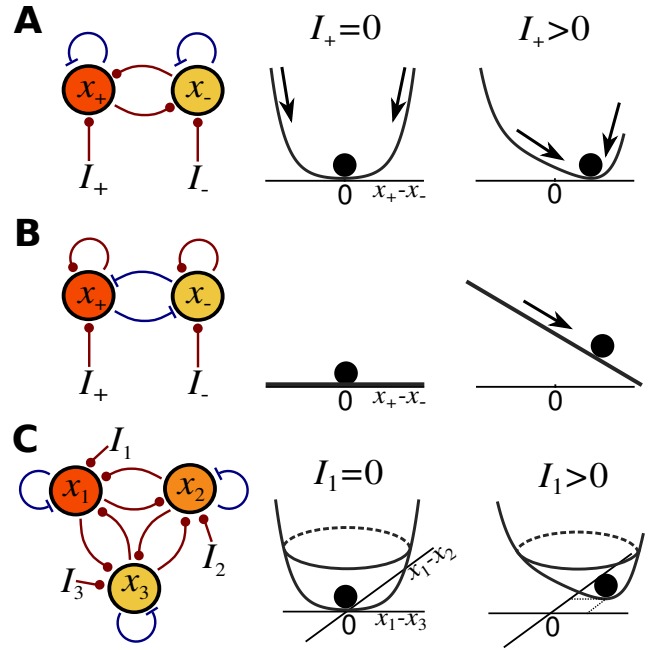

FIG. 7. Neural population models of evidence accumulation. (**A**) Two populations $u_\pm$ receive a fluctuating stimulus with mean $I_\pm$; they are mutually coupled by excitation (circles) and locally coupled by inhibition (flat ends). When $I_+ > 0$, the fixed point of the system has coordinates satisfying $x_+ > x_-$ as shown in the plots of the associated potentials. (**B**) Taking $\epsilon_\pm \to 0$ in Eq. (16) 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.

the log probabilities, $x_\theta(t)$, over the environmental states. This leads to fluctuations in the estimate $\widehat{\theta}(t)$ of the most probable alternative (Fig. 6**C**,**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, $\widehat{\theta}(t)$, 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 [19].

VI. A NEURAL IMPLEMENTATION OF AN OPTIMAL OBSERVER

Previous neural models of decision making typically relied on mutually inhibitory neural networks [12, 20, 33], with each population representing one alternative. In contrast, inference in dynamic environments with two states, H_+ and H_-, can be optimally performed by mutually excitatory neural populations with activities (firing

rates) r_+ and r_-,

$$dr_+ = [I_+(t) - \alpha r_+ + F_+(r_- - r_+)] \, dt + dW_+, \quad (16a)$$
$$dr_- = [I_-(t) - \alpha r_- + F_-(r_+ - r_-)] \, dt + dW_-, \quad (16b)$$

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, $W_\pm$ are Wiener processes representing the variability in the input signal with covariance defined as in Eq. (14) (See Appendix D). Thus, $I_\pm(t)dt + dW_\pm$ represents the total input to population $r_\pm$. When $\alpha > 0$ and sufficiently small, population activities are modulated by self-inhibition, and mutual excitation (Fig. 7**A**). Note, the parameter α determines the leak in the activity of each individual population, which depends on both the time constants and recurrent architecture of the local network [33]. Taking $y = r_+ - r_-$ reduces Eq. (16) to the SDE for the log likelihood ratio, Eq. (6). In the limit of a stationary environment, $\epsilon_\pm \to 0$, we obtain a linear integrator $dr_\pm = [I_\pm dt + dW_\pm] - \alpha(r_+ + r_-)dt/2$, as in previous studies [2, 20].

To show that the populations mutually excite each other, we set $W_+ = W_- = 0$, and linearize Eqs. (16). When the environment has not changed for a long time, Eq. (16) 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, note that the Jacobian matrix of (F_+, F_-) has the form:

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

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

The model given by Eq. (16) is matched to the timescale of the environment determined by $\epsilon_\pm$. Solutions approach stationary distributions if input is constant. Due to mutual excitation, they are very sensitive

to changes in inputs, a feature absent in previous connectionist models [34]. Even when ϵ is small, Eq. (16) has a single attracting state determined by the mean inputs $I_\pm^0$. We illustrate the response of the model to inputs using potentials (Fig. 7**A**). In contrast to the single attractor of Eq. (16), mutually inhibitory models can possess a neutrally stable line attractor that integrates inputs ($\epsilon_\pm \equiv 0$, Fig. 7**B**) [35].

We can extend our results for the $N = 2$ case by deriving a neural population model of decision making in changing environments. In [13], 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 number of possible states, $\{H_1, ..., H_N\}$, decisions can be performed optimally by neural populations $x_1, ..., x_N$ coupled by mutual excitation

$$dr_i = \left[I_i(t) - \alpha r_i + \sum_{j \neq i} F_{ij}(r_j - r_i) \right] dt + d\widehat{W}_i(t),$$
$$(17)$$

where the mean input is $I_i(t) = I_i^0$ when $H(t) = H_i$ and 0 otherwise and $(d\widehat{W}_1(t), ..., d\widehat{W}_N(t))^T = \Lambda(t)d\mathbf{W}_t$ describes input noise with $\Lambda(t)$ defined as in Eq. (13). Population firing rates are again determined by inhibition 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 (Fig. 7**C**).

Note that, as in the case of $N = 2$ alternatives, taking the limit of Eq. (17) as $\epsilon_{ij} \to 0$, we obtain linear integrators [20]

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

VII. 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 [12, 22, 24]. 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 an extended time.

We have made several assumptions about the model to simplify these initial derivations. Our ideal observer is assumed to be aware both of the uncertainty of their own

measurements and about the frequency with which the environment changes. A more realistic model would require that a naive observer learn the underlying volatility of the environment. Efforts to model the case of initially unknown transition rates produced hierarchical models that identify the location of change-points [36]. However, this approach quickly grows in computational complexity, since the probability of change points is determined by accounting for all possible transition histories [27]. 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 [37]. Likelihood update procedures have also been proposed for multiple alternative tasks in the limit $\epsilon_{ij} \to 0$ [32, 38]. Furthermore, Eq. (12) for the case $N = 2$ was derived in [39], but its dynamics were not analyzed. 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 not commonly discussed in the physics literature, but 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 quasistatic 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 [2, 3]. Dynamic environments offer a promising future direction for theory and experiments to probe the biophysical mechanisms that underlie decisions.

APPENDIX

In this appendix, 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 a 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 [2]. 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 integro–differential equation that represents the log probability for a continuum of possible environmental states, $\theta \in [a, b]$.

Note that throughout the appendix, 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 $\Delta t \to 0$. 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 we can relate the current probability, $L_{n,+}$, to the conditional probabilities at the time of the previous observation, t_{n-1}:

$$L_{n,+} = \frac{\Pr(H(t_n) = H_+)}{\Pr(\xi_{1:n})} \Pr(\xi_{1:n}|H(t_n) = H_+)$$

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

To derive a recursive equation, with probabilities that are not conditioned on the state at t_n, we first use Bayes' rule again to write

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

and

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

Plugging these formulas into our expression for $L_{n,+}$, we can then write

$$L_{n,+} = \frac{1}{\Pr(\xi_{1:n})} \times \Big((1 - \epsilon_{\Delta t,+})\Pr(\xi_{1:n}|H(t_n) = H_+; H(t_{n-1}) = H_+)\Pr(H(t_{n-1}) = H_+)$$

$$+ \epsilon_{\Delta t,-}\Pr(\xi_{1:n}|H(t_n) = H_+; H(t_{n-1}) = H_-)\Pr(H(t_{n-1}) = H_-)\Big).$$

The observation ξ_n is independent from the sequence of observations $\xi_{1:n-1}$ when conditioned on the states $H(t_n) = H_+$ and $H(t_{n-1}) = H_\pm$, respectively. Thus, we obtain

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

$$+ \epsilon_{\Delta t,-}\Pr(\xi_{1:n-1}|H(t_{n-1}) = H_-)\Pr(H(t_{n-1}) = H_-)\Big)$$

$$= \frac{\Pr(\xi_n|H(t_n) = H_+)}{\Pr(\xi_{1:n})} \times \Big((1 - \epsilon_{\Delta t,+})\Pr(H(t_{n-1}) = H_+|\xi_{1:n-1})\Pr(\xi_{1:n-1})$$

$$+ \epsilon_{\Delta t,-}\Pr(H(t_{n-1}) = H_-|\xi_{1:n-1})\Pr(\xi_{1:n-1})\Big)$$

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

Thus, by using our definition of the probabilities $L_{n,\pm}$, we can write an update equation for $L_{n,+}$ in terms of the probabilities $L_{n-1,\pm}$ at the previous time, t_{n-1},

$$L_{n,+} = \frac{\Pr(\xi_{1:n-1})\Pr(\xi_n|H(t_n) = H_+)}{\Pr(\xi_{1:n})} \times$$

$$((1 - \epsilon_{\Delta t,+})L_{n-1,+} + \epsilon_{\Delta t,-}L_{n-1,-}), \qquad (A1)$$

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

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

$$L_{n,-} = \frac{\Pr(\xi_{1:n-1})\Pr(\xi_n|H(t_n) = H_-)}{\Pr(\xi_{1:n})} \times$$
$$(\epsilon_{\Delta t,+}L_{n-1,+} + (1 - \epsilon_{\Delta t,-})L_{n-1,-}), \qquad \text{(A2)}$$

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

From Eqs. (A1) and (A2), the ratio $R_n = L_{n,+}/L_{n,-}$ is readily seen to satisfy the recursive equation

$$R_n = \frac{f_{\Delta t,+}(\xi_n)}{f_{\Delta t,-}(\xi_n)} \frac{(1 - \epsilon_{\Delta t,+})R_{n-1} + \epsilon_{\Delta t,-}}{\epsilon_{\Delta t,+}R_{n-1} + 1 - \epsilon_{\Delta t,-}}, \qquad \text{(A3)}$$

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 a SDE. To begin, we divide both sides of Eq. (A3) by R_{n-1} and take logarithms to yield

$$y_n - y_{n-1} = \ln\frac{f_{\Delta t,+}(\xi_n)}{f_{\Delta t,-}(\xi_n)} + \ln\frac{1 - \epsilon_{\Delta t,+} + \epsilon_{\Delta t,-}e^{-y_{n-1}}}{1 - \epsilon_{\Delta t,-} + \epsilon_{\Delta t,+}e^{y_{n-1}}}. \qquad \text{(B1)}$$

Following [2, 40], 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, Eq. (B1) 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

$$\Delta y_n \approx \ln\frac{f_{\Delta t,+}(\xi_n)}{f_{\Delta t,-}(\xi_n)} + \Delta t(\epsilon_-(e^{-y_n} + 1) - \epsilon_+(1 + e^{y_n}))$$
$$= E_\xi\left[\ln\frac{f_{\Delta t,+}(\xi_n)}{f_{\Delta t,-}(\xi_n)}\middle| H(t_n)\right] +$$
$$\left(\ln\frac{f_{\Delta t,+}(\xi_n)}{f_{\Delta t,-}(\xi_n)} - E_\xi\left[\ln\frac{f_{\Delta t,+}(\xi_n)}{f_{\Delta t,-}(\xi_n)}\middle| H(t_n)\right]\right) +$$
$$\Delta t(\epsilon_-(e^{-y_n} + 1) - \epsilon_+(1 + e^{y_n})). \qquad \text{(B2)}$$

where we 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

$$\Delta y_t \approx \Delta t g_{\Delta t}(t) + \sqrt{\Delta t}\rho_{\Delta t}(t)\eta +$$
$$\Delta t(\epsilon_-(e^{-y_t} + 1) - \epsilon_+(1 + e^{y_t})), \qquad \text{(B3)}$$

where η is random variable with standard normal distribution, and

$$g_{\Delta t}(t) := \frac{1}{\Delta t}E_\xi\left[\ln\frac{f_{\Delta t,+}(\xi)}{f_{\Delta t,-}(\xi)}\middle| H(t)\right] \qquad \text{and}$$
$$\rho_{\Delta t}^2(t) := \frac{1}{\Delta t}\text{Var}_\xi\left[\ln\frac{f_{\Delta t,+}(\xi)}{f_{\Delta-}(\xi)}\middle| H(t)\right]. \qquad \text{(B4)}$$

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 $\Delta t \to 0$ limit. We discuss different ways of introducing such a scaling in the next section.

Assuming that we have well-defined limits $g(t) := \lim_{\Delta t\to 0} g_{\Delta t}(t)$ and $\rho^2(t) := \lim_{\Delta t\to 0} \rho_{\Delta t}^2(t)$, the discrete-time stochastic process, Eq. (B3), approaches the SDE

$$dy = g(t)dt + \rho(t)dW_t + (\epsilon_-(e^{-y} + 1) - \epsilon_+(1 + e^y))dt, \qquad \text{(B5)}$$

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 Eq. (B1), and Eq. (B5), respectively. This correspondence can be made precise using the Donsker Invariance Principle [23].

In sum, Eq. (B5), can be viewed as an approximation of the logarithm of the likelihood ratio whose evolution is given exactly by Eq. (A3). For a fixed interval Δt, the parameters of the two equations are related via Eq. (B4), and $\epsilon_{\Delta t,\pm}/\Delta t = \epsilon_\pm$.

Appendix C: Precise correspondence.

We now discuss two approaches in which the correspondence between Eqs. (B1) and (B5) 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 [3]. 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 (B5) 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 Eq. (B4) it is then straightforward to compute [2, 40],

$$g_{\Delta t}(t) = \pm\frac{(\mu_+ - \mu_-)^2}{2\sigma^2} = g_\pm,$$

$$\rho_{\Delta t}^2(t) = \frac{(\mu_+ - \mu_-)^2}{\sigma^2} = \rho^2,$$

and note that $g(t) \in \{g_+, g_-\}$ is a telegraph process [25] with the probability masses $P(g_+, t)$ and $P(g_-, t)$ 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 sub-observations, each with statistics that scale with the length of the interval and the number of sub-observations. We define a family of stochastic processes parameterized by k, the number of sub-observations 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 contain 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 Eq. (B2) by the family of stochastic processes parameterized by k [2]

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

The scaling in this approximation guarantees that the drift is given by the limit $g(t) = \lim_{\Delta t \to 0} g_{\Delta t}(t) = \lim_{\Delta t \to 0} \frac{1}{\Delta t} E_\xi\left[\ln\frac{f_{\Delta t,+}(\xi)}{f_{\Delta t,-}(\xi)}\middle| H(t)\right] = E_\xi\left[\ln\frac{f_+(\xi)}{f_-(\xi)}\middle| H(t)\right]$ and the variance $\rho^2(t) = \lim_{\Delta t \to 0} \rho_{\Delta t}^2 = \lim_{\Delta t \to 0} \frac{1}{\Delta t}\text{Var}_\xi\left[\ln\frac{f_{\Delta t,+}(\xi)}{f_{\Delta t,-}(\xi)}\middle| H(t)\right] = \text{Var}_\xi\left[\ln\frac{f_+(\xi)}{f_-(\xi)}\middle| H(t)\right]$. Furthermore, as $k \to \infty$, by the

Central Limit Theorem,

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

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 $\Delta t \to 0$ yields Eq. (B5).

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, $i = 1, .., N$. 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 Eqs. (A1) and (A2)), 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

$$L_{n,i} = f_{\Delta t, i}(\xi_n) \times$$
$$\left(\left[1 - \sum_{j \neq i} \epsilon_{\Delta t, ji}\right] L_{n-1,i} + \sum_{j \neq i} \epsilon_{\Delta t, ij} L_{n-1,j}\right),$$
(D1)

where $f_{\Delta t, i}(\xi_n) = \Pr(\xi_n|H_i, t_n)$. From Eq. (D1), it follows that 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

$$\Delta x_{n,i} \approx \ln f_{\Delta t,i}(\xi_n) + \Delta t \sum_{j \neq i} \left(\epsilon_{ij} e^{x_{n,j} - x_{n,i}} - \epsilon_{ji} \right)$$
$$= \mathrm{E}_\xi \left[\ln f_{\Delta t,i}(\xi) | H(t_n) \right] +$$
$$\left(\ln f_{\Delta t,i}(\xi_n) - \mathrm{E}_\xi \left[\ln f_{\Delta t,i}(\xi) | H(t_n) \right] \right) +$$
$$\Delta t \sum_{j \neq i} \left(\epsilon_{ij} e^{x_{n,j} - x_{n,i}} - \epsilon_{ji} \right), \quad (D2)$$

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

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

$$\Delta x_{t,i} \approx \Delta t g_{\Delta t,i}(t) + \sqrt{\Delta t} \rho_{\Delta t,i}(t) \eta_i +$$
$$\Delta t \sum_{j \neq i} \left(\epsilon_{ij} e^{x_{t,j} - x_{t,i}} - \epsilon_{ji} \right), \quad (D3)$$

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

$$g_{\Delta t,i}(t) := \frac{1}{\Delta t} \mathrm{E}_\xi \left[\ln f_{\Delta t,i}(\xi) | H(t) \right] \quad \text{and}$$
$$\rho_{\Delta t,i}^2(t) := \frac{1}{\Delta t} \mathrm{Var}_\xi \left[\ln f_{\Delta t,i}(\xi) | H(t) \right].$$

The correlation of η_i's is given by

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

Note that Eq. (D3) is the multiple-alternative version of Eq. (B3). Equivalently, we can write Eq. (D3) as

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

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

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

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

$$g_i(t) := \lim_{\Delta t \to 0} g_{\Delta t,i}(t), \quad \text{and} \quad \Sigma_{ij}(t) := \lim_{\Delta t \to 0} \Sigma_{\Delta t,ij}(t),$$
$$(D4)$$

are well defined, we obtain the system of SDEs

$$\mathrm{d}x_i = g_i(t)\mathrm{d}t + \mathrm{d}\widehat{W}_i(t) + \sum_{j \neq i} \left(\epsilon_{ij} e^{x_j - x_i} - \epsilon_{ji} \right) \mathrm{d}t, \quad (D5)$$

or equivalently as the vector system

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

where $\mathbf{g}(t) = (g_1(t), ..., g_N(t))^T$ and $\Lambda(t)\Lambda(t)^T = \Sigma(t)$ are defined using the limits in Eq. (D4), $K_i(\mathbf{x}) = \sum_{j \neq i} \left(\epsilon_{ij} e^{x_j - x_i} - \epsilon_{ji} \right)$, and the components of $\mathbf{W}_t$ are independent Wiener processes. We can recover Eq. (B5) 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, Eq. (D5) can be viewed as an approximation of the logarithm of the probability whose evolution is given exactly by Eq. (D1). For a fixed interval Δt, the parameters of these equations are related via Eq. (D5), and $\epsilon_{\Delta t,ij}/\Delta t = \epsilon_{ij}$.

The limits $g_i(t) := \lim_{\Delta t \to 0} g_{\Delta t,i}(t)$ and $\Sigma_{ij}(t) := \lim_{\Delta t \to 0} \Sigma_{\Delta t,ij}(t)$ 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 sub-observations, $(\xi_1, \ldots, \xi_k)$, with mean and variance scaled by $\frac{\Delta t}{k}$. To do so we approximate $\ln f_{\Delta t,i}(\xi_n)$ in Eq. (D2) by

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

Appendix E: Log likelihood ratio for multiple alternatives.

We can also derive a continuum limit for the log likelihood ratio for any two choices $i, j \in \{1, 2, ..., N\}$. From Eq. (D1), 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

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

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 Eq. (D5) we obtain

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

or

$$
\mathrm{d}y_{ij} = g_{ij}(t)\mathrm{d}t + \mathrm{d}\widehat{W}_{ij} +
$$
$$
\left(\sum_{k \neq j} \epsilon_{kj} - \sum_{k \neq i} \epsilon_{ki} + \sum_{k \neq i} \epsilon_{ik}\mathrm{e}^{y_{ki}} - \sum_{k \neq j} \epsilon_{jk}\mathrm{e}^{y_{kj}} \right)\mathrm{d}t,
$$
$$(E1)$$

where $g_{ij}(t) = \mathrm{E}_\xi \left[\ln \frac{f_i(\xi)}{f_j(\xi)} \Big| H(t) \right]$ and $\widehat{W}$ is a Wiener process with covariance matrix given by $\mathrm{Cov}_\xi \left[\widehat{W}_{ij}, \widehat{W}_{i'j'} \Big| H(t) \right] = \mathrm{Cov}_\xi \left[\ln \frac{f_i(\xi)}{f_j(\xi)}, \ln \frac{f_{i'}(\xi)}{f_{j'}(\xi)} \Big| H(t) \right]$. We can also write Eq. (E1) in vector form

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

where $\mathbf{K}_{ij}(\mathbf{y}) = \sum_{k \neq j} \epsilon_{kj} - \sum_{k \neq i} \epsilon_{ki} + \sum_{k \neq i} \epsilon_{ik}\mathrm{e}^{y_{ki}} - \sum_{k \neq j} \epsilon_{jk}\mathrm{e}^{y_{kj}}$, $\Lambda(t)\Lambda(t)^T = \Sigma(t)$ 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 $L_{n,\theta} = \mathrm{Pr}(H(t_n) = H_\theta | \xi_{1:n})$. 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

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

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 $\mathrm{Pr}(H(t_n) = H_\theta | H(t_{n-1}) = H_\theta) = 1 - \int_a^b \epsilon_{\Delta t, \theta'\theta}\mathrm{d}\theta'$. 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

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

Notice that the sum from the $N < \infty$ case, as in Eq. (D1), 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 $\frac{\mathrm{Pr}(\xi_{1:n-1})}{\mathrm{Pr}(\xi_{1:n})}$, since we wish to compare the magnitudes of the probabilities. We obtain

$$
L_{n,\theta} = f_{\Delta t, \theta}(\xi_n)\left(\left[1 - \int_a^b \epsilon_{\Delta t, \theta'\theta}\mathrm{d}\theta' \right]L_{n-1,\theta} + \int_a^b \epsilon_{\Delta t, \theta\theta'}L_{n-1,\theta'}\mathrm{d}\theta' \right), \quad (F1)
$$

where $f_{\Delta t, \theta}(\xi_n) = \mathrm{Pr}(\xi_n | H(t_n) = H_\theta)$. From Eq. (F1), 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

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

To approximate this discrete-time stochastic process with a 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,

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

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

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

$$= \mathrm{E}_\xi \left[f_{\Delta t,\theta}(\xi)| H(t_n)\right] + $$
$$\left(\ln f_{\Delta t,\theta}(\xi_n) - \mathrm{E}_\xi \left[\ln f_{\Delta t,\theta}(\xi)| H(t_n)\right]\right) + $$
$$\Delta t \int_a^b \left(\epsilon_{\theta\theta'}\mathrm{e}^{x_{n,\theta'}-x_{n,\theta}} - \epsilon_{\theta'\theta}\right) \mathrm{d}\theta',$$

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

$$\Delta x_{t,\theta} \approx \Delta t g_{\Delta t,\theta}(t) + \sqrt{\Delta t}\rho_{\Delta t,\theta}(t)\eta_\theta + $$
$$\Delta t \int_a^b \left(\epsilon_{\theta\theta'}\mathrm{e}^{x_{t,\theta'}-x_{t,\theta}} - \epsilon_{\theta'\theta}\right) \mathrm{d}\theta', \quad (F2)$$

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

$$g_{\Delta t,\theta}(t) := \frac{1}{\Delta t}\mathrm{E}_\xi \left[\ln f_{\Delta t,\theta}(\xi)| H(t)\right], \qquad \text{and}$$

$$\rho^2_{\Delta t,\theta}(t) := \frac{1}{\Delta t}\mathrm{Var}_\xi \left[\ln f_{\Delta t,\theta}(\xi)| H(t)\right].$$

The correlation of η_i's is given by

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

Equivalently, we can write Eq. (F2) as

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

where $\widehat{W}_{\Delta t} := (\widehat{W}_{\Delta t,\theta})_{\theta\in[a,b]}$ follows a multivariate Gaussian distribution with mean zero and covariance function $\Sigma_{\Delta t,\theta\theta'}$ given by

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

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

$$g_\theta(t) := \lim_{\Delta t\to 0} g_{\Delta t,\theta}(t), \quad \text{and} \quad \Sigma_{\theta\theta'}(t) := \lim_{\Delta t\to 0} \Sigma_{\Delta t,\theta\theta'}(t),$$
$$(F3)$$

are well defined, we obtain the system of SDEs

$$\mathrm{d}x_\theta = g_\theta(t)\mathrm{d}t + \mathrm{d}\widehat{W}_\theta(t) + \int_a^b \left(\epsilon_{\theta\theta'}\mathrm{e}^{x_{\theta'}-x_\theta} - \epsilon_{\theta'\theta}\right) \mathrm{d}\theta'\mathrm{d}t,$$

or equivalently as the system of SDEs

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

where $\mathbf{g}(t) = \left(g_\theta(t)\right)_{\theta\in[a,b]}$ and $\Lambda(t)\Lambda(t)^T = \Sigma(t)$ are defined using the limits in Eq. (F3), $K(\mathbf{x}) = \int_a^b \left(\epsilon_{\theta\theta'}\mathrm{e}^{x_{\theta'}-x_\theta} - \epsilon_{\theta'\theta}\right) \mathrm{d}\theta'$, and the components of $\mathbf{W}_t$ are independent Wiener processes.

[1] J. R. Busemeyer and J. T. Townsend, Psychological review **100**, 432 (1993).

[2] R. Bogacz, E. Brown, J. Moehlis, P. Holmes, and J. D. Cohen, Psychol Rev **113**, 700 (2006).

[3] J. I. Gold and M. N. Shadlen, Annu Rev Neurosci **30**, 535 (2007).

[4] B. W. Brunton, M. M. Botvinick, and C. D. Brody, Science **340**, 95 (2013).

[5] T. D. Seeley, S. Camazine, and J. Sneyd, Behavioral Ecology and Sociobiology **28**, 277 (1991).

[6] I. D. Couzin, J. Krause, N. R. Franks, and S. A. Levin, Nature **433**, 513 (2005).

[7] G. Balázsi, A. van Oudenaarden, and J. J. Collins, Cell **144**, 910 (2011).

[8] J. Neyman and E. S. Pearson, *On the problem of the most efficient tests of statistical hypotheses* (Springer, 1992).

[9] A. Wald and J. Wolfowitz, The Annals of Mathematical Statistics **19**, 326 (1948).

[10] J. I. Gold and M. N. Shadlen, Neuron **36**, 299 (2002).

[11] R. Ratcliff and P. L. Smith, Psychological review **111**, 333 (2004).

[12] M. Usher and J. L. McClelland, Psychol Rev **108**, 550 (2001).

[13] J. M. Beck, W. J. Ma, R. Kiani, T. Hanks, A. K. Churchland, J. Roitman, M. N. Shadlen, P. E. Latham, and A. Pouget, Neuron **60**, 1142 (2008).

[14] R. Bogacz, M. Usher, J. Zhang, and J. L. McClelland, Philos Trans R Soc Lond B Biol Sci **362**, 1655 (2007).

[15] P. Domenici and R. W. Blake, J Exp Biol **200**, 1165 (1997).

[16] K. M. Passino, T. D. Seeley, and P. K. Visscher, Behavioral Ecology and Sociobiology **62**, 401 (2008).

[17] T. E. Behrens, M. W. Woolrich, M. E. Walton, and M. F. Rushworth, Nature neuroscience **10**, 1214 (2007).

[18] M. R. Nassar, K. M. Rumsey, R. C. Wilson, K. Parikh, B. Heasly, and J. I. Gold, Nature neuroscience **15**, 1040

(2012).

[19] J. T. McGuire, M. R. Nassar, J. I. Gold, and J. W. Kable, Neuron **84**, 870 (2014).

[20] T. McMillen and P. Holmes, Journal of Mathematical Psychology **50**, 30 (2006).

[21] M. Usher and J. L. McClelland, Psychological review **111**, 757 (2004).

[22] P. L. Smith and R. Ratcliff, Trends Neurosci **27**, 161 (2004).

[23] P. Billingsley, *Probability and Measure* (Wiley, 1995).

[24] R. Ratcliff, Psychological review **85**, 59 (1978).

[25] M. Kac, *Probability and Related Topics in Physical Sciences* (American Mathematical Society, 1959).

[26] J. Y. Angela, in *Advances in neural information processing systems* (2006), pp. 1545–1552.

[27] R. P. Adams and D. J. MacKay, arXiv preprint arXiv:0710.3742 (2007).

[28] A. N. Shiryaev, *Optimal stopping rules*, vol. 8 (Springer Science & Business Media, 2007).

[29] L. P. Sugrue, G. S. Corrado, and W. T. Newsome, science **304**, 1782 (2004).

[30] C. W. Gardiner, *Handbook of stochastic methods for physics, chemistry, and the natural sciences* (Springer-Verlag, Berlin, 2004), 3rd ed.

[31] R. Bogacz and K. Gurney, Neural computation **19**, 442 (2007).

[32] V. Draglia, A. G. Tartakovsky, and V. V. Veeravalli, Information Theory, IEEE Transactions on **45**, 2448 (1999).

[33] X.-J. Wang, Neuron **36**, 955 (2002).

[34] K.-F. Wong, A. C. Huk, M. N. Shadlen, and X.-J. Wang, Front Comput Neurosci **1**, 6 (2007).

[35] C. K. Machens, R. Romo, and C. D. Brody, Science **307**, 1121 (2005).

[36] R. C. Wilson, M. R. Nassar, and J. I. Gold, Neural Comput **22**, 2452 (2010).

[37] S. Deneve, Neural Comput **20**, 91 (2008).

[38] P. Armitage, Journal of the Royal Statistical Society. Series B (Methodological) **12**, 137 (1950).

[39] Q. Zhao and J. Ye, Signal Processing, IEEE Transactions on **58**, 5994 (2010).

[40] S. Bitzer, H. Park, F. Blankenburg, and S. J. Kiebel, Frontiers in human neuroscience **8** (2014).