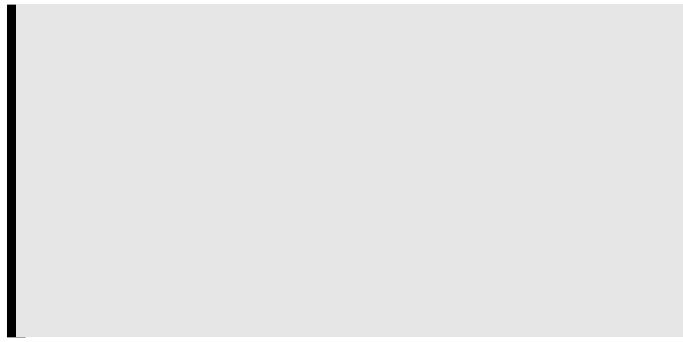
arXiv:1902.01535v3 [q-bio.NC] 9 Oct 2019

ORIGINAL ARTI



ear (Piet et al., 2018; Brunton et al., 2013). Each click train is generated by a Markov-modulated Poisson process (Fischer

sensitive to the mode of evidence accumulation they use in fluctuating environments. We also show how using different models and different data fitting methods can lead to divergent results, especially in the presence of sensory noise. We argue that similar issues can arise whenever we try to interpret data from decision-making tasks.

2 | NORMATIVE MODEL FOR THE DYNAMIC CLICKS TASK

In the dynamic clicks task an observer is presented with two Poisson click streams, $s^L(t)$ and $s^R(t)$ ($0 < t < T$), and needs to decide which of the two has a higher rate (Brunton et al., 2013). The rates of the two streams are not constant, but change according to a hidden, continuous-time Markov chain, $x(t)$, with binary state space $\{x^R, x^L\}$. The frequency of the switches is determined by the *hazard rate*, h, so that $P(x(t+dt) \neq x(t) \mid x(t) = x^R) = h\,dt + o(dt)$. The left and right rates, $\lambda^L(t)$ and $\lambda^R(t)$, can each take on one of two values, $\{\lambda_{\text{high}}, \lambda_{\text{low}}\}$ with $\lambda_{\text{high}} > \lambda_{\text{low}} > 0$. When $x(t) = x^L$, we have $(\lambda^L(t), \lambda^R(t)) = (\lambda_{\text{high}}, \lambda_{\text{low}})$, and when $x(t) = x^R$ the opposite is true. Therefore $x(t) = x^k$ means that stream k has the higher rate at time t: $\lambda^k(t) = \lambda_{\text{high}}$ (Fig. 1A). The observer is prompted to identify the side of the higher rate stream, $x(T)$, at a random time T. The interrogation time, T, is sampled ahead of time by the experimenter for each trial and is unknown to the subject. We refer the reader to Piet et al. (2018) and Brunton et al. (2013) for more details about the experimental setup.

This task is closely related to the *filtering of a Hidden Markov Model* studied in the signal processing literature (Cappé et al., 2005; Rabiner and Juang, 1986). For a single, 2-state Markov-modulated Poisson process (Fischer and Meier-Hellstern, 1992), the filtering problem was solved by Rudemo (1972) – see also (Snyder, 1975) for review and extensions. This filtering problem corresponds to a task in which a single, variable rate click stream is presented to the observer who has to report whether at some time T the rate is high or low. In the present case, the observer is presented with two coupled Markov-modulated Poisson processes. The normative model reduces to that considered by Rudemo (1972)

FIGURE 1 A: Schematic of the dynamic clicks task from Piet et al. (2018). B: A single trajectory of the log-likelihood ratio (LLR), y_t, during a trial. The click streams and environment state are shown above the trajectory. C: Response accuracy of the ideal observer as a function of interrogation time for two distinct SNR values, S• $\bar{h}$, defined in Eq. (6). Two distinct pairs of click rates used in simulations (low = „• and ‡• Hz) were chosen to match each SNR at hazard rate h = , Hz, resulting in overlaying dashed (low = „•Hz) and solid (low = ‡•Hz) lines. For S• $\bar{h}$ = f, we take ¹ high, low º = ¹†‰. , ˆ, „•º and ¹ ß Á

3 | THE SIGNAL-TO-NOISE RATIO OF DYNAMIC CLICKS

Four parameters characterize the dynamic clicks task: the hazard rate, h, duration of a trial, *i.e.* interrogation time, T, the low click rate, $_{\text{low}}$, and the high click rate, $_{\text{high}}$. However, we next show that only two effective parameters typically govern an ideal observer's performance (Fig. 1C,D): the product of the interrogation time and the hazard rate, hT, and the signal-to-noise ratio (SNR) of the dynamic stimulus. The former corresponds to the mean number of switches in a trial, and the latter combines the click rates $_{\text{low}}$ and $_{\text{high}}$ into a Skellam–type SNR (Eq. (4) below), scaled by the hazard rate h (Eq. (6)).

To motivate our definition, consider first the case of a static environment, $h = \bullet$ Hz, for which the normative model is given by Eq. (2) without the nonlinear term. Since does not affect the sign of y, response accuracy depends entirely on the difference in click counts $N^{R}{}_{1}t^{o}$ $N^{L}{}_{1}t^{o}$, where $N^{j}{}_{1}t^{o}$ are the counting processes associated with each click stream. Thus we can define the difference in click counts as the *signal*, and the SNR as the ratio between the signal mean and standard deviation at time T (Skellam, 1946),

$$\text{SNR}^{T}_{\bullet} := \frac{\text{E}»N^{R}{}_{1}T^{o}{}_{\%} \quad \text{E}»N^{L}{}_{1}T^{o}{}_{\%}}{\sqrt{\text{Var}»N^{R}{}_{1}T^{o}{}_{\%} + \text{Var}»N^{L}{}_{1}T^{o}{}_{\%}}} = \frac{T_{\text{high}} \quad T_{\text{low}}}{\sqrt{T_{\text{high}} + T_{\text{low}}}} = S \quad \sqrt{T}, \tag{3}$$

where

$$S := \frac{\text{high} \quad \text{low}}{\sqrt{\text{high} + \text{low}}}. \tag{4}$$

In a dynamic environment, the volatility of the environment, governed by the hazard rate, h, also affects response accuracy. The environment can switch states immediately before the interrogation time, T, decreasing response accuracy. This suggests that accuracy will not only be determined by the click rates, but also by the length of time the environment remains in the same state prior to interrogation. Using this fact and the definition of SNR in a static environment, we determine the statistics for the difference in the number of clicks between the high- and low-rate streams during the final epoch preceding interrogation (for derivation details see Appendix B). Averaging over the Poisson distributions characterizing the click numbers, and the epoch length distribution yields a nonlinear expression representing the SNR that involves $S\bullet\sqrt{h}$

WHERE *%*

signal (See Appendix C).

Fig. 1C shows examples in which the ideal observer's response accuracy is constant when SNR and hT are fixed. Accuracy is computed as the fraction of trials at which the observer's belief, y_t, matches the underlying state, $x(t)$, at the interrogation time, T, that is the fraction of trials for which $\text{sign}(y_T) = x(T)$. The accuracy as a function of T and h = , remains constant if we change $_{\text{high}}$ and $_{\text{low}}$, but keep S fixed. As the interrogation time T is increased, the accuracy saturates to a value below 1 (Fig. 1C), consistent with previous modeling studies of decision-making in dynamic environments (Glaze et al., 2015; Veliz-Cuba et al., 2016; Radillo et al., 2017; Piet et al., 2018). Evidence discounting limits the magnitude of the LLR, y_t. Hence a sequence of low rate clicks can lead to errors, especially for low SNR values. Moreover, on some trials the state, $x(t)$, switches close to the interrogation time T. As it may take multiple clicks for y_t to change sign after a change point (See Fig. 1B), this can also lead to an incorrect response.

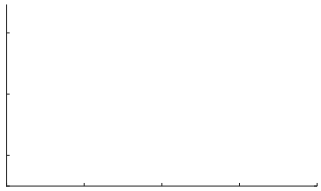
In Fig. 1D we show that the maximal accuracy (obtained for T sufficiently large) as a function of $_{\text{high}}$ and $_{\text{low}}$ (colormap), is approximately constant along SNR level sets (black oblique curves). This correspondence is not exact when $_{\text{high}}$ and $_{\text{low}}$ are small (Fig. 1D inset), and we conjecture that this is because higher order statistics of the signal determine response accuracy in this parameter range. As discussed in Appendix C, for large $_{\text{high}}$ and $_{\text{low}}$ we can use a diffiusion approximation for the dynamics of Eq. (2). When $_{\text{high}}$ and $_{\text{low}}$ are small, the diffusion approximation does not apply, and response accuracy is characterized by features of the signal beyond its mean and variance. Since the SNR only describes the ratio between the mean and standard deviation of the stimulus, it cannot capture the impact of higher order statistics on accuracy at low click rates. Nonetheless, the SNR predicts response accuracy well.

The consequences of these observations are twofold: Two parameter combinations determine optimal performance, potentially simplifying experiment design. To ensure coverage of different response accuracy regimes, we can initially vary SNR and hT. To increase the accuracy of an ideal observer, it is not sufficient to increase both click rates, for instance, since the SNR stays constant if $_{\text{high}}$ and $_{\text{low}}$ follow the parabolas shown in Fig. 1D. Second, this approach makes testable predictions about the accuracy of an optimal observer: If we change parameters so that SNR and hT are fixed, and a subject's accuracy is affected, this indicates that the subject may not have learned the hazard rate, h or is using a suboptimal discounting model.

4 | POST CHANGE-POINT DECISIONS DEPEND ON SNR

To understand how an optimal observer adapts to environmental changes, we next ask how their fraction of correct responses depends on the final time, T_{fin}, between the last change point preceding a decision and the decision itself (Fig. 2A). Overall accuracy again depends on both SNR and rescaled trial time hT. In addition, for sufficiently long trials, accuracy as a function of time since the last change point depends only on the rescaled time since the change point, hT_{fin} and the SNR.

If the click rates, $_{\text{high}}$ and $_{\text{low}}$, are varied, but S and h are held fixed, the accuracy as a function of T_{fin} remains unchanged (Fig. 2B, for h = ,, S = f). On the other hand, accuracy changes if we fix S• $\sqrt{h}$ (SNR) but vary h (Fig. 2B, leiNYWP 1 @pi



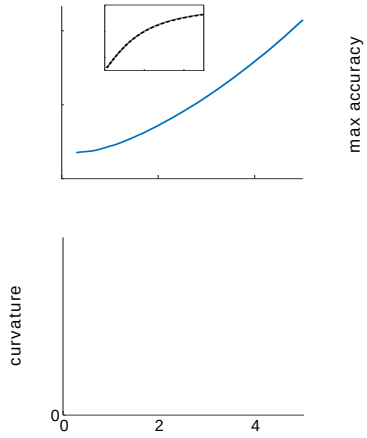


FIGURE 3 A: Optimal linear discounting rate, in Eq. (7), as a function of S• $\bar{h}$. A(inset): The accuracy of the linear and nonlinear model are nearly identical over a wide range of SNR values, S• $\bar{h}$, when the linear discounting rate is set to (here h = , is fixed). B: Response accuracy near the optimal discounting rate for the linear model (dashed), and assumed hazard rates for the nonlinear model (solid) for several SNR values (h = ,). The linear model is more sensitive to relative changes in the discounting rate. The relative error is defined as , ••¹ $\hat{h}$ hº•h for the nonlinear model and , ••¹ $\hat{\theta}$ º• for the linear model. C: Curvature (absolute value of the second derivative) of the accuracy profiles in panel B, evaluated at their peak, as a function of S• $\bar{h}$. The curvature, and hence sensitivity, of the nonlinear model is higher for intermediate and large values of S• $\bar{h}$. Since the functions in panel B do not depend on the actual values of $\hat{h}$ and $\hat{\theta}$, but rather the relative distance of these parameters from reference values, what we show in this plot are *relative* curvatures. We compare relative curvatures as $\hat{h}$ and $\hat{\theta}$ do not have the same units. D: Ratio of the accuracy of the linear model to that of the normative model, as SNR is varied. Along each curve, the discounting rate $\hat{\theta}$ of the linear model is held fixed at the value that would maximize accuracy at the reference SNR indicated by the legend.

5 | A LINEAR APPROXIMATION OF THE NORMATIVE MODEL

Following Piet et al. (2018) we next show that an approximation of the normative model given by Eq. (2) can be tuned to give near optimal accuracy, but the accuracy of the approximation tends to be sensitive to the changes in the discounting parameter. This approximate, linear model is given by,

$$\frac{dy_t}{dt} = \left[\sum_{i=,} \Theta\left(\quad ¹t \quad t_i^{R\,o}\right) \quad \sum_{j=,} \Theta\left(\quad ¹t \quad t_j^{L\,o}\right)\right] \quad y_t\,. \tag{7}$$

In particular, here the nonlinear sinh term in Eq. (2) is replaced by a linear term proportional to the accumulated evidence.

When tuned appropriately, Eq. (7) closely approximates the dynamics and accuracy of the optimal model (Fig. 3A) (Piet et al., 2018; Veliz-Cuba et al., 2016). Moreover, it also provides a good fit to the responses of rats on a dynamic clicks task (Piet et al., 2018). As the normative and linear models exhibit similar dynamics, it appears that they are difficult to distinguish. However, as we show next, the linear model is more sensitive to changes in its discounting parameter, providing a potential way to distinguish between the models.

We assume that T is large enough so that accuracy has saturated (as in Fig. 1C), and compare the maximal accuracy of the nonlinear and linear model. For the linear discounting rate that maximizes accuracy, = , the linear model

obtains accuracy nearly equal to the normative model (Fig. 3A, inset). The optimal linear discounting rate, , increases with SNR (Fig. 3A), whereas the discounting term in the normative, nonlinear model remains constant when the hazard rate, *h*, is fixed. When SNR is large, evidence discounting in the linear model can be stronger (larger), since each evidence increment is more reliable and can be given more weight. When SNR is lower, linear evidence discounting is weaker (smaller) resulting in the averaging of noisy evidence across longer timescales.

and $Acc_{lin}{}^{1}S^{o}$ when $h =$, is the red curve in Fig. 3D. Since the ratio between $Acc_{norm}{}^{1}S^{ref\,o}$ and $Acc_{lin}{}^{1}S^{ref\,o}$ is near 1, the linear and normative models cannot be distinguished at S^{ref}. However, a subject using the normative model tuned at S^{ref}, will still perform optimally at $S \neq S^{ref}$, if and h are held constant. On the other hand, a linear model optimized at S^{ref}, will no longer be optimal at $S \neq S^{ref}$. This distinction is captured by the drop in the accuracy ratio along the red curve in Fig. 3D.

We can quantify the distinction between the two models by their relative difference:

$$\frac{Acc_{norm}{}^{1}S^{new\,o} \quad Acc_{lin}{}^{1}S^{new\,o}}{Acc_{norm}{}^{1}S^{new\,o}} = \bullet.\bullet\dagger.$$

More generally, for TWH@UNWWX@RNXYS@td@[HrefI]Y1@tm@[HhI]tWoreIM1YSHgenerI1YHallyaisW1IMVPHsI]tjoTbgH

$$Acc_{norm}$$

Before fitting these models to choice data, we note that an increase in sensory noise, , decreases the value of the discounting parameters that maximize accuracy in both models (Piet et al., 2018): Noisier observations require integration of information over longer timescales (Fig. 4A,B). Thus, adaptivity to change points is sacrificed in order to pool over larger sets of observations . This, in turn, leads to larger biases, particularly after change points. A similar trade-off between adaptivity and bias has been observed in models and human subjects performing a related dynamic decision task (Glaze et al., 2018).

We next fit the discounting parameters in both models using synthetic choice data, treating the other parameters of the models as known. To do so we produced responses using a fixed *reference model* from both classes, and fit a model from each class to the resulting datasets. Specifically, let $m_{\text{ref}} \in \{\text{L}, \text{NL}\}$ (L = linear, NL = nonlinear) denote the reference model used to produce the choice data, and let $m_{\text{fit}} \in \{\text{L}, \text{NL}\}$ denote the model that was fit to the resulting data. We independently studied the four possible model pairs $(m_{\text{fit}}, m_{\text{ref}})$. In what follows, refers to the discounting parameter that was fit to data in any given class, so that := when $m_{\text{fit}} = \text{L}$ and := h when $m_{\text{fit}} =$

the reference choice data with the nonlinear and linear models, respectively. To pick these constants in our simulations, we took the values that would maximize accuracy in the corresponding noise-free systems. That is, $h_{\text{ref}} := h_{\text{stim}} =$, and $_{\text{ref}} :=$ ‡.ˆ† (See Appendix F for more details on the simulations).

During a single fit, we generated stimulus data for N i.i.d. trials,

$$\mathsf{D} := \mathsf{f}^{1}\mathsf{T}_k, d_k{}^{\mathsf{o}} : , \quad k \quad N\mathsf{g}, \tag{10}$$

where $\mathsf{T}_k :=$ $t_i^R \;{}^{N_{R,k}}_{i=,}$, $t_j^L \;{}^{\mathsf{o}N_{L,k}}_{j=,}$ is the sequence of N

hand, the linear model fits – ^{1}L, NL^o and ^{1}L, L^o – converge more rapidly, likely because the linear model is sensitive to changes in its discounting parameter (See Fig. 3B,C).

In anticipation of our next section, we point out that computing the MLE can be treated as a statistical learning problem in which we minimize a negative log-likelihood loss function over the dataset D (See Eq. 7.8 in Friedman et al. (2001)):

$$L_{LL}(d_k, m_{\text{fit}}(T_k)|\ ,\ ^o) := -\log P(d_k = m_{\text{fit}}(T_k)|\ ,\ ^o). \tag{12}$$

Here d_k and $m_{\text{fit}}(T_k)$ are the choices generated by the reference and fit models, respectively, on the k^{th} trial. As before the discounting parameter, , and the level of sensory noise, , parametrize the fitted model. Fitted model responses $m_{\text{fit}}(T_k)$ are non-deterministic only because of sensory noise. The likelihood $P(d_k = m_{\text{fit}}(T_k)|\ ,\ ^o)$ is the probability that the response generated by the fit model on trial k matches the response observed in the data (See Appendix F for details on how this likelihood was computed for each model class), which must be obtained from many realizations of $m_{\text{fit}}(T_k)$ subject to click noise of amplitude . The MLE, , for m_{fit} is then found by minimizing the expected loss across all trials,

:= argmin *WNOf1XS@WNYMRSVHFitteIS1HxpectedIM1YUHlossITP@uY1@TWNTXU@TVVNSVW@

of the loss function $L_{0/1}$ over the data samples (T_k, d_k) and across realizations, Z_j, of sensory noise,

$$\theta := \operatorname{argmin}_{\theta} E\left[L_{0/1}(d_k, m_{\text{fit}}(T_k, Z_j; \theta); \theta)\right] = \operatorname{argmin}_{\theta} \frac{1}{QN} \sum_{j=1}^{Q} \sum_{k=1}^{N} L_{0/1}(d_k, m_{\text{fit}}(T_k, Z_j; \theta); \theta).$$

For a binary decision model, this involves finding the parameter θ that minimizes the expected number of mismatches (or probability of a mismatch) between the choices of the model and those observed by the data (minimizing 0/1-loss), or maximizes the expected number of matches (or probability of a match) between the data and fit model (maximizing 0/1-prediction accuracy). In our fits, we used $Q = 1$, sampling a single realization of click noise perturbations per click stream. As we sampled from a large number of click streams, this was sufficient to average the loss function.

BotgNUtgNUtgNM 1 YSHtoI@td@[HkI]tjOd@[HkI]tjOdS

information can then be used to tease apart candidate model classes the experimental subject might be employing. Here we have focused on properties of a normative, nonlinear model, and its differences with a close, linear approximation. We found that the linear model is more sensitive to changes in the discounting parameter compared to the nonlinear model, and suggest this is why fitting a linear model to choice data requires fewer trials than fitting a nonlinear model.

In dynamic environments, task parameters may have predictable effects on subjects' overall accuracy and accuracy relative to change points. We have shown that there is a range of intermediate to high SNR in which the linear model is sensitive to changes in its discounting parameter, but the nonlinear model is not. This suggests this range could be probed to distinguish the evidence accumulation strategies subjects are using. These strategies may also be fit by other approximate models, like accumulators with no-flux boundaries or sliding-window integrators (Wilson et al., 2013; Glaze et al., 2015; Barendregt et al., 2019), which can also be sensitive to changes in their discounting parameters.

Psychophysical tasks used to infer subjects' decision-making strategies can require extensive training and data collection (Hawkins et al., 2015b,a). Normative and approximately normative decision-making models diverge most in their response accuracy when tasks are of intermediate difficulty. As we have shown, task difficulty may be controlled by combinations of task parameters representing fewer dimensions than the total number of parameters. Identifying these parameter combinations may be possible by computing the signal-to-noise (SNR) ratio of the stimulus produced by a particular parameter set. However, subjects' responses are also susceptible to noise from sensing and processing

characterizing discounting between clicks, when evidence arrives discretely. Many different functions could lead to the same amount of evidence discounting between clicks, leading to ambiguity in the model selection process.

Parameter identification for evidence accumulation models can be sensitive to the method chosen to fit model responses to choice data (Ratcliff and Tuerlinckx, 2002). Glaze et al. (2015) used the approach of minimizing the cross-entropy error function, which measures the dissimilarity between binary choices in the model and the data. Piet et al. (2018) used a maximum likelihood approach to identify model parameters that most closely matched choice data. This is related to the Bayesian estimation approach we used to fit parameters of the nonlinear and linear models. We obtained similar results by minimizing the expected 0/1-loss, which biases towards less variable models, especially for models with strong sensory noise (Fig. 5). A more careful approach to fitting model parameters should also consider

for all $n \in \mathbb{N}$ and all $n\Delta t < t < (n+1)\Delta t$. In the following, our discrete-time evidence accumulation equations are embedded in continuous-time via the correspondence given by Eq. (14). As $\Delta t \to 0$ the resulting equations apply to the original state process x_t in virtue of the sampled-time approximation just described.

Just as in Eq. (1), the log-posterior odds ratio in discrete-time is:

$$y_n := \log \frac{P(x_n = x^R \mid s^R(n), s^L(n))}{P(x_n = x^L \mid s^R(n), s^L(n))}.$$

Hence, equations (A.3) and (B.1) from the appendix of Veliz-Cuba et al. (2016) hold in our context:

$$y_n - y_{n-1} = \log \frac{f_t^R(n)}{f_t^L(n)} + \log \frac{1 + h(t) + e^{y_{n-1}}}{1 + h(t) + e^{y_n}}.$$

In addition, we use the approximation $\log(1+z) \approx z$ for small $|z|$, since $0 < t < 1$, so that:

$$y_n = \log \frac{f_t^R(n)}{f_t^L(n)} + h(t) \sinh(y_n).$$

Taking the limit $\Delta t \to 0$ yields the ODE:

$$\frac{dy_t}{dt} = \left[\sum_{i=1} \delta(t - t_i^R) - \sum_{j=1} \delta(t - t_j^L)\right] + h(t)\sinh(y_t),$$

or the equivalent rescaled version

$$\frac{dy_t}{dt} = \sum_{i=1} \delta(t - t_i^R) - \rho$$

Therefore to obtain the unconditional expectation and variance for N, we must marginalize using the laws of total expectation and variance with respect to the distribution of epoch times given in Eq. (15). This yields

$$\mathrm{E}[N] = \left(\frac{1}{\lambda_{\text{high}}} - \frac{1}{\lambda_{\text{low}}}\right) \int_0^T p(t)\,dt = \frac{\left(e^{-hT} - 1\right)}{h} \left(\frac{1}{\lambda_{\text{high}}} - \frac{1}{\lambda_{\text{low}}}\right) \tag{16}$$

for the total expectation. Notice that as $T \to \infty$, the expected number of clicks is limited from above by $\lim_{T\to\infty} \mathrm{E}[N] = \left(\frac{1}{\lambda_{\text{high}}} - \frac{1}{\lambda_{\text{low}}}\right) h$. Using the law of total variance we can thus compute

$$\begin{aligned}\mathrm{Var}[N] &= \mathrm{Var}[\mathrm{E}[N \mid \lambda]] + \mathrm{E}[\mathrm{Var}[N \mid \lambda]] = \left(\frac{1}{\lambda_{\text{high}}} - \frac{1}{\lambda_{\text{low}}}\right)^2 \mathrm{Var}[\lambda] + \mathrm{E}[\lambda] \left(\frac{1 - e^{-hT}}{h}\right) \left(\frac{1}{\lambda_{\text{high}}} - \frac{1}{\lambda_{\text{low}}}\right) \\ &= \frac{\left(1 - e^{-hT}(hT + 1)\right)}{h^2} \left(\frac{1}{\lambda_{\text{high}}} - \frac{1}{\lambda_{\text{low}}}\right)^2 + \frac{\left(1 - e^{-hT}\right)}{h} \left(\frac{1}{\lambda_{\text{high}}} - \frac{1}{\lambda_{\text{low}}}\right).\end{aligned} \tag{17}$$

Plugging Eq. (16) and (17) into the expression for $\mathrm{SNR}_h^T = \mathrm{E}[N] \Big/ \sqrt{\mathrm{Var}[N]}$ yields

$$\mathrm{SNR}_h^T = \sqrt{\frac{\left(1 - e^{-hT}\right)^2 \left(\frac{1}{\lambda_{\text{high}}} - \frac{1}{\lambda_{\text{low}}}\right)^2}{\left(1 - e^{-hT}(hT + 1)\right) \left(\frac{1}{\lambda_{\text{high}}} - \frac{1}{\lambda_{\text{low}}}\right)^2 + h\left(1 - e^{-hT}\right) \left(\frac{1}{\lambda_{\text{high}}} - \frac{1}{\lambda_{\text{low}}}\right)}}. \tag{18}$$

low 1 TX@tRUNRP 1@tf[@SNXSV@P@td@[H hI 1 WHtIUU@tf@UNWUW@P@tTYW@tXPY@t 1@[HVIUXHyieY 1@gIM 1 f@WNtjOfSSY@Wd@[H»I]tjOfS

Recalling our de

C | DIFFUSION APPROXIMATION

Here we demonstrate the diffusion approximation of the normative model for the dynamic clicks task, Eq. (2) in the limit of large Poisson rates $_{high}$ and $_{low}$. Diffusion approximations for jump processes have been addressed by Lánskỳ (1997), and Richardson and Swarbrick (2010) who studied the impact of shot noise and pulsatile synaptic inputs on integrate-and-fire models. Following this work, we note that the difference of the click streams in Eq. (2) can be approximated by a drift-diffusion process with matched mean, ^{1} hi $_{low}$º, and variance, ^{f1} hi ý $_{low}$º. This results in the following stochastic differential equation (SDE) for the approximation $\hat{y}_t$:

$$d\hat{y}_t = g^{1}t^{o1}\,{}_{high}\;{}_{low}{}^{o}dt\;\acute{y}\;\sqrt{\;{}_{high}\;\acute{y}\;{}_{low}}\,dW_t\;\;fh\sinh^{1}\hat{y}_t{}^{o}dt, \tag{21}$$

where $g^{1}t^{o} = \text{sign}\,\text{»}\,{}_{R}{}^{1}t^{o}\;\;{}_{L}{}^{1}t^{o\%}$ and dW_t is the increment of a Wiener process. Note the resulting nonlinear drift-diffusion model is similar to the normative models presented in (Glaze et al., 2015; Veliz-Cuba et al., 2016). The SNR of the signal in Eq. (21) can be associated with the mean divided by the standard deviation in an average-length epoch. Fixing this SNR leads to the relations in Eq. (6). Importantly, the signal in Eq. (21) is characterized entirely by its mean and variance, so we expect that the performance of the model can be directly associated with the SNR. Note, however, that Eq. (21) will only be valid for $_{high,\ low}$,. Otherwise, one must consider the effects of higher order moments of the click streams, and a prediction of performance purely based on the SNR will break down (Fig. 1D, Inset), since higher order statistics likely shape response accuracy in these cases.

D | MODEL IDENTIFICATION

We fi

where $_i$, $_j$ N¹ , º, revealing y_{Tk} is simply the sum of i.i.d. normal random variables scaled by exponential decay. Conditioning on the clicks T_k, then y_{Tk} is normally distributed y_{Tk} j T_k, with expectation and variance

$$E_k := E\ y_{Tk}\,\mathrm{j}\,T_k, \quad = \begin{bmatrix} \sum_{i=,}^{N_R^k} e^{\ {}_1T^k\ t_R^i\ {}^o} \\ \sum_{j=,}^{N_L^k} e^{\ {}_1T^k\ t_L^j\ {}^o} \end{bmatrix},$$

$$E$$
$$V$$

Our method focuses on exploiting the likelihood function $P(D \mid \theta)$. We have,

$$P(D \mid \theta) = P(T_{,:N}, d_{,:N} \mid \theta) = P(d_{,:N} \mid T_{,:N}, \theta) P(T_{,:N} \mid \theta) = P(d_{,:N} \mid T_{,:N}, \theta) P(T_{,:N}),$$

where the last step comes from the fact that the clicks trains are independent of the discounting parameter used by the decision-making model[8]. From there, we remark that the choice data are conditionally independent on the clicks stimulus and the discounting parameter. Thus,

$$P(d_{,:N} \mid T_{,:N}, \theta) = \prod_{k=1}^{N} P(d_k \mid T_k, \theta).$$

Therefore we can rewrite Eq. (25) as:

$$P(\theta \mid D) \propto P(\theta) \prod_{k=1}^{N} P(d_k \mid T_k, \theta). \tag{26}$$

We use uniform priors for θ, over a finite interval $[0, a]$. In this context, the problem of computing the posterior distribution of θ reduces to assessing the likelihoods of the decision data on each trial, $P(d_k \mid T_k, \theta)$, for a range of θ-values spanning the interval $[0, a]$. In practice, we picked $a = 10$ when fitting the linear model and $a = 1$ when fitting the nonlinear model. Finally, note that for numerical stability reasons, our algorithms actually sum log-likelihood values, as opposed to multiplying probability values. Relegating the θ-independent prior into a normalization constant C, Eq. (26) becomes, in the log-domain:

$$\log P(\theta \mid D) = C + \sum_{k=1}^{N} \log P(d_k \mid T_k, \theta), \qquad \theta \in [0, a]. \tag{27}$$

E | MINIMIZING 0/1-LOSS IN A BERNOULLI RANDOM VARIABLE

Consider a simple stochastic binary decision-making model in which we ignore the specifics of evidence sources, as in Pesaran and Timmermann (1992). We that in this case the 0/1-loss function also leads to biased estimates. This result has been pointed out in previous work in which parameter fitting results have been compared between Bernoulli random variables fit with the 0/1-loss function as opposed to maximum likelihood estimators (Friedman, 1997; Friedman et al., 2001).

Consider a Bernoulli random variable B_* with success probability p_* generating the reference choices, and the fit Bernoulli model B_f with success probability p_f. Minimizing the log-likelihood loss function recovers $p_f = p_*$ in the limit of a large number of trials $N \to \infty$. In this limit, given p_f, we have that the expected loss measured by the negative log-likelihood is

$$L_{LL}(d \mid p_f) = -[p_* \log p_f + (1 - p_*) \log(1 - p_f)], \tag{28}$$

which is minimized[9] at $p_f = p_*$, the mean of B_*. Thus, the parameter from the reference model is recovered, as the

[8] We remind the reader that we operate a distinction between the discounting parameter of the decision maker and the hazard rate used to produce the data.

[9] Note Eq. (28) is the cross-entropy between B_* and B_f.

Bernoulli random variable satisfies the requirements for the MLE to be consistent (Wald, 1949).

On the other hand, if we fit the parameter p_f by minimizing the expected 0/1-loss function, in the limit of N ! 1 trials, the expected loss is

$$\mathrm{L}^{\acute{Y}}\bullet\bullet_{,}\,{}^{1}d\,\mathrm{j}p_f{}^{\mathrm{o}} = \;, \quad \mathrm{P}^{1}B_{,} = B$$

10,000 trials. The fitting algorithm NNf@1NP@1NPNiP1W@VSRn1PVNkglementatioVSRn1PVofPNiP1Wt1YXRn1PVB

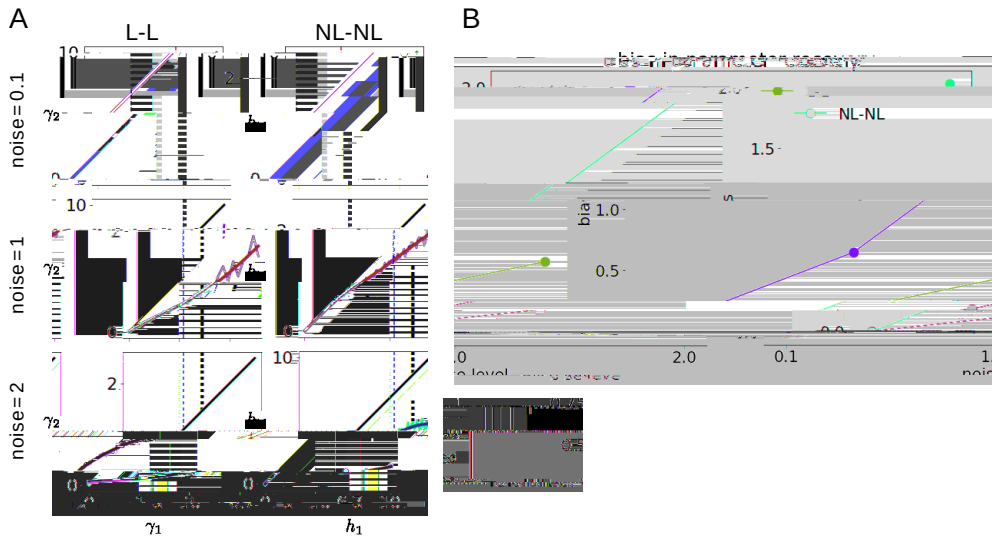


FIGURE 6 Bias in parameter recovery as a function of sensory noise. A: Recovered discounting parameter from the fits as a function of the reference discounting parameter used to produce the initial decision data. Top and bottom rows are reproductions of Fig. 5 while the middle row is for an intermediate level of sensory noise. The actual fit parameters (golden) were smoothed (green) in order to compute the bias in panel B for the reference discounting parameters indicated by the red dotted lines. The black diagonal indicates the identity line, which would correspond to perfect parameter recovery. B: Bias in parameter recovery Eq. (29) as a function of sensory noise, for the two model pairs (L-L in blue and NL-NL in golden).

ACKNOWLEDGEMENTS

We thank Gaia Tavoni, Alex Filipowicz, and Alex Piet for helpful feedback on an earlier version of this manuscript. Some computations for this manuscript were done using the Ohio Supercomputer Center (1987).

REFERENCES

Adelson, E. H. and Bergen, J. R. (1985) Spatiotemporal energy models for the perception of motion. *JOSA A*, 2, 284–299.

Balasubramanian, V. (1997) Statistical inference, Occam's razor, and statistical mechanics on the space of probability distributions. *Uier.,*

Born, R. T. and Bradley, D. C. (2005) Structure and function of visual area mt. *Annu. Rev. Neurosci.*, 28, 157–189.

Britten, K. H., Shadlen, M. N., Newsome, W. T. and Movshon, J. A. (1992) The analysis of visual motion: a comparison of neuronal and psychophysical performance. *J Neurosci.*, 12, 4745–4765.

Brody, C. D. and Hanks, T. D. (2016) Neural underpinnings of the evidence accumulator. *Curr. Op. Neurobiol.*, 37, 149–157.

Brunton, B. W., Botvinick, M. M. and Brody, C. D. (2013) Rats and humans can optimally accumulate evidence for decision-making. *Science*, 340, 95–8.

Buesing, L., Bill, J., Nessler, B. and Maass, W. (2011) Neural dynamics as sampling: a model for stochastic computation in recurrent networks of spiking neurons. *PLoS Comput. Biol.*, 7, e1002211.

Cappé, O., Moulines, E. and Rydén, T. (2005) Inference in hidden Markov models. *Springer Series in Statistics*, 652.

Dai, B., Ding, S., Wahba, G. et al. (2013) Multivariate bernoulli distribution. *Bernoulli*, 19, 1465–1483.

Faisal, A. A., Selen, L. P. and Wolpert, D. M. (2008) Noise in the nervous system. *Nat. Rev. Neurosci.*, 9, 292–303.

Fischer, W. and Meier-Hellstern, K. (1992) The Markov-modulated Poisson process (MMPP) cookbook. *Perform. Eval.*, 18, 149–171.

Friedman, J., Hastie, T. and Tibshirani, R. (2001) *The elements of statistical learning*, vol. 1, chap. 7: Model Assessment and Selection. Springer series in statistics New York, NY, USA.

Friedman, J. H. (1997) On bias, variance, 0/1—loss, and the curse-of-dimensionality. *Data Min. Knowl. Discov.*, 1, 55–77.

Gardiner, C. (2009) *Stochastic methods*, vol. 4. Springer Berlin.

Geisler, W. S. (2003) Ideal observer analysis. In *The Visual Neurosciences* (eds. L. Chalupa and J. Werner), 825–837. Boston: MIT Press.

Glaze, C. M., Filipowicz, A. L., Kable, J. W., Balasubramanian, V. and Gold, J. I. (2018) A bias–variance trade-off governs individual differences in on-line learning in an unpredictable environment. *Nat. Hum. Behav.*, 2, 213.

Glaze, C. M., Kable, J. W. and Gold, J. I. (2015) Normative evidence accumulation in unpredictable environments. *Elife*, 4.

Gold, J. I. and Shadlen, M. N. (2002) Banburismus and the brain: decoding the relationship between sensory stimuli, decisions, and reward. *Neuron*, 36, 299–308.

— (2007) The neural basis of decision making. *Ann. Rev. Neurosci.*, 30.

Hanks, T. D., Kopec, C. D., Brunton, B. W., Duan, C. A., Erlich, J. C. and Brody, C. D. (2015) Distinct relationships of parietal and prefrontal cortices to evidence accumulation. *Nature*, 520, 220–223.

Hawkins, G., Wagenmakers, E., Ratcliff, R. and Brown, S. (2015a) Discriminating evidence accumulation from urgency signals in speeded decision making. *J Neurophysiol.*, 114, 40–47.

Hawkins, G. E., Forstmann, B. U., Wagenmakers, E.-J., Ratcliff, R. and Brown, S. D. (2015b) Revisiting the evidence for collapsing boundaries and urgency signals in perceptual decision-making. *J Neurosci.*, 35, 2476–2484.

Huk, A. C. and Shadlen, M. N. (2005) Neural activity in macaque parietal cortex reflects temporal integration of visual motion signals during perceptual decision making. *J Neurosci.*, 25, 10420–10436.

Krajbich, I. and Rangel, A. (2011) Multialternative drift-diffusion model predicts the relationship between visual fixations and choice in value-based decisions. *Proc. Natl. Acad. Sci. U.S.A.*, 108, 13852–13857.

Langhaar, H. L. (1980) *Dimensional analysis and theory of models*. R.E. Krieger Publishing: Huntingdon, NY.

Lánskỳ, P. (1997) Sources of periodical force in noisy integrate-and-fire models of neuronal dynamics. *Phys. Rev. E*, 55, 2040.

Latimer, K. W., Yates, J. L., Meister, M. L., Huk, A. C. and Pillow, J. W. (2015) Single-trial spike trains in parietal cortex reveal discrete steps during decision-making. *Science*, 349, 184–187.

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

Odoemene, O., Pisupati, S., Nguyen, H. and Churchland, A. K. (2018) Visual evidence accumulation guides decision-making in unrestrained mice. *J Neurosci.*, 38, 10143–10155.

Ohio Supercomputer Center (1987) Ohio Supercomputer Center. `http://osc.edu/ark:/,Š…Š†/f†s,phˆ„`.

Park, H., Lueckmann, J.-M., Von Kriegstein, K., Bitzer, S. and Kiebel, S. J. (2016) Spatiotemporal dynamics of random stimuli

Snyder, D. L. (1975) *Random point processes*. John Wiley and Sons.

Reviews and Responses

Publication Decision 1 from Neurons, Behavior, Data analysis, and Theory on July 27, 2019

Editorial board's determination: Revise and resubmit

Comments from the editor. Sorry again for the long delay. The reviewer is more or less happy with the manuscript, there is a list of suggestions, which I would like to ask you to pay careful attention to before the paper can be accepted.

Reviewers' comments are italicized. Our responses are in plain text. Changes to the manuscript are in blue.
Comments to the author.

as discussed in Piet et al (2018). In general, the linear model will not adapt well to interspersed trials with different parameters, whereas the normative model will be more robust. This is now explained in the following paragraphs we have added to the end of Section 4:

The preceding point is illustrated by the following thought experiment. Assume a subject is extensively trained on a fixed set of task parameters: $S^{ref} = f, {}^{1}{}_{low}, {}_{high}{}^{o} = {}^{1}f, ‰.†{}^{o}\text{Hz}, h = , \text{Hz}$ (peak of the red curve in Fig. 3D). We then introduce some trials with different click rates, say $S^{new} = †$ with ${}^{1}{}_{low}^{new}, {}_{high}^{new o} = {}^{1}, f, †„{}^{o}\text{Hz}$ and $h = , \text{Hz}$, chosen so that $= \log{}^{1}{}_{high} \bullet {}_{low}{}^{o}$ is constant across the two conditions. We denote by $\text{Acc}_{lin}{}^{1}S^{o}$ and $\text{Acc}_{norm}{}^{1}S^{o}$ the accuracies of an observer using the linear and normative models on trials with a given S. Since the subject was trained on click rates that correspond to S^{ref}, their discounting strategy will be adapted to these values. Note that the ratio between $\text{Acc}_{norm}{}^{1}S^{o}$ and $\text{Acc}_{lin}{}^{1}S^{o}$ when $h = ,$ is the red curve in Fig. 3D. Since the ratio between $\text{Acc}_{norm}{}^{1}S^{ref o}$ and $\text{Acc}_{lin}{}^{1}S^{ref o}$ is near 1, the linear and normative models cannot be distinguished at S^{ref}. However, a subject using the normative model tuned at S^{ref}, will still perform optimally at $S \neq S^{ref}$, if and h are held constant. On the other hand, a linear model optimized at S^{ref}, will no longer be optimal at $S \neq S^{ref}$. This distinction is captured by the drop in the accuracy ratio along the red curve in Fig. 3D.

We can quantify the distinction between the two models by their relative difference:

$$\frac{\text{Acc}_{norm}{}^{1}S^{new o} \quad \text{Acc}_{lin}{}^{1}S^{new o}}{\text{Acc}_{norm}{}^{1}S^{new o}} = \bullet.\bullet†.$$

More generally, for any decision making model, we may define the quantity

$$D_{model}{}^{1}S^{o} := \frac{\text{Acc}_{norm}{}^{1}S^{o} \quad \text{Acc}_{model}{}^{1}S^{o}}{\text{Acc}_{norm}{}^{1}S^{o}}$$

which will equal 0 if the model used is the normative one. If we compute $D_{model}{}^{1}S^{o}$ using responses from a real subject, one can generate curves such as those in Fig. 3D. If the curves are not constant (equal to 1), this would suggest the subject is not using an optimal model. Furthermore, a single value of S^{new} for which $D_{model}{}^{1}S^{new o} \neq \bullet$ provides evidence that the model is not optimal.

Minor comments:

Introduction: when the term 'nonlinear model' is first used, the authors haven't yet clarified that here the term is synonymous with 'normative model.' Also, some readers might not be familiar with what a 0/1 loss function is, making that part of the Intro a little unclear for them.

It is important to note that the nonlinear model is only normative when the discounting parameter is tuned exactly, and this is why we used this specific phrasing. To clarify this, we have added the word 'nonlinear' in parentheses when mentioning the normative model in the prior sentence. We also added the following footnote: The 'nonlinear' model here refers to the family of models obtained by tuning the discounting rate away from the value defining the normative model. This detuning results in a model that is not normative.

We describe the 0/1-loss function now in the sentence following its introduction: The 0/1-loss function gives a one unit penalty on trials in which the decision predicted by the model and the data disagree, and no penalty when they agree. Therefore, minimizing this loss function leads to models that best match the trial-to-trial responses in the data rather than the response accuracy.

Page 4, last paragraph: You need the word 'alone': the phrase 'kappa does not predict the response accuracy' should probably be

Thanks for pointing this out. We have modified the text as suggested.

Although the confidence intervals do not seem to contain the true parameter value, are the mismatches statistically significant in Figure 4CD? That is, are the parameter values significantly greater than the true parameter value?

We would like to note that the whiskers of each box in Fig. 4C,D do not represent confidence intervals. Instead, they represent the interval of values that are not considered outliers. More specifically, if q_1, q_3 are the first and third quartiles respectively, then the whiskers define the interval: $[q_1 - 1.5(q_3 - q_1), q_3 + 1.5(q_3 - q_1)]$.

Instead of testing the hypothesis that the MAP estimates are different than the $_{true}$ value, we provide a summary statistic for the training datasets of size 500 trials. We have added the following sentences to the end of section 6: For reference datasets of size 500, 98% of the 500 MAP estimates in the L-NL fits lie strictly above $_{true}$, versus 50.4% for the corresponding L-L fits. Similarly, 86.6% of the estimates in the NL-L fits lie strictly below h_{true}, versus 44.2% for the corresponding NL-NL fits. We believe this is more informative than a p-value.

Publication Decision 2 from Neurons, Behavior, Data analysis, and Theory on August 26, 2019

Editorial board's determination: Accept

Comments from the editor.