https://doi.org/10.1007/s10827-019-00733-5

subjects use approximately normative decision strategies, and when and how they fail to do so, can be computationally challenging. For instance, one may wish to study how a subject's estimate of the environmental timescale impacts their response accuracy, or how heuristic evidence-discounting strategies compare to optimal ones (Glaze et al. 2018; Radillo et al. 2019). To address these questions, previous work has primarily relied on Monte Carlo simulations (Veliz-Cuba et al. 2016; Piet et al. 2018), which can be computationally expensive.

Here, we show how to reframe dynamic decision models by deriving corresponding differential Chapman-Kolmogorov (CK) equations (See Eq. (6)). This approach allows us to quickly compute observer beliefs and performance, and compare models. Realizations of our models are described by stochastic differential equations with a drift term that switches according to a two-state Markov process, and leak terms that discount evidence. To describe these models using CK equations, we treat the switching process as a source of dichotomous noise, and condition on its state to track conditional belief densities. These methods allow us to quickly answer questions about how characteristics of optimal models and their approximations vary across ranges of task parameters.

Nonlinear, normative models can thus be compared to approximate linear and cubic discounting models, models with internal noise, and explicitly solvable bounded accumulation models with no flux boundaries. These models all can obtain near-optimal response accuracy, but each has very different belief distributions. This suggests that subject confidence reports could be used to distinguish subject decision strategies in data.

Detailed analyses, including belief distribution calculations, can be performed rapidly and accurately with our methods, allowing us to see *why* each approximate model performs better at different task difficulty levels. Monte Carlo methods fare much worse in terms of computation time and accuracy (See Fig. 9). Our methods also extend to tasks with pulsatile evidence, where drift and diffusion are replaced by jump terms. Our work thus demonstrates how partial differential equation descriptions of stochastic decision models, previously successful in understanding decision making in static environments (Busemeyer and Townsend 1992; Moehlis et al. 2004; Bogacz et al. 2006), can be extended to dynamic environments.

2

We begin by considering the dynamic RDMD task (Glaze et al. 2015; Veliz-Cuba et al. 2016); an observer looks at a screen of dots which move, on average, right or left. The average direction of motion, which we call the state $s(t)$, switches in time between states s_+ (right-moving) and s_- (left-moving) as a two-state continuous time Markov process with hazard rate h, so $P(s(t+\Delta t) \neq s(t)) = h\Delta t + o(\Delta t)$.[1]

an observer whose belief is represented by Eq. (2). These quantities can be changed by varying psychophysical task parameters (Glaze et al. 2015, 2018

How does the response accuracy of an observer whose belief is described by Eq. (2) change when $\tilde{h}$ is mistuned? Veliz-Cuba et al. (2016) addressed this question using Monte Carlo sampling, but computational costs prevented a complete answer. Since Eq. (2) is rescaled, we take $h = 1$ for the remainder of our investigation; all other cases can be recovered by rescaling time. Before asking how changing $\tilde{h}$ alters accuracy, we first briefly mention how accuracy varies with *evidence strength*, fixing $\tilde{h} = h = 1$. The density $p_S(y, t)$ computed using Eq. (6) rapidly converges to the stationary solution, with most of its mass above zero (Fig. 2a). As m, increases, more mass of the stationary distribution moves to positive values (Fig. 2b), but the total mass, equal to $\lim_{T \to \infty} \mathrm{Acc}(T)$, always saturates at a value less than 1 due to discounting and state switching.

When the observer misestimates the hazard rate, $\tilde{h} \neq h$, we expect the long term accuracy to suffer. Effects on accuracy are subtle, but do follow a general pattern: overestimating the hazard rate ($\tilde{h} > h$) causes the observer to discount prior evidence too strongly, resulting in more errors driven by observation noise (Fig. 2c). On the other hand, observers that underestimate the hazard rate ($0 < \tilde{h} < h$) discount evidence too slowly and are less adaptive to change points. Change point triggered response accuracy plots show both of these trends (Fig. 2d). Accuracy obtains a lower ceiling value during longer epochs without environmental changes when the discounting rate $\tilde{h}$ is too high. On the other hand, accuracy recovers more slowly following changes when the discounting rate $\tilde{h}$ is too low. This bias-variance tradeoff is common to binary choice experiments in dynamic environments (Glaze et al. 2015, 2018): Low and aim to tune λ so $\mathrm{Acc}_{\infty}(\lambda)$ is maximized. Second, to quantify the distance between the belief distributions, we compute the Kullback-Leibler (KL) divergence

$$D_{\mathrm{KL}}\left(\overline{p}_s^N \,||\, \overline{p}_s^L\right) = \int_{-\infty}^{\infty} \overline{p}_s^N(y) \ln\left[\frac{\overline{p}_s^N(y)}{\overline{p}_s^L(y;\lambda)}\right] \mathrm{d}y \qquad (10)$$

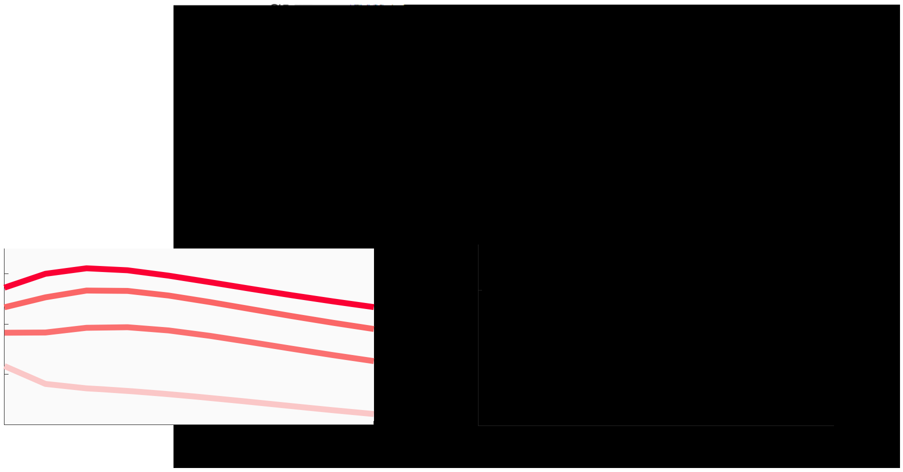
between the stationary normative distribution, $\overline{p}_s^N(y)$, obtained from Eq. (6), and the stationary distribution of the linear approximation, $\overline{p}_s^L(y;\lambda)$, obtained from Eq. (8). While it is possible for models to have nearby

the value of $\tilde{h}$ that maximizes response accuracy decreases as D

To compute steady state accuracy of the bounded accumulator model, we can integrate Eq. (17) to obtain a formula that depends on m and β:

$$\begin{aligned}\mathrm{Acc}_{\infty}(\beta) &= \int_0^{\beta} \bar{p}_s(y)\,\mathrm{d}y = C_1\beta \\ &+ \frac{C_2\left(1 - e^{-q\beta}\right)}{q}\left[e^{q\beta} + (mq - (m +\right.\end{aligned}$$

probability density $\bar{p}_s^C(y)$ and that of the normative model, $\bar{p}_s^N(y)$, though the model is more complex (Friedman et al. 2001



insights into how and why organisms fail to perform optimally (Geisler 2003). Investigating optimal models and their approximations requires simulations across large parameter spaces; these necessarily require rapid simulation techniques to obtain refined results. Efficient computational methods are therefore essential for the analysis of evidence accumulation models, and their application to experiment design.

Using differential CK equations to describe ensembles of decision model realizations speeds up computation and describes the time-dependent probability density of an observer's belief. Thus, traditional metrics of performance (e.g., accuracy) and other less common model comparison metrics (KL divergence) can be computed rapidly. This opens new avenues for comparing normative and heuristic decision making models, and for determining task parameter ranges to distinguish models. There is also hope that in high throughput experiments, sufficient data could be collected to specify subject confidence distributions, which could be fit, or compared to model predictions (Piet et al. 2019).

Doubly stochastic and jump-diffusion models appear in a number of other contexts in neuroscience and beyond (Hanson 2007; Horsthemke and Lefever 2006). For instance, dichotomous and white noise have been included in linear integrate and fire (LIF) models to model voltage or channel fluctuations (Droste and Lindner 2014; 2017; Salinas and Sejnowski 2002). The interspike interval statistics of these models can be analyzed directly by work considered more general forms of non-stationarity, their mathematical treatments focused on decision time statistics for single trials, rather than trial ensembles, as we studied here using our differential CK approach. Our mathematical approach allowed us to compare the performance of a broad array of evidence accumulation models across task parameter space.

Our study has focused on models of an observer has a fixed estimate of the discounting rate, and does no further learning of the change rate. Previous studies by Radillo et al. (2017) and Glaze et al. (2018) derived ideal observer models capable of inferring the change rate of a dynamic environment, and showed approximations can perform nearly as well in some circumstances. It is possible to formulate the ensemble dynamics of such models using differential CK equations, but the state space can be high-dimensional as the observer must track probabilities over possible change rate values h. In such cases, numerical methods for solving high-dimensional partial differential equations are needed to make solving the ensemble equation in this way worthwhile. In ongoing work, we have developed ways of quantifying the rate at which learning occurs in these models (Eissa et al. 2019), and also identified when it is useful to apply this differential CK equation approach to analyzing model performance. These results will be reported elsewhere.

In recent years, decision-making models and experiments have been developed to incorporate more naturalistic scenarios in which the environment changes in fluid yet predictable ways. The associated normative models can be complex, and efficient simulation techniques are important for evaluating performance across different models and so we can compute the limits of Eq. (29) as

$$g(t) = \lim_{t \to 0} g_t(t) \square \pm {}^{2\mu^2}$$