Decision-time statistics of nonlinear diffusion models: Characterizing long sequences of subsequent trials

Sebastian Vellmer[a,b], Benjamin Lindner[a,b]

[a]*Bernstein Center for Computational Neuroscience, Berlin 10115, Germany*
[b]*Physics Department, Humboldt University Berlin, Berlin 12489, Germany*

Abstract

Cognitive models of decision making are an important tool in studies of cognitive psychology and have been successfully used to fit experimental data and to relate them to neurophysiological mechanisms in the brain. One of the most important models for binary decision making is the diffusion-decision model (DDM) in which a diffusion process that models the accumulation of perceived evidence yields the decisions upon reaching one of two thresholds. Due to its simplicity, the model is analytically tractable and has been used to bridge the gap between implementations of decision making in neurobiologically plausible neural networks and experiments. However, biologically realistic network models exhibit nonlinear dynamics that yield via mean-field-reduction techniques a nonlinear DDM for which analytical solutions and proper numerical tools in general are not known. Furthermore, although often agents have to make a number of subsequent decisions, the statistics of such sequences of decisions (containing information on whether the decisions are correct or incorrect and on their timing) are so far poorly understood. Here we introduce the decision trains, sequences of negative or positive spikes at the decision times with the sign corresponding to the correctness of the decision. The decision trains enable a proper characterization of experiments in which many trials are performed consecutively. For the principal reference case of independent decisions (renewal statistics), we derive relations between the second-order statistics of the decision trains (i.e. their power spectra) and the response-time densities. Most importantly, we extend an efficient numerical procedure for spiking neuron models, the threshold-integration method, to determine the temporal statistics of nonlinear DDMs. The threshold-integration method provides the temporal statistics, i.e. decision rates, decision-time densities and the decision-train power spectra. Moreover, the procedure is used for the calculation of the linear response to a sinusoidal modulation. We compare all results with direct simulations of the stochastic model.

Keywords: Decision making, diffusion-decision model, stochastic decision models

1. Introduction

Real life tasks often require sequences of decisions to be made and, thus, decision making on a low level of cognition is a crucial ability for higher animals to survive. Many studies in cognitive psychology are devoted to the question how animals, and particularly primates, make decisions and how the process is implemented in the brain [22, 26, 8, 21]. In combination with decision making experiments, abstract cognitive models are a useful tool to understand how accumulated noisy perceptual representations yield decisions [16].

In the standard model for binary decision making, the diffusion-decision model (DDM) that was introduced by Ratcliff [48] (see also Ratcliff and Smith [52] for comparison with other models), evidence for a decision is modeled as temporally uncorrelated Gaussian noise and a constant drift. Thus, within this model, the accumulation of perceived evidence is a diffusion process. Decisions are made when the process exceeds one of the thresholds that restrict the process. The model and some variations of it were extensively used in studies of decision making of primates and humans (see e.g. Ratcliff et al. [55], Ratcliff [49], Palmer et al. [44], Ratcliff and McKoon [50], Cisek et al. [10], McKoon and Ratcliff [37], Johanson et al. [28]) and furthermore, combined with direct measurements of neural activity in monkeys [63, 62]. It was also used to gain theoretical insights into the optimal decision process with reward maximization [4, 3]. Due to the simplicity of the model, it is analytically tractable and the response-time distributions are given by the solution of a first-passage-time problem [48].

Several extensions of the model were introduced in order to account for additional features of decision making and obtain more realistic modeling e.g. trial-to-trial variability of the starting point [51] or of the non-decision time [54], multi-alternative tasks [38, 31] or urgency signals and collapsing boundaries [10, 25]; see the detailed review by Ratcliff et al. [53] for a history of the model. Many of these extensions yielded temporally inhomogeneous but spatially homogeneous diffusion processes for which the-

Email addresses: sebastian.vellmer@bccn-berlin.de (Sebastian Vellmer), benjamin.lindner@physik.hu-berlin.de (Benjamin Lindner)

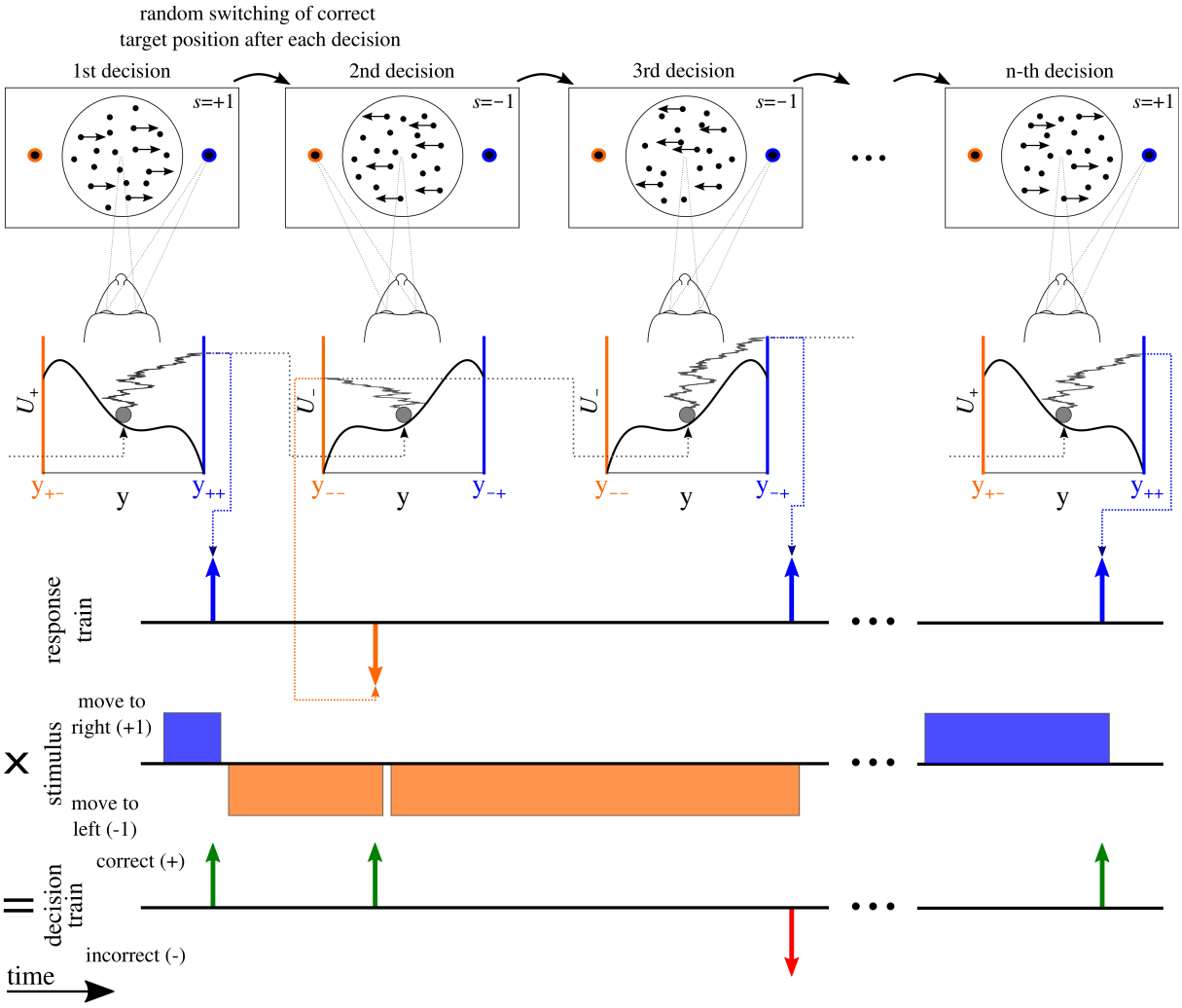

Figure 1: **Experimental paradigm related to model and decision train** (color online). In this paper novel statistics for a two-alternative forced-choice task are suggested, here illustrated for a motion discrimination task experiment (used in e.g. [62]). The subject, here a monkey, is in front of a screen that shows a random cloud of dots of which some perform random walks while the other dots move coherently towards one of the two opposing targets on the left (stimulus $s = -1$) and right ($s = +1$) hand sides indicated by the orange and blue circles, respectively. As soon as the subject perceives the direction of the coherent motion, it makes the decision by a saccade towards the respective target. The variable y accumulates perceived evidence towards the left or right target with positive and negative increments, respectively. For a given stimulus, i.e. a certain ratio of dots moving to the left or right target with a certain speed, this stochastic process of evidence accumulation may analogously described by an overdamped Brownian particle (grey circle) in potentials $U_-(y)$ and $U_+(y)$, respectively as sketched below the monkey's head. A decision is made when $y(t)$ hits one of the thresholds y_{s-} or y_{s+} indicated by vertical orange and blue lines, respectively. After a decision and possibly a break, the experiment is repeated with a randomly picked coherent motion direction and a relocation of all dots; in our model this would correspond to the reset of x' (here shown from left to right). Experimental parameters such as coherence and motion strength are fixed during the whole experiment (in contrast to e.g. [62]). The sequence of responses may be captured by the response train in which times of the saccades towards the left and right targets are represented by δ-peaks with positive and negative signs, respectively. By multiplication of the response train with the stimulus encoded by +1 for right and -1 for left movement shown below, the decision train is obtained (bottom) in which correct and incorrect decisions are encoded. In this paper we exclusively consider the statistics of this decision train.

oretical tools can be applied in order to calculate the first-passage-time densities, also referred to as response-time distributions in this context (see for instance [58] and [72] for review on integral equation methods).

DDMs are also interesting from another perspective: they may bridge the gap between experiments of decision making and neurobiologically motivated models that describe how the decision making process is implemented in the brain. In such models, evidence, in form of sensory information, enters competing neural networks. For instance, information encoded by (approximately Poissonian) spike trains of neurons is accumulated by a neural population; this accumulation can be approximated by an Ornstein-Uhlenbeck process [73, 74]. A decision is made following the winner-take-all principle in which one population is very active and inhibits the other one that is silent [81, 82, 35, 47]. The dynamics of such networks can be reduced to one-dimensional DDMs by the use of mean-field

techniques [82]. In contrast to the original DDM with only constant drift and noise, the nonlinear network dynamics also yield nonlinear dynamics of the DDM [64]. Apart from that, also the consideration of a changing environment yields a nonlinear term in the DDM [78]. For the nonlinear diffusion processes no analytical solutions are known and it is also not straight forward how to apply the integral equation techniques developed for processes with constant or linear drift functions [58]. Here we present an efficient numerical procedure to calculate first-passage-time densities for general nonlinear DDMs that is based on the method of threshold integration by Richardson [59, 60] for white-noise-driven integrate-and-fire neurons.

Experiments related to binary decision times often measure response-time distributions by performing sequential trials with randomly varying setups (see e.g. [41, 23, 62, 49]). In such a procedure, slow transients that may occur due to the loss of the subject's attention and also possible correlations between trials are neglected. As a second innovation in this paper, somewhat in contrast to the usual approach in decision modeling, we consider the statistics of sequences of consecutive decisions, the decision trains. This approach includes the information that is usually captured in response-time densities by the distributions of the inter-decision intervals (IDI) but also goes beyond it.

We explain the novel statistics by means of a thought experiment, a motion discrimination task. In such two-alternative forced-choice experiments, randomly distributed dots are presented on a screen to a subject as shown in Fig. 1. A number of these dots perform one-dimensional random walks in the horizontal direction with predetermined variance of step sizes. The remaining dots are moving coherently towards either the left or the right hand side of the screen. The experimental control parameters are the fraction of coherently moving dots, speed of the motion and step-size variance of the random walks. The subject is placed in front of the screen and has to determine the direction of the coherently moving points and, subsequently, make a decision by a saccade or by pressing a button. Here, one trial of the experiment consists of a sequence of consecutive decisions in which the parameters are fixed but the physical meaning of correct and incorrect, here direction of coherent motion, switches randomly. After each decision all the dots are randomly relocated, but the coherence and motion strength as well as the variance of the random-walk step size remains the same.

Results of such an experiment may be captured in the response train that displays the saccades to the left and right targets by positive and negative spikes at the times at which the respective responses were made. By multiplying the response train with +1 and -1 corresponding to the stimuli to the right and left targets, respectively, the decision train is obtained in which the sign of the spike encodes correct or incorrect decisions. In order to introduce a simple reference model, in this paper we focus only on the decision train and make two simplifying assumptions for the decision making process.

Firstly, we assume that both types of decisions are made symmetrically with respect to their statistics. In other words, we assume that there is no measurable difference in the statistics of the decision times, which implies that, for instance, the probability and the statistics of the decision times for a *left* decision given a left-going stimulus are the same as those for a *right* decision given a right-going stimulus. Secondly, we make the simplifying assumption that all decisions are made independently of each other. If both assumptions are true, the explicit sequence of stimuli is not relevant for the decision-train statistics.

In contrast to the standard statistics of response times, the decision trains can describe the decision process also on time scales much longer than the mean response-times. If breaks between a decision and reset are included, for instance emerging from a random foreperiod (see [23]) or from getting a reward, it would be possible to subtract the duration of the break from the respective IDI; in this case, the time scale of the decision train would not represent the absolute physical time anymore.

If we characterize the decision process as a time series in the form of the trains of correct and incorrect decisions, the most important statistics are the mean values of these trains and their variance distribution in the frequency domain, i.e. their power spectra [30]. Generally, the decision-train power spectra might reveal novel insights into the correlation structure of subsequent decision making processes.

For the important reference case of a decision sequence with independent IDIs, we provide analytical formulas that connect its statistics, namely, the interval distributions from reset to the ensuing decision, between consecutive decisions of the same kind and the power spectra of the decision trains. The decision-train statistics are also suitable to describe how the decisions in such a sequence are influenced by time-dependent signals imposed by the experimenter and lasting over the whole sequence, e.g. a periodic modulation of the velocity in a motion discrimination task. This can be characterized by the linear response of the mean decision train to a known stimulus. Furthermore, the decision train is also useful to describe slow fluctuations in the experiments that emerge due to slow intrinsic changes. Slow noise or non-stationary behavior could be captured in extensions of the model considered here.

The paper is organized as follows: in Sec. 2 we introduce the general nonlinear diffusion-decision model and define the decision trains and their important statistics that are the IDI statistics and the decision-train power spectra; we briefly discuss an analytically tractable case, the original DDM. In Sec. 3 we present how the threshold-integration method by Richardson [60] can be adapted to the nonlinear DDM in order to efficiently determine various statistics: the stationary density of the decision variable, rates, IDI probability densities, the power spectra of the decision trains, and the linear response of the decision trains upon periodic stimulation. We conclude in Sec. 6 with a

summary of our results and a brief outlook on possible extensions of the model and methods.

2. Diffusion-decision model, spike-train statistics, and a first analytically tractable example

2.1. A stimulus-dependent model and its transformation to a correct-incorrect-decision model

For the introduced thought experiment, a stimulus is given by dots moving either to the left ($s = -1$) or to the right ($s = +1$) target (cf. Fig. 1). The subject perceives noisy evidence for the movement direction that is accumulated by the variable y (starting at position $y = 0$ at $t = t_0$) that obeys the stochastic dynamics:

$$\tau_y \dot{y} = f_s(y) + \sigma\sqrt{2\tau_y}\xi(t). \tag{1}$$

Here $\xi(t)$ is Gaussian white noise with zero mean and $\langle\xi(t')\xi(t)\rangle = \delta(t - t')$ (angular brackets denote an ensemble average), representing either distracting fluctuations in the input (set by the experimenter), intrinsic noise in neural receptors and in the neural networks, or a mixture of these noise sources. The time scale of the process is determined by the time constant τ_y. For the sake of generality, the stimulus dependent drift $f_s(y)$ ($f_+(y)$ for $s = +1$ and $f_-(y)$ for $s = -1$) may also depend on the evidence y in a nonlinear fashion.

Decisions of the motion direction are made when sufficient evidence is accumulated and $y(t)$ exceeds one of the thresholds y_{s+} and y_{s-} for right or left motion, respectively, with $y_{s-} < y_{s+}$. Thinking about decision making as an accumulation process, one may suggest that, for instance, the critical amount of accumulated negative evidence for a left decision should not depend on the actual signal, i.e. thresholds should be signal-independent ($y_{--} = y_{+-}$ and $y_{-+} = y_{++}$). However, considering the neurophysically motivated approach, outlined for instance in Ref. [64], the DDM in Eq. (1) arises from a mean-field approximation of two competing neural networks by reduction to a single-variable dynamics. Decisions in these neural-network models correspond to attractor states (which are not explicitely described in the decision model): one population fires at a high rate whereas the other one is silent. In the reduced description of the DDM the boundaries (thresholds) represent points of no return on the way into these attractor states: once the system has reached one of these points, it will go to the attractor state in a very short time and is very unlikely to go back to the starting point. In Appendix A we give a simple example how these points of no return may indeed be affected by the signal (which is a bias signal in the effective potential). Hence, somewhat in contrast to one standard assumption of the DDM framework, we allow for signal-dependent boundaries in our model. Of course, nothing keeps us from considering the standard assumption of signal-independent boundaries. The decision occurring at time t_1, i.e. the response of the model, can be marked by a δ-function $\delta(t-t_1)$ with a prefactor $+1$ or -1 for absorption at the boundary y_{s+} or y_{s-}, respectively. We note that crossing the thresholds y_{++} or y_{--} corresponds to correct decisions whereas crossing y_{+-} or y_{-+} indicates incorrect decisions.

After a non-decision time of size Δ (a refractory period during which y is undefined), the evidence variable is reset to $y = 0$ and a new stimulus s is presented; strictly speaking, the function $f_s(y)$ in Eq. (1) is thus time-dependent due to its dependence on the stimulus. The stimulus however is binary, it may only change in direction but not in strength. The next decision is again made by crossing one of the boundaries and registered by a new response spike at decision time t_2. By repeating this procedure and adding up the response spikes, we obtain the response train shown in Fig. 1. Moreover, by multiplying with the stimulus at the decision, we obtain a train of correct (prefactor $+1$) and incorrect (prefactor -1) decisions, called the *decision train*, that characterizes the *sequence* of decisions.

Under certain symmetry assumptions, the model can still be strongly simplified and in particular the statistics of the decision train do not depend on the specific stimulus applied. Before we come to this, we would like to outline a useful mechanical analogy for the dynamics of the decision model.

The dynamics Eq. (1) can also be viewed as that of an overdamped Brownian particle in a potential $U_s(y)$ with $f_s(y) = -\partial_y U_s(y)$. This instructive analogy reveals where a noiseless evidence variable would go to and whether the evidence variable can reach the boundaries only by fluctuations. The standard DDM would correspond to a Brownian particle sliding down an inclined plane (the inclination determined by the bias signal) and being perturbed by noise such that the particle can also go uphill occasionally. More general potentials (nonlinearities) may correspond to a particle that has to overcome potential barriers or may be accelerated by an increasing slope. In these cases, noise may play a more prominent role than just to somewhat randomize the decision process by introducing the possibility of erroneous decisions and adding jitter to the decision time. For a metastable potential, for instance, the decision variable $y(t)$, initially started close to the potential minimum, will not be able to escape from this attractor at all if there is no noise in the system. We emphasize that the potentials $U_s(y)$ have no further physical meaning beyond their illustrative purpose.

Here we are exclusively interested in the statistics of the decision train. To this end, our model can be considerably simplified if we assume that the statistics for both types of decisions are indistinguishable and the potentials and thresholds are symmetric ($U_+(y) = U_-(-y)$, $y_{++} = -y_{--}$, $y_{+-} = -y_{-+}$). In other words, we assume that there is no statistical difference in making correct decisions for a right stimulus and for a left stimulus. We note that a more basic assumption was already implicit in the original model Eq. (1), namely, that the current decision does not depend on stimuli and decisions made previously

in the sequence.

Instead of the accumulated evidence for left and right motion, we now consider the evidence accumulation for the correct choice x that we define as:

$$x = \begin{cases} y & (s = +1) \\ -y & (s = -1), \end{cases} \tag{2}$$

for which we only need to consider one potential $U(x) = U_+(x) = U_-(-x)$, independent of the applied stimulus.

According to Eq. (1), the dynamics of x are given by:

$$\tau_x \dot{x} = f(x) + \sigma\sqrt{2\tau_x}\xi(t) \tag{3}$$

(the time constant is unchanged $\tau_x = \tau_y$). The above reset rules for $y(t)$ translate to the following decision rule: when $x(t)$ exceeds one of the thresholds $x_c = y_{++} = -y_{--}$ or $x_i = y_{+-} = -y_{-+}$ at time t, a correct or incorrect decision is made, respectively. Subsequently, x is undefined during the non-decision time Δ, afterwards reset to zero, $x_{\rm st} = 0$, and starts evolving again according to Eq. (3). The reset rule can be summarized as follows:

$$\text{if } x(t) < x_i \text{ or } x(t) > x_c: \quad x(t+\Delta) \to 0, \qquad x_i < 0 < x_c. \tag{4}$$

The correct and incorrect decisions at times $t_{c,j}$ and $t_{i,k}$, respectively, correspond to stochastic point processes $d_i(t)$ and $d_c(t)$. To distinguish between both decision trains, we endow the incorrect decision train with a negative and correct decision train with a positive sign. The decision train $D(t)$ introduced above corresponds to the sum of both:

$$\begin{aligned} d_i(t) &:= -\sum_{i=0}^{n_i} \delta(t - t_{i,k}) \\ d_c(t) &:= \sum_{j=0}^{n_c} \delta(t - t_{c,j}) \\ D(t) &:= d_i(t) + d_c(t). \end{aligned} \tag{5}$$

2.2. Statistics of the decision trains

The ensemble average over a decision train equals the number of decisions per unit time, i.e. the decision rates:

$$\begin{aligned} r_c(t) &= \langle d_c(t) \rangle \\ r_i(t) &= -\langle d_i(t) \rangle \\ R(t) &= r_c(t) + r_i(t), \end{aligned} \tag{6}$$

where also the rate of incorrect decisions is defined to be positive. To avoid the repetition, in the following all statistics are introduced only for correct decisions. Due to the symmetry, all formulas hold true if the kind of decision is swapped ($c \leftrightarrow i$).

By definition, the introduced model generates a renewal point process [11], i.e. after each reset the system has the same properties, thus, different response times are uncorrelated. As a consequence, decision trains are fully characterized by the statistics of time intervals between two

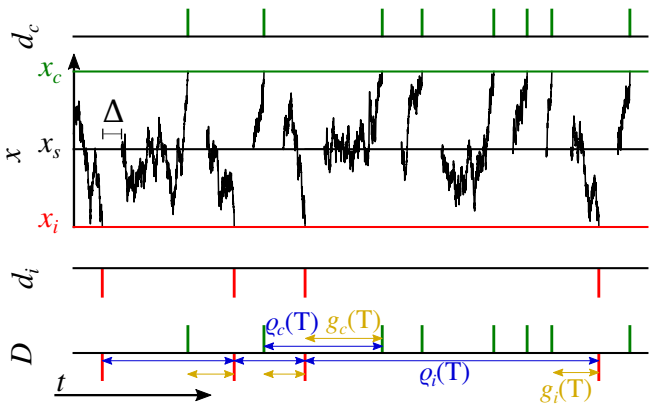

Figure 2: **Diffusion-decision model** (color online). The accumulation of noisy evidence $x(t)$ (black trajectory) follows the nonlinear stochastic dynamics in Eq. (3). If the trajectory exceeds one of the thresholds x_c or x_i, the correct or incorrect decision is made, respectively, and the trajectory is reset to zero.After a decision was made, x is undefined during the non-decision time Δ and, subsequently, evolves again according to Eq. (3). The decisions are represented as Dirac delta-functions at the decision times in the corresponding decision train $d_c(t)$ or $d_i(t)$. The combined decision train is given by $D(t)$ in which the intervals that correspond to $g_i(T)$ and $\varrho_i(T)$ are presented in yellow and blue, respectively.

subsequent decisions. Here we distinguish between the response-time density $g_c(T)$ for correct decisions occurring at a time interval T after the last decision (regardless of its type) and the inter-decision interval (IDI) density $\varrho_c(T)$ for the time between a correct decision and the preceding correct decision. Put differently, $g_c(T)$ is the interval distribution for $D(t)$, taking into account only intervals with a correct decision at the end, and $\varrho_c(T)$ is the IDI density for $d_c(t)$ (see Fig. 2). Note that $g_c(T)$ is not normalized (its integral over time yields the probability of a correct decision) but $\varrho_c(T)$ is normalized. Incidentally, the calculations of the densities are the solutions of first-passage-time problems that have a long history in science and a broad range of applications (see for instance [13, 12, 27, 57]).

The probability density for the time between two subsequent decisions of *any kind* is given by the sum of the individual densities:

$$g(T) = g_c(T) + g_i(T) \tag{7}$$

(this sum is normalized). Given these decision-time densities, one may ask of the connection between $g_c(T)$, $g_i(T)$ and $\varrho_c(T)$, $\varrho_i(T)$. Exploiting the renewal character of the model, we obtain for the Fourier-transformed densities:

$$\tilde{\varrho}_c(\omega) = \frac{\tilde{g}_c(\omega)}{1 - \tilde{g}_i(\omega)} \tag{8}$$

(see Appendix B for the detailed derivation). Here we used the Fourier transform

$$\tilde{F}_T(\omega) = \int_0^T F(t) e^{i\omega t} dt, \tag{9}$$

where the absence of the index T indicates that we take the limit $T \to \infty$.

Despite the independence of the response-time (intervals), the decision trains still have a nontrivial correlation structure that can be captured by their correlation functions, or equivalently in the frequency domain by their power spectra. The power spectrum of a time series $F(t)$ is defined by

$$S_F(\omega) = \lim_{T\to\infty} \frac{\langle |\tilde{F}_T(\omega)|^2 \rangle}{T}. \tag{10}$$

For a renewal spike train, the power spectrum can be expressed by the Fourier transform of the inter-spike intervals [see 75] resulting in our case in the following formulas:

$$s_c(\omega) = r_{c,0} \frac{1 - |\tilde{\varrho}_c(\omega)|^2}{|1 - \tilde{\varrho}_c(\omega)|^2} = r_{c,0} \frac{|1 - \tilde{g}_i|^2 - |\tilde{g}_c|^2}{|1 - \tilde{g}_i - \tilde{g}_c|^2}. \tag{11}$$

Here $r_{c,0}$ and $r_{i,0}$ indicate the stationary rates. The decision-train power spectrum of the combined decision train $D(t)$ describes the second order fluctuation statistics of both, correct and incorrect decisions and, thus, might be helpful for the analysis of decision making experiments. It can be calculated from the individual decision-train power spectra and the decision rates (see Appendix C for the detailed derivation):

$$\begin{aligned} S(\omega) &= s_c(\omega)\left(1 - \frac{r_{i,0}}{r_{c,0}}\right) + s_i(\omega)\left(1 - \frac{r_{c,0}}{r_{i,0}}\right) + R_0 \\ &= \left(1 - \frac{r_{c,0}}{r_{i,0}}\right)\left[s_i(\omega) - \frac{r_{i,0}}{r_{c,0}} s_c(\omega)\right] + R_0, \end{aligned} \tag{12}$$

where we have introduced the total stationary decision rate $R_0 = r_{c,0} + r_{i,0}$. In the second line of Eq. (12), we have recast the terms such a way that the power spectrum appears as a difference of the two single decision-train spectra. Remarkably, in case of equal rates of the decisions ($r_{c,0} = r_{i,0}$), the spectrum of the combined decision train is spectrally flat ($S(\omega) = R_0$), i.e. it does not depend at all on the frequency. This surprising property of a white spectrum is independent of the shape of the individual decision trains and of the chosen nonlinearity $f(x)$, the noise intensity and the points x_c, x_i as long as both decisions have equal probability. Note that the white spectrum only arises because correct and incorrect decisions enter the combined decision train with opposite amplitude. A flat spectrum would generally not be observed for instance if correct and incorrect decisions would contribute with spikes of the same amplitude.

2.3. An analytically tractable case

For the simple case of a linear potential ($f(x) = \mu$), the subthreshold dynamics of the model are a Wiener process given by:

$$\tau_x \dot{x} = \mu + \sigma\sqrt{2\tau_x}\xi(t). \tag{13}$$

Such a process is a simplified version of Ratcliff's original model (in the latter there is also trial-to-trial variability of the drift rate μ considered) and can be treated analytically in order to determine the stationary decision rates and the inter-decision interval densities (see for instance [48, 20, 71]). For $\mu \neq 0$, the stationary decision rate for correct decisions in our notation is given by:

$$\begin{aligned} r_{c,0} &= \frac{\mu e_i}{\tau_x[x_c e_i - x_i e_c] + \Delta(e_i - e_c)}, \\ e_c &= 1 - \exp\left(-\frac{\mu}{\sigma^2} x_c\right),\ e_i = 1 - \exp\left(-\frac{\mu}{\sigma^2} x_i\right). \end{aligned} \tag{14}$$

This rate describes how many correct decisions per time unit can be expected and together with the rate of incorrect decisions $r_{i,0}$ (obtained as we remind the reader by swapping $c \leftrightarrow i$ in Eq. (14)) provides the probability of a correct decision by $r_{c,0}/(r_{c,0} + r_{i,0})$. The response-time density for correct decisions reads in our notation [cf. [48] Eq. (A9) where the left boundary is set to zero and the reset is at a non-vanishing point]:

$$\begin{aligned} g_c(T) &= \frac{2\pi\sigma^2 \exp\left[\mu x_c/(2\sigma^2)\right]}{\tau_x (x_c - x_i)^2} \Theta(T - \Delta) \\ &\times \sum_{k=1}^{\infty} k \sin\left[\frac{k\pi x_c}{x_c - x_i}\right] \exp\left[-\frac{(T-\Delta)}{\tau_x}\left(\frac{\mu^2}{4\sigma^2} + \frac{k^2\pi^2\sigma^2}{(x_c - x_i)^2}\right)\right], \end{aligned} \tag{15}$$

where $\Theta(T - \Delta)$ denotes the Heaviside function and incorporates the non-decision time. Instead of calculating the response-time densities as functions of time as in [48, 20, 71], we may calculate them in Fourier domain. In our notation, the Fourier-transformed response-time density for correct decisions reads:

$$\begin{aligned} \tilde{g}_c(\omega) &= \frac{\exp\left(\frac{\mu x_c}{2\sigma^2} + i\omega\Delta\right)\sinh\left(x_i\kappa(\omega)\right)}{\sinh\left([x_i - x_c]\kappa(\omega)\right)}, \\ \kappa(\omega) &= \sqrt{\frac{\mu^2}{4\sigma^4} - \frac{i\omega\tau_x}{\sigma^2}} \end{aligned} \tag{16}$$

(see Appendix D for detailed derivation). The connection between the density in frequency domain and the density time domain is given by the inverse Fourier transformation:

$$g_c(T) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \tilde{g}_c(\omega) e^{-i\omega T}\, d\omega \tag{17}$$

Following Eq. (8) the Fourier-transformed IDI density of the correct decision train is given by:

$$\tilde{\varrho}_c(\omega) = \frac{e^{i\omega\Delta + \mu x_c/(2\sigma^2)} \sinh[x_i\kappa]}{\sinh[(x_i - x_c)\kappa] + e^{i\omega\Delta + \mu x_i/(2\sigma^2)} \sinh[x_c\kappa]}, \tag{18}$$

where, for brevity, we suppressed the frequency dependence of $\kappa(\omega)$. Lengthy expressions for the power spectra can be obtained via Eq. (11) and Eq. (12) but are not explicitly stated here. We have tested all statistics by a comparison with numerical simulations and found excellent agreement (see Fig. 3). Many features of the statistics for nonlinear decision models (see below) can be already

observed in the analytically tractable case. For instance, correct decisions (green spikes in panel a) are made more often but incorrect ones (red downward spikes in a) are observed too; the probability density of the decision variable shows the typical kink at the reset point $x_{\rm st} = 0$ (see b); the IDI densities of the individual decision trains can be unimodal or multimodal (see d); power spectra of the decision trains (e) display maxima at frequencies that correspond to the inverse of the peak times of the respective IDI density ϱ_c or ϱ_i (as revealed in c and d). Peaks in the power spectrum thus show that the decision process is temporally more regular than a Poisson process and the different frequencies at which maxima are attained reflect the typical times in which the thresholds are reached.

The analytical method that led to the explicit formulas for densities, rates, and spectra above cannot be applied for the general nonlinear model Eq. (3). In the following we adapt an efficient numerical procedure, the threshold-integration method, originally introduced for nonlinear integrate-and-fire neurons by Richardson [59, 60] and generalize it here to the case of two absorbing boundaries. The method is simple to implement and yields a rapid numerical solution for all statistics of interest.

Fig.	τ_x [s]	σ	x_i	x_c	$f(x)$
3	0.1	0.5	-1	2	0.2
4	0.1	0.5	-1	1	-x+0.2
5	0.1	0.4	-1	1	$2x^3$-x+0.2
6	0.1	0.7	-1.4	1.4	$-16x^3 + 18x + 2.5$
7	1.0	2	-3	1	1.085-$2x^2$-x-0.5e^x-8sin($2\pi x$)
8	0.1	0.4	-1	1	$2x^3$-x+0.2

Table 1: Parameters of the nonlinear DDMs. In all cases $\Delta = 0.2$.

2.4. Trial-to-trial variability of stimulus difficulty

In many experiments (for instance [51]) a sequence of stimuli is presented in which the difficulty, in our example experiment in Fig. 1 the coherence of dots, is chosen randomly for each decision. Here we consider a sequence of stimuli with n levels of difficulty for the motion to the right $s = +j$ and to the left $s = -j$ with $j \in \{1, 2, \ldots, n\}$. After a decision has been made, the difficulty of the subsequent stimulus is randomly chosen with the probability Prob(j). In order to consider varying stimulus difficulties in the DDM for correct and incorrect decisions in Eq. (3), the subthreshold dynamics can be extended to stimulus-dependent drift function and noise intensity:

$$\tau_x \dot{x} = f_j(x) + \sigma_j \sqrt{2\tau_x}\xi(t). \tag{19}$$

Also the thresholds $x_{j,c}$ and $x_{j,i}$ may depend on the stimulus difficulty as explained in Sec. 2.1. Corresponding to each stimulus difficulty j, an individual response-time density for correct and incorrect decisions $\mathfrak{g}_{j,c}(T)$ and $\mathfrak{g}_{j,i}(T)$, respectively, can be calculated. For the simple reference case of independent decisions, the correct response-time density for the entire decision train with varying difficulty is given by the weighted average:

$$g_c(T) = \sum_{j=1}^{n} \mathrm{Prob}(j)\mathfrak{g}_{j,c}(T). \tag{20}$$

With the response-time densities for correct and incorrect decisions, the other statistics of the decision train are given by Eqs. (8), (11) and (12). The corresponding rates $r_{c,0}$ and $r_{i,0}$ are given by the inverse of the mean value of ϱ_c and ϱ_i given in Eq. (8), respectively. Alternatively, we could calculate, for instance, $r_{c,0}$, from the stationary rates $\mathfrak{r}_{j,c,0}$ of the single states j as follows

$$r_{c,0} = \left[\sum_{j=1}^{n} \mathrm{Prob}(j)\mathfrak{r}_{j,c,0}^{-1}\right]^{-1}. \tag{21}$$

3. Threshold-integration method

In order to calculate the statistics of interest, i.e. the decision-train statistics, we have to determine the probability density of the evidence accumulation and the associated probability currents through the thresholds. The probability density of an ensemble of decision-making processes is denoted by $P(x,t)$; the corresponding probability flux will be referred to as $J(x,t)$. Due to the reset mechanism, $P(x,t)$ and $J(x,t)$ are only nonzero in between the thresholds. The white noise that enters the x-dynamics results in absorbing boundary conditions (cf. [12]):

$$P(x_c, t) = P(x_i, t) = 0. \tag{22}$$

Between the boundaries $P(x,t)$ is continuous but $J(x,t)$ exhibits discontinuities at sources and sinks of probability that are located at the reset and the thresholds, respectively. Rates, densities and power spectra can be computed from the solutions of $P(x,t)$ and $J(x,t)$ under different reset and/or initial conditions all of which can be represented by the source-and-sink function $\nu(x,t)$. The latter function may take into account the initial placement of probability (initial condition), the reset of probability and the absorption of probability at the boundaries. By means of this function, the continuity equation can be formulated as:

$$\partial_t P(x,t) + \partial_x J(x,t) = \nu(x,t) \tag{23}$$

(see the equivalent treatment of an integrate-and-fire dynamics by Richardson [60]). The probability flux $J(x,t)$ is given by [61]:

$$\tau_x J(x,t) = f(x)P(x,t) - \sigma^2 \partial_x P(x,t). \tag{24}$$

Inserting Eq. (24) into Eq. (23) yields the Fokker-Planck equation, however, for our purposes keeping two equations of first order is beneficial. First, we generalize the threshold-integration method to find the stationary solution following closely the work by Richardson [60].

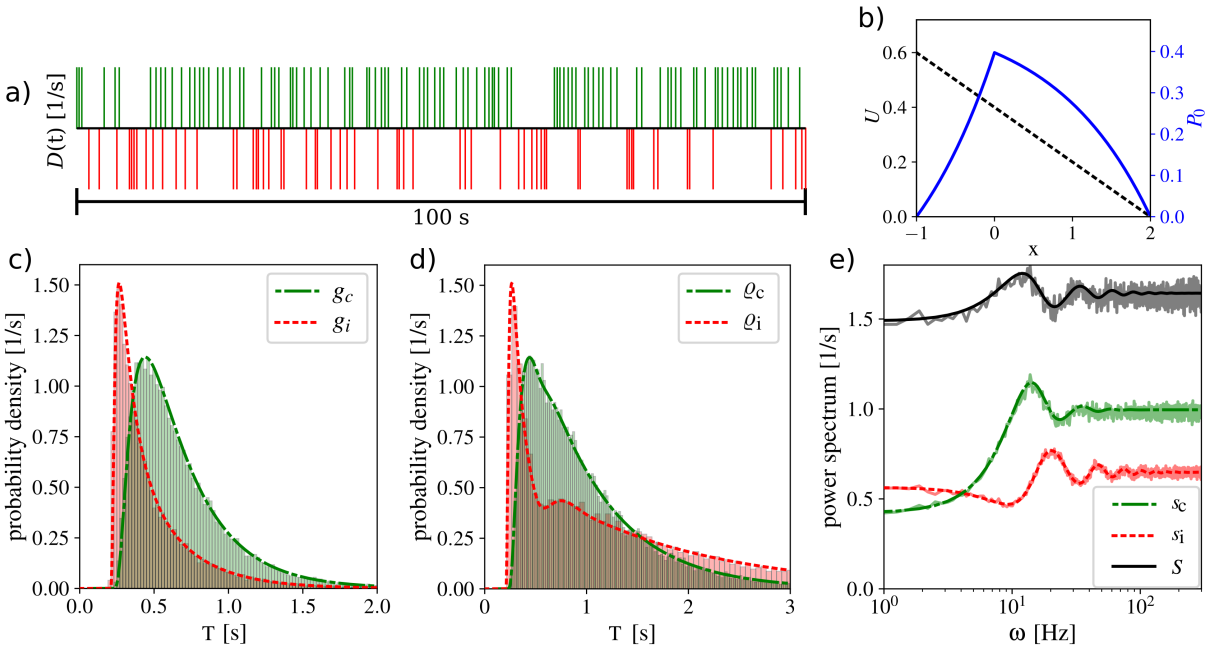

Figure 3: **Temporal statistics of classic DDM as Wiener process between two boundaries** (color online). Decision train (a), potential (dashed black line) and stationary solution (solid blue line, see Appendix D) (b), response-time densities $g_c(T)$ (green) and $g_i(T)$ (red) from analytical theory (dashed and dash-dotted lines) and direct simulations of Eqs. 3 and 4 (histograms) (c), IDI densities of individual decision trains, $\varrho_c(T)$ and $\varrho_i(T)$ (d) and decision-train power spectra of individual (red and green) and combined (solid) trains from theory (dashed, dash-dotted and solid lines) and simulations (bright lines). See Tab. 1 for parameters.

3.1. Stationary solution

The stationary probability density $P_0(x)$ does not depend on time ($\partial_t P_0(x) = 0$) and the probability flux $J_0(x)$ is piecewise constant between reset and the corresponding threshold. The entire probability escaping through the thresholds is reinserted such that the source-and-sink function $\nu_0(x)$ reads:

$$\begin{aligned}\nu_0(x) = (r_{c,0} + r_{i,0})\delta(x) - r_{c,0}\delta(x - x_c) \\ -r_{i,0}\delta(x - x_i).\end{aligned} \tag{25}$$

Here, the rates $r_{c,0}$ and $r_{i,0}$ refer to the stationary decision rates for correct and incorrect decisions, respectively; these are determined below. From Eqs. (23) and (24) we obtain two ordinary differential equations of first order:

$$\begin{aligned}\partial_x J_0 =& - r_{c,0}\delta(x - x_c) - r_{i,0}\delta(x - x_i) \\ &+ (r_{c,0} + r_{i,0})\delta(x) \\ \partial_x P_0 =& - \frac{1}{\sigma^2}(\tau_x J_0 - f(x)P_0).\end{aligned} \tag{26}$$

The key to determine the solutions $P_0(x)$ and $J_0(x)$ is the introduction of the two unnormalized densities $p_i(x) = P_0(x)/r_{i,0}$ and $p_c(x) = P_0(x)/r_{c,0}$ and currents $j_{i,0}(x) = J_0(x)/r_{i,0}$ and $j_{c,0}(x) = J_0(x)/r_{c,0}$ that range from each threshold to the reset. The values of these probability densities and fluxes at the thresholds are known as

$$p_{c,0}(x_c) = p_{i,0}(x_i) = 0 \tag{27}$$

and

$$j_{c,0}(x_c) = -j_{i,0}(x_i) = 1. \tag{28}$$

The whole density and current arises from numerical integration from the threshold x_i to the reset of the ordinary differential equations:

$$\begin{aligned}\partial_x j_{i,0} =& - \delta(x - x_i), \\ \partial_x p_{i,0} =& - \frac{1}{\sigma^2}(\tau_x j_{i,0} - f(x)p_{i,0}).\end{aligned} \tag{29}$$

Here a simple Euler integration is used (see for instance [46]) with N integration steps for the whole probability density and the step size $\Delta x = (x_c - x_i)/N$. The k-th step of the numerical integration reads:

$$\begin{aligned}p_{i,0}(x_i + k\Delta x) \approx p_{i,0}(x_i + [k-1]\Delta x) \\ + \frac{\Delta x}{\sigma^2}[\tau_x + f(x_i + [k-1]\Delta x)p_{i,0}(x_i + [[k-1]\Delta x)].\end{aligned} \tag{30}$$

We used that $j_{i,0} = -1$ for the considered values of x. From the threshold to reset $N_i = -x_i/\Delta x$ integration steps are performed to calculate $p_{i,0}(0) \approx p_{i,0}(x_i + N_i\Delta x)$, where N_i is rounded to the next integer. In the same way, $p_{c,0}(0) \approx p_{c,0}(x_c - N_c\Delta x)$ is determined by performing $N_c = N - N_i$ integration steps of:

$$\begin{aligned}\partial_x j_{c,0} =& - \delta(x - x_c), \\ \partial_x p_{c,0} =& - \frac{1}{\sigma^2}(\tau_x j_{c,0} - f(x)p_{c,0})\end{aligned} \tag{31}$$

from x_c to the reset of which the k-th step reads:

$$p_{c,0}(x_c - k\Delta x) \approx p_{c,0}(x_c - [k-1]\Delta x) + \frac{\Delta x}{\sigma^2}(\tau_x - f(x_c - [k-1]\Delta x)p_{c,0}(x_c - [k-1]\Delta x)). \tag{32}$$

The stationary decision rates, probability density and flux are determined by exploiting the continuity and normalization of $P_0(x)$:

$$r_{c,0}p_{c,0}(0) = r_{i,0}p_{i,0}(0), \tag{33}$$

$$\int_{x_i}^{x_c} P_0\, dx = 1 - R_0\Delta. \tag{34}$$

Here, the term $R_0\Delta$ represents the fraction of probability that does not evolve due to the non-decision time period (this part of the ensemble is not included in the density $P_0(x)$). Both conditions uniquely determine the rates:

$$r_{c,0}^{-1} = \int_0^{x_c} p_{c,0}\, dx + \frac{p_{c,0}(0)}{p_{i,0}(0)}\left(\Delta + \int_{x_i}^{0} p_{i,0}\, dx\right) + \Delta. \tag{35}$$

The rate of incorrect decisions can be obtained by switching c to i or by using Eq. (33). Examples of stationary solutions and the corresponding potentials $U(x)$ are presented in Figs. 4-7 b.

3.2. IDI densities and decision-train power spectra

Since the partial differential equations (23) and (24) are difficult to treat numerically, we apply the Fourier transformation to obtain two ordinary differential equations in the frequency domain:

$$\begin{aligned}\partial_x \tilde{J}(x,\omega) &= i\omega\tilde{P}(x,\omega) + \tilde{\nu}(x,\omega)\\ \partial_x \tilde{P}(x,\omega) &= -\frac{1}{\sigma^2}(\tau_x\tilde{J} - f(x)\tilde{P}(x,\omega)).\end{aligned} \tag{36}$$

Similarly to the stationary case, we introduce new variables for which we know the values at the thresholds. These are given by the ratios of the probability density and the efflux at the corresponding thresholds, and the current and those effluxes:

$$\begin{aligned}\tilde{p}_c(x,\omega) &= \frac{\tilde{P}(x,\omega)}{\tilde{J}(x_c,\omega)}, \tilde{j}c(x,\omega) = \frac{\tilde{J}(x,\omega)}{\tilde{J}(x_c,\omega)},\\ \tilde{p}_i(x,\omega) &= \frac{\tilde{P}(x,\omega)}{-\tilde{J}(x_i,\omega)}, \tilde{j}_i(x,\omega) = \frac{\tilde{J}(x,\omega)}{-\tilde{J}(x_i,\omega)}\\ &\text{with } \tilde{p}_c(x_c) = \tilde{p}_i(x_i) = 0\\ &\text{and } \tilde{j}_c(x_c) = -\tilde{j}_i(x_i) = 1.\end{aligned} \tag{37}$$

With these initial values, we numerically integrate the ordinary differential equations that we obtain from Eq. (36):

$$\begin{aligned}\partial_x \tilde{j}_c &= i\omega\tilde{p}_c - \delta(x - x_c)\\ \partial_x \tilde{p}_c &= -\frac{1}{\sigma^2}(\tau_x\tilde{j}_c - f\tilde{p}_c)\end{aligned} \tag{38}$$

from threshold to reset. Starting at x_c we perform N_c Euler steps of which the k-th is given by:

$$\begin{aligned}&\tilde{j}_c(x_c - k\Delta x,\omega) = \tilde{j}(x_c - [k-1]\Delta x,\omega)\\ &- \Delta x\, i\omega\tilde{p}_c(x_c - [k-1]\Delta x),\\ &\tilde{p}_c(x_c - k\Delta x,\omega) = \tilde{p}_c(x_c - [k-1]\Delta x,\omega)\\ &+ \frac{\Delta x}{\sigma^2}\tau_x\tilde{j}(x_c - [k-1]\Delta x,\omega)\\ &- \frac{\Delta x}{\sigma^2}f(x_c - [k-1]\Delta x)\tilde{p}_c(x_c - [k-1]\Delta x,\omega).\end{aligned} \tag{39}$$

From x_i we also perform N_i steps of Euler integration:

$$\begin{aligned}&\tilde{j}_i(x_i + k\Delta x,\omega) = \tilde{j}(x_i + [k-1]\Delta x,\omega)\\ &+ \Delta x\, i\omega\tilde{p}_i(x_i + [k-1]\Delta x),\\ &\tilde{p}_i(x_i + k\Delta x,\omega) = \tilde{p}_i(x_i + [k-1]\Delta x,\omega)\\ &- \frac{\Delta x}{\sigma^2}\tau_x\tilde{j}_i(x_i + [k-1]\Delta x,\omega)\\ &+ \frac{\Delta x}{\sigma^2}f(x_i + [k-1]\Delta x)\tilde{p}_c(x_i + [k-1]\Delta x,\omega).\end{aligned} \tag{40}$$

The resulting values at the reset $\tilde{p}_{c,s} = \tilde{p}_c(0,\omega) \approx \tilde{p}_c(x_c - N_c\Delta x,\omega)$, $\tilde{p}_{i,s} = \tilde{p}_i(0,\omega) \approx \tilde{p}_c(x_i + N_i\Delta x,\omega)$, $\tilde{j}_{c,s} = \tilde{j}_c(0,\omega) \approx \tilde{j}_c(x_c - N_c\Delta x,\omega)$ and $\tilde{j}_{i,s} = \tilde{j}_i(0,\omega) \approx \tilde{j}_i(x_i + N_i\Delta x,\omega)$ yield the statistics introduced above by exploiting the continuity of $\tilde{P}(x,\omega)$ and the jump condition due to the reinsertion of probability of $\tilde{J}(x,\omega)$ at zero. The values of $\tilde{p}_{c,s}$, $\tilde{p}_{i,s}$, $\tilde{j}_{c,s}$ and $\tilde{j}_{i,s}$ are determined at equidistant frequencies up to ω_{cut} in steps of $\Delta\omega$ to determine the Fourier-transformed IDI densities as explained below and, subsequently, the IDI densities in time domain by fast Fourier transformation. The used parameters are presented in Tab. 2. In order to give an impression about the required numerical effort for the solution on a common laptop, we list in the last column the computation time for the determination of the stationary probability density and the response-time densities in time domain (including the inverse Fourier transform).

Fig.	ω_{cut} [Hz]	$\Delta\omega$ [Hz]	N	time [s]
4	200	1	1000	2.8
5	200	1	1000	2.8
6	300	1.5	4000	3.1
7	300	1.5	1000	2.8
8			1000	

Table 2: Numerical parameters used for the threshold integration method. We also list the needed run times for the determination of the stationary probability density and the response-time densities (laptop with intel core i7 8550u CPU).

In order to calculate the response-time densities $g_c(T)$ and $g_i(T)$, we have to initially inject unit probability at the reset point (this is the initial condition after a decision has been made) but switch off any reset of probability. This corresponds to a classical first-passage-time (FPT) problem for two absorbing boundaries. Accordingly, the

sink-and-source function in the time domain reads:

$$\begin{aligned}\nu_{\mathrm{FPT}}(T) &= \delta(T-\Delta)\delta(x)\\ &\quad - g_c(T)\delta(x-x_c) - g_i(T)\delta(x-x_i),\end{aligned} \tag{41}$$

where the effluxes of probability at the two thresholds represent the FPT densities [20] that are equivalent to the response-time densities $g_c(T)$ and $g_i(T)$. The Fourier-transformed function reads:

$$\begin{aligned}\tilde{\nu}_{\mathrm{FPT}}(\omega) &= e^{i\omega\Delta}\delta(x)\\ &\quad - \tilde{g}_i(\omega)\delta(x-x_i) - \tilde{g}_c(\omega)\delta(x-x_c).\end{aligned} \tag{42}$$

The continuity of $\tilde{P}(x,\omega)$ and the jump of $\tilde{J}(x,\omega)$ at zero that incorporates the initial injection of probability yield the conditions:

$$\begin{aligned}\tilde{g}_c(\omega)\tilde{p}_{c,s} &= \tilde{g}_i(\omega)\tilde{p}_{i,s},\\ \tilde{g}_i(\omega)\tilde{\jmath}_{i,s} + e^{i\omega\Delta} &= \tilde{g}_c(\omega)\tilde{\jmath}_{c,s}\end{aligned} \tag{43}$$

from which we may identify the Fourier-transformed response-time densities as:

$$\tilde{g}_c(\omega) = \frac{\tilde{p}_{i,s}e^{i\omega\Delta}}{\tilde{p}_{i,s}\tilde{\jmath}_{c,s} - \tilde{p}_{c,s}\tilde{\jmath}_{i,s}}, \qquad \tilde{g}_i(\omega) = \tilde{g}_c(\omega)\frac{\tilde{p}_{c,s}}{\tilde{p}_{i,s}}. \tag{44}$$

The densities in time domain are determined by the inverse Fourier transformation in Eq. (17) that is numerically determined by a fast Fourier transformation, specifically the Python library *numpy.fft*.

For the individual decision trains, the IDI distributions and their power spectra can be calculated using Eq. (8) and Eq. (11), respectively. Alternatively, the statistics directly follow under the assumption of other sink-and-source functions $\nu(x,t)$, a way that we pursue in the following; we have verified that the final expressions for both approaches are the same. In order to calculate the IDI density of the individual decision train $\varrho_c(T)$, we have to reinsert continuously the probability that crosses the other (opposing) threshold x_i at x_s. Consequentially, the Fourier-transformed sink-and-source function is given by:

$$\begin{aligned}\tilde{\nu}_c(\omega) =&(1+\tilde{\beta}_i(\omega))\delta(x-x_s)e^{i\omega\Delta}\\ &-\tilde{\varrho}_c(\omega)\delta(x-x_c) - \tilde{\beta}_i(\omega)\delta(x-x_i).\end{aligned} \tag{45}$$

Here $\tilde{\beta}_i(\omega)$ is the Fourier-transformed efflux at x_i. While the continuity condition is the same as in the previous case, the jump condition has to incorporate the additional reset of probability:

$$\begin{aligned}\tilde{\varrho}_c(\omega)\tilde{p}_{c,s} &= \tilde{\beta}_i(\omega)\tilde{p}_{i,s},\\ \tilde{\beta}_i(\omega)\tilde{\jmath}_{i,s} + (1+\tilde{\beta}_i(\omega))e^{i\omega\Delta} &= \tilde{\varrho}_c(\omega)\tilde{\jmath}_{c,s},\end{aligned} \tag{46}$$

and we obtain:

$$\tilde{\varrho}_c(\omega) = \frac{\tilde{p}_{i,s}e^{i\omega\Delta}}{\tilde{p}_{i,s}\tilde{\jmath}_{c,s} - \tilde{p}_{c,s}(\tilde{\jmath}_{i,s} + e^{i\omega\Delta})}. \tag{47}$$

For $\tilde{\varrho}_i$ the analogous calculation with corresponding sink-and-source function yields us:

$$\tilde{\varrho}_i(\omega) = \frac{\tilde{p}_{c,s}e^{i\omega\Delta}}{\tilde{p}_{i,s}(\tilde{\jmath}_{c,s} - e^{i\omega\Delta}) - \tilde{p}_{c,s}\tilde{\jmath}_{i,s}}. \tag{48}$$

To calculate the conditional decision rates $\tilde{m}_c(\omega)$ and $\tilde{m}_i(\omega)$, that are the essential part of the individual decision-train power spectra (see Appendix C), all the probability absorbed at both thresholds is reinserted at x_s. The corresponding sink-and-source function is given by:

$$\begin{aligned}\tilde{\nu}_S(\omega) =&(1+\tilde{m}_c(\omega)+\tilde{m}_i(\omega))\delta(x-x_s)e^{i\omega\Delta}\\ &-\tilde{m}_c(\omega)\delta(x-x_c) - \tilde{m}_i(\omega)\delta(x-x_i).\end{aligned} \tag{49}$$

In this case, the continuity and jump conditions read:

$$\begin{aligned}\tilde{m}_c(\omega)\tilde{p}_{c,s} &= \tilde{m}_i(\omega)\tilde{p}_{i,s},\\ \tilde{m}_i(\omega)\tilde{\jmath}_{i,s} + (1+\tilde{m}_i(\omega)+\tilde{m}_c(\omega))e^{i\omega\Delta} &= \tilde{m}_c(\omega)\tilde{\jmath}_{c,s},\end{aligned} \tag{50}$$

with which we obtain:

$$\begin{aligned}\tilde{m}_c(\omega) &= \frac{\tilde{p}_{i,s}e^{i\omega\Delta}}{\tilde{p}_{i,s}(\tilde{\jmath}_{c,s} - e^{i\omega\Delta}) - \tilde{p}_{c,s}(\tilde{\jmath}_{i,s} + e^{i\omega\Delta})},\\ \tilde{m}_i(\omega) &= \frac{\tilde{p}_{c,s}e^{i\omega\Delta}}{\tilde{p}_{i,s}(\tilde{\jmath}_{c,s} - e^{i\omega\Delta}) - \tilde{p}_{c,s}(\tilde{\jmath}_{i,s} + e^{i\omega\Delta})}.\end{aligned} \tag{51}$$

The decision-train power spectrum can be calculated by [27]:

$$s_c(\omega) = r_{c,0}(1+\tilde{m}_c(\omega)). \tag{52}$$

4. Results of the threshold-integration method

The decision-time statistics will depend on all the parameters of the system, but particularly, on the specific shape of the function $f(x)$ or, alternatively, on the shape of the corresponding potential $U(x)$ [where $-\partial_x U(x) = f(x)$]. In the following, we test the numerical routines presented for some interesting cases with nonlinear potential; for better comparison, we will show all statistics of interest in the manner exemplified in Fig. 3.

4.1. Ornstein-Uhlenbeck process (leaky accumulator model)

In the example presented in Fig. 4, we choose the simplest deviation from a Wiener process: a linear force, $f(x) = -x+\mu$, (or quadratic potential $U(x) = (x-\mu)^2/2$), for which the stochastic differential equation corresponds to an Ornstein-Uhlenbeck process (see Tab. 1 for parameters and Fig. 4b for the potential shape). This model, has been used in studies of decision making and is also known as leaky accumulator model; analytical solutions for the response-time densities are unknown (see for instance [72, 52]).

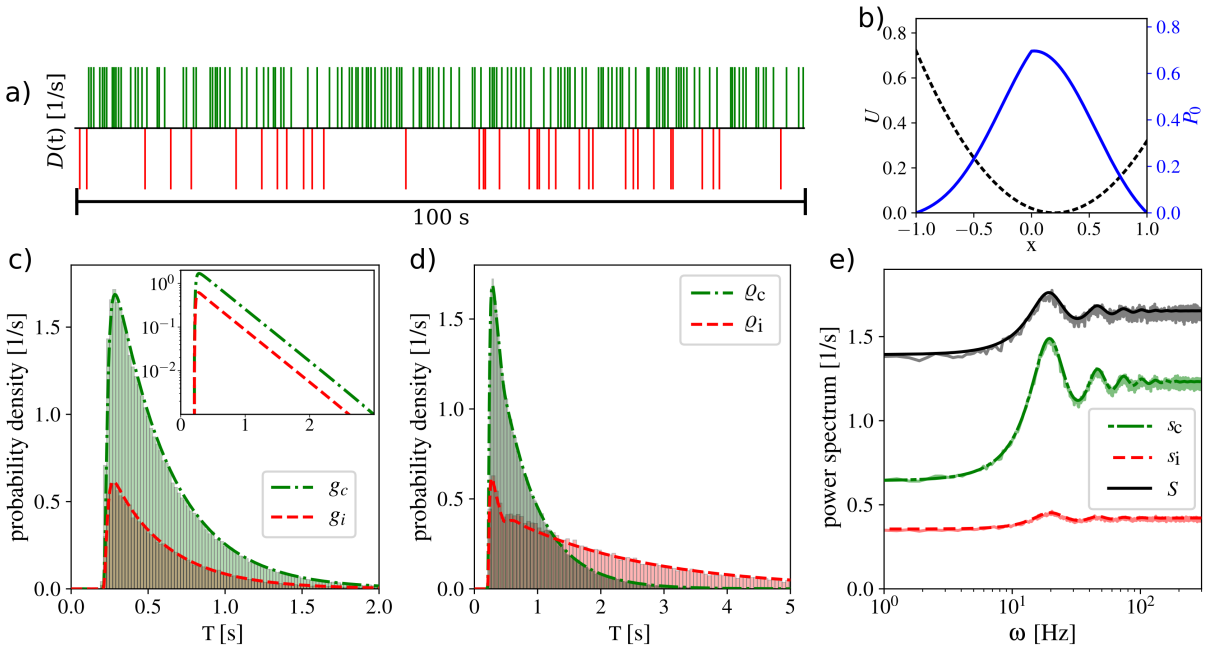

Figure 4: **Temporal statistics of the Ornstein-Uhlenbeck DDM.** See caption in Fig. 3 for detailed description and Tab. 1 for parameters. Inset in panel c with logarithmic scale shows exponential tail of the response-time densities.

We obtain an excellent agreement between the results of the threshold-integration method and from direct stochastic simulations of Eqs. 3 and 4. The asymmetry due to a positive drift $\mu > 0$ models the signal for correct decisions, that are thus more likely than incorrect decisions. The resulting response-time densities $g_c(T)$ and $g_i(T)$ exhibit a shape similar to a gamma distribution (cf. Fig. 4 c) with an exponential tail (see inset).

For the individual decision trains, the correct decisions are quite regular and the distribution $\varrho_c(T)$ is similar to $g_c(T)$ while the low probability of incorrect decisions yields a long-tailed IDI distribution $\varrho_i(T)$ as presented in Fig. 4d. Furthermore, a second peak in $\varrho_i(T)$ arises due to the case of a correct and a subsequent incorrect decision. Due to the long and irregular time between subsequent incorrect decisions, the corresponding spectra $s_i(\omega)$ is almost flat, similar to the spike-train spectrum of a Poisson process. In contrast, $s_c(\omega)$ exhibits peaks at the decision rate and its higher harmonics due to the regularity of $d_c(t)$. These properties of the individual spectra also emerge in the spectrum of the (combined) decision train $S(\omega)$ (cf. Fig. 4 e). Note that the peaks observed in the combined spectrum $S(\omega)$ are a consequence of the nonvanishing bias for one of the decisions, because according to Eq. (12), we could expect a perfectly flat spectrum in the absence of any bias.

4.2. Fourth-order polynomial potentials

As a nonlinear example (Fig. 5), we use a fourth-order polynomial for $U(x)$, based on a mean-field calculation for competing neural populations in [64] (see Tab. 1 for parameters and Fig. 5 b for the potential shape). The resulting decision-train statistics are similar to the statistics from the Ornstein-Uhlenbeck example, which is not so surprising at second thought: For the former the stochastic process has to reach one of two possible thresholds, for the latter, the process has to overcome one of two potential barriers (and then to quickly proceed to the corresponding threshold) - in either cases fluctuations are needed to make a decision and they will also shape the statistics of the decision times in a qualitatively similar way (in particular, at large times). Consequently, response-time densities are unimodal and have an exponential tail, the power spectrum of the correct-decision train displays peaks at multiples of the rate of correct decisions, and the power spectrum of incorrect decisions is almost flat because escapes to the left boundary x_i or, equivalently, over the left barrier, are rare, resulting almost in Poisson statistics for the incorrect decisions. Hence, a nonlinearity in the drift function does not necessarily lead to qualitatively novel features in the decision-train statistics.

However, to demonstrate that qualitatively different features become possible with nonlinear drift functions, we use once more a fourth-order polynomial for $U(x)$ but with different parameters such that the reset is near the maximum of the potential and two minima are located in proximity of the two thresholds as indicated by the olive green and cyan lines in Fig. 6 b. This is a truly bistable dynamics with additional absorbing boundaries at the left and right corresponding here still to incorrect and correct decisions. As in the previous examples, parameters are chosen such that incorrect decisions are very unlikely to happen.

In contrast to the previous examples, the response-time

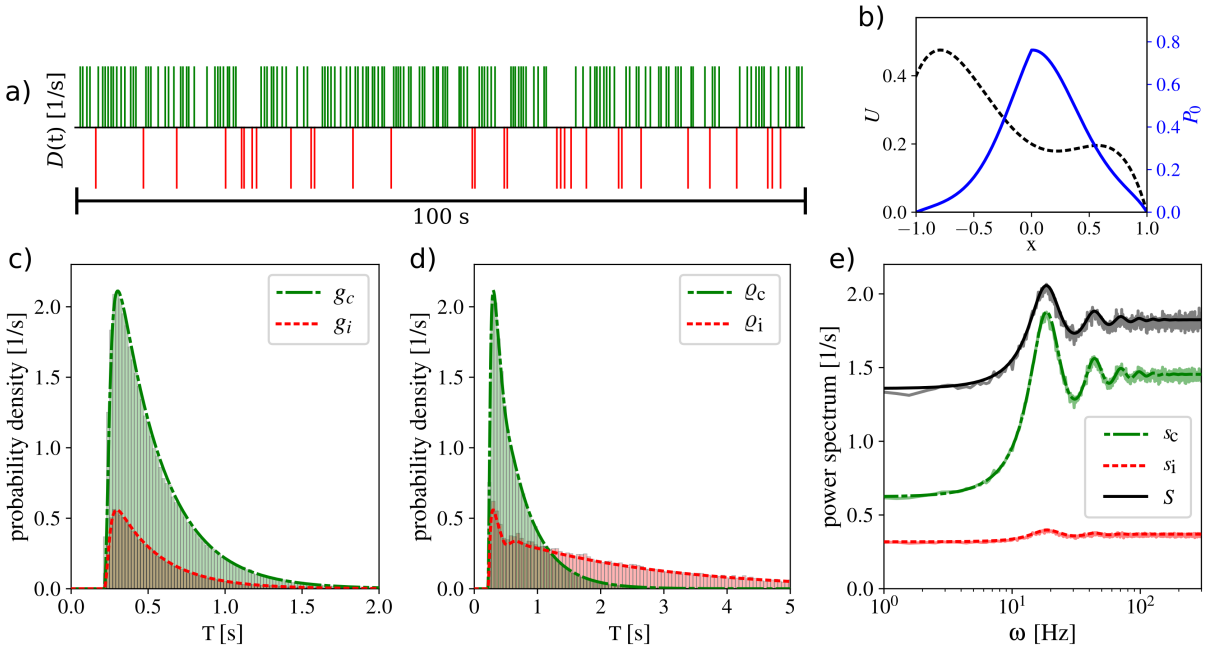

Figure 5: **Temporal statistics of nonlinear DDM with polynomial potential.** See caption in Fig. 3 for detailed description and Tab. 1 for parameters.

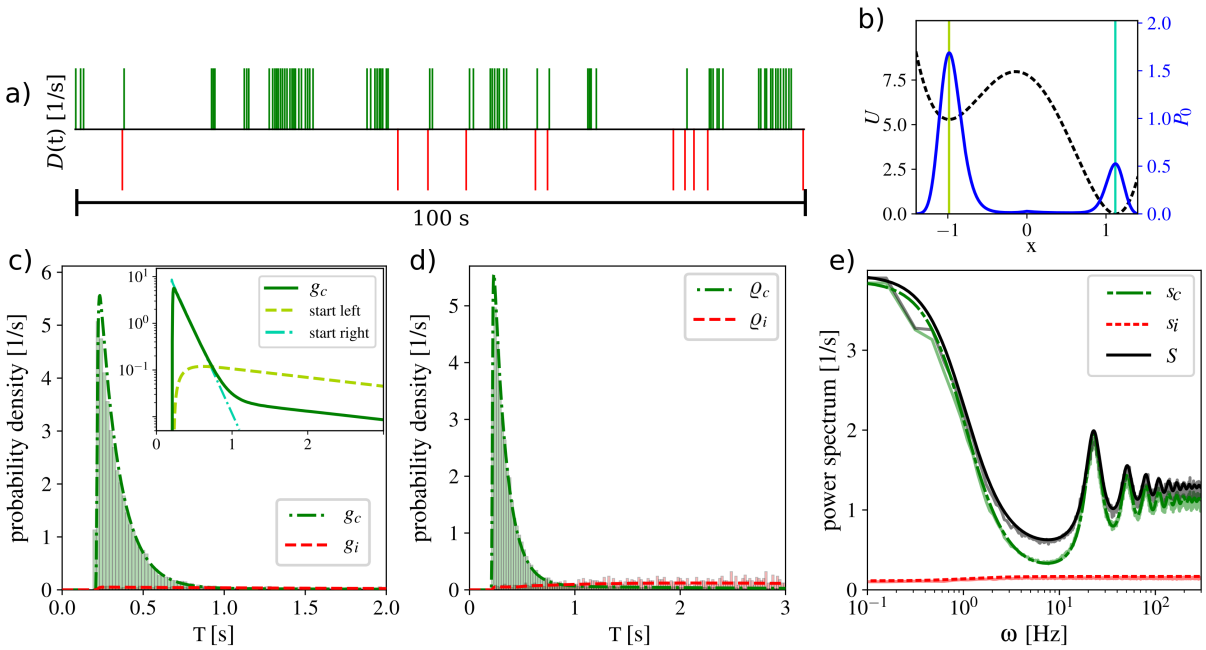

Figure 6: **Temporal statistics of nonlinear DDM with bistable polynomial potential.** See caption in Fig. 3 for detailed description and Tab. 1 for parameters. Inset with logarithmic scale in panel c shows biexponential tail of the response-time density: large exponent shown belongs to the first-passage-time density from the minimum close to the boundary as cyan dot-dashed line, low exponent belongs to first-passage time density from minimum close to the threshold indicated as olive green dashed line. The corresponding minima are also indicated in panel b as vertical lines in the respective colors.

density for correct decisions exhibits a bi-exponential tail as presented in the inset of Fig. 6 c. The underlying mechanism for the two apparent time scales is the following: in a short period of time after the reset, even low evidence pushes x towards one of the two minima with similar probability in a preliminary decision. If x is located near the minimum near x_c, the threshold is relatively quickly crossed within a short time causing the first part of the bi-exponential tail. Note that the escape rate of this part is equal to the inverse of the first-passage time for a start at the minimum near x_c (see inset of Fig. 6 c cyan dash-dotted line). If x first moves towards the minimum far away from x_c, the time that is required to cross x_c is much higher because now *two* potential barriers have to be crossed by the assistance of the driving noise. The exponential tail at long times fits to the first-passage time density if the trajectory starts at the minimum close to x_c (see olive green line in inset). The presence of the pre-decision process leads to a rather irregular decision train as presented in Fig. 6 a. The corresponding decision-train power spectrum exhibits increased power at low frequencies (see Fig. 6 e) and is thus qualitatively different to the cases considered before in Fig. 4 and Fig. 5.

4.3. An example with equal decision rates

For the last example, we use a highly nonlinear asymmetric potential with several local minima and asymmetric boundaries $x_i = -3x_c$ (see Tab. 1 for parameters and Fig. 7 b for the potential shape). This potential has no meaning for the decision making process but is shown to demonstrate the general applicability of the method.

We also use this case to illustrate the peculiar consequence of equal decision rates: a flat decision-train power spectrum. Specifically, we adapted the constant drift in order to obtain equal decision rates (vanishing decision bias). Again, results from the threshold-integration method are confirmed by the direct simulations.

The longer distance from reset to the threshold x_i yields a more regular train $d_i(t)$ compared to $d_c(t)$ (cf Fig. 7 a), and also, longer IDIs for incorrect than for correct decisions, cf. locations of the maxima of $g_c(T)$ and $g_i(T)$ (cf. Fig. 7 c). The individual decision spectra presented in Fig. 7 d are different and exhibit peaks at different locations. However, despite the changes in timing, we emphasize that the probability of correct and incorrect decisions are equal in this example. The equal decision rates result in a flat spectrum of the combined decision train as predicted by Eq. (12) and also found in our simulations (cf. Fig. 7 e, black line).

5. Linear response to modulation of input

We can also imagine that, in addition to the switching between two constant stimuli, the overall strength of the bias is deliberately modulated over time by the experimenter, e.g. in order to probe the system in comparison with different models (see discussion). If the bias is weakly modulated in time, this will result in a likewise small modulation of the decision rates (and the resulting decision train will not be a renewal process anymore). Here we generalize the method by Richardson [59] to compute the linear response of the decision rates to a parameter modulation of the nonlinear DDM. We assume a weak modulation of a parameter α (which could be, for instance, modulation of the constant input signal, see below)

$$\alpha(t) = \alpha_0 + \varepsilon \sin(\omega t + \phi) \tag{53}$$

with a small modulation amplitude ε and an initial phase ϕ.

5.1. Threshold-integration method for the linear response

We expand the probability density and current to linear order in ε:

$$\begin{aligned} P(x,t) &\approx P_0(x) + \varepsilon P_1(x,\omega) e^{i\omega t}, \\ J(x,t) &\approx J_0(x) + \varepsilon J_1(x,\omega) e^{i\omega t}. \end{aligned} \tag{54}$$

Here, $P_1(x,\omega)$ and $J_1(x,\omega)$ denote the complex linear response of the probability density and current, respectively. $P_1(x,\omega)$ is also subject to absorbing boundary conditions at both thresholds. The parameter modulation causes modulations of the decision rates $\tilde{r}_c(\omega)$ and $\tilde{r}_i(\omega)$:

$$r_c(t) = r_{c,0} + \varepsilon \tilde{r}_c(\omega) e^{i\omega t}, \tag{55}$$

$$r_i(t) = r_{i,0} + \varepsilon \tilde{r}_i(\omega) e^{i\omega t}, \tag{56}$$

from which we can extract the amplitudes ($|\tilde{r}_A(\omega)|$) and phases ($\Phi(\tilde{r}_c) = \arctan(\mathrm{Im}(\tilde{r}_c)/\mathrm{Re}(\tilde{r}_c))$. In general, the linear response terms for probability and current obey the ordinary differential equations:

$$\begin{aligned} \partial_x J_1 =& i\omega P_1 + (\tilde{r}_c + \tilde{r}_i) e^{i\omega\Delta} \delta(x - x_s) \\ & - \tilde{r}_c \delta(x - x_c) - \tilde{r}_i \delta(x - x_i), \\ \partial_x P_1 =& -\frac{1}{\sigma^2} \left(\tau_x J_1 - f P_1 - \partial_\alpha (f - \sigma^2 \partial_x) \Big|_{\alpha=\alpha_0} P_0 \right). \end{aligned} \tag{57}$$

To determine the solution, we split probability flux and density into three parts

$$\begin{aligned} J_1 &= \tilde{r}_c \tilde{j}_c + \tilde{r}_i \tilde{j}_i + \tilde{j}_\varepsilon, \\ P_1 &= \tilde{r}_c \tilde{p}_c + \tilde{r}_i \tilde{p}_i + \tilde{p}_\varepsilon. \end{aligned} \tag{58}$$

Here the densities $\tilde{p}_c$ and $\tilde{p}_i$ and the corresponding currents are only nonzero in the interval $[x_i, x_s]$ and $(x_s, x_c]$, respectively. These parts take into account the reinsertion of probability, such that they are given by the ordinary differential equations:

$$\begin{aligned} \partial_x \tilde{j}_c &= i\omega \tilde{p}_c - \delta(x - x_c), \\ \partial_x \tilde{p}_c &= -\frac{1}{\sigma^2} (\tau_x \tilde{j}_c - f(x) \tilde{p}_c), \\ \partial_x \tilde{j}_i &= i\omega \tilde{p}_i - \delta(x - x_i), \\ \partial_x \tilde{p}_i &= -\frac{1}{\sigma^2} (\tau_x \tilde{j}_i - f(x) \tilde{p}_i), \end{aligned} \tag{59}$$

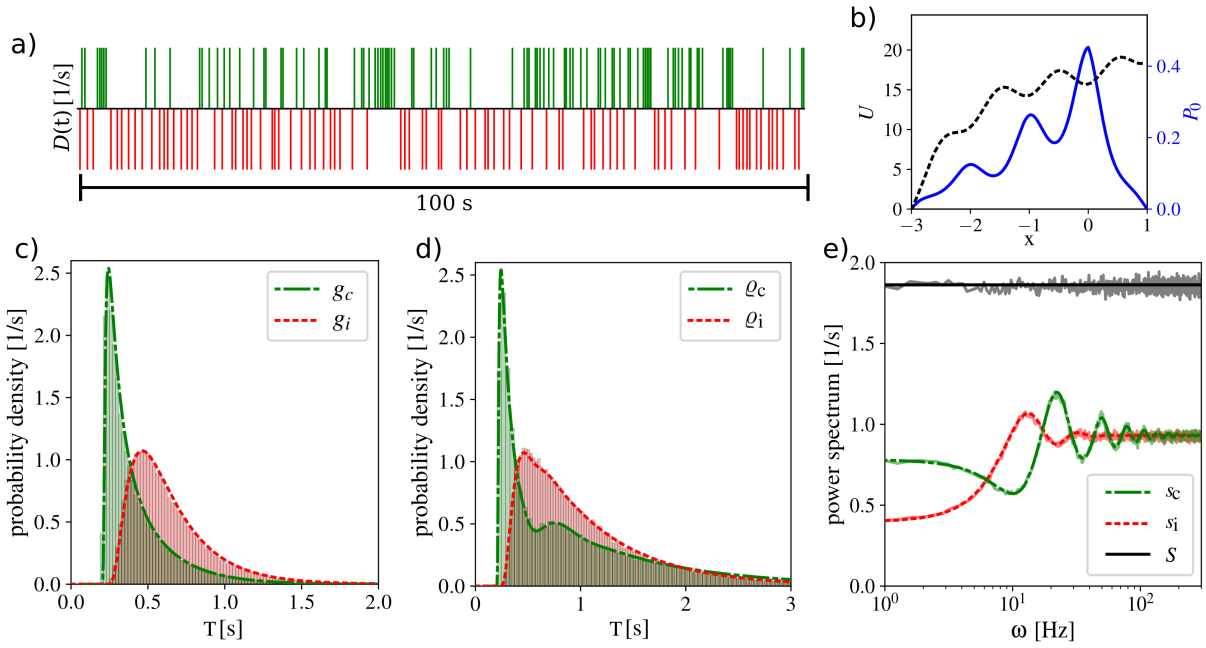

Figure 7: **Temporal statistics of nonlinear DDM with equal decision rates and asymmetric potential.** See caption in Fig. 3 for detailed description and Tab. 1 for parameters.

for which we again know the values at the thresholds as

$$-\tilde{\jmath}_i(x_i,\omega) = \tilde{\jmath}_c(x_c,\omega) = 1 \tag{60}$$

and

$$\tilde{p}_c(x_c,\omega) = \tilde{p}_i(x_i,\omega) = 0. \tag{61}$$

Numerical integration from the respective threshold to the reset provides the values $\tilde{p}_{i,s} = \tilde{p}_i(x_s,\omega)$, $\tilde{p}_{c,s} = \tilde{p}_c(x_s,\omega)$, $\tilde{\jmath}_{i,s} = \tilde{\jmath}_i(x_s,\omega)$ and $\tilde{\jmath}_{c,s} = \tilde{\jmath}_c(x_s,\omega)$ which will be important in the following calculation. The third terms in probability density and current in Eq. (58) take into account the linear response term that includes the derivative in α and the stationary density. Thus, $\tilde{\jmath}_\varepsilon$ and $\tilde{p}_\varepsilon$ obey the ordinary differential equations:

$$\begin{aligned}\partial_x \tilde{\jmath}_\varepsilon =& i\omega\tilde{p}_\varepsilon + \lambda_{\jmath,\varepsilon}\delta(x-x_s),\\ \partial_x \tilde{p}_\varepsilon =& -\frac{1}{\sigma^2}\left(\tau_x\tilde{\jmath}_\varepsilon - f\tilde{p}_\varepsilon - \partial_\alpha(f-\sigma^2\partial_x)\Big|_{\alpha=\alpha_0} P_0\right)\\ &+ \lambda_{p,\varepsilon}\delta(x-x_s).\end{aligned} \tag{62}$$

Because the particular part $\tilde{p}_\varepsilon$ of P_1 does not generate any further efflux of probability, we know that $\tilde{\jmath}_\varepsilon(x_i,\omega) = \tilde{\jmath}_\varepsilon(x_c,\omega) = \tilde{p}_\varepsilon(x_i,\omega) = \tilde{p}_\varepsilon(x_c,\omega) = 0$. In order to calculate the jumps in probability density $\lambda_{p,\varepsilon}$ and current $\lambda_{\jmath,\varepsilon}$, we may integrate Eq. (62) numerically from the thresholds to the reset i.e.,

$$\begin{aligned}\lambda_{p,\varepsilon} &= \tilde{p}_\varepsilon(x\to x_s^+,\omega) - \tilde{p}_\varepsilon(x\to x_s^-,\omega),\\ \lambda_{\jmath,\varepsilon} &= \tilde{\jmath}_\varepsilon(x\to x_s^+,\omega) - \tilde{\jmath}_\varepsilon(x\to x_s^-,\omega),\end{aligned} \tag{63}$$

where the index $+$ $(-)$ indicates the limit taken from above (below). Finally, we exploit the continuity of the probability density P_1 and the jump condition of J_1 that read:

$$\begin{aligned}\tilde{r}_c\tilde{p}_{c,s} - \tilde{r}_i\tilde{p}_{i,s} + \lambda_{p,\varepsilon} &= 0,\\ \tilde{r}_c\tilde{\jmath}_{c,s} - \tilde{r}_i\tilde{\jmath}_{i,s} + \lambda_{\jmath,\varepsilon} &= (\tilde{r}_i+\tilde{r}_c)e^{i\omega\Delta},\end{aligned} \tag{64}$$

to determine the rate modulations:

$$\begin{aligned}\tilde{r}_i &= \frac{\lambda_{p,\varepsilon}(\tilde{\jmath}_{c,s} - e^{i\omega\Delta}) - \lambda_{\jmath,\varepsilon}\tilde{p}_{c,s}}{\tilde{p}_{i,s}(\tilde{\jmath}_{c,s} - e^{i\omega\Delta}) - \tilde{p}_{c,s}(\tilde{\jmath}_{i,s} + e^{i\omega\Delta})},\\ \tilde{r}_c &= \frac{\tilde{r}_i\tilde{p}_{i,s} - \lambda_{p,\varepsilon}}{\tilde{p}_c(x_s)}.\end{aligned} \tag{65}$$

As an example, we calculate the linear response of a sinusoidal modulation of the constant input with the polynomial potential used in Fig. 5 for which the corresponding Langevin equation reads:

$$\tau_x\dot{x} = f(x) + \varepsilon\cos(\omega t) + \sigma\sqrt{2\tau_x}\xi(t), \tag{66}$$

with $\varepsilon = 0.1$ and vanishing initial phase of the signal. All other parameters are chosen as in the third example (see Tab. 1). With this simple additive signal, the second equation in Eq. (57) attains the form:

$$\partial_x P_1 = -\frac{1}{\sigma^2}(\tau_x J_1 - f(x)P_1 - P_0). \tag{67}$$

5.2. Example for the linear response of a nonlinear DDM

The results of the linear response calculation for a specific case (the polynomial potential studied in Fig. 5) are presented in Fig. 8 and compared to direct simulations of Eq. (66). The modulations determined with the threshold-integration method for $\omega = 13.54\,\mathrm{Hz}$ (red and green

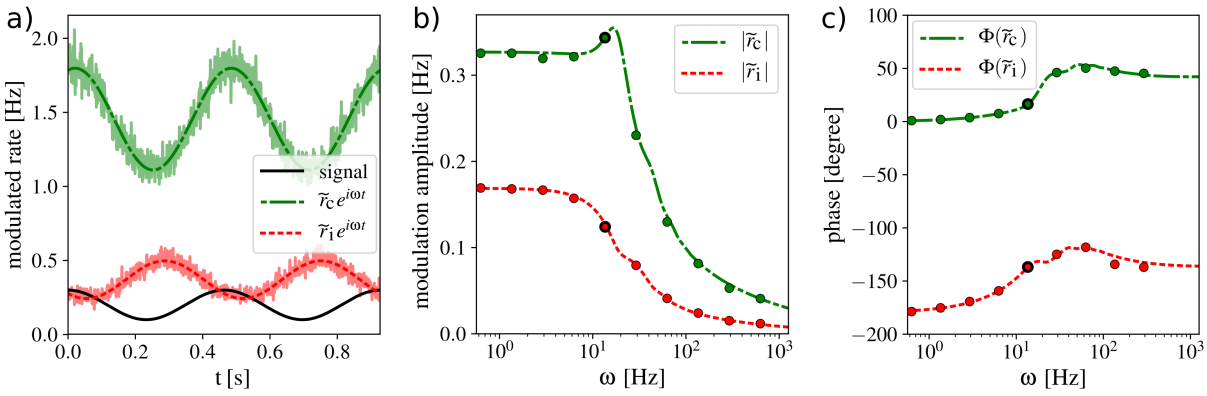

Figure 8: **Linear response to sinusoidal modulation of input.** a: Modulations of the decision rates $r_i(t)$ and $r_c(t)$ (red and green) from the threshold-integration method (dashed and dot-dashed solid) and direct simulations (jagged lines) and modulated signal (black) with $\varepsilon = 0.1$ and $\omega = 13.54\,\mathrm{Hz}$. b and c: amplitude and phase of rate modulations from threshold-integration method (dashed and dot-dashed lines) and direct simulations (circles) where emphasized circles belong to panel a. See Tab. 1 for all parameters.

solid lines in Fig. 8 a) match the direct simulations (jagged lines). Also for other frequencies, the introduced method is capable to adequately predict the amplitude and the phase of the rate modulations as shown in Fig. 8 b and c, respectively. For increasing frequency, the amplitude of both modulations decays, but for the correct decisions, it first reaches its maximum at 16 Hz while the decay is monotonic for the incorrect decisions. Note, that the small resonant peak in the correct modulation is found at the same frequency where also the main peak in the power spectrum (Fig. 5 e) is observed. Generally, the absolute modulation amplitude is larger for the decision that is preferred in the unperturbed state. However, remarkably, the *relative* modulation amplitude in the specific case considered is larger for the unpreferred decision.

The modulation of incorrect decisions is characterized by a high negative phase for all frequencies. It is simple to understand, what the limit of the phase shift is for arbitrarily low frequencies. Here a slow positive signal ($\varepsilon \cos(\omega t) > 0$) pushes the decision variable towards the correct decision whereas it does the opposite and enhances escapes via the boundary for incorrect decisions during its negative half cycle ($\varepsilon \cos(\omega t) < 0$).

6. Discussion and outlook

In this paper we have introduced the decision trains in which sequences of consecutive binary decisions are represented as the sum of delta-spikes at the decision times. We derived the connections between their basic statistics that are the interdecision interval densities and the power spectra under the assumption of uncorrelated decisions and stationarity. Surprisingly, we find that, in case of equal stationary decision rates, the power spectrum of the combined decision train is always constant, i.e. it is a white noise in the sense of a vanishing second-order correlation. This result is independent of the spectra of the individual decision trains. For the evaluation of related experiments, taking into account not only the decision time densities, as it is the case in most studies, but also the power spectrum might reveal novel features of the underlying dynamics of binary decision making, since the spectrum includes correlations between decision times that are neglected otherwise. Thus, specifically Eqs. (8) and (11) can be used as a test of the renewal property for the decision times in a sequence of consecutive decisions. If a nonrenewal behavior is observed, e.g. due to the subject's loss of attention, the representation of the binary decisions by the decision trains may turn out to be a useful characteristics of transient changes (see below for some thoughts on how this could be modelled).

We also introduced a general nonlinear DDM for which the subthreshold nonlinearity can be for instance chosen to approximate binary decision making in neural networks [64] or changing environments during decision making [78]. For such models, no analytical solution is known. Here we generalized the threshold-integration method introduced by Richardson for nonlinear integrate-and-fire neurons in [59, 60] for the nonlinear DDM. If implemented on a standard personal computer or laptop, this simple method provides a rapid means of calculating the model's statistics within seconds. Hence, it is a proper tool to fit experimental data to different model systems. In particular, theoretical predictions such as [64, 78] leading to distinct subthreshold nonlinearities can be tested against experimental data. The introduced procedure might be a first step to tackle the more general inverse problem, i.e. to constrain nonlinear models by the IDI statistics and the here introduced power spectra of the decision trains.

Besides the model's stationary statistics (corresponding to the standard setup of a constant signal plus a distracting noise), the threshold-integration method can be used to compute the linear response of the decision rates, as we have shown in the last part of the results section. The linear response approach could be instrumental to understand experiments in which a prescribed modulation of

evidence results in modulations of decisions in a long sequence of subsequent trials. Although we are not aware of such experiments, we think they are feasible and would provide an excellent means to further test nonlinear decision models. We are encouraged to think so because it has been shown for the related class of spiking neuron models that their spontaneous activity can be very similar irrespective of the subthreshold nonlinearity, whereas their response properties drastically differ [80]. Put differently, neuron models that can hardly be distinguished when looking at their spontaneously generated spike trains reveal their differences in their response to and correlation with a prescribed time-dependent stimulus. The same can be expected for the decision-making model studied here.

Although the model considered in this paper is certainly more general than the standard DDM, we still made a number of drastically simplifying assumptions. We assumed for instance that the decision process is symmetric for both stimuli (e.g. coherent motion towards the left or right direction), i.e. if we only ask for the statistics of correct and incorrect decisions, the DDM does not depend on the specific stimulus of a certain trial. This symmetry may not be given and we may extend our analysis for a model in which a stimulus dependent nonlinear potential (and equivalently noise intensity or other parameters) is used. This would mean that we have to resort to the more explicit stimulus-driven model Eq. (1). However, if the experimenter uses independent draws for the binary signal, we still expect renewal statistics for the decision trains and simple relations between the power spectrum of the decision train and the corresponding interval densities.

We also assumed that the intrinsic fluctuations of the decision process can be represented by uncorrelated Gaussian input noise. Real fluctuations, however, either distraction noise (stimulus components not belonging to the signal) or intrinsic noise, e.g. originating in the involved neural networks [2, 5, 15, 45], are always temporally correlated. One way to incorporate temporal correlations is a Markovian embedding which was developed first in statistical physics, see e.g. [42, 24, 70], but has also been frequently used in computational neuroscience for correlated Gaussian [6, 39, 40, 65] and non-Gaussian noise [14, 43]. We have recently applied this method for integrate-and-fire neurons subject to an arbitrarily correlated Gaussian noise [79]. A similar generalization for the decision model seems to be feasible. More difficult is the question what kind of colored noise should be used. The application of the decision-train statistics developed here will be helpful to detect indirectly the effects of correlated fluctuations in the form of nonrenewal features in the decision trains as discussed above.

The most severe assumption we have made concerns the renewal character of our model. We have completely neglected that experimentally measured response times often depend on the used sequence of stimuli (see for instance [36, 9, 29]). For a simple two-alternative forced choice task, repetitions of the stimulus seem often to come along with a shortening of the mean response time, an effect known as facilitation.

Our continuous model for the accumulation variable could be extended as follows to capture this first-order effect:

$$\tau_y \dot{y} = f_s(y) + a + \sigma\sqrt{2\tau_y}\xi(t), \tag{68}$$

$$\tau_a \dot{a} = -a, \tag{69}$$

$$\text{if at time } t \text{ decision to the right } \; a(t) \to a(t) + \varepsilon,$$

$$\text{if at time } t \text{ decision to the left } \; a(t) \to a(t) - \varepsilon.$$

The variable a implements a positive feedback: if (for a stimulus going to the right) a decision to the right has been made, a increases by ε and will furthermore accelerate the decision if the next stimulus goes to the right as well. A stimulus repetition leads thus to a reduction of the mean response time as observed in the experiment. In between decision times, a decays exponentially (according to Eq. (69)), hence previous decisions are forgotten after some time. Even if the stimuli are completely randomly drawn (as a Bernoulli process), the response train is not a renewal process anymore because the dynamics for the accumulation variable keeps a memory about previous decisions by virtue of the variable $a(t)$. We note that alternatively to modifying the drift, the variable a could modify the reset point with a similar facilitation effect and the additional feature of a variable reset point somewhat similar but not identical to other generalizations of the DDM [51, 55, 54].

We do not state this extension of our model here because it is most realistic or capable of capturing all the stimulus-induced facilitation and expectation effects that are observed in the various experiments; indeed, more detailed models have been successfully fit to experimental data (see, for instance, [9, 19, 29]). Rather we would like to use the extension of our model to point out a further analogy with stochastic models of neural spike trains. Spike-frequency adaptation has been studied by different authors in a very similar way as sketched above with one important difference: in neuron models $a(t)$ would represent the strength of a slow inhibitory current, which acts as a *negative* feedback (achieved by changing $+a$ to $-a$ in Eq. (68)). One feature observed in adapting neurons [56, 7, 17] as well as in neuronal models [34, 7, 67] are correlations among the interspike intervals illustrating that the associated point process is not a renewal process. Analytical approaches to calculate ISI correlations for adapting neurons [66, 76, 68, 67, 69, 77] could inspire similar approaches to the correlations of response times in DDMs. Moreover, the Fokker-Planck equation that we treated here specifically with the threshold-integration method can also be used for multi-dimensional setups, for instance, with an adaptation (or facilitation) variable [32, 1, 79]. In particular, the methods of Ref. [79] seem to be extendable to explore the effect of facilitation-induced sequential memory on the decision-train power spectrum.

The study of such generalized models endowed with colored-noise and facilitation variables and their compari-

son to experimental data present exciting topics for future investigations.

7. Acknowlegdegements

This paper was developed within the scope of the IRTG 1740 / TRP 2015/50122-0, funded by the DFG / FAPESP. We acknowledge inspiring discussions with Magnus Richardson (University of Warwick, UK).

Conflict of interest

The authors declare that they have no conflict of interest.

Appendix A. Signal-dependent thresholds in neurophysically motivated models

Nonlinear DDMs can be motivated neurophysically as the reduction of the mean-field approximation of two competing neural networks to a single variable (see for instance [82, 64]).

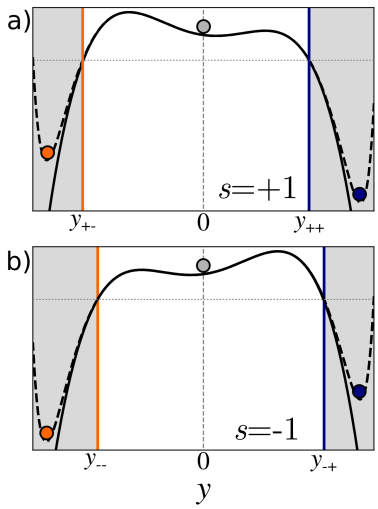

Figure A.9: **Signal dependence of threshold points.** The subthreshold dynamics of the decision variable corresponds to the overdamped motion in a quartic potential (solid lines), signal-biased to the right (a) or left (b) [64]. Attractor states (dash-lined minima, corresponding to left or right decisions) are not part of the potential. Within the DDM model, the decision is made if a threshold, i.e. a point of no return on the way into one of the attractors, is reached. These points are defined by a certain potential depth (dotted line) and depend on the signal [blue, orange lines shift between (a),(b)].

The neurons in each population are coupled such that each population amplifies its own activity and inhibits the other one. The asymptotic behavior of the system is bistable; it is characterized by two stable fixed points in each one of which one population is firing at a high rate and the other one is silent. Decision making is in this picture equivalent to the system reaching one of these fixed points. By the application of mean-field techniques, the dynamics can be reduced to a two-dimensional rate model. Considering only the difference of the rates, the subthreshold dynamics can further be simplified to the one-dimensional dynamics given by the nonlinear DDM in Eq. (1). The final attractor states are not part of the DDM description; instead threshold points are defined at which it becomes almost certain that the dynamics will reach (in a short time) the respective fixed point.

In Fig. A.9 we illustrate the emergence of the threshold points in the DDM model for a specific example and explain why there is a dependence of these *points of no return* on the signal. Taking the example from Ref. [64] that corresponds to a quartic potential resulting from a Taylor expansion of the dynamics around the initial point, we see that the true attractor states of the network dynamics (indicated by dash-lined minima) are not included. In our simplified description, we associate decisions with reaching the thresholds $y_{\pm,\pm}$ that can be defined differently. A reasonable definition is that the threshold points are attained where the quartic potential reaches a certain depth (dotted lines in Fig. A.9). The so-defined threshold points will depend on the tilt of the potential, i.e. on the signal. Put differently, in general $y_{++} \neq y_{-+}$ and $y_{+-} \neq y_{--}$ as can also be seen by comparing their positions (blue and orange lines) in Fig. A.9a and b.

Appendix B. Relation between response-time densities and inter-decision-interval densities

Here we show the detailed derivation of the connection between $g_c(T)$ and $\varrho_c(T)$ which also holds true for incorrect decisions if we exchange $c \leftrightarrow i$. Given that a correct decision occurred at $T = 0$, the probability density for the next correct or incorrect decision is determined by $g_c(T)$ or $g_i(T)$, respectively. If an incorrect decision occurs at T', we have to take into account the reset such that the next decision is either correct with the probability $g_c(T - T')$ or incorrect with $g_i(T - T')$, such that the probability to obtain n incorrect decisions followed by an correct decision is given by an n-fold convolution of $g_c(T)$ with $g_i(T)$ [11]. Taking into account that the number of incorrect decisions in between two sequential correct decisions can be arbitrarily high we obtain:

$$\begin{aligned}\varrho_c(T) &= g_c(T) + \int_0^T dT' g_c(T-T') g_i(T') \\ &\quad + \int_0^T dT' \int_0^{T'} dT'' g_c(T-T') g_i(T'-T'') g_i(T'') \\ &\quad + \ldots \\ &= g_c * \left(\delta(T) + \sum_{n=1}^{\infty} \underbrace{g_i * \ldots * g_i}_{n-\text{terms}} \right),\end{aligned} \tag{B.1}$$

where the asterisk denotes convolution. The Fourier transform of the expression yields Eq. (8):

$$\tilde{\varrho}_c(\omega) = \tilde{g}_c(\omega)\left(\sum_{n=0}^{\infty}[\tilde{g}_i(\omega)]^n\right) = \frac{\tilde{g}_c(\omega)}{1-\tilde{g}_i(\omega)}. \tag{B.2}$$

Appendix C. Power spectra of combined decision train

In order to derive the power spectrum of the combined decision train, we use the autocorrelation function of the decision train for correct decisions that is given by:

$$c_c(\tau) = \langle d_c(t)d_c(t+\tau)\rangle - \langle d_c(t)\rangle^2 \tag{C.1}$$
$$= r_{c,0}(\delta(\tau) + m_c(\tau) - r_{c,0}). \tag{C.2}$$

The conditional decision rate $m_c(\tau)$, the probability density for a decision at τ under the condition of a reference decision at $\tau = 0$, is the essential part of the autocorrelation function [27, 18]. This probability density is determined by the summed (distinct) probability to have one, two, and so on interdecision intervals T between the reference time t and the time $t+\tau$. Since all intervals are independent, these probabilities are simple convolutions of the density ϱ_c and one obtains

$$m_c(\tau) = \sum_{n=1}^{\infty}\underbrace{\varrho_c * ... * \varrho_c}_{n-\text{terms}}(\tau) \tag{C.3}$$

$$\tilde{m}_c(\omega) = \frac{\tilde{\varrho}_c(\omega)}{1-\tilde{\varrho}_c(\omega)}. \tag{C.4}$$

The autocorrelation function for the combined decision train is given by:

$$\begin{aligned} C(\tau) =& \langle D(t)D(t+\tau)\rangle - \langle D(t)\rangle^2 \\ =& c_i(\tau) + c_c(\tau) + 2r_{i,0}r_{c,0} \\ &+ \langle d_i(t)d_c(t+\tau)\rangle + \langle d_c(t)d_i(t+\tau)\rangle \\ =& c_i(\tau) + c_c(\tau) + 2r_{i,0}r_{c,0} - r_{i,0}m_c(\tau) - r_{c,0}m_i(\tau) \\ =& c_i(\tau)(1-\frac{r_{c,0}}{r_{i,0}}) + c_c(\tau)(1-\frac{r_{i,0}}{r_{c,0}}) + R_0\delta(\tau). \end{aligned} \tag{C.5}$$

Since, according to the Wiener-Khinchin theorem [20], the power spectrum is given by the Fourier transform of the autocorrelation function, we obtain:

$$S(\omega) = s_c(\omega)\left(1-\frac{r_{i,0}}{r_{c,0}}\right) + s_i(\omega)\left(1-\frac{r_{c,0}}{r_{i,0}}\right) + R_0. \tag{C.6}$$

corresponding to Eq. (12) in the main text.

Appendix D. Wiener process

The simplest version of the DDM, the Wiener process in Eq. (13) with the reset condition in Eq. (4), is analytically tractable. Here the derivation of the analytical formulas for the stationary rates and the Fourier-transformed response-time densities are presented in our notation. The stationary probability density obeys the stationary Fokker-Planck equation and the normalization condition

$$\begin{aligned} &\left[-\frac{\mu}{\tau_x}\partial_x + \frac{\sigma^2}{\tau_x}\partial_x^2\right]P_0(x) + (r_{c,0}+r_{i,0})\delta(x) = 0, \\ &\int_{x_i}^{x_c} P_0\,dx = 1 - (r_{c,0}+r_{i,0})\Delta. \end{aligned} \tag{D.1}$$

The solution reads:

$$\begin{aligned} P_0(x) &= \theta^{-1}\begin{cases} e_c\left[\exp\left(\frac{\mu}{\sigma^2}[x-x_i]\right)-1\right] & \text{for } x \le 0, \\ e_i\left[\exp\left(\frac{\mu}{\sigma^2}[x-x_c]\right)-1\right] & \text{for } x > 0 \end{cases}, \\ \theta &= x_i e_c - x_c e_i - \frac{\mu}{\tau_x}\Delta(e_i - e_c), \\ e_c &= 1-\exp\left(-\frac{\mu}{\sigma^2}x_c\right),\ e_i = 1-\exp\left(-\frac{\mu}{\sigma^2}x_i\right). \end{aligned} \tag{D.2}$$

The time-dependent Fokker-Planck equation that corresponds to Eq. (13) is given by:

$$\partial_t P(x,t) = \left[-\frac{\mu}{\tau_x}\partial_x + \frac{\sigma^2}{\tau_x}\partial_x^2\right]P(x,t) + \delta(x)\delta(t-\Delta), \tag{D.3}$$

with absorbing boundary conditions at the thresholds and the continuity condition at the reset point

$$P(x_i,t) = P(x_c,t) = 0,\ \lim_{\varepsilon\to 0} P(x_s+\varepsilon,t) - P(x_s-\varepsilon,t) = 0. \tag{D.4}$$

The last term in Eq. (D.3) represents the insertion of all probability at reset $x_s = 0$ after the non-decision time has elapsed.

We next consider the Fourier transformation of Eq. (D.3) for $\tilde{P}(x,\omega)$ and change to a new function $q(x,\omega)$ by

$$\tilde{P}(x,\omega) = \exp\left(\frac{\mu x}{2\sigma^2}\right)q(x,\omega), \tag{D.5}$$

which leads us to the following ordinary differential equation for $q(x,\omega)$:

$$\frac{\sigma^2}{\tau_x}\frac{d^2}{dx^2}q(x,\omega) + \left(\frac{\mu^2}{4\sigma^2\tau_x} + i\omega\right)q(x,\omega) = -e^{i\omega\Delta}\delta(x) \tag{D.6}$$

(for the general method, applied to an integrate-and-fire model, see [33], ch. 2.4). Homogeneous solutions of Eq. (D.6) (setting the right hand side to zero) are given by

$$q_1(x,\omega) = \exp[\kappa(\omega)x],\ \ q_2(x,\omega) = \exp[-\kappa(\omega)x] \tag{D.7}$$

with $\kappa(\omega)$ in Eq. (16). The inhomogenous solution $q(x,\omega)$ can be constructed by the linear combination

$$q(x,\omega) = \begin{cases} a_1q_1 + a_2q_2 \text{ for } x_i \le x \le x_s \\ b_1q_1 + b_2q_2 \text{ for } x_s \le x \le x_c \end{cases} \tag{D.8}$$

The coefficients a_1, a_2, b_1 and b_2 are determined by conditions at the boundaries and the reset point: at x_i and

x_c the density must vanish (absorbing boundaries), at x_s the density must not suffer a jump (continuity of the density in the entire interval) but the probability current does, reflecting the insertion of probability after a decision has been made and the non-decision time has passed. These conditions can be summarized as follows

$$q(x_i,\omega) = q(x_c,\omega) = 0, \quad \lim_{\varepsilon\to 0} q(x_s+\varepsilon,\omega) - q(x_s-\varepsilon,\omega) = 0$$
$$\lim_{\varepsilon\to 0} \frac{d}{dx}\, q(x,\omega)|_{x_s-\varepsilon} - \frac{d}{dx}\, q(x,\omega)|_{x_s+\varepsilon} = \frac{\exp(i\omega\Delta)\tau_x}{\sigma^2}. \tag{D.9}$$

Using these equations to calculate the coefficients a_1, a_2, b_1, b_2, we know $q(x,\omega)$ and can determine the Fourier transforms of the response-time densities $\tilde{g}_i$ and $\tilde{g}_c$ as the Fourier transforms of the probability currents at the thresholds, i.e. by

$$\begin{aligned}\tilde{g}_i(\omega) &= \frac{\sigma^2}{\tau_x}\exp\left(\frac{\mu x_i}{2\sigma^2}\right)\frac{d}{dx}q(x,\omega)\Big|_{x=x_i},\\ \tilde{g}_c(\omega) &= -\frac{\sigma^2}{\tau_x}\exp\left(\frac{\mu x_c}{2\sigma^2}\right)\frac{d}{dx}q(x,\omega)\Big|_{x=x_c},\end{aligned} \tag{D.10}$$

which lead to the expression for the IDI density, Eq. (16) in the main text.

[1] Augustin, M., Ladenbauer, J., Baumann, F., Obermayer, K., 2017. Low-dimensional spike rate models derived from networks of adaptive integrate-and-fire neurons: Comparison and implementation. PLoS Comput. Biol. 13, e1005545.

[2] Bair, W., Koch, C., Newsome, W., Britten, K., 1994. Power spectrum analysis of bursting cells in area MT in the behaving monkey. Journal of Neuroscience 14, 2870.

[3] Bogacz, R., 2007. Optimal decision-making theories: linking neurobiology with behaviour. Trends in Cognitive Sciences 11 (3), 118–125.

[4] Bogacz, R., Brown, E., Moehlis, J., Holmes, P., Cohen, J. D., 2006. The physics of optimal decision making: a formal analysis of models of performance in two-alternative forced-choice tasks. Psychological review 113 (4), 700.

[5] Brunel, N., 2000. Dynamics of sparsely connected networks of excitatory and inhibitory spiking neurons. Journal of Computational Neuroscience 8, 183.

[6] Brunel, N., Sergi, S., 1998. Firing frequency of leaky integrate-and-fire neurons with synaptic current dynamics. Journal of Theoretical Biology 195, 87.

[7] Chacron, M. J., Longtin, A., Maler, L., 2001. Negative interspike interval correlations increase the neuronal capacity for encoding time-dependent stimuli. J. Neurosci. 21, 5328.

[8] Chittka, L., Skorupski, P., Raine, N., 2009. Speed–accuracy tradeoffs in animal decision making. Trends in Ecology & Evolution 24 (7), 400–407.

[9] Cho, R. Y., Nystrom, L. E., Brown, E. T., Jones, A. D., Braver, T. S., Holmes, P. J., Cohen, J. D., 2002. Mechanisms underlying dependencies of performance on stimulus history in a two-alternative forced-choice task. Cognitive, Affective, & Behavioral Neuroscience 2 (4), 283–299.

[10] Cisek, P., Puskas, G. A., El-Murr, S., 2009. Decisions in changing conditions: the urgency-gating model. Journal of Neuroscience 29 (37), 11560–11571.

[11] Cox, D. R., 1962. Renewal Theory. Methuen, London.

[12] Cox, D. R., Miller, H. D., 1965. The Theory of Stochastic Processes. Methuen & Co, London.

[13] Darling, D. A., Siegert, A. J. F., 1953. The 1st passage problem for a continuous markov process. Ann. Math. Stat. 24, 624.

[14] Droste, F., Lindner, B., 2014. Integrate-and-fire neurons driven by asymmetric dichotomous noise. Biological Cybernetics 108, 825.

[15] Dummer, B., Wieland, S., Lindner, B., 2014. Self-consistent determination of the spike-train power spectrum in a neural network with sparse connectivity. Frontiers in Computational Neuroscience 8, 104.

[16] Dutilh, G., Annis, J., Brown, S. D., Cassey, P., Evans, N. J., Grasman, R. P. P. P., Hawkins, G. E., Heathcote, A., Holmes, W. R., Krypotos, A.-M., Kupitz, C. N., Leite, F. P., Lin, V. L. . Y. S., Logan, G. D., Palmeri, T. J., Starns1, J. J., Trueblood, J. S., van Maanen, L., van Ravenzwaaij, D., Vandekerckhove, J., Visser, I., Voss, A., White, C. N., Wiecki, T. V., Rieskamp, J., Donkin, C., 2019. The quality of response time data inference: A blinded, collaborative assessment of the validity of cognitive models. Psychonomic Bulletin & Review 26, 1051.

[17] Farkhooi, F., Strube-Bloss, M. F., Nawrot, M. P., 2009. Serial correlation in neural spike trains: Experimental evidence, stochastic modeling, and single neuron variability. Phys. Rev. E. 79, 021905.

[18] Gabbiani, F., Koch, C., 1998. Principles of spike train analysis. In: Methods in Neuronal Modeling: From Synapses to Networks. Cambridge, MA: MIT Press, p. 313.

[19] Gao, J., Wong-Lin, K., Holmes, P., Simen, U., J. D. Cohen, J. D., 2009. Sequential effects in two-choice reaction time tasks: decomposition and synthesis of mechanisms. Neural Computation 21 (9), 2407–2436.

[20] Gardiner, C. W., 1985. Handbook of Stochastic Methods. Springer-Verlag, Berlin.

[21] Gerstner, W., Kistler, W. M., Naud, R., Paninski, L., 2014. Neuronal Dynamics: From single neurons to networks and models of cognition. Cambridge University Press, Cambridge.

[22] Gold, J., Shadlen, M., 2007. The neural basis of decision making. Annual Review of Neuroscience 30.

[23] Green, D. M., Smith, A. F., Gierke, S. M., 1983. Choice reaction time with a random foreperiod. Perception & Psychophysics 34 (3), 195–208.

[24] Guardia, E., Marchesoni, F., Miguel, M. S., 1984. Escape times in systems with memory effects. Physics Letters A 100, 15.

[25] Hawkins, G. E., Forstmann, B. U., Wagenmakers, E. J., Ratcliff, R., Brown, S. D., 2015. Revisiting the evidence for collapsing boundaries and urgency signals in perceptual decision-making. Journal of Neuroscience 35 (6), 2476–2484.

[26] Heekeren, H., Marrett, S., Ungerleider, L., 2008. The neural systems that mediate human perceptual decision making. Nature Reviews Neuroscience 9 (6), 467.

[27] Holden, A. V., 1976. Models of the Stochastic Activity of Neurones. Springer-Verlag, Berlin.

[28] Johanson, D., Cesario, J., Pleskac, T., 2018. How prior information and police experience impact decisions to shoot. Journal of Personality and Social Psychology 115 (4), 601.

[29] Jones, M., Curran, T., Mozer, M. C., Wilder, M. H., 2013. Sequential effects in response time reveal learning mechanisms and event representations. Psychological review 120 (3), 628.

[30] Kampen, N. G. V., 1993. Short first-passage times. J. Stat. Phys. 70, 15.

[31] Krajbich, I., Rangel, A., 2011. Multialternative drift-diffusion model predicts the relationship between visual fixations and choice in value-based decisions. Proceedings of the National Academy of Sciences 108 (33), 13852–13857.

[32] Ladenbauer, J., Augustin, M., Obermayer, K., 2014. How adaptation currents change threshold, gain, and variability of neuronal spiking. J. Neurophysiol. 111, 939.

[33] Lindner, B., 2002. Coherence and Stochastic Resonance in Nonlinear Dynamical Systems. Logos-Verlag, Berlin.

[34] Liu, Y. H., Wang, X. J., 2001. Spike-frequency adaptation of a generalized leaky integrate-and-fire model neuron. J. Comput. Neurosci. 10, 25.

[35] Lo, C. C., Wang, X. J., 2006. Cortico–basal ganglia circuit mechanism for a decision threshold in reaction time tasks. Nature Neuroscience 9 (7), 956.

[36] Luce, R. D., et al., 1986. Response times: Their role in inferring elementary mental organization. Oxford University Press on Demand.

[37] McKoon, G., Ratcliff, R., 2018. Adults with poor reading skills, older adults, and college students: The meanings they understand during reading using a diffusion model analysis. Journal of Memory and Language 102, 115–129.
[38] McMillen, T., Holmes, P., 2006. The dynamics of choice among multiple alternatives. Journal of Mathematical Psychology 50 (1), 30–57.
[39] Moreno, R., de la Rocha, J., Renart, A., Parga, N., 2002. Response of spiking neurons to correlated inputs. Physical Review Letters 89, 288101.
[40] Moreno-Bote, R., Parga, N., 2010. Response of integrate-and-fire neurons to noisy inputs filtered by synapses with arbitrary timescales: Firing rate and correlations. Neural Comput. 22, 1528.
[41] Morgan, M. J., Ward, R., 1980. Conditions for motion flow in dynamic visual noise. Vision research 20 (5), 431–435.
[42] Mori, H., 1965. A continued-fraction representation of time-correlation functions. Progress of Theoretical Physics 34, 399.
[43] Müller-Hansen, F., Droste, F., Lindner, B., 2015. Statistics of a neuron model driven by asymmetric colored noise. Physical Review E 91, 022718.
[44] Palmer, J., Huk, A. C., Shadlen, M. N., 2005. The effect of stimulus strength on the speed and accuracy of a perceptual decision. Journal of Vision 5 (5), 1–1.
[45] Pena, R. F. O., Vellmer, S., Bernardi, D., Roque, A. C., Lindner, B., 2018. Self-consistent scheme for spike-train power spectra in heterogeneous sparse networks. Frontiers in Computational Neuroscience 12 (9).
[46] Press, W. H., Teukolsky, S. A., Vetterling, W. T., Flannery, B. P., 2007. Numerical Recipes 3rd Edition: The Art of Scientific Computing. Cambridge University Press, Cambridge.
[47] Rao, R. P. N., 2010. Decision making under uncertainty: a neural model based on partially observable markov decision processes. Frontiers in Computational Neuroscience 4, 146.
[48] Ratcliff, R., 1978. A theory of memory retrieval. Psychological Review 85 (2), 59.
[49] Ratcliff, R., 2002. A diffusion model account of response time and accuracy in a brightness discrimination task: Fitting real data and failing to fit fake but plausible data. Psychonomic Bulletin & Review 9 (2), 278.
[50] Ratcliff, R., McKoon, G., 2008. The diffusion decision model: theory and data for two-choice decision tasks. Neural Computation 20 (4), 873–922.
[51] Ratcliff, R., Rouder, J. N., 1998. Modeling response times for two-choice decisions. Psychological Science 9 (5), 347–356.
[52] Ratcliff, R., Smith, P. L., 2004. A comparison of sequential sampling models for two-choice reaction time. Psychological Review 111 (2), 333.
[53] Ratcliff, R., Smith, P. L., Brown, S. D., McKoon, G., 2016. Diffusion decision model: current issues and history. Trends in Cognitive Sciences 20 (4), 260–281.
[54] Ratcliff, R., Tuerlinckx, F., 2002. Estimating parameters of the diffusion model: Approaches to dealing with contaminant reaction times and parameter variability. Psychonomic Bulletin & Review 9 (3), 438.
[55] Ratcliff, R., Zandt, T. V., McKoon, G., 1999. Connectionist and diffusion models of reaction time. Psychological Review 106 (2), 261.
[56] Ratnam, R., Nelson, M. E., 2000. Nonrenewal statistics of electrosensory afferent spike trains: Implications for the detection of weak sensory signals. J. Neurosci. 20, 6672.
[57] Redner, S., 2001. A Guide to First-Passage Processes. Cambridge University Press, Cambridge, UK.
[58] Ricciardi, L. M., Sacerdote, L., Sato, S., 1984. On an integral equation for first-passage-time probability densities. Journal of Applied Probability 21 (2), 302–314.
[59] Richardson, M. J. E., 2007. Firing-rate response of linear and nonlinear integrate-and-fire neurons to modulated current-based and conductance-based synaptic drive. Physical Review E 76, 021919.
[60] Richardson, M. J. E., 2008. Spike-train spectra and network response functions for non-linear integrate-and-fire neurons. Biological Cybernetics 99, 381.
[61] Risken, H., 1984. The Fokker-Planck Equation. Springer, Berlin.
[62] Roitman, J. D., Shadlen, M. N., 2002. Response of neurons in the lateral intraparietal area during a combined visual discrimination reaction time task. Journal of Neuroscience 22 (21), 9475–9489.
[63] Romo, R., Salinas, E., 2001. Touch and go: decision-making mechanisms in somatosensation. Annual Review of Neuroscience 24 (1), 107–137.
[64] Roxin, A., Ledberg, A., 2008. Neurobiological models of two-choice decision making can be reduced to a one-dimensional nonlinear diffusion equation. PLoS Computational Biology 4 (3), e1000046.
[65] Schwalger, T., Droste, F., Lindner, B., 2015. Statistical structure of neural spiking under non-poissonian or other non-white stimulation. Journal of Computational Neuroscience 39, 29.
[66] Schwalger, T., Lindner, B., 2010. Theory for serial correlations of interevent intervals. Eur. Phys. J. Spec. Topics 187, 211.
[67] Schwalger, T., Lindner, B., 2013. Patterns of interval correlations in neural oscillators with adaptation. Front. Comp. Neurosci. 7, 164.
[68] Schwalger, T., Tiana-Alsina, J., Torrent, M. C., Garcia-Ojalvo, J., Lindner, B., 2012. Interspike-interval correlations induced by two-state switching in an excitable system. Epl-Europhys. Lett. 99, 10004.
[69] Shiau, L., Schwalger, T., Lindner, B., 2015. Interspike interval correlation in a stochastic exponential integrate-and-fire model with subthreshold and spike-triggered adaptation. J. Comput. Neurosci. 38, 589.
[70] Siegle, P., Goychuk, I., Talkner, P., Hänggi, P., 2010. Markovian embedding of non-markovian superdiffusion. Physical Review E 81, 011136.
[71] Smith, P., 1990. A note on the distribution of response times for a random walk with gaussian increments. Journal of Mathematical Psychology 34 (4), 445–459.
[72] Smith, P. L., 2000. Stochastic dynamic models of response time and accuracy: A foundational primer. Journal of mathematical psychology 44 (3), 408–463.
[73] Smith, P. L., 2010. From poisson shot noise to the integrated ornstein–uhlenbeck process: Neurally principled models of information accumulation in decision-making and response time. Journal of Mathematical Psychology 54 (2), 266–283.
[74] Smith, P. L., 2015. The poisson shot noise model of visual short-term memory and choice response time: Normalized coding by neural population size. Journal of Mathematical Psychology 66, 41–52.
[75] Stratonovich, R. L., 1967. Topics in the Theory of Random Noise. Gordon and Breach, New York.
[76] Urdapilleta, E., 2011. Onset of negative interspike interval correlations in adapting neurons. Phys. Rev. E. 84, 041904.
[77] Urdapilleta, E., 2016. Noise-induced interspike interval correlations and spike train regularization in spike-triggered adapting neurons. Europhys. Lett. 115 (6), 68002.
[78] Veliz-Cuba, A., Kilpatrick, Z. P., Josic, K., 2016. Stochastic models of evidence accumulation in changing environments. SIAM Review 58 (2), 264–289.
[79] Vellmer, S., Lindner, B., 2019. Theory of spike-train power spectra for multidimensional integrate-and-fire models. Physical Review Research 1, 023024.
[80] Vilela, R. D., Lindner, B., 2009. A comparative study of three different integrate-and-fire neurons: spontaneous activity, dynamical response, and stimulus-induced correlation. Phys. Rev. E. 80, 031909.
[81] Wang, X. J., 2002. Probabilistic decision making by slow reverberation in cortical circuits. Neuron 36 (5), 955–968.
[82] Wong, K. F., Wang, X. J., 2006. A recurrent network mechanism of time integration in perceptual decisions. Journal of Neuroscience 26 (4), 1314–1328.